

**STATISTICAL INVESTIGATION
INTO
ACADEMIC PERFORMANCE
IN THE FACULTY OF SCIENCE
AT THE UNIVERSITY OF CAPE TOWN
IN THE PERIOD 1990-1997**

**Prepared by:
Katarzyna Ronda**

**Supervised by:
Prof. T. T. Dunne and Prof. C. G. Troskie**

Thesis prepared in fulfilment of the requirements for the degree of Master of Science
in the Department of Statistical Sciences at the University of Cape Town

September 1999

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

**STATISTICAL INVESTIGATION
INTO
ACADEMIC PERFORMANCE
IN THE FACULTY OF SCIENCE
AT THE UNIVERSITY OF CAPE TOWN
IN THE PERIOD 1990-1997**

**Prepared by:
Katarzyna Ronda**

**Supervised by:
Prof. T. T. Dunne and Prof. C. G. Troskie**

Thesis prepared in fulfilment of the requirements for the degree of Master of Science
in the Department of Statistical Sciences at the University of Cape Town

September 1999

**STATISTICAL INVESTIGATION
INTO
ACADEMIC PERFORMANCE
IN THE FACULTY OF SCIENCE
AT THE UNIVERSITY OF CAPE TOWN
IN THE PERIOD 1990-1997**

PART ONE

**Prepared by:
Katarzyna Ronda**

**Supervised by:
Prof. T. T. Dunne and Prof. C. G. Troskie**

Thesis prepared in fulfilment of the requirements for the degree of Master of Science
in the Department of Statistical Sciences at the University of Cape Town

September 1999

DECLARATION

The project discussed in this dissertation has been completed in the Department of Statistical Sciences at the University of Cape Town, during the period March 1998 to September 1999. It has been supervised by Prof. T. T. Dunne and co-supervised by Prof. C. G. Troskie.

This study represents the original work by the author except where the use of the work of others has been duly acknowledged in the text. The content of the thesis has not been submitted in any form to any other university.

Katarzyna Ronda

“ Non scholae, sed vitae discimus ”

ACKNOWLEDGEMENTS

I am indebted to a large number of people who have extensively contributed to the development and completion of this project.

First of all, I would like to thank most gratefully Professor Tim Dunne, the project's supervisor, for his excellent guidance, expertise advice and the much appreciated encouragement throughout the entire duration of this project and also for his many efforts and hours devoted to ensure that the project could be on track despite the difficulties and that this dissertation could eventually reach the best structure and format in the given circumstances.

A sincere word of thanks goes to Professor Cas Troskie, the project's co-supervisor, for his excellent advice in the writing of the dissertation and for providing the financial support for my M. Sc. studies.

I wish to thank Professor Cliff Moran, the former Dean of the Faculty of Science at UCT, for accepting the proposal from the Department of Statistical Science that led to developing of my M.Sc. thesis and greatly appreciated sponsorship of the data re-arrangement procedures performed in this project.

A special gratitude goes to Judith Du Toit, the Manager of the Faculty of Science at UCT, for her interest in my research work and her involvement in organising the student data in the initial stages of the project.

I would also like to acknowledge Daniel Franco of the Information Technology Services Unit (ITS) at UCT for writing and implementing the computer programme to perform the necessary re-arrangement procedures of the original data set.

I wish to express my appreciation to Jaqui Sommerville of the User Support Services at UCT for all her support and advice in designing the new database structure and her continuous help with many aspects of the analyses in *Statistica*.

A special word of thanks goes to Francesca Little of the Department of Statistical Sciences for her advice and input on the logistic regression analysis.

A particular word of thanks is extended to Fred Bowers, the Departmental Assistant in the Department of Statistical Sciences for his help in developing the final shape of the dendrograms presented in the body of this dissertation.

I would also like to thank the Department of Statistical Sciences for allowing me to use their computer facilities and advice of many staff members throughout the duration of the project.

Sincere and heartfelt gratitude goes to my family and friends for their words of support and encouragement.

Finally, I would like to dedicate this work to my late mother, Bozena Rek-Ronda.

SYNOPSIS

Ultimate academic success at any tertiary institution is affected and partially determined by many factors related to various aspects of individual's life. These factors could be separated into the following distinct categories, namely, educational, biographical, environmental and personal factors. Some of these determinants are used in the admission procedures adopted at tertiary institutions.

In South Africa, the results of different final matriculation examinations (referred to as *matric or matric exams*) written in several educational departments throughout the country are employed to assess the individual's potential to succeed. However, effectiveness of matric results as predictors of successful academic performance has always been controversial.

Expressing these concerns and desiring to explore them, the Faculty of Science at the University of Cape Town (UCT) accepted a proposal from the Department of Statistical Sciences to investigate several issues affecting students' performance in the Faculty. The proposal has led to developing this M.Sc. thesis. The major issue of concern in this study is to describe, on a retrospective basis, the extent to which the current selection criteria based on the matric results may have predicted various types of academic performance in the Faculty amongst those selected and admitted.

The thesis also exhibits a coherent and fairly complete methodology that is applicable at general or at particular levels of student performance data analysis, on a continuing year-to-year basis. The particular statistical methods and techniques used in this study have been summarised and discussed in the three Appendices.

The original structure of the data extracted from the university source had a particularly awkward format and complex configuration. Preliminary investigation revealed that unifying and then manipulating the files in any statistical package would be particularly troublesome as multiple records occurred for each student, one record per matric subject or university course in each file. Thus to deal with this problem a new format for the database had to be designed and implemented in order to perform the required analyses in the study.

Retrospectively, out of the chosen set of available predictors, the matric and faculty point scores (particular methods of summarising and evaluating the matric exams) have emerged as the most efficient but competing predictors in distinguishing between the potentially successful and unsuccessful students in many of the investigated definitions of academic performance.

The educational departments appear to be not strictly comparable with respect to their apparent predictive value for academic performance over the period studied.

Additionally, significant differences have been noted between the White and other students with the latter performing substantially worst in many of the examined definitions of academic performance. This finding, not surprising and rather expected, has distinctly emphasised the issue of insufficient resources and substandard education provided to various groups within the Black South African population in the past.

CONTENTS

<u>CHAPTERS</u> ¹	<u>PAGE</u>
0. INTRODUCTION	1.
0.1 Introduction	
0.2 Technical Aspects	
1. DATABASE DESIGN AND FORMAT	10.
1.1 Data set	
1.2 Original format of data	
1.3 Predicament	
1.4 Reconciliation between the old and new data formats	
1.5 Revised format of unified data set	
1.6 Creation of the unified data set: summary	
2. RELATIONSHIPS BETWEEN ACADEMIC PERFORMANCE IN MATRIC AND AT UNIVERSITY	37.
2.1 Variables	
2.2 Issues in the analysis of the relationships between academic performance in matric and at university	
2.3 Analyses in the project	
3. RELATIONSHIPS BETWEEN POINT SCORES AND UNIVERSITY FIRST YEAR PERFORMANCE	65.
3.1 Association between faculty point score and performance in first academic year at UCT	
3.2 Association between matric point score and performance in first academic year at UCT	
3.3 Inference-based recursive modelling: FIRM	
3.4 Technical Comments on methods used	
4. ASSOCIATIONS BETWEEN POINT SCORES AND FIRST YEAR PERFORMANCE	144.
4.1 Relationships within categories between faculty point score and academic performance	
4.2 Relationships within categories between matric point score and academic performance	

¹ Only main sections are listed.

5. PREDICTORS OF PERFORMANCE IN FIRST YEAR UNIVERSITY COURSES	155.
5.1 Biology related university courses	
5.2 Maths related university courses	
5.3 Physical Science related university courses	
5.4 SFP university courses	
6. RELATIONSHIPS BETWEEN ACADEMIC YEARS	323.
6.1 Relationships between academic performance in first and second years	
6.2 Relationship between academic performance in first and third years	
7. GRADUATION ISSUES	329.
7.1 Ordinal response analysis	
7.2 Binary response analysis	
8. CONCLUSIONS	348.
8.1 Summary of most important findings	
8.2 Recommendations for further research	
BIBLIOGRAPHY	352.

APPENDICES²

APPENDIX A: MEASURES OF ASSOCIATION

- SECTION A1: Introduction
- SECTION A2: Measures of Association: Pearson and Spearman
- SECTION A3: Measures of Association: Kedall's Tau, Somers' *D* and Gamma
- SECTION A4: Other Measures of Association
- SECTION A5: Summary
- SECTION A6: *BMDP* Manual: Discrepancies and Corrections
- SECTION A7: Example

APPENDIX B: GENERALISED LINEAR MODELLING

- SECTION B1: Exponential Family of Distributions
- SECTION B2: Generalised Linear Models: Formal Definition
- SECTION B3: Information and Variance
- SECTION B4: Inference
- SECTION B6: Parameters Estimation
- SECTION B5: Log-Linear Models

APPENDIX C: FIRM: FORMAL INFERENCE-BASED RECURSIVE MODELLING

- SECTION C1: Introduction
- SECTION C2: Mode of Operation
- SECTION C3: Types of Predictors
- SECTION C4: Analysis of a Set of *N* Observations
- SECTION C5: Test Statistics
- SECTION C6: Inference
- SECTION C7: Sample Considerations

ADDENDA

ADDENDUM 1

CONFIRM summary file for the ordinal response analysis

ADDENDUM 2

CATFIRM summary file for the binary response analysis

² Both Appendices and Addenda do not have specific page numbers. They contain only references to the specific Appendix or Addenda and the particular section they belong to.

LIST ³ OF TABLES AND DIAGRAMS

PAGE

CHAPTER 0

Table 0: Content of chapters and appendices	8.
---	----

CHAPTER 1

Diagram 1.6: Stages in producing unified data set	34.
---	-----

CHAPTER 2

Table 2.1.1: Explanatory Variables	39.
Table 2.1.2: Response Variables	44.
Table 2.2: Issues identified in the analysis of the relationships. between academic performance in matric and at university	47.
Table 2.3: Specific Analyses	52.
Diagram 2.2: Possible multivariate analyses	45.
Diagram 2.3: Graphical representation of the problem studied	50.

³ The list includes only the tables and diagrams that have been numbered in the body of the thesis for clarity.

CHAPTER 0: INTRODUCTION

0.1 INTRODUCTION

Ultimate academic success at any tertiary institution (a social organisation with explicit racial and cultural features) is affected and partially determined by many factors related to various aspects of individual's life. In South African context, these factors could be separated into the following distinct categories, namely ¹:

a) Educational factors, for example:

- Type of secondary school attended
- Quality of secondary schooling tuition received
- Academic progress after enrolment at a tertiary institution

b) Biographical/Environmental factors, for example:

- Gender
- Race and Culture
- Family history
- Financial constraints

c) Personal factors, for example:

- Individual's disposition
- Ability to cope in new circumstances
- Dedication and perseverance
- Potential and motivation to study and acquire further knowledge

All of these factors, to a greater or lesser extent, impact on students' university or technikon ² career. However, assessing their relative contributions and importance would in some circumstances be scientifically questionable or even socially undesirable. For instance, defining and quantifying the variables associated with individual's personality or intelligence would be rather difficult and troublesome to manage.

Most of these factors could be used in the admission procedures. However the granting or refusal of entry to students into tertiary institutions based solely on the particular attributes of their origins or financial circumstances, for example, would be unethical.

¹ It should be noted, however, that the aim of this separation is not to exhaust or complete all the topics associated with academic performance but rather to present an overall picture in an ordered manner.

² Technikon is a type of tertiary institution in South Africa, organised on similar lines to any South African university but offering diplomas mainly in technical disciplines.

Thus, these concerns introduce the question of what features a most convenient set of admission criteria into a tertiary institution should possess. The criteria ought to be fair in isolating those factors that can most precisely identify the required academic abilities for prospective students to succeed at university or technikon. At the same time, these factors must also to be measurable in a relatively easy and unbiased way.

Thus, pursuing both the issues of impartiality and objectivity, selection procedures employed throughout South African tertiary institutions take into account only certain of the educational factors in distinguishing between the potentially successful and unsuccessful students. The results of different final matriculation examinations (later in the thesis this set of examinations is referred to as *matric* or *matric exams*) written in distinct educational departments throughout the country are employed to assess the individual's potential to succeed. However, effectiveness of examination results as predictors of successful performance has always been controversial with the discussion being intensified in the wake of the recent changes in the country's educational system³. Notwithstanding efforts by the South African Certification Council (SAFCERT)⁴ to adjust and standardise matriculation results across provincial departments, many of the tertiary institutions suspect that the standards set by the educational departments for the matriculation examinations diverge. Thus the final marks and grades awarded may not always be sufficiently comparable across the Educational Departments.

The fact that some of the educational departments (such as the Department of Education and Training (DET)) used to be racially based and under-resourced, thus routinely producing few students achieving results allowing entry into a tertiary institution, has added a further distortion to the already complex situation. To deal with this issue of unjustifiable inequality in the past, students with sub-standard results have been admitted into special programmes⁵ offered by some of the South African universities. Creators of these programmes have expressed the view that the disadvantaged students would eventually succeed at a tertiary institution when provided with additional help at the beginning of their studies.

Having outlined the matters associated with the academic performance in South African context, the major issues of concern in this study are: the extent to which the current selection criteria based on the results of matriculation examinations apparently predict various types of success at a tertiary institution amongst those

³ The changes relate to the rearrangement of the old segregated departments and the establishment of the new provincial educational departments.

⁴ The South African Certification Council (SAFCERT), continuation of the Joint Matriculation Board after the formation of the new nine provinces in 1996, verifies the process of setting up and evaluating the matriculation examinations written in the different educational departments. The organisation is responsible for moderating the examination papers and ensuring that matriculation standards remain consistent from year to year. There is also of a committee of professional statisticians, who are responsible for making the appropriate adjustment to the examination results based on norms from the previous 5 years.

⁵ These programmes usually provide an additional year during which students attend introductory courses to upgrade their skills. During or after completing the first year, the students enrol into regular programmes with typical curricula and thus if successful eventually obtain the standard undergraduate degrees.

selected and admitted, and the effectiveness of the special programmes addressing the needs of disadvantaged students, in terms of success rates and predictive factors for conventional courses.

Thus, expressing these concerns and desiring to explore them, the Faculty of Science at the University of Cape Town (UCT) accepted a proposal from the Department of Statistical Sciences to investigate several issues affecting students' academic performance in the Faculty. The proposal has led to developing this M.Sc. thesis.

Before the discussion of the various issues associated with the academic performance is closed, the researcher believes that other issues require some mention although not addressed in this thesis. Ideally, tertiary institutions should anticipate a certain level of competence from prospective students. Thus it is not clear that tertiary institutions per se should be required to rectify the substantial inadequacies, mistakes and wrongdoings of the secondary schooling system at high tertiary-level costs of courses and additional academic years. A student should not only have the necessary potential⁶ but also be equipped with some elementary skills acquired at a secondary school level despite the student's particular origins or history. This attainment of proper objectives by the secondary school system would enable the tertiary institutions to make optimal use of their resources.

Thus the thesis examines the data on students at admission and their academic trajectories at a single institution, the University of Cape Town, in order to ascertain any patterns within the current approaches adopted in respect of inequities in education and student needs. The emergence of any patterns will therefore be restricted in interpretation to the historical and current strategy of admission to regular and academic development programmes. There will be little debate offered in the thesis on the implications of the admission criteria within a social context of unchanged or improving quality of education at secondary level.

⁶ The results of matriculation examinations could be used to assess that potential. However this assessment may not always be reliable or satisfactory.

However attention will be given, on a retrospective basis, to the manner in which the current selection criteria ⁷ and new criteria adopted can be monitored and where necessary adapted to suit the purpose of admission procedures, namely

- to give access in principle to all who appear able to graduate within a demonstrable time horizon
- to decline applications from any persons who could not reasonably be expected to graduate timeously
- to categorise successful applications and to channel an admitted student into an academic programme that is appropriate for that person's apparent profile of needs and abilities.

Admission procedures have ethical and resource objectives that are essentially multidimensional. Optimal procedures will be sought by tertiary institutions but will essentially involve making trade-offs between objectives that compete with one another in the context of constrained resources. It is a goal of this thesis to interpret and present historical data (1990-1997) in a manner that will allow more conscious and informed choices in the trade-off between the objectives.

A secondary goal of the thesis is to exhibit a coherent and fairly complete methodology that is applicable at general or at particular levels of student performance data analysis, on a continuing year-to-year basis.

⁷ At UCT the symbols obtained by students for the individual matric subjects are converted into types of point scores, namely the matric and the faculty point scores. The reader may refer to Table 2.1.1 in chapter 2 for a detailed description in which the two point scores are calculated.

Admission criteria into the Faculty of Science are based on the faculty point scores and they are briefly summarised in the table below.

certain admission	likely admission	waitlisted	minimum subject requirements
faculty point scores of 48 and more	faculty point scores between 47 and 45	faculty point scores between 44-43	E (HG) or B(SG) Mathematics or E (HG) or D (SG) Physical Science or Biology or Physiology

At this stage we may point out that students had been admitted into the Faculty with the faculty point scores lower than 43.

0.2 TECHNICAL ASPECTS

In the set ⁸ of students analysed in this project the following two subgroups may be distinguished: the regular ⁹ B.Sc. students and the Science Foundation Programme ¹⁰ (SFP) students. Results obtained using the appropriate statistical techniques will be presented and explored separately for the two programmes.

At this stage we present several descriptions of the relative performance at a university at various stages of students' academic career. Some of these definitions will be later incorporated in certain aspects of the current admission and re-admission criteria in the Faculty of Science at UCT. The list on the following page provides a concise summary of the performance definitions ¹¹ identified in the project. These definitions will be incorporated later in the exploration of the response variables ¹².

⁸See Chapter 1 for detailed description of the set of students under analysis.

⁹ The regular B.Sc. programme is a minimum three calendar year undergraduate degree offered in the Faculty of Science at UCT. The current set of the Faculty's graduation rules states that a student is awarded a B.Sc. degree upon completion of the following nine full-year courses or the equivalent:

- four first-year full-year courses or the equivalent
- three second-year full-year courses or the equivalent
- two third-year full-year courses or the equivalent

Up to 1993 students registering in the first year were required to complete only 8 full-year courses i.e. they had to complete only one third year full-year course or the equivalent.

¹⁰ The Science Foundation Programme (SFP) is a special academic support programme in the Faculty of Science at UCT offered to students from disadvantaged backgrounds (mostly from former DET schools). The programme provides introductory courses in maths, physics, chemistry and biology.

¹¹ These performance definitions incorporate some of the admission criteria applicable in the Faculty of Science at UCT. For example, a student is allowed to register in the Faculty of Science for a second year if he or she has passed (i.e. obtained 50% or more) at least two full first-year courses or the equivalent in the first calendar year of study.

¹² See Chapter 2 for explanations and definitions of the response variables.

LIST OF POSSIBLE DEFINITIONS OF PERFORMANCE AT UCT

1)

Specific courses: actual number of first-year courses passed by a first-year student in a particular calendar year

(Modification could apply to students with either second or third-year registration status for second or third-year courses respectively completed in any given calendar year)

2)

Specific courses: actual marks/percentages obtained by a first-year student for particular first-year courses in a given calendar year

(Modification could apply to students with either second or third-year registration status for second or third-year courses respectively completed in any given calendar year)

3)

Success = passing 2 or more full first-year courses or the equivalent for a first-year student in a given calendar year

Failure = passing less than 2 full first-year courses or the equivalent for a first-year student in a given calendar year

(Modification could apply to students with either second or third-year registration status for second or third-year courses respectively completed in any given calendar year; the cut-off point of 2 any year full- year courses could be changed to any other desired number of courses passed i.e.: 0; 0.5; 1; 1.5; 2; 2.5; 3; 3.5 and 4 or more)

4)

Success = average of 50% or more for a minimum of 2 full first-year course equivalents passed by a first-year student in a given calendar year

Failure = average of less than 50% for a minimum of 2 full first-year course equivalents passed by a first-year student in a given calendar year

(Modification could apply to students with either second or third-year registration status for second or third-year courses respectively completed in any given calendar year; the cut-off point could be changed to any desired percentage)

5)

Success = 50% or more for any of the individual first-year courses completed by a first-year student in a given calendar year

Failure = less than 50% for any of the individual courses completed by a first-year student in a given calendar year

(Modification could apply to students with either second or third-year registration status for second or third-year courses respectively completed in any given calendar year; the cut-off point could be changed to any desired percentage)

6)

Success =graduating after 3 years for a student in first year in a particular calendar year

Failure =graduating after 4 or more years for a student in first year in a particular calendar year

or

Success =graduating after 3 or 4 years for a student in first year in a particular calendar year

Failure =graduating after more than 4 years for a student in first year in a particular calendar year

All these definitions are calendar year specific but they could be suitable modified if the analysis was to be made on the data in all the eight years of study combined.

As the Faculty of Science is greatly interested in possible inferences and implications of the project's data analysis, the main body of the thesis explores elements that may be useful to the Faculty's decision making process. These elements are:

- description of necessary data-rearrangement procedures
- development and examination of statistical summaries
- identification of various interactions between academic performance in matric and university levels

The theory of statistical methods applied to examine the issues of concern is presented and explained in the appendices.

This deliberate division of developments, results and statistical theory has been made to address the multipurpose nature of this applied project, namely to shed light on the issues of students' matric and university performance as specified by the Faculty of Science and to fulfil the requirements for the M.Sc. degree. Thus the appendices should not be viewed as only additional or supplementary elements but fully integral components of the thesis.

Furthermore to facilitate the project's multipurpose nature, many footnotes have included in the thesis. They not only complement the text by providing additional information and insight on the particular issues but also in some instances include explanation and development of ideas, concepts and results. In that respect the footnotes are not just supplements to the thesis but integral parts of it.

The single notational convention N is used to denote the *sample size* through the entire thesis. This is to ensure presentational comparability and continuity with other statistical notation used in the thesis ¹³.

Table 0 on the following page summarises the contents of the chapters and appendices of the thesis.

¹³ For example, in Appendix A on measures of association R and C represent the row and column totals in a contingency table respectively.

TABLE 0
CONTENT OF CHAPTERS AND APPENDICES

CHAPTERS	TITLE	CONTENT
CHAPTER 0	INTRODUCTION	Presentation of the problem at stake and the issues of major concern
CHAPTER 1	DATABASE DESIGN AND FORMAT	Description of the old and new data formats; Explanation of the ways in which the two sets were reconciled
CHAPTER 2	RELATIONSHIPS BETWEEN ACADEMIC PERFORMANCE IN MATRIC AND AT UNIVERSITY	Comprehensive summary of the various links and interactions between academic performance at matric and university levels
CHAPTER 3	RELATIONSHIPS BETWEEN POINT SCORES AND UNIVERSITY FIRST YEAR PERFORMANCE	Investigation of the relationships between the faculty and the matric point scores and the academic performance in first year
CHAPTER 4	ASSOCIATION BETWEEN POINT SCORES AND FIRST YEAR PERFORMANCE	Examination of the presence of relationships for the faculty and for the matric point scores against the academic performance in the first year within categories of the biographical explanatory variables
CHAPTER 5	PREDICTORS OF PERFORMANCE IN FIRST YEAR UNIVERSITY COURSES	Investigation of the likely predictors of academic performance in first year university courses
CHAPTER 6	RELATIONSHIPS BETWEEN ACADEMIC YEARS	Investigation of the mutual interdependencies amongst performance in the different academic years
CHAPTER 7	GRADUATION ISSUES	Investigation of the potential predictors of the duration of the B.Sc. degree
CHAPTER 8	CONCLUSIONS AND RECOMMENDATIONS	Summary of most important findings; Recommendations for further research
APPENDICES	TITLE	CONTENT
APPENDIX A	MEASURES OF ASSOCIATION	DESCRIPTION OF THE STATISTICAL THEORY OF THE MEASURES OF ASSOCIATION.
Section A1	Introduction	General consideration of statistical aspects of the measures of association
Section A2	Measures of Association: Pearson and Spearman	Description and discussion on the following measures of association: <ul style="list-style-type: none"> • Pearson's product moment correlation coefficient r • Spearman's rank-order correlation coefficient r_s
Section A3	Measures of Association: Kendall's T , Somers' D and Gamma G	Presentation of the important definitions related to the topic discussed and the notation adopted throughout Appendix A Description and discussion on the following measures of association: <ul style="list-style-type: none"> • Kendall's rank-order coefficient T • Somers' D Statistic • Gamma Statistic G
Section A4	Other Measures of Association	Discussion on the classical chi-square test of association and variations thereof Description of the Kappa coefficient K
Section A5	Summary	Concise summary and comparison of the five measures of association (Sections A2 and A3)
Section A6	<i>BMDP</i> Manual: Discrepancies and Corrections	Presentation of the discrepancies between the formulas in <i>BMDP</i> manual and those in the original sources and addition of the necessary corrections
Section A7	Example	Numerical example presenting the calculations of the five measures of association

<i>APPENDIX B</i>	GENERALISED LINEAR MODELS	DESCRIPTION OF THE STATISTICAL THEORY OF THE GENERALISED LINEAR MODELS
Section B1	Exponential Family of Distributions	Presentation of Exponential Family of Distributions and its relevance to the Generalised Linear Models
Section B2	Generalised Linear Models: Formal Definition	Description of the main components of the Generalised Linear Models
Section B3	Information and Variance	Presentation of some results required for further development of the Generalised Linear Models
Section B4	Parameters Estimation	Development of techniques (Newton-Raphson method and Fisher method of scoring) required for parameter estimation
Section B5	Inference	Description of three methods for hypothesis testing (Score test, Maximum Likelihood Estimates test and Likelihood Ratio Test); Assessing adequacy of and redundancy within a model; Presentation of confidence intervals for the model and parameter estimates
Section B6	Log-linear Models	Summary of the theory applied to log linear modelling in contingency tables
<i>APPENDIX C</i>	FIRM: FORMAL INFERENCE-BASED RECURSIVE MODELLING	DESCRIPTION OF THE METHOD OF RECURSIVE MODELLING
Section C1	Introduction	Overview of the method
Section C2	Mode of Operation	Description of the procedures implemented in FIRM
Section C3	Types of predictor variables	Description of the various types of predictor variables used in FIRM
Section C4	Analysis of a set of N observations	Specification of the methods used to split and merge data items
Section C5	Test Statistics	Specification of the different test statistics implemented in FIRM
Section C6	Inference	Application of the Bonferroni inequality and multiple comparisons to the operation of FIRM
Section C7	Sample Considerations	Issues of concern relating to the sizes of the samples under analysis

CHAPTER 1: DATABASE DESIGN AND FORMAT

1.1 DATA SET

The data set analysed in the study covers a period of eight years, namely 1990 to 1997. The study population comprises only those **first-year students** who have **passed the South African school leaving matriculation examinations** ¹ and are the **first-time entrants** (into any university / tertiary institution programme) in the Science Faculty ² at the University of Cape Town (UCT) during the review period.

¹ Students who have passed foreign school leaving matriculation examinations and also those who have been registered in other faculties at UCT or at any other tertiary institution in South Africa or around the world are excluded from the present study.

² The data set includes the following types of students:

- Those who were registered in the Science Faculty in their first academic year and remained in the Faculty during the review period.
- Those who were registered in the Science Faculty in their first academic year and moved to other Faculties in the subsequent years.

Students who took courses offered in the Science Faculty but were formally registered, as students in other Faculties at UCT are not included in the current data set analysed in this study.

1.2 ORIGINAL FORMAT OF DATA SETS

1.2.1 GENERAL REMARKS

The data was extracted from a database (*set in MS Office Access format*) made available from the Academic Development Programme (ADP) Unit at UCT. The database was prepared by the university's administrative staff and extracted from students' records stored on "Heritage"³.

Due to the format of the ADP Unit's database the data necessary for the current analysis had to be extracted in the following way:

For each calendar year in the study there were **two types of Excel files**.

The **first type of files** referred to as **school details files** contained names in the form of unique student ID numbers and personal details of first-time students of a given calendar year registering in the Faculty of Science at UCT and their matric subjects, grades and aggregate profiles.

The **second type of files** referred to as **university courses files** contained the student ID numbers and the subjects completed by students of any academic year of study with the courses grades, in a specified calendar year.

The tables below display the names and descriptions of the columns appearing in the **school details** and **university courses files** respectively, for each record. Many records occur for each case, one record per subject or course taken under each student ID number. The essential linking element between multiple records per case within a file for one student's profile and between profiles for the same student across the other files, is the student ID number. The student ID number is a nine-character label. It is composed of the first three consonants of surname (plus X's for any missing consonants), first three letters of personal name and a three digit internal UCT serial number.

³ "Heritage" is a computerised support system at UCT that stores biographical and academic progress details of all students registered at the university. At the time of writing it was being phased out and replaced by a new student records module within U.C.T.'s PRISM using SAP/R3.

1.2.2 FILE STRUCTURES

a) SCHOOL DETAILS FILE

There was one school details file for each calendar year of the analysis. Each file contained only the personal details and matric profiles of those students who registered at UCT for their first year of study in the particular calendar year.

Column Name	Column Description	Type of Information ⁴	Within Multiple records per case
✓ <i>Student Number</i>	University ID, unique to student	Label	Repeated
✓ <i>Gender</i>	Male or Female	Nominal	Repeated
<i>Colour</i>	White, Black, Coloured, Indian	Nominal	Repeated
✓ <i>Language</i>	Student's preferred language (including one or more of the official languages or other)	Nominal	Repeated
<i>Exam Date</i>	Date of Matric Exams	Year/Month	Repeated
<i>Education Department</i>	Name of Department issuing Matric Certificate	Nominal	Repeated
<i>Subject</i>	Name of Matric subject	Nominal	Variable
<i>Subject Grade</i>	Symbol attained	Ordinal	Variable
<i>Subject Score</i>	Actual number of points scored on Matric Exam	Integer	Variable
<i>School Number</i>	Number assigned to school attended	Number-Label	Repeated
<i>Matric Point-Score</i>	Unweighted Point-Score	Integer	Repeated
<i>Faculty Point-Score</i>	Weighted Point-Score (the Science Faculty doubles points for Maths and the better of Biology and Physical Science)	Integer	Repeated
<i>Academic Year</i>	Student's Year of Study (in these school files the Academic Year value shown is 1)	Ordinal	Repeated
<i>Promotion Code</i>	Student's Status at the end of Academic Year 1	Nominal	Repeated

Extracts of contents of typical school details files are shown on the following page.

⁴ Description to labels used in "type of information" column:

Nominal-information recorded is specific to each student but must fall into one of pre-specified mutually disjoint categories e.g. gender: either male or female

Ordinal-information recorded is specific to each student and it indicates the order or position in ordered series of categories e.g. subject grade: one of the symbols from A to F is included (with A being the highest grade)

Label-information recorded is unique to each student e.g. student number specific to each student, school number unique to each school

Extract from 1993-school details file

Student number	Sex	Cls.	Lang.	Exam Date	Educ. Dept.	Subject	Grade	Score	School No.	Fac.Pt.	Mat.Pt.	Ac Yr	Prom Code
ADMSHU001	M	C	EN	1992/11	CO	MATS	A	241	146	55	41	1	CON
ADMSHU001	M	C	EN	1992/11	CO	HISH	A	338	146	55	41	1	CON
ADMSHU001	M	C	EN	1992/11	CO	BIOH	A	320	146	55	41	1	CON
ADMSHU001	M	C	EN	1992/11	CO	PSCH	B	286	146	55	41	1	CON
ADMSHU001	M	C	EN	1992/11	CO	EGMH	B	281	146	55	41	1	CON
ADMSHU001	M	C	EN	1992/11	CO	AFSH	D	159	146	55	41	1	CON
ADMSHU001	M	C	EN	1992/11	CO	MATS	A	241	146	55	41	1	CON
ADMSHU001	M	C	EN	1992/11	CO	HISH	A	338	146	55	41	1	CON
ADMSHU001	M	C	EN	1992/11	CO	BIOH	A	320	146	55	41	1	CON
ADMSHU001	M	C	EN	1992/11	CO	PSCH	B	286	146	55	41	1	CON
ADMSHU001	M	C	EN	1992/11	CO	EGMH	B	281	146	55	41	1	CON
ADMSHU001	M	C	EN	1992/11	CO	AFSH	D	159	146	55	41	1	CON
AHMMUN003	F	I	EN	1992/11	CO	EGMH	C	264	686	48	36	1	REN
AHMMUN003	F	I	EN	1992/11	CO	AFSH	C	198	686	48	36	1	REN
AHMMUN003	F	I	EN	1992/11	CO	MATH	D	236	686	48	36	1	REN
AHMMUN003	F	I	EN	1992/11	CO	PSCH	C	259	686	48	36	1	REN
AHMMUN003	F	I	EN	1992/11	CO	BIOH	B	285	686	48	36	1	REN
AHMMUN003	F	I	EN	1992/11	CO	HISH	C	266	686	48	36	1	REN
AHMMUN003	F	I	EN	1992/11	CO	EGMH	C	264	686	48	36	1	REN
AHMMUN003	F	I	EN	1992/11	CO	AFSH	C	198	686	48	36	1	REN
AHMMUN003	F	I	EN	1992/11	CO	MATH	D	236	686	48	36	1	REN
AHMMUN003	F	I	EN	1992/11	CO	PSCH	C	259	686	48	36	1	REN
AHMMUN003	F	I	EN	1992/11	CO	BIOH	B	285	686	48	36	1	REN
AHMMUN003	F	I	EN	1992/11	CO	HISH	C	266	686	48	36	1	REN

Extract from 1996-school details file

Student number	Sex	Cls.	Lang.	Exam Date	Educ. Dept.	Subject	Grade	Score	School No.	Fac.Pt.	Mat.Pt.	Ac Yr	Prom Code
NDLROS001	F	B	ZU	1995/11	ET	MATH	C	0	11877	46	35	1	COI
NDLROS001	F	B	ZU	1995/11	ET	BIOH	C	0	11877	46	35	1	COI
NDLROS001	F	B	ZU	1995/11	ET	PSCH	C	0	11877	46	35	1	COI
NDLROS001	F	B	ZU	1995/11	ET	ZUMH	C	0	11877	46	35	1	COI
NDLROS001	F	B	ZU	1995/11	ET	EGSH	D	0	11877	46	35	1	COI
NDLROS001	F	B	ZU	1995/11	ET	AFSH	C	0	11877	46	35	1	COI
NDXSIR001	M	I	EN	1995/11	IN	EGMH	C	271	11689	45	34	1	REN
NDXSIR001	M	I	EN	1995/11	IN	GEOH	B	298	11689	45	34	1	REN
NDXSIR001	M	I	EN	1995/11	IN	BIOH	C	240	11689	45	34	1	REN
NDXSIR001	M	I	EN	1995/11	IN	PSCH	D	212	11689	45	34	1	REN
NDXSIR001	M	I	EN	1995/11	IN	AFSH	D	155	11689	45	34	1	REN
NDXSIR001	M	I	EN	1995/11	IN	MATH	D	213	11689	45	34	1	REN
NGMKGO001	M	B	SS	1994/11	IE	GEOH	C	0	666	38	29	1	COI
NGMKGO001	M	B	SS	1994/11	IE	EGMH	C	0	666	38	29	1	COI
NGMKGO001	M	B	SS	1994/11	IE	AFSH	E	0	666	38	29	1	COI
NGMKGO001	M	B	SS	1994/11	IE	MATH	E	0	666	38	29	1	COI
NGMKGO001	M	B	SS	1994/11	IE	BIOH	D	0	666	38	29	1	COI
NGMKGO001	M	B	SS	1994/11	IE	PSCH	E	0	666	38	29	1	COI

b) UNIVERSITY COURSES FILE

There were several university courses files for each calendar year of the analysis. Each file contained the courses completed in a particular calendar year of the analysis by all students who shared a common first-year of registration.

Column Name	Column Description	Type of Information	Within Multiple records per case
<i>Student Number</i>	University ID, unique to student	Label	Repeated
<i>Faculty</i>	Degree for which student registered in the Calendar Year of Study	Nominal	Repeated
<i>Academic Year</i>	Student's Year of Study (in these files the Academic Year values depend on student' academic progress and not the actual number of years a student has been registered for at the University)	Ordinal	Repeated
<i>Promotion Code</i>	Student's Status at the end of Calendar Year	Nominal	Repeated
<i>Course</i>	Code of the University Course registered	Nominal	Variable
<i>Course Score %</i>	Percentage obtained for a particular Course	Integer	Variable
<i>Symbol</i>	Grade of Percentage obtained i.e. first (1), upper second(2+), lower second(2-), third(3), failure(F), DP ⁵ refused (DPR) and absent(AB)	Ordinal	Variable

Extracts of typical university courses files are presented on the next page.

⁵ DP is an abbreviation for "Duly Performed Certificate", i.e. a minimum requirement in respect of course work that a student ought to fulfil before he or she is allowed to write a final examination.

Extract from 1993 university courses file for students admitted in 1993

Student number	Faculty	Ac. Yr.	Prom Code	Course	Score %	Symbol
ADMSHU001	SCIB01	1	CON	ZOO103S	68	2-
ADMSHU001	SCIB01	1	CON	MTH101S	0	DPR
ADMSHU001	SCIB01	1	CON	PHY100W	0	DPR
ADMSHU001	SCIB01	1	CON	CEM101W	51	3
ADMSHU001	SCIB01	1	CON	MTH101F	0	DPR
AHMMUN003	SCIB01	1	REN	PHY100W	43	F
AHMMUN003	SCIB01	1	REN	MTH105W	0	AB
AHMMUN003	SCIB01	1	REN	CEM100W	36	F
AHMREA001	SCIB01	1	REN	PHY100W	0	AB
AHMREA001	SCIB01	1	REN	BIO100S	0	AB
AHMREA001	SCIB01	1	REN	CEM109S	0	AB
AHMREA001	SCIB01	1	REN	STA100S	0	AB
AHMREA001	SCIB01	1	REN	MTH101F	63	2-
AHMREA001	SCIB01	1	REN	BIO100F	32	F

Extract from 1997 university courses file for students admitted in 1996

student number	Faculty	Ac. Yr.	Prom Code	Course	Score %	Symbol
LNDAND004	SCIB03	1	CON	PHY107F	62	2-
LNDAND004	SCIB03	1	CON	BIO100F	53	3
LNDAND004	SCIB03	1	CON	MAM106H	53	3
LNDAND004	SCIB03	1	CON	CEM110H	43	F
LNDAND004	SCIB03	1	CON	MAM142S	68	2-
LNDAND004	SCIB03	1	CON	PHY204S	54	3
LNDTHE003	SCIB03	2	CON	MIC201F	70	2+
LNDTHE003	SCIB03	2	CON	BOT205S	61	2-
LNDTHE003	SCIB03	2	CON	MIC202S	67	2-
LNDTHE003	SCIB03	2	CON	BCH201F	57	3
LNDTHE003	SCIB03	2	CON	BOT202F	62	2-
LNDTHE003	SCIB03	2	CON	BCH202S	62	2-
LNRJOH001	SCIB03	2	CON	MAM200W	59	3
LNRJOH001	SCIB03	2	CON	CRU100W	60	2-
LNRJOH001	SCIB03	2	CON	CSC200W	60	2-

The academic progress of students registered for the first year of study in one of the calendar years was traced through entries in several university courses files. It has been established that students who have graduated have taken at most 6 years to complete their undergraduate degrees in the Science Faculty. Therefore, where the complete data was available, the academic progress for a particular student was traced for 6 years at most. This 6-year tracing was performed for students registered for the first year of study in 1990, 1991 and 1992. Since the exam results for 1998 and the later years were not available at the time of extraction, the academic progress for students registered for the first year of study in any of the years 1993, 1994, 1995, 1996 and 1997 was traced for shorter periods of time.

The table below summarises the type and the number of files available for each year of the study.

<i>First Year of Study</i>	<i>School Details File</i>	<i>Number of Calendar Years of University Courses</i>	<i>Calendar Years for which the Academic Progress was traced</i>
<i>UCT intake year</i>	<i>UCT intake year</i>	<i>Files</i>	
1990	1 ('90)	6	90, 91, 92, 93, 94, 95
1991	1 ('91)	6	91, 92, 93, 94, 95, 96
1992	1 ('92)	6	92, 93, 94, 95, 96, 97
1993	1 ('93)	5	93, 94, 95, 96, 97
1994	1 ('94)	4	94, 95, 96, 97
1995	1 ('95)	3	95, 96, 97
1996	1 ('96)	2	96, 97
1997	1 ('97)	1	97

1.3 PREDICAMENT

Both school details and university courses files exhibited common and multiple segments across records for each case or student, i.e. in the school details files there was a set of separate records/rows (one for each matric subject that a student completed); similarly in the university courses files there were separate records for each university subject completed by a student, i.e. multiple records per student.

The following extracts from the school details and university courses files from 1990 illustrate the structure.

School details file (1990)

Student number	Sex	Cls.	Lang.	Examdate	Educ Dept	School Subject	Grade	Score	Sch No.	Mat Pt	FacPt	Ac. Yr.	Prom Code
ABRDAW001	M	C	EA	1989/11	CO	EGMH	D	219	845	32	44	1	REN
ABRDAW001	M	C	EA	1989/11	CO	PSCS	D	174	845	32	44	1	REN
ABRDAW001	M	C	EA	1989/11	CO	BIOH	E	184	845	32	44	1	REN
ABRDAW001	M	C	EA	1989/11	CO	GEOH	C	244	845	32	44	1	REN
ABRDAW001	M	C	EA	1989/11	CO	MATH	A	320	845	32	44	1	REN
ABRDAW001	M	C	EA	1989/11	CO	AFSH	C	199	845	32	44	1	REN
Label	Some categorical variables: gender, colour, language, exam date, educational department					One subject + grade + score per row			Information is repeated across rows/records sharing common label				

University Courses File (1990)

Student number	Faculty	Ac. Yr.	Prom Code	Course	Score %	Symbol
ABRDAW001	SCIB01	1	REN	AMA101F	51	3
ABRDAW001	SCIB01	1	REN	EGS100W	48	F
ABRDAW001	SCIB01	1	REN	STA100S	31	F
ABRDAW001	SCIB01	1	REN	MTH105W	33	F
Label + categorical variables (faculty + promotion code) + integer (academic year)				Subject Information		

The original format of the data makes it difficult to combine the school details and university courses files in Access without having to increase the records for each student and each year of further registration. Some of these combination problems are exacerbated for students who wrote matric more than once.

The original data format makes the analyses unnecessarily more difficult and complex. It requires also that each time we wish to carry on a new analysis, a separate query be designed in Access. Then the query needs to be extracted for analysis in *Statistica* (the statistical package used in this project). Following such a route we would eventually produce a large number of files with duplicated information requiring huge storage space.

1.4 RECONCILIATION BETWEEN THE OLD AND NEW DATA FORMATS

In order to eliminate the combination problems in analysis it was decided that the initial format of the data set had to be changed. Thus the data was transposed into a new format ⁶. The data transformation was accomplished in such a way as to ensure the analyses would be much simpler and facilitate future updates to the database thus allowing other similar investigations from time to time.

The purpose of the modification was to convert the data from multiple records to one record per student containing all the relevant biographical, matric subjects and university courses details. The diagram detailing the stages involved in creating the unified data set is presented section 1.6. A reader may wish to become familiar with the diagram's content to understand the complexity of the various steps and efforts made to obtain the unified data set.

Before a full long record per student could be implemented several issues had to be considered. The paragraphs below summarise the decisions taken in the unifying process.

⁶ The change of the data format was accomplished with the approval and financial support of the Dean of the Science Faculty, Prof. VC Moran. The actual computer programme that performed the transformation was written by Daniel Franco of the Information Technology Services (ITS) Unit at UCT.

1.4.1 MODIFICATION OF THE SCHOOL DETAILS FILES

a) Issues of Repeated Matric and Repeated Records

The original database included the record of all matric exams completed by the students.

Within the group of students who had written matric more than once, there were some students who each had duplicate records for one or more matric subject either with the same or different symbols. Some symbols were clearly spurious zeros, but some involved genuine marks and symbols from rewrites.

The table below illustrates such a situation (only the most recent matric exam written is shown here).

Student number	Sex	Cls	Lang	Examdate	Educ Dept	Subject	Grade	Score
DCKWAY001	M	W	EN	1991/11	CA	ACTH	D	0
DCKWAY001	M	W	EN	1991/11	CA	ACTH	C	269
DCKWAY001	M	W	EN	1991/11	CA	AFSH	D	0
DCKWAY001	M	W	EN	1991/11	CA	AFSH	C	196
DCKWAY001	M	W	EN	1991/11	CA	EGMH	C	273
DCKWAY001	M	W	EN	1991/11	CA	EGMH	C	0
DCKWAY001	M	W	EN	1991/11	CA	GEOH	B	303
DCKWAY001	M	W	EN	1991/11	CA	GEOH	B	0
DCKWAY001	M	W	EN	1991/11	CA	MATH	B	281
DCKWAY001	M	W	EN	1991/11	CA	MATH	E	0
DCKWAY001	M	W	EN	1991/11	CA	PSCH	C	257
DCKWAY001	M	W	EN	1991/11	CA	PSCH	D	0

In order to deal with these students the better or best of their results in each subject have selected to be present in the new database. Thus, inter alia, spurious zeros are dropped.

Within the group of students who had written matric once, there were some students with duplicate records for the same matric subjects with the same symbol for each subject.

The table on the following page illustrates the situation.

Student number	Sex	Cls	Lang	Examdate	Educ Dept	Subject	Grade	Score
ADMDOR001	F	W	EN	1992/11	CA	ACTH	B	313
ADMDOR001	F	W	EN	1992/11	CA	ACTH	B	313
ADMDOR001	F	W	EN	1992/11	CA	AFSH	D	174
ADMDOR001	F	W	EN	1992/11	CA	AFSH	D	174
ADMDOR001	F	W	EN	1992/11	CA	BIOH	B	318
ADMDOR001	F	W	EN	1992/11	CA	BIOH	B	318
ADMDOR001	F	W	EN	1992/11	CA	EGMH	B	292
ADMDOR001	F	W	EN	1992/11	CA	EGMH	B	292
ADMDOR001	F	W	EN	1992/11	CA	MATH	A	320
ADMDOR001	F	W	EN	1992/11	CA	MATH	A	320
ADMDOR001	F	W	EN	1992/11	CA	PSCH	B	306
ADMDOR001	F	W	EN	1992/11	CA	PSCH	B	306

The duplicate profile was dropped in respect of those students.

b) Subject Score: Points scored for Matric exams

The actual number of points that students obtain for matric exams was excluded from the new database. In the view of the ADP unit they were not reliable and not comparable across different matric subjects⁷ and different educational departments. Many of the points were missing.

⁷ Exams in different matric subjects were often written out of different totals.

c) Language

In school details files from eight years there were some 24 different home languages. It had been suggested that in the new database there should be only 12 different language categories i.e. eleven for the official languages and one for Afrikaans and English combined. The remaining languages (such as Bulgarian, Dutch, Flemish, French, German, Greek, Hebrew, Polish, Portuguese, Spanish, Tamil and Chinese) would be included under the English language category. This move was justified as people speaking those languages are presumed to function as the English section of the South African population. However, since Chinese speaking students were expected to form the largest group⁸ of students speaking a language different from the eleven official languages at home, it was proposed that a thirteenth language category for the Chinese language would be created. Thus in the new database the following students' home languages are distinguished:

- Afrikaans
- English
- English and Afrikaans combined
- Chinese
- Xhosa
- Zulu
- Tswana
- Tsonga
- Northern Sotho
- Southern Sotho
- Swazi
- Venda
- Ndebele

⁸ In the final database there are only 30 Chinese-speaking students. However it is believed that a larger number of first-language Chinese speakers is present amongst the Science Faculty students. It is presumed that many of them had put English as their home language to emphasise their fluency in the English language. A similar comment may also perhaps applicable to the other foreign language groups.

d) Matric Exam Dates

There were 38 different matric exam dates recorded. In order to reduce the number of matric exam dates the following procedure was adopted. The dates were allocated to 9 distinct matric exam year categories. The core for each of these categories was formed by matric exams written in November in each of the calendar years of the analysis. The remaining dates were assigned to one of the specified categories on the following basis: all exams dates with a particular calendar year but with an earlier month than November were allocated to matric exam category in the previous calendar year. This move was justified as the dates corresponded to supplementary or deferred matric exams written in respect of the previous calendar years.

The following table gives the names and the structure of the matric exam dates categories adopted in the new database

Category Name	Dates included
Earlier than 1989/11	1980/11; 1983/11;1987/12; 1988/07;1988/11; 1989/03
1989/11	1989/11; 1990/02; 1990/03;
1990/11	1990/11; 1991/03; 1991/05; 1991/06
1991/11	1991/11; 1992/04; 1992/05; 1992/06
1992/11	1992/11; 1992/12; 1993/03; 1993/06
1993/11	1993/11; 1994/02; 1994/03;
1994/11	1994/11; 1995/02; 1995/03; 1995/04; 1995/05
1995/11	1995/11; 1996/03;
1996/11	1996/11

e) Matric Subjects

The initial data set contained many matric subjects (more than 100). Only certain school subjects, believed to be most directly or strongly associated with academic performance in the Faculty of Science, were chosen for inclusion in the new database.

The table below lists the chosen matric subjects.

Subject Code	Subject Name
MATH	Mathematics(HG)
MATS	Mathematics(SG)
ADMH	Additional Mathematics(HG)
BIOH	Biology(HG)
BIOS	Biology(SG)
PSCH	Physical Science(HG)
PSCS	Physical Science(SG)
GEOH	Geography(HG)
GEOS	Geography(SG)
HISH	History (HG)
HISS	History(SG)
CSTH	Computer Science(HG)
AFMH	Afrikaans (1st.)(HG)
AFSH	Afrikaans (2nd.)(HG)
AFSS	Afrikaans (2nd.)(SG)
EGMH	English (1st.)(HG)
EGMS	English (1st.)(HG)
EGSH	English (2nd.)(HG)
NSMH	Northern Sotho(HG)
SSMH	Southern Sotho(HG)
SZMH	Swazi (HG)(?)
TNMH	Tsonga(HG)
TSMH	Tswana(HG)
VEMH	Venda(HG)
ZUMH	Zulu(HG)
XHMH	Xhosa(HG)

f) Education Departments

There did not seem to be any justifiable or meaningful way of combining the old educational departments under the framework of the new departments. Therefore the old and new educational departments are all present in the new database.

The following table lists the names and the codes for the old and the new Educational Departments.

Department Code	Department Name
Old Departments	
CA	Cape Education Department
ET	Department of Education and Training
NE	DEC ⁹ : House of Assembly
IN	DEC: House of Delegates
CO	DEC:House of Representatives
JM	Joint Matriculation Board
NA	Natal Education Department
OF	Orange Free State Education Department
TK	Transkei Education Department
TR	Transvaal Education Department
IE	
New Departments	
EC	Eastern Cape Education Department
FS	Free State Education Department
GT	Gauteng Education Department
KN	KwaZulu-Natal Education Department
MP	Mpumalanga Education Department
NC	Northern Cape Education Department
NP	Northern Province Education Department
NW	North West Province Education Department
WC	Western Cape Education Department

⁹ DEC is an abbreviation for Department of Education and Culture.

1.4.2 MODIFICATION OF THE UNIVERSITY COURSES FILES

In the university courses files there were some 1390 different courses codes. In order to reduce the number of courses to be included in the new database the following issues had to be addressed.

a) Courses in Other Faculties

Since the Science Faculty's students often registered for subjects offered in other faculties some of the popular courses have been included in the new database, e.g.

b) First year, Second Year and Third Year Courses

The new database includes the university courses offered at first, second and third academic year levels. It also allows one to determine how many years a particular student takes to complete his or her undergraduate degree.

c) Repeated University Courses

The database holds the record of all occasions and percentages attained by a student who repeated a particular university course.

d) University Course Codes

Over the period from 1990 to 1997 some of the course codes changed (especially those offered by the Mathematics Department and Applied Mathematics Department as the two departments merged). The subjects with the old and the new codes are included separately in the new database. Later in the actual analysis these course codes may be combined and considered to represent essentially the same course.

e) Course Symbols

The university courses achievement symbols (i.e. identifying the grade of percentage attained) are excluded from the final database. The actual percentage is recorded where appropriate. The symbols AB (absent) and DPR (DP refused) are combined and thus it is not possible to distinguish between them in the new database. These symbols appear in the same columns as the percentages and are logically distinct from a zero mark.

f) Degree/Faculty of Registration

Due to the way in which the university courses files were originally extracted, we have information about students who were initially registered in the Faculty of Science and then transferred to other faculties or started with postgraduate degrees in the Faculty of Science. There are 69 different degree codes across the previous ten faculties. Each degree code represents a stream within a degree that was offered by one of the UCT faculties i.e. Arts, Commerce, Education, Engineering, Fine Arts and Architecture, Law, Medicine, Music, Science and Social Science and Humanities. Thus in respect of students transferring from the Science Faculty to other faculties after first year of study the new degree code is recorded and no further examination history is preserved in respect of that student.

The table below lists degree codes found in the university courses files.

Faculty	Degree Code
ARTS	ARTB01
	ARTH01
	ARTM01
	ARTZ01
COMMERCE	COMB01
	COMB02
	COMB05
	COMG07
	COMG10
	COMG12
	COMG13
	COMG14
	COMG15
	COMG16
	COMG18
	COMH01
	COMM01
	COMM02
	COMM03
COMZ01	
EDUCATION	EDUB01
	EDUG02
	EDUG03
	EDUG04
ENGINEERING	ENGB01
	ENGB02
	ENGB03
	ENGB05
	ENGB06
	ENGB07
	ENGB08
	ENGB09
	ENGB10
	ENGH01
ENGM01	

	ENGM04
	ENGZ01
FINE ART AND	FAAB01
ARCHITUTURE	FAAB03
	FAAB05
	FAAB07
	FAAM03
	FAAZ01
LAW	LAWP01
MEDICINE	MEDB03
	MEDB04
	MEDB07
	MEDB08
	MEDD01
	MEDH01
	MEDM95
MUSIC	MUSB95
	MUSM01
	MUSU01
	MUSZ01
SCIENCE	SCIB01
	SCIB03
	SCID01
	SCIH01
	SCIM01
	SCIZ01
	SCIZ09
SOCIAL	SSHB01
SCIENCE AND	SSHB03
HUMANITIES	SSHH01
	SSHH03
	SSHM04
	SSHZ01
	SSHZ05

1.5 REVISED FORMAT OF UNIFIED DATA SET

1.5.1 GENERAL COMMENTS

To facilitate the process of creating one record per student the data comprising the school and university courses files was transformed into four DBASE or text files. These files were combined to produce a final unified data file ready for the analysis in *Statistica*.

The table below summarises the file names and the part of the original information available in each of them.

Files in the unified data set	Information available
Matrici	Student details and "important" matric subjects in the years 1990-1997
Univi	Student details and "important" university courses in the years 1990-1997
Matricu	Student number and "unimportant" matric subjects in the years 1990-1997
Univu	Student number and "unimportant" university courses in the years 1990-1997
MatriciUnivi	Student details and both "important" matric subjects and "important" university courses in the years 1990-1997

Brief descriptions and extracts from each of these files are presented in the following sections.

1.5.2 UNIFIED MATRIC AND UNIVERSITY DETAILS FILES

a) "IMPORTANT" MATRIC DETAILS FILES

MATRICI file has 41 distinct fields.

The first nine fields contain biographical details:

- student number
- gender
- colour
- home language (as listed in 1.4.1.c)
- exam data (as per table in 1.4.1.d)
- educational department (as per table in 1.4.1.f)
- school attended
- weighted matric point-score
- unweighted matric point-score.

The next 26 fields indicate important matric subjects as per table in 1.4.1.e

The remaining 6 fields are the following indicators:

- missing unweighted point-score
- missing weighted matric point-scores
- other matric subjects i.e. those included in "unimportant" matric details files
- number of matric subjects present in this file
- number of records for each student in the original matric details files.

The file contains 2496 students' records.

An extract from this file is presented on the following page. The extract has been made from the file in *Statistica* in such way as to best fit the MS Word format of this document. An original row-column structure arranged by cases and variables is transposed and printed in portrait on the page while the original structure is preserved in the file.

b) "IMPORTANT" UNIVERSITY DETAILS FILES

UNIVI file has 627 fields.

The majority of the fields document "important" university courses completed by students in each of the calendar years of the analyses.

The other indicator fields for each of the calendar years are:

- academic year of student's registration
- promotion code at the end of the year
- faculty of registration during the year
- new degree code if student has changed degrees
- number of credits at the end of the year

There is a field titled "first year" that indicates the calendar year in which students were registered for the first time at the university.

The file contains 2496 students records.

An extract of this file displaying only *information relevant to 1990* is shown below: Similarly to the file in a), the extract has been made from the file in *Statistica* in such way as to best fit the MS Word format of this document. An original row-column structure arranged by cases and variables is transposed and printed in portrait on the page while the original structure is preserved in the file.

StudentIdnt	abrhir001	abrter001	afrjud001	agnian001	ahtlem001	allben001
Acad_yr0	1	1	1	1	1	1
Ama104w0						
Bot102s0						
Bot103f0	62		60		55	
Cem001w0				68		
Cem100w0	56					
Cem101w0			60		47	
Cem102w0						45
Credit0	4	3..5	4	0	4	3..5
Csc105w0		56				
Eco102w0		54				
Egs100w0						
Faculty0	SCIB01	SCIB01	SCIB01	SCIB01	SCIB01	SCIB01
Gly105f0						
Mth101f0			55		59	
Mth101s0						
Mth103h0				67		
Mth105w0	74	50				58
Mth106w0						
Newdeg0	NONE	NONE	NONE	NONE	NONE	NONE
Phy001w0				61		
Phy100w0	70		70		55	
Phy104w0						54
Prncde0	CON	CON	CON	CON	CON	CON
Psy101w0						
Sms100w0						
Sta100s0			57		66	
Univu0	0	1	0	1	0	1
Zoo001h0				61		
Zoo103s0	65		61		50	
Zoo200f0						
Zoo201s0						

c) COMBINED UNIFIED FILE

This file contains the combined and matched 2496 records of the preceding sections. Extra fields have been created in this file incorporating details already present in the file to hold additional information required for particular analyses e.g. number of years taken to graduate or combining specific courses information across all the years of analysis. The file was also saved in both *Statistica* and *ASCII* formats for analyses in *Statistica* and *FIRM*.

d) "UNIMPORTANT" MATRIC AND UNIVERSITY DETAILS FILES

Both MATRICU and UNIVU files essentially preserved the format of the original *Excel* files (as described in section 1.2). They were created to ensure that none of the original data would be irretrievably lost in the transformation process, and if required eliminated data would be readily available at a later stage.

1.6. CREATION OF THE UNIFIED DATA SET: SUMMARY

DIAGRAM 1.6 STAGES IN PRODUCING UNIFIED DATA SET

STAGE 1: INITIAL EXTRACTION OF DATA

DATA SOURCE

Excel files of data on Science Faculty's students in the years 1990 to 1997 were extracted from the ADP Unit database. The ADP Unit database was the most reliable and readily available source of student data at UCT.

The ADP Unit database had a particularly awkward format and complex configuration, which dictated the final form of the *Excel* files.

TYPES OF FILES

Two types of *Excel* files were extracted:

- 8 files with biographical and matric subjects data
- 33 files with university courses data

PROBLEMS

Preliminary investigation revealed that unifying and then manipulating the files in any statistical package would be particularly troublesome as multiple records occurred for each student, one record per subject or course in each file.

STAGE 2: DESIGNING UNIFIED DATA SET

MAIN CONSIDERATIONS

A unified data set ought to serve the following purposes: ensure ease of data manipulation and thus manageability of statistical analyses; facilitate updates to data set, and thereby permit similar investigations in the future.

Thus there ought to be one record per student containing all the relevant biographical, matric subjects and university courses details.

DECISIONS TAKEN

Due to variety and volume of details stored in the original files, resulting from uniqueness of students' profiles, decisions were made on reconciling the original and unified data formats i.e. inclusion/retention of specific matric and university details in the unified data set.

These decisions were taken after consultation with the Science Faculty Officer and were mostly influenced by either the number of students taking a particular matric subject or university course or Faculty's likely desire to address a particular issue.

ADDITIONAL COMMENTS

The design was a complex issue and once completed, ensured that the coding was relatively simple.

Checking for inconsistencies and inaccuracies was a time consuming process.

STAGE 3: COMPUTER PROGRAM TO UNIFY DATA SET

COMPUTER PROGRAM

Computer program to unify the data in the original 41 files to 4 files and one combined "important" details file was written by Daniel Franco of ITS at UCT.

The coding and database platforms were DOS based dBASE III+.

This development was financially supported by the Faculty of Science with the approval of the Dean of the Faculty Prof. V.C Moran.

FILES DEVELOPED

Upon consultation with ITS Department at UCT, the issues from STAGE 2 were addressed by developing the following 4 files:

- 2 files storing "important" matric and "important" university details exhibiting one record per student
- 2 files storing "unimportant" matric and "unimportant" university details exhibiting the original files' format

COMMENTS ON DEVELOPMENT

The development was undertaken on a Pentium MMX 166 MHz PC.

Matric data could be regenerated easily for additional years of study, unless there have been changes in the matric subject/language/educational department codes used.

University data could also be regenerated relatively easily. However it is a time consuming process (being both Input / Output intensive) taking at least 36 hours on a PC which was entirely dedicated to the task. This time period is the function of the number of records needing to be processed and the way in which they are processed and combined from multiple records into a single record, allowing all students to be represented in the data for all the years included in the study.

STAGE 4: ULTIMATE FILE USED FOR ANALYSIS IN STATISTICA AND FIRM

UNIFIED FILE

The "important" matric and "important university details files produced in STAGE 3 were eventually combined in *Statistica* to make one unified file ready for statistical analysis.

Since the two components of the unified file were essentially text files some additional (sometimes even manual) changes had to be made to the file in *Statistica* and *Firm* to make it operational in these statistical packages. This manipulation was mostly performed using *Excel*.

For certain analyses performed in either in *Statistica* or *Firm* only some of the fields from the unified file were used. This focus resulted in forming several smaller versions of the unified file, essentially for the ease of manipulation and in order to speed up calculations.

OTHER CONCERNS

Despite the fact that particular care was taken in the design stages of this exercise to ensure that all details of interest were readily available from the unified file, the only way of obtaining some of that additional information was by non-trivial process of creating extra variables in *Statistica* based on details from the file e.g.

- number of years taken to graduate
- fields with marks for particular university courses completed in all the eight years of study (as opposed to the calendar year specific information readily available in the unified file)
- creating unified data file with information specific to SFP Programme students only

UNIFIED DATA SET: FINAL COMMENTS

If one wished to include additional years of study (beyond 1997) in the current database, the database tables would have to be modified. The major problem is that the data currently consists of over 620 fields and most platforms do not support that number of fields. To overcome this obstacle, a strategy to combine fields into one was adopted. Each additional year of study would increase the number of fields by approximately 110 fields. The overall number of fields is dependent on the number of university courses of specific interest to the study.

Should any subsequent studies require a different set of course, this requirement would affect the code and the database structure.

To develop a holistic system independent of the number of years of study, the number of courses of specific interest, or any other specific requirements, would take many more hours than were used in the current exercise. To create (in particular) a system which is independent of any programmer intervention may require several months (full time).

CHAPTER 2: RELATIONSHIPS BETWEEN ACADEMIC PERFORMANCE IN MATRIC AND AT UNIVERSITY

The nature and scope of the relationships between students' performance in matric and at university is multi-featured. Thus isolating the *variables*¹ of interest in the current data set will assist in recognising those features. This procedure will also facilitate identifying the appropriate statistical techniques required for the analysis of related issues.

The variables of interest in this research can be assigned, according to the roles they are assigned in the analysis, into two distinct groups, namely, *Explanatory* and *Response variables*.

¹ Variables represent features that can assume a range of categorical or numerical outcomes (or values) in an observed sample. The observed values are recorded and manipulated in any statistical research. The variables may differ in the roles played in the investigation (i.e. explanatory vs. response variables) and the types of measurement applied to them. Thus, each variable can be classified into one of the following types:

- Nominal-each outcome is allocated one of several specified mutually exclusive categories
- Binary-special type of a nominal variable with only two categories
- Ordinal-each outcome is allocated to one of a natural order or ranking of mutually exclusive categories
- Interval-if each outcome is on a numerical scale which admits meaningful differences
- Ratio-if each outcome is on a numerical scale which admits zero and meaningful ratios

Within numerical variables of interval and ratio types we may distinguish the following variable forms:

- *Discrete*- if measurements can assume only distinct specified values.
- *Continuous*- if measurements can assume any value on a continuum.

2.1 VARIABLES

2.1.1 EXPLANATORY VARIABLES

Explanatory variables² (also called independent or predictor variables) are those variables whose realisable outcomes are treated as predictors for outcomes of other variables. The explanatory variables are free to vary in response to other variables; they can be viewed as non-random observations, fixed by observation in the study design.

In this research, the explanatory variables reflect factors related to different matric and personal characteristics and factors associated with academic performance at university. Thus, the following variables can be assigned an explanatory role in the current data set:

1. Matric (Unweighted) Point-Score
2. Faculty (Weighted) Point-Score
3. Performance in individual matric subjects
4. Gender
5. Colour
6. School Attended
7. Matric Department
8. Home Language
9. Performance in "important" university courses³ (as precursor of performance in other university courses)

Table 2.1.1 on the following page summarises these nine explanatory variables presenting their types and the range of possible outcomes or values.

² Within the family of explanatory variable we may distinguish a group of variables referred to as covariates. This nomenclature is frequently used in the variance components analysis where covariates denote those explanatory continuous variables likely to be correlated with the dependent variable of interest.

³ "Important" subjects are those that are compulsory for completion of the B.Sc. degree and any particular major in the Faculty of Science at UCT.

TABLE 2.1.1**EXPLANATORY VARIABLES**

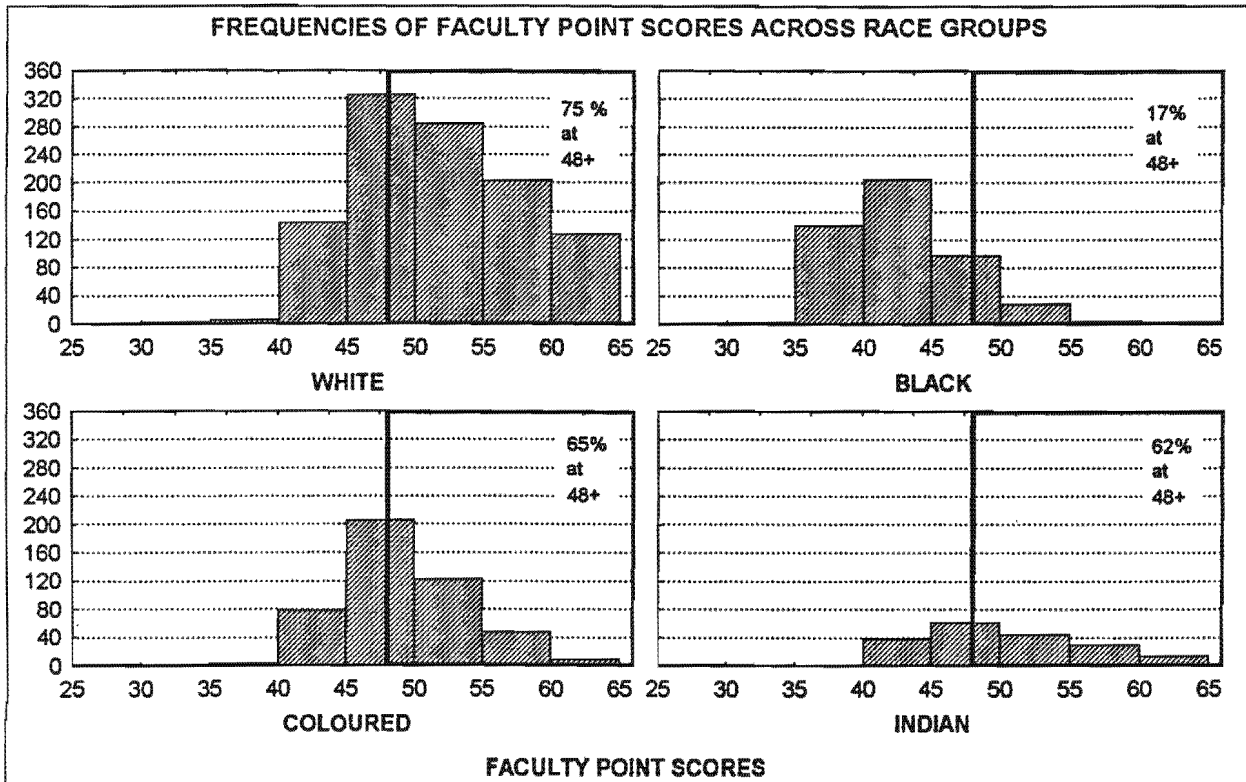
EXPLANATORY VARIABLES	VARIABLE TYPE	ADMISSIBLE VALUES																		
<p>Matric (Unweighted) Point-Score</p> <p><i>(sum of scores for six best matric subjects)</i></p>	Ordinal	<p>Counts (from 12 to 48; calculated according ⁴ to TABLE A)</p> <p><u>TABLE A</u></p> <table border="1"> <thead> <tr> <th>symbol</th> <th>HG</th> <th>SG</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>8</td> <td>6</td> </tr> <tr> <td>B</td> <td>7</td> <td>5</td> </tr> <tr> <td>C</td> <td>6</td> <td>4</td> </tr> <tr> <td>D</td> <td>5</td> <td>3</td> </tr> <tr> <td>E</td> <td>4</td> <td>2</td> </tr> </tbody> </table>	symbol	HG	SG	A	8	6	B	7	5	C	6	4	D	5	3	E	4	2
		symbol	HG	SG																
		A	8	6																
B	7	5																		
C	6	4																		
D	5	3																		
E	4	2																		
<p>Faculty (Weighted) Point-Score</p> <p><i>(sum of scores for six best matric subjects with the scores for maths and the best of physical science or biology doubled)</i></p>	Ordinal	<p>Counts (from 16 to 64; calculated according to TABLE A with specified doublings)</p>																		
<p>Performance in Individual Matric Subjects</p> <p><i>(maths, physical science, biology, history, geography, computer science and languages: English, Afrikaans, Sotho, Swazi, Tsonga, Tswana, Zulu, Xhosa)</i></p>	Ordinal	Symbols (A to E) or Counts																		

⁴ Table A has been adapted from the Faculty of Science booklet and represents the way in which the Faculty converts matric symbols into points.

<p>Performance in important university courses (as precursors of other university courses)</p>	Ordinal	Actual percentages for single courses obtained by students
	Binary	<p>Two categories only:</p> <p>1=success: if a student scored 50% or more</p> <p>0=failure: if a student scored less than 50%</p> <p>The cut-off point may be changed to any desired percentage</p>

As mentioned in chapter 0, admission into the Faculty of Science is based on the faculty point scores. The minimum faculty point score of 48 is required for certain admission. However students had been admitted into the Faculty with lower faculty point scores.

The histograms below display the frequencies⁷ of the faculty point scores across the four race groups. In each of the histograms a block has been drawn to mark off the number of students with the faculty point scores of 48 and more.



We note that most of the White students had been admitted with the faculty point score of 48 and more. The proportions of both the Coloured and Indian students admitted with the faculty point scores of 48 and more are also high. However most of the Black students had been admitted with the faculty point scores of less than 48.

Additionally we note that the set of students analysed is clearly dominated by the White students.

⁷ The histograms have been produced across the four race groups, as the frequencies are important for understanding and discussion of results presented in later chapters.

2.2.2 RESPONSE VARIABLES

Response variables (also called dependent variables) are those in which the study interests lies, particularly in determining the extent to which the observed and observable responses are affected by explanatory variables.

In this research, the response variables are factors reflecting academic performance at university. Thus, it is possible to distinguish the following response variables in the data set:

1. Number of university courses passed
2. Performance in "important" university courses
3. Number of years taken to qualify

Table 2.2.2 on the next page gives a summary of the identified response variables, their types and the possible values of these variables.

TABLE 2.2.2

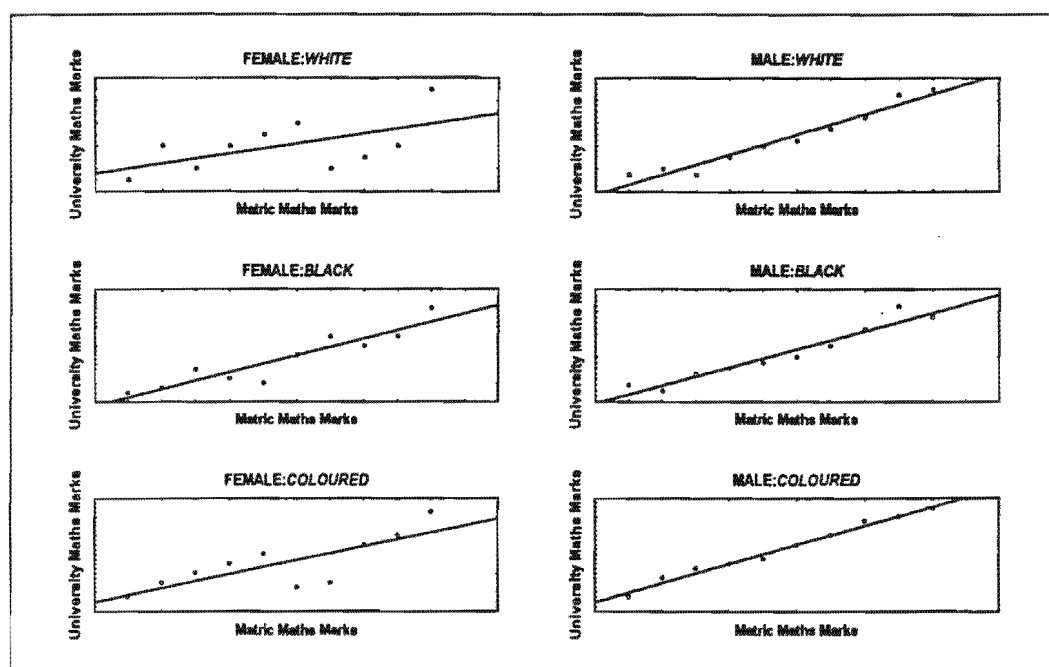
RESPONSE VARIABLES

RESPONSE VARIABLES	VARIABLE TYPE	RANGE OF VALUES (for fuller performance definitions refer to chapter 0)
<p><i>Number of courses passed (for a student with a specific registration year status)</i></p>	Ordinal	Actual number of courses passed
	Binary	<p>Two categories only:</p> <p>1=Success: if a student passed 2 or more full year courses or the equivalent 0=Failure: if a student passed less than 2 full-year courses or the equivalent</p> <p>The cut of point of 2 full-year courses or the equivalent could be changed to any desired number of exams passed i.e.: 0; 0.5; 1; 1.5; 2; 2.5; 3; 3.5; 4 or more</p> <p>or</p> <p>Two categories only:</p> <p>1=Success: average of 50% or more for a minimum of 2 full-year courses or the equivalent 0=Failure: average of less than 50% for a minimum of 2 full-year courses or the equivalent</p>
<p><i>Performance in "important" subjects (in a specific academic year)</i></p>	Ordinal	Actual percentages for single courses obtained by students
	Binary	<p>Two categories only:</p> <p>1=Success: if a student scored 50% or more 0=Failure: if a student scored less than 50%</p> <p>The cut-off point could be changed to any desired percentage</p>
<p><i>Number of years taken to qualify</i></p>	Ordinal	Actual number of years taken to qualify
	Binary	<p>Two categories:</p> <p>1=Success: graduation after 3 years 0=Failure: graduation after 4 or more years</p> <p>or</p> <p>1=Success: graduation after 3 or 4 years 0=Failure: graduation after more than 4 years</p>

Analyses may use these variables (both explanatory and response variables in Table 2.1.1 and Table 2.2.2 respectively) at several levels of complexity: univariate, bivariate or multivariate.

It is always possible to choose from an array of statistical techniques that will allow multivariate combinations of different variables across the groups in the tables. Some of these multivariate analyses will be constructed⁸. For instance, it is possible to examine contrasts of the relationship between the Faculty point-score and first-year university maths scores, over the subgroups constituted by gender and colour. A graphical representation of possible multivariate relationships is presented in Diagram 2.2 below, on the assumption that the maths marks are available for each student. Such a set of graphs is most directly meaningful when the graphs rendered easily comparable by the use of common origins and common scales.

DIAGRAM 2.2
POSSIBLE MULTIVARIATE ANALYSES



The import of such analyses is a description of a set of observed relationships. The underlying mechanisms by which such patterns arise (e.g. differential historical resource access across racial groups) are matters of interpretation and contextual inference. The existence of contrasting patterns may indicate the desirability of some policy to attain specific goals of common achievement across all groups. Policy is likely to be most appropriate when it is formulated to address the plausible causes of the contrasting phenomena presented in the graphs.

⁸ See Table 2.3 for the particular issues investigated in this project and the appropriate techniques used.

2.2 ISSUES IN THE ANALYSIS OF THE RELATIONSHIPS BETWEEN ACADEMIC PERFORMANCE IN MATRIC AND AT UNIVERSITY

X | In the context of this study having nominated the explanatory and response variables in the available data, we are in a position to identify various possible links between matric and university performance. It is important to point out that connections between observed variables can be explored not only between matric and any of the three academic years of Science student's academic career, but also between the consecutive university years themselves.

Table 2.2 presents in detail the various associations chosen as a structure of analysis for this study between the four levels of academic performance (matric and three university years).

TABLE 2.2

ISSUES IDENTIFIED IN THE ANALYSIS OF THE RELATIONSHIPS BETWEEN ACADEMIC PERFORMANCE IN MATRIC AND AT UNIVERSITY

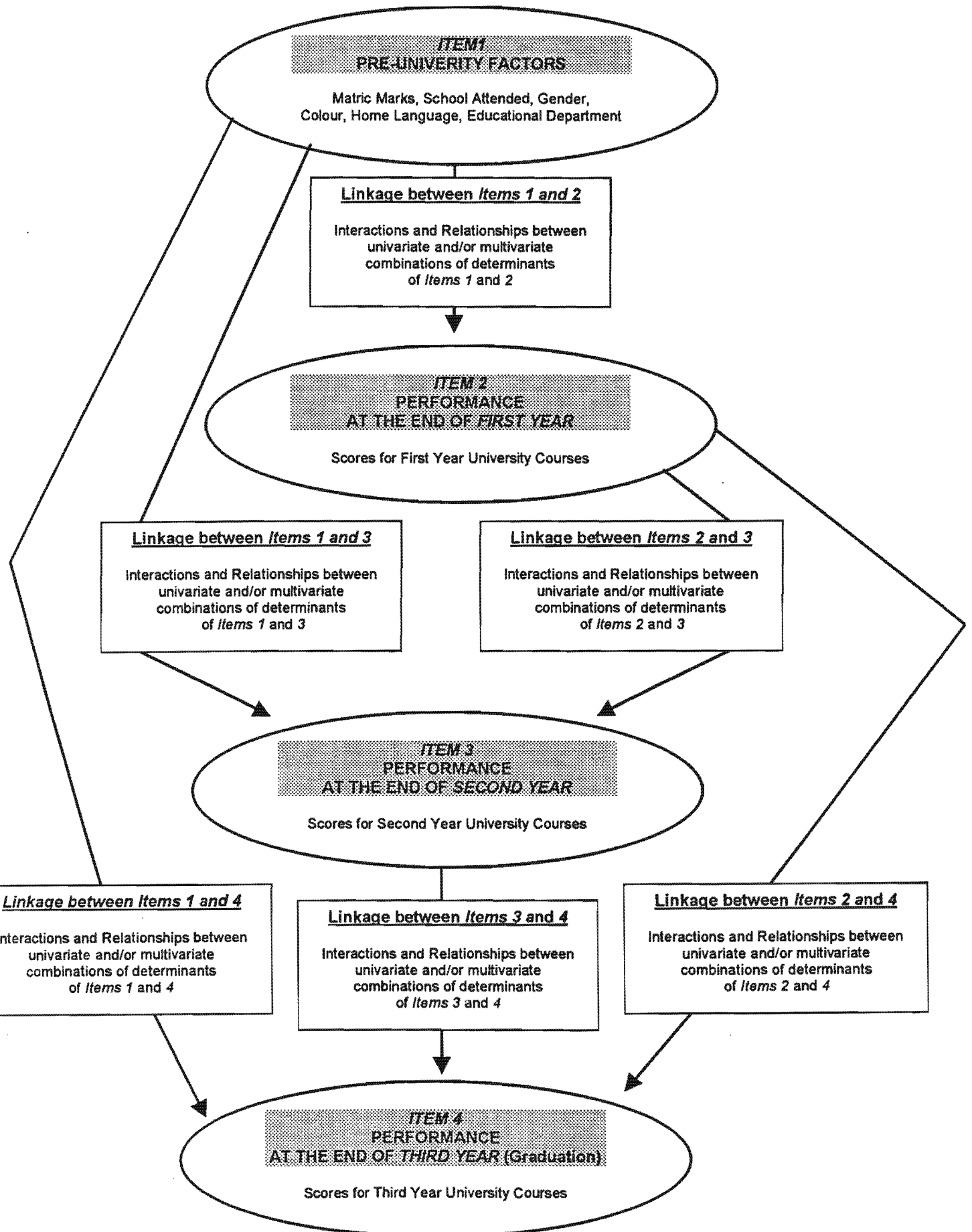
ISSUES IDENTIFIED ON THE BASIS OF ANNUAL INTAKE AND AGGREGATED OVER ALL YEARS				
OVERALL AND AT PARTICULAR COURSES		FIRST YEAR	SECOND YEAR	THIRD YEAR
		Interactions and relationships between <i>pre-university factors</i> and <i>success at UCT</i> in first academic year	Interactions and relationships between <i>pre-university factors</i> and <i>success at UCT</i> in second academic year	Interactions and relationships between <i>pre-university factors</i> and <i>success at UCT</i> in third academic year
	1.	1.1. Association between the <i>weighted matric point-score</i> and <i>success at UCT</i> in first academic year (different definitions possible) (i.e. how well does weighted matric point-score predicts success at university?)	1.2. Association between the <i>weighted matric point-score</i> and <i>success at UCT</i> in second academic year (different definitions possible)	1.3. Association between the <i>weighted matric point-score</i> and <i>success at UCT</i> in third academic year (different definitions possible)
	2.	2.1. Relevance of <i>doubling the scores for maths and the better of physical science or biology</i> and its relationship to <i>success at UCT</i> in first academic year (different definitions possible) (i.e. having doubled the scores for these subjects, are we in a better position in distinguishing between prospective successful and unsuccessful students)	2.2. Relevance of <i>doubling the scores for maths and the better of physical science or biology</i> and its relationship to <i>success at UCT</i> in second academic year (different definitions possible)	2.2. Relevance of <i>doubling the scores for maths and the better of physical science or biology</i> and its relationship to <i>success at UCT</i> in third academic year (different definitions possible)

OVERALL AND AT PARTICULAR COURSES	3.	3.1. Association between <i>individual matric subjects</i> (such as maths, physical science, biology, English and other language (school subject) and <i>success at UCT</i> in first academic year (different definitions possible)	3.2. Association between <i>individual matric subjects</i> (such as maths, physical science, biology, English and other language (school subject) and <i>success at UCT</i> in second academic year (different definitions possible)	3.3. Association between <i>individual matric subjects</i> (such as maths, physical science, biology, English and other language (school subject) and <i>success at UCT</i> in third academic year (different definitions possible)
	4.	4.1. Association between <i>other pre-university factors</i> (such as gender, colour, matric dept.(old and new), school attended) and <i>success at UCT</i> in first academic year (different definitions possible)	4.2. Association between <i>other pre-university factors</i> (such as gender, colour, matric dept.(old and new), school attended) and <i>success at UCT</i> in second academic year (different definitions possible)	4.3. Association between <i>other pre-university factors</i> (such as gender, colour, matric dept.(old and new), school attended) and <i>success at UCT</i> in third academic year (different definitions possible)
	5.	_____	5.1. Association between <i>success</i> in first academic year (different definitions possible) and <i>success</i> in second academic year (different definitions possible)	5.2. Association between <i>success</i> in first academic year (different definitions possible) and <i>success</i> in third academic year (different definitions possible)
	6.	_____	_____	6.1. Association between <i>success</i> in second academic year (different definitions possible) and <i>success</i> in third academic year (different definitions possible)

	<p>7.</p> <p>7.1. Association within categories (gender, race, matric departments, home language) between matric point-scores and <i>success</i> in first academic year (different definitions possible)</p>	<p>7.2. Association within categories (gender, race, matric departments, home language) between matric point-scores and <i>success</i> in second academic year (different definitions possible)</p>	<p>7.3. Association within categories (gender, race, matric departments, home language) between matric point-scores and <i>success</i> in third academic year (different definitions possible)</p>
--	--	--	---

The diagram on the following page is a systematic overview of the postulated relationships between the four levels of academic performance.

DIAGRAM 2.3
GRAPHICAL REPRESENTATION OF THE PROBLEM STUDIED



2.3. ANALYSES IN THE PROJECT

Issues specified in Table 2.1.1, Table 2.1.2 and Table 2.2 provide a wide range of plausible analyses to be performed with different choices of explanatory and response variables. Table 2.3 specifies the set of expected analyses to be embarked on in the project.

For each issue identified in Table 2.2, a corresponding entry in Table 2.3 lists:

- appropriate statistical methods ⁸
- chosen explanatory and response variables

The reader may note that Table 2.3 was prepared in the initial stage of the project as guide to be used for analysing the questions of interest after the required data structure had been made readily accessible and manageable. However due to the time constraints and other complications mentioned in the body of the dissertation, the analyses in the following sections were not explicitly addressed in this thesis 1.2, 2.2, 3.2, 3.3, 6, 7.2 and 7.3.

The reader may also note that the issues explicitly described in section 4 are addressed in the analyses of the previous sections.

⁸ Three statistical methods: Measures of Association, Generalized Linear Modelling Techniques and Formal Inference-Based Recursive Modelling "FIRM" are used in this project. They are described in detail in Appendices A, B and C respectively.

TABLE 2.3
SPECIFIC ANALYSES

ISSUES INVESTIGATED	METHODS	EXPLANATORY VARIABLE(S)	RESPONSE VARIABLE(S)
1.1 Association between the weighted matric point-score and success at UCT in first academic year	1.1.1 Measures of Association - data for individual years and data combined across all the years 90-97 - data arranged in contingency tables - output: 5 measures ⁹ of association and associated variances	1.1.1 Faculty Point Score	1.1.1 Number of exams passed in first academic year (i.e. number of credits obtained) (separate analyses for ordinal and binary variable types)
	1.1.2 GLIM ¹⁰ : - data for individual years and data combined across all the years 90-97 - output: regression equation with significant explanatory variables	1.1.2 Faculty Point Score; Individual Matric Subjects; Matric Dept ¹¹ ; Gender; Colour Language	1.1.2 Number of exams passed in first academic year (separate analyses for ordinal and binary variable types)
	1.1.3 Firm - data combined across all the years 90-97 - output: dendrogram	1.1.3 Faculty Point Score; Individual Matric Subjects; Matric Dept.; Gender; Colour Language; Year of Enrolment	1.1.3 Number of exams passed in first academic year (separate analyses for ordinal and binary variable types)

⁹ The five measures of association are Pearson product-moment correlation coefficient, Spearman correlation coefficient, Kendall's Tau, Somers' *D* and Gamma Statistics.

¹⁰ The utilised technique is a part of the Generalized Linear Modelling Procedures (GLIM).

¹¹ Also referred to as educational department.

	<p>significant explanatory variables</p> <p>1.3.3 Firm</p> <ul style="list-style-type: none"> - data combined across all the years 90-97 - output: dendrogram 	<p>1.3.3</p> <p>Faculty Point Score; Individual Matric Subjects; Matric Dept.; Gender; Colour; Language; Year of Enrolment</p>	<p>1.3.3</p> <p>Number of years taken to qualify</p> <p>(separate analyses for ordinal and binary variable types)</p>
--	---	--	---

<p>2.1 Relevance of doubling the scores for maths and the better of physical science or biology and its relationship to success at UCT in first academic year</p>	<p>2.1.1 Measures of Association</p> <ul style="list-style-type: none"> - data for individual years and data combined across all the years 90-97 - data arranged in contingency tables - output: 5 measures of association and associated variances <p>2.1.2 GLIM:</p> <ul style="list-style-type: none"> - data for individual years and data combined across all the years 90-97 - output: regression equation with significant explanatory variables <p>2.1.3. Firm</p> <ul style="list-style-type: none"> - data combined across all the years 90-97 - output: dendrogram 	<p>2.1.1</p> <p>Matric Point Score</p> <p>2.1.2</p> <p>Matric Point Score; Individual Matric Subjects; Matric Dept.; Gender; Colour; Language</p> <p>2.1.3</p> <p>Matric Point Score; Individual Matric Subjects; Matric Dept.; Gender; Colour; Language; Year of Enrolment</p>	<p>2.1.1</p> <p>Number of exams passed in first academic year</p> <p>(separate analyses for ordinal and binary variable types)</p> <p>2.1.2</p> <p>Number of exams passed in first academic year</p> <p>(separate analyses for ordinal and binary variable types)</p> <p>2.1.3</p> <p>Number of exams passed in first academic year</p> <p>(separate analyses for ordinal and binary variable types)</p>
<p>2.2 Relevance of doubling the scores for maths and the better of physical science or biology and its relationship to success at UCT in second academic year</p>	<p>2.2.1 Measures of Association</p> <ul style="list-style-type: none"> - data for individual years and data combined across all the years 90-97 - data arranged in contingency tables - output: 5 measures of association and associated variances <p>2.2.2 GLIM:</p> <ul style="list-style-type: none"> - data for individual years and data combined across all the years 90-97 - output: regression equation with 	<p>2.2.1</p> <p>Matric Point Score</p> <p>2.2.2</p> <p>Matric Point Score; Individual Matric Subjects; Matric Dept.; Gender; Colour; Language;</p>	<p>2.2.1</p> <p>Number of exams passed in second academic year</p> <p>(separate analyses for ordinal and binary variable types)</p> <p>2.2.2</p> <p>Number of exams passed in second academic year</p> <p>(separate analyses for ordinal and binary</p>

	<p>significant explanatory variables</p> <p>2.2.3 Firm</p> <ul style="list-style-type: none"> - data combined across all the years 90-97 - output: dendrogram 	<p>2.2.3</p> <p>Matric Point Score; Individual Matric Subjects; Matric Dept.; Gender; Colour; Language; Year of Enrolment</p>	<p>variable types)</p> <p>2.2.3</p> <p>Number of exams passed in second academic year</p> <p>(separate analyses for ordinal and binary variable types)</p>
<p>2.3 Relevance of doubling the scores for maths and the better of physical science or biology and its relationship to success at UCT in third academic year</p>	<p>2.3.1 Measures of Association</p> <ul style="list-style-type: none"> - data for individual years and data combined across all the years 90-97 - data arranged in contingency tables - output: 5 measures of association and associated variances <p>2.3.2 GLIM:</p> <ul style="list-style-type: none"> - data for individual years and data combined across all the years 90-97 - output: regression equation with significant explanatory variables <p>2.3.3. Firm</p> <ul style="list-style-type: none"> - data combined across all the years 90-97 - output: dendrogram 	<p>2.3.1</p> <p>Matric Point Score</p> <p>2.3.2</p> <p>Matric Point Score; Individual Matric Subjects; Matric Dept.; Gender; Colour; Language</p> <p>2.3.3</p> <p>Matric Point Score; Individual Matric Subjects; Matric Dept.; Gender; Colour; Language; Year of Enrolment</p>	<p>2.3.1</p> <p>Number of years taken to qualify</p> <p>(separate analyses for ordinal and binary variable types)</p> <p>2.3.2</p> <p>Number of years taken to qualify</p> <p>(separate analyses for ordinal and binary variable types)</p> <p>2.3.3</p> <p>Number of years taken to qualify</p> <p>(separate analyses for ordinal and binary variable types)</p>

Methods

Explan Var

Resp Var

X

	Methods	Explan Var	Resp Var
3.1 Association between performance individual matric subjects (such as maths, physical science, biology , English and other language (school subject) and success at UCT in first academic year	<p>3.1.1 Measures of Association</p> <ul style="list-style-type: none"> - data for individual years and data combined across all the years 90-97 - data arranged in contingency tables - output: 5 measures of association and associated variances <p>3.1.2 GLIM:</p> <ul style="list-style-type: none"> - data for individual years and data combined across all the years 90-97 - output: regression equation with significant explanatory variables <p>3.1.3 Firm</p> <ul style="list-style-type: none"> - data combined across all the years 90-97 - output: dendrogram 	<p>3.1.1</p> <p>Performance in a particular matric subject (maths, physical science, chemistry, biology)</p> <p>3.1.2</p> <p>Performance in a particular matric subject (maths, physical science, chemistry, biology) Matric Dept.; Gender; Colour; Language</p> <p>3.1.3</p> <p>Performance in a particular matric subject (maths, physical science, chemistry, biology); Matric Dept.; Gender; Colour; Language; Calendar Year in which a university course was completed</p>	<p>3.1.1</p> <p>Performance in a particular university first year course (maths, physical science, chemistry, biology, statistics)</p> <p>(separate analyses for ordinal and binary variable types)</p> <p>3.1.2</p> <p>Performance in a particular university first year course (maths, physical science, chemistry, biology, statistics)</p> <p>(separate analyses for ordinal and binary variable types)</p> <p>3.1.3</p> <p>Performance in a particular university first year course (maths, physical science, chemistry, biology, statistics)</p> <p>(separate analyses for ordinal and binary variable types)</p>
3.2 Association between individual matric subjects (such as maths, physical science, biology , English and other language (school subject) and success at UCT in second academic year	<p>3.2.1 Measures of Association</p> <ul style="list-style-type: none"> - data for individual years and data combined across all the years 90-97 - data arranged in contingency tables - output: 5 measures of association and associated variances 	<p>3.2.1</p> <p>Performance in a particular matric subject (maths, physical science, chemistry, biology)</p>	<p>3.2.1</p> <p>Performance in a particular university second-year course</p> <p>(separate analyses for ordinal and binary variable types)</p>

	<p>3.2.2 GLIM:</p> <ul style="list-style-type: none"> - data for individual years and data combined across all the years 90-97 - output: regression equation with significant explanatory variables <p>3.2.3 Firm</p> <ul style="list-style-type: none"> - data combined across all the years 90-97 - output: dendrogram 	<p>3.2.2</p> <p>Performance in a particular matric subject (maths, physical science, chemistry, biology, statistics); Matric Dept.; Gender; Colour; Language</p> <p>3.2.3</p> <p>Performance in a particular matric subject (maths, physical science, chemistry, biology); Matric Dept.; Gender; Colour; Language; Calendar Year in which a university course was completed</p>	<p>3.2.2</p> <p>Performance in a particular university second-year course</p> <p>(separate analyses for ordinal and binary variable types)</p> <p>3.2.3</p> <p>Performance in a particular university second-year course</p> <p>(separate analyses for ordinal and binary variable types)</p>
<p>3.3 Association between individual matric subjects (such as maths, physical science, biology, English and other language (school subject) and success at UCT in third academic year</p>	<p>3.3.1 Measures of Association</p> <ul style="list-style-type: none"> - data for individual years and data combined across all the years 90-97 - data arranged in contingency tables - output: 5 measures of association and associated variances <p>3.3.2 GLIM:</p> <ul style="list-style-type: none"> - data for individual years and data combined across all the years 90-97 - output: regression equation with significant explanatory variables 	<p>3.3.1</p> <p>Performance in a particular matric subject (maths, physical science, chemistry, biology)</p> <p>3.3.2</p> <p>Performance in a particular matric subject (maths, physical science, chemistry, biology) Matric Dept.; Gender; Colour; Language</p>	<p>3.3.1</p> <p>Performance in a particular university third-year course</p> <p>(separate analyses for ordinal and binary variable types)</p> <p>3.3.2</p> <p>Performance in a particular university third-year course</p> <p>(separate analyses for ordinal and binary variable types)</p>

	<p>3.3.3 Firm</p> <ul style="list-style-type: none"> - data combined across all the years 90-97 - output: dendrogram 	<p>3.3.3</p> <p>Performance in a particular matric subject (maths, physical science, chemistry, biology) Matric Dept.; Gender; Colour; Language; Calendar Year in which a university course was completed</p>	<p>3.3.3</p> <p>Performance in a particular university third- year course</p> <p>(separate analyses for ordinal and binary variable types)</p>
--	---	--	---

<p>4.1 Association between other pre-university factors (such as gender, colour, matric dept.(old and new), school attended) and success at UCT in first academic year</p>	<p>4.1.1 GLIM:</p> <ul style="list-style-type: none"> - data combined across all the years 90-97 - output: regression equation with significant explanatory variables <p>4.1.2 Firm</p> <ul style="list-style-type: none"> - data combined across all the years 90-97 - output: dendrogram 	<p>4.1.1</p> <p>Gender, Colour, Matric Dept.; School Attended</p> <p>4.1.2</p> <p>Gender, Colour, Matric Dept.; School Attended</p>	<p>4.1.1</p> <p>Number of courses passed in first academic year</p> <p>(separate analyses for ordinal and binary variable types)</p> <p>4.1.2</p> <p>Number of courses passed in first academic year</p> <p>(separate analyses for ordinal and binary variable types)</p>
<p>4.2 Association between other pre-university factors (such as gender, colour, matric dept.(old and new), school attended) and success at UCT in second academic year</p>	<p>4.2.1 GLIM:</p> <ul style="list-style-type: none"> - data combined across all the years 90-97 - output: regression equation with significant explanatory variables <p>4.2.2 Firm</p> <ul style="list-style-type: none"> - data combined across all the years 90-97 - output: dendrogram 	<p>4.2.1</p> <p>Gender, Colour, Matric Dept.; School Attended</p> <p>4.2.2</p> <p>Gender, Colour, Matric Dept.; School Attended</p>	<p>4.2.1</p> <p>Number of courses passed in second academic year</p> <p>(separate analyses for ordinal and binary variable types)</p> <p>4.2.2</p> <p>Number of courses passed in second academic year</p> <p>(separate analyses for ordinal and binary variable types)</p>
<p>4.3 Association between other pre-university factors (such as gender, colour, matric dept.(old and new), school attended) and success at UCT in second academic year</p>	<p>4.3.1 GLIM:</p> <ul style="list-style-type: none"> - data combined across all the years 90-97 - output: regression equation with significant explanatory variables 	<p>4.3.1</p> <p>Gender, Colour, Matric Dept.; School Attended</p>	<p>4.3.1</p> <p>Number of courses passed in third academic year</p> <p>(separate analyses for ordinal and binary variable types)</p>

	<p>4.3.2 Firm</p> <ul style="list-style-type: none"> - data combined across all the years 90-97 - output: dendrogram 	<p>4.3.2</p> <p>Gender, Colour, Matric Dept.; School Attended</p>	<p>4.3.2</p> <p>Number of courses passed in third academic year</p>
--	---	--	--

<p>5.1 Association between success in first academic year and success in second academic year</p>	<p>5.1.1 Measures of Association</p> <ul style="list-style-type: none"> - data for individual years and data combined across all the years 90-97 - data arranged in contingency tables - output: 5 measures of association and associated variances 	<p>5.1.1</p> <p>Number of exams passed in first academic year</p> <p>(separate analyses for ordinal and binary variable types)</p>	<p>5.1.1</p> <p>Number of courses passed in second academic year</p> <p>(separate analyses for ordinal and binary variable types)</p>
<p>5.2 Association between success in first academic year and success in third academic year</p>	<p>5.2.1 Measures of Association</p> <ul style="list-style-type: none"> - data for individual years and data combined across all the years 90-97 - data arranged in contingency tables - output: 5 measures of association and associated variances 	<p>5.2.1</p> <p>Number of exams passed in second academic year</p> <p>(separate analyses for ordinal and binary variable types)</p>	<p>5.2.1</p> <p>Number of courses passed in third academic year</p> <p>(separate analyses for ordinal and binary variable types)</p>

<p>6. Association between success in second academic year and success in third academic year</p>	<p>6. Measures of Association</p> <ul style="list-style-type: none"> - data for individual years and data combined across all the years 90-97 - data arranged in contingency tables - output: 5 measures of association and associated variances 	<p>6.</p> <p>Number of exams passed in second academic year</p> <p>(separate analyses for ordinal and binary variable types)</p>	<p>6.</p> <p>Number of exams passed in third academic year</p> <p>(separate analyses for ordinal and binary variable types)</p>
--	---	--	---

<p>7.1 Association within categories (gender, race, matric departments, home language) between matric point-scores and success in first academic year</p>	<p>7.1.1 GLIM: Log-linear models</p> <ul style="list-style-type: none"> - data for individual years and data combined across all the years 90-97 - data cross-classified by explanatory variables and arranged in contingency tables - output: log-linear model 	<p>7.1.1</p> <p>Faculty Point Score and</p> <p>Gender; Race or Gender; Matric Dept. or Gender; Home Language</p>	<p>7.1.1</p> <p>Number of exams passed in first academic year</p> <p>(separate analyses for ordinal and binary variable types)</p> <p>or</p> <p>Performance in first year university courses: maths, physics, biology, statistics (separate analyses for ordinal and binary variable types)</p>
<p>7.2 Association within categories (gender, race, matric departments, home language) between matric point-scores and success in second academic year</p>	<p>7.2.1 GLIM: Log-linear models</p> <ul style="list-style-type: none"> - data for individual years and data combined across all the years 90-97 - data cross-classified by explanatory variables and arranged in contingency tables - output: log-linear model 	<p>7.2.1</p> <p>Faculty Point Score and</p> <p>Gender; Race or Gender; Matric Dept. or Gender; Home Language</p>	<p>7.2.1</p> <p>Number of exams passed in second academic year</p> <p>(separate analyses for ordinal and binary variable types)</p>
<p>7.3 Association within categories (gender, race, matric departments, home language) between matric point-scores and success in third academic year</p>	<p>7.3.1 GLIM: Log-linear models</p> <ul style="list-style-type: none"> - data for individual years and data combined across all the years 90-97 - data cross-classified by explanatory variables and arranged in contingency tables - output: log-linear model 	<p>7.3.1</p> <p>Faculty Point Score and</p> <p>Gender; Race or Gender; Matric Dept. or Gender; Home Language</p>	<p>7.3.1</p> <p>Number of years taken to qualify</p> <p>(separate analyses for ordinal and binary variable types)</p>

CHAPTER 3: RELATIONSHIPS BETWEEN POINT SCORES AND UNIVERSITY FIRST YEAR PERFORMANCE

In this chapter we examine the statistical relationships between the weighted *faculty* and the unweighted *matric* point scores and the academic performance of the incoming first year students in the first calendar year of their studies in the Faculty of Science at UCT. Presentation of these results follows the manner and order specified ¹ in Table 2.3 from chapter 2.

Ordinal ² and binary ³ definitions of the academic performance variable were used. Individual analyses were performed for the two definitions of the response variable under each of the three statistical methods ⁴ used, separately for the faculty and matric point scores. For each method, results are presented separately for each of the eight calendar years (1990-1997) and also for the eight years combined ⁵ data set.

¹ Refer to sections 1.1 and 2.1 in Table 2.3 in chapter 2.

² The following *ordinal definition* of the academic performance variable is used:

- number of full-year first year courses passed or the equivalents (i.e. the actual number of credits obtained) by an incoming first year student in the first calendar year of his or her study

³ The following *binary definition* of the academic performance variable is used:

two categories:

- an incoming first year student passed the 2 or more full-year first year courses in the first calendar year of his or her study (i.e. obtained at least 2 credits)
- an incoming first year student passed fewer than the 2 full-year first year courses in the first calendar year of his or her (i.e. obtained fewer than 2 credits)

⁴ The results have been developed using the following statistical methods:

- Measures of Association
- Generalised Linear Modelling
- Formal Inference-based Recursive Modelling (as implemented in the package FIRM).

These statistical methods are discussed in the Appendices. The reader is advised to become familiar with those discussions before examining the chapters presenting the statistical results.

⁵ The eight years of data were combined on the assumption that all the variables used in the statistical methods are comparable across the eight years in the study.

Predicament

Initially all the analyses presented in this and subsequent chapters were performed on the entire data set consisting of 2488 students (i.e. the set of the students in the regular B.Sc. and the SFP Programmes combined⁶). These initial results indicated clearly that the matric point score was a more significant and reliable predictor than the faculty point score in relation to the number of credits obtained by the first year students.

This finding prompted the researcher to compare the two types of scores to establish whether or not this difference in predictive strength was attributable to inappropriate weights being assigned to either matric maths or physical science or biology when structuring the faculty point score. In that exploration it was found that for some 239 students the recorded matric and faculty point scores were actually identical. This feature was contrary to the information provided from the ADP Unit and indicated a possible fault with some of the data recording procedures at UCT.

In the remaining time frame allocated to completing of the project it was not feasible to establish which of the point scores reflected correctly the matric profiles of the particular students. Thus, as an immediate solution to the encountered problem, these students were excluded from any further investigation involving either or both of the point scores. Thus, all the necessary analyses were repeated on the set consisting of only the 2249 students whose point scores had positive differences (faculty point score – matric point score). The outcomes of these analyses are discussed in the remaining sections of this chapter.

At a later point in the analysis output from FIRM (refer to sections 3.1.1 c)) signalled the presence of lower matric point score and high faculty point score for a small subset of students. These data cases were investigated. A maximal difference of 16 between the faculty and matric point scores occurs for students with distinctions in matric Mathematics, Physical Science and Biology. It is possible that original matric point scores were derived before the effect of re-marks was recorded by U.C.T. in the faculty point score alone. Thus some difference values higher than 16 were possible, but we report⁷ some eight cases for which the anomaly was severe. In general we had no further information on which to base corrections, these cases were not corrected and remained in the database.

⁶ Note that the separation of the B.Sc. and the SFP Programmes is taken into account only in the analysis of the individual first year level university courses.

⁷ A list of students in alphabetical order for whom the differences between point scores were severe.

STUDENT NUMBER	FACULTY POINT SCORE	MATRIC POINT SCORE	DIFFERENCE IN POINT SCORES
ADMREH001	50	27	23
ASTWEN001	56	31	25
DMNMIC006	50	13	37
HRXSUN001	45	24	21
MTLLEB002	47	27	21
NDLMGI001	44	22	22
NKSINN001	47	23	24

3.1 ASSOCIATION BETWEEN FACULTY POINT SCORE AND PERFORMANCE IN FIRST ACADEMIC YEAR AT UCT

3.1.1 MEASURES OF ASSOCIATION

a) Ordinal Response Analysis

Eight annual and one overall bivariate data sets were formed by the *faculty point-scores* of the incoming first-year Science students and the numbers of credits obtained by these students in the first calendar year of their study.

A contingency table view of the data is implicit in several measures of association presented in the nine tables below. Some of the measures of association presented here used a contingency table view of the bivariate data.

FACULTY POINT SCORE (X) and CREDITS IN FIRST YEAR (ordinal) (Y)

YEAR : 1990, valid cases: 236		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.66	< 0.05
Spearman correlation coefficient	0.66 (t=13.50)	< 0.05
Kendall's Tau	0.51 (Z=11.78)	< 0.05
Somers' D Statistics	X/Y=0.55 Y/X=0.38	< 0.05 < 0.05
Gamma Statistic	0.57 (Z=11.78)	< 0.05
Pearson Chi-square Statistic	128.17	0.1412

YEAR : 1991, valid cases: 267		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.49	< 0.05
Spearman correlation coefficient	0.50 (t=9.45)	< 0.05
Kendall's Tau	0.38 (Z=9.24)	< 0.05
Somers' D Statistics	X/Y=0.45 Y/X=0.34	< 0.05 < 0.05
Gamma Statistic	0.43 (Z=9.24)	< 0.05
Pearson Chi-square Statistic	155.08	0.0044

YEAR : 1992, valid cases: 253		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient		< 0.05
Spearman correlation coefficient	0.56 (t=10.83)	< 0.05
Kendall's Tau	0.42 (Z=9.96)	< 0.05
Somers' D Statistics	X/Y=0.51 Y/X=0.37	< 0.05 < 0.05
Gamma Statistic	0.48 (Z=9.96)	< 0.05
Pearson Chi-square Statistic	158.60	0.0011

YEAR : 1993, valid cases: 243		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.48	< 0.05
Spearman correlation coefficient	0.51 (t=9.22)	< 0.05
Kendall's Tau	0.38 (Z=8.95)	< 0.05
Somers' D Statistics	X/Y=0.45 Y/X=0.35	< 0.05 < 0.05
Gamma Statistic	0.43 (Z=8.95)	< 0.05
Pearson Chi-square Statistic	143.71	0.0415

YEAR : 1994, valid cases: 307		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.46	< 0.05
Spearman correlation coefficient	0.48 (t=9.62)	< 0.05
Kendall's Tau	0.38 (Z=9.59)	< 0.05
Somers' D Statistics	X/Y=0.45 Y/X=0.34	< 0.05 < 0.05
Gamma Statistic	0.41 (Z=9.59)	< 0.05
Pearson Chi-square Statistic	172.85	0.0002

YEAR : 1995, valid cases: 338		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.54	< 0.05
Spearman correlation coefficient	0.54 (t=11.20)	< 0.05
Kendall's Tau	0.41 (Z=11.31)	< 0.05
Somers' D Statistics	X/Y=0.52 Y/X=0.39	< 0.05 < 0.05
Gamma Statistic	0.46 (Z=11.31)	< 0.05
Pearson Chi-square Statistic	266.46	< 0.05

YEAR : 1996, valid cases: 303		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.58	< 0.05
Spearman correlation coefficient	0.56 (t=11.79)	< 0.05
Kendall's Tau	0.43 (Z=11.14)	< 0.05
Somers' D Statistics	X/Y=0.52 Y/X=0.42	< 0.05 < 0.05
Gamma Statistic	0.47 (Z=11.14)	< 0.05
Pearson Chi-square Statistic	247.36	< 0.05

YEAR : 1997, valid cases: 302		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.64	< 0.05
Spearman correlation coefficient	0.63 (t=14.33)	< 0.05
Kendall's Tau	0.49 (Z=12.76)	< 0.05
Somers' D Statistics	X/Y=0.60 Y/X=0.47	< 0.05
Gamma Statistic	0.54 (Z=12.76)	< 0.05
Pearson Chi-square Statistic	271.39	< 0.05

YEARS: 1990-1997 valid cases: 2249		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.55	<0.05
Spearman correlation coefficient	0.57 (t=32.78)	<0.05
Kendall's Tau	0.43 (Z=30.67)	<0.05
Somers' D Statistics	X/Y=0.51 Y/X=0.39	<0.05
Gamma Statistic	0.48 (Z=30.67)	<0.05
Pearson Chi-square statistic	849.60	<0.05

Examining the first eight tables of the individual calendar years, we find all the measures to be statistically significant⁸. They are positive and moderately strong. Most of them lie in the interval between 0.40 and 0.50 and in several instances they are even greater than 0.60.

Despite the fact that in each of the considered tables the measures are of similar magnitude, we note the following:

- The Spearman and the Pearson correlation coefficients are usually of similar magnitudes indicating a somewhat linear⁹ monotonic relationship between the variables under investigation.

⁸ A result is statistically significant if there appears to be evidence against some implicit or explicit hypothesis.

The word "significant" is mostly used in the thesis when referring to statistical significance as opposed to the meaning of significance in the general sense (having a special or suggestive meaning, being important or consequential).

⁹ For example, as the faculty point score becomes higher, the average number of credits obtained by the corresponding subset of students increases linearly.

- Kendall's Tau and the Gamma statistic are of similar magnitude with the latter being slightly greater for all the time-periods. This feature arises because the denominator of the Gamma statistic is always lower than that of Kendall's Tau ¹⁰ while the numerators of both measures have the same form and are, thus, equal in any given set of data.
- The Somers' D statistic Y/X is always smaller than the statistic X/Y . Here the first is usually the smallest measure. This feature may be partially attributed to the large number of point scores and the small number of possible credits ¹¹.
- The Somers' D X/Y is always higher than the Gamma statistics ¹².
- The z-test statistics for Kendall's Tau and the Gamma statistic are identical ¹³.

¹⁰ The denominator of Kendall's Tau is equal to the square root of the product of the number of pairs not tied on variable X and the number of pairs not tied on variable Y . The denominator of the Gamma statistic is equal to the exact number of pairs, which either agree or disagree i.e. the number of untied pairs. Thus the denominator of the Gamma statistic is slightly smaller than the denominator of Kendall's Tau and the Gamma statistic is slightly larger than Kendall's Tau. For more details see Appendix A where the measures of association are discussed.

¹¹ The denominator of Somers' D statistic corrects for the number of ties in the level of the given variable. Thus the measure is an index of ordinal association between pairs of observations that are not tied on the given variable.

Since there are more distinct faculty point scores (variable X) than there are distinct numbers of credits obtained (variable Y), fewer observations are tied on the X variable than on the Y variable. When the faculty point score is given, the denominator of the statistic is greater than when the given variable represents the number of credits obtained. Thus the statistic Y/X may be smaller than the statistic X/Y . For more details see Appendix A where the measures of association are discussed.

¹² A similar explanation arises from inspection of denominators as applies to Kendall's Tau and Gamma statistic.

¹³ The numerators of Kendall's Tau, Gamma and Somers' D statistics are identical, namely $(P-Q)$.

Also the numerators of their s_0^2 variances are identical: $\left(4 \left[\sum \sum a_{ij} (A_{ij} - D_{ij})^2 - \frac{(P-Q)^2}{N} \right] \right)$.

The expression for the z-statistic in all four cases is therefore,

$$z = \frac{(P-Q)}{\left(4 \left[\sum \sum a_{ij} (A_{ij} - D_{ij})^2 - \frac{(P-Q)^2}{N} \right] \right)^{1/2}}$$

- The Pearson chi-square statistic indicates a statistically significant association between the rows and columns in the contingency table formed by the two variables under investigation (except ¹⁴ in 1990).

However we note that the use of the chi-square statistic alone is sub-optimal for the ordinal categories. For further limitation of the Pearson chi-square statistic, the reader may refer to Section B4 in Appendix B.

Similar comments may be made about all the measures determined for the eight years combined data set.

Based on these statistical results, we may conclude:

- the incoming first year students admitted into the Faculty of Science with the *higher* faculty point scores are likely to pass *more* full-year first-year courses in the first calendar year of their study
- the incoming first year students admitted into the Faculty of Science with the *lower* faculty point score are likely to pass *fewer* full-year first year courses in the first calendar year of their study

¹⁴ The Pearson chi-square statistics may be unreliable when some expected frequencies are lower than 5. This assumption was not fulfilled in any year for contingency tables constructed from the raw data.

b) Binary Response Analysis ¹⁵

Bivariate data sets were formed by the actual *faculty point-scores* and the performance variable with two categories: category 1: incoming first year student obtained 2 or more credits and category 2 :incoming first year student obtained fewer than 2 credits in the first calendar year of his/her study.

The results are displayed in the following nine tables below.

FACULTY POINT SCORE (X) and CREDITS IN FIRST YEAR (binary) (Y)

YEAR : 1990, valid cases: 236		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.59	<0.05
Spearman correlation coefficient	0.61 (t=11.96)	<0.05
Kendall's Tau	0.51 (Z=11.75)	<0.05
Somers' D Statistics	X/Y=0.75 Y/X=0.35	<0.05 <0.05
Gamma Statistic	0.77 (Z=11.75)	<0.05
Pearson Chi-square Statistic	107.91	<0.05

YEAR : 1991, valid cases: 267		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.49	<0.05
Spearman correlation coefficient	0.39 (t=6.91)	<0.05
Kendall's Tau	0.33 (Z=7.97)	<0.05
Somers' D Statistics	X/Y=0.52 Y/X=0.20	<0.05 <0.05
Gamma Statistic	0.55 (Z=7.97)	<0.05
Pearson Chi-square Statistic	63.27	<0.05

YEAR : 1992, valid cases: 253		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.44	<0.05
Spearman correlation coefficient	0.46 (t=8.16)	<0.05
Kendall's Tau	0.38 (Z=9.07)	<0.05
Somers' D Statistics	X/Y=0.57 Y/X=0.25	<0.05 <0.05
Gamma Statistic	0.60 (Z=9.07)	<0.05
Pearson Chi-square Statistic	64.00	<0.05

¹⁵ Note that the binary analysis was performed to address the utility of the faculty point scores in relation to the current re-admission criteria of the Faculty of Science at UCT after the first year.

YEAR : 1993, valid cases: 243		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.34	<0.05
Spearman correlation coefficient	0.35 (t=5.84)	<0.05
Kendall's Tau	0.29 (Z=6.83)	<0.05
Somers' D Statistics	X/Y=0.45 Y/X=0.19	<0.05 <0.05
Gamma Statistic	0.47 (Z=6.83)	<0.05
Pearson Chi-square Statistic	50.89	0.0072

YEAR : 1994, valid cases: 307		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.35	<0.05
Spearman correlation coefficient	0.35 (t=6.46)	<0.05
Kendall's Tau	0.29 (Z=7.58)	<0.05
Somers' D Statistics	X/Y=0.44 Y/X=0.19	<0.05 <0.05
Gamma Statistic	0.46 (Z=7.58)	<0.05
Pearson Chi-square Statistic	73.73	<0.05

YEAR : 1995, valid cases: 338		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.37	<0.05
Spearman correlation coefficient	0.37 (t=7.36)	<0.05
Kendall's Tau	0.31 (Z=8.54)	<0.05
Somers' D Statistics	X/Y=0.49 Y/X=0.20	<0.05 <0.05
Gamma Statistic	0.51 (Z=8.54)	<0.05
Pearson Chi-square Statistic	65.63	<0.05

YEAR : 1996, valid cases: 303		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.45	<0.05
Spearman correlation coefficient	0.46 (t=8.97)	<0.05
Kendall's Tau	0.38 (Z=9.95)	<0.05
Somers' D Statistics	X/Y=0.54 Y/X=0.27	<0.05 <0.05
Gamma Statistic	0.56 (Z=9.95)	<0.05
Pearson Chi-square Statistic	81.23	<0.05

YEAR : 1997, valid cases: 302		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.39	<0.05
Spearman correlation coefficient	0.39 (t=7.46)	<0.05
Kendall's Tau	0.33 (Z=8.54)	<0.05
Somers' D Statistics	X/Y=0.49	<0.05
	Y/X=0.22	<0.05
Gamma Statistic	0.51 (Z=8.54)	<0.05
Pearson Chi-square Statistic	80.94	<0.05

YEARS: 1990-1997 valid cases: 2249		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.42	<0.05
Spearman correlation coefficient	0.43 (t=22.43)	<0.05
Kendall's Tau	0.36 (Z=25.35)	<0.05
Somers' D Statistics	X/Y=0.53	<0.05
	Y/X=0.24	<0.05
Gamma Statistic	0.56 (Z=25.35)	<0.05
Pearson Chi-square statistic	468.07	<0.05

We find all the measures to be significant in the nine time periods considered. They are positive and mostly lie in the interval between 0.30 and 0.40 with only few being lower than 0.30. They are usually lower than the corresponding measures for the ordinal response analysis in 3.1.1 a).

Despite the fact that in each of the considered tables the measures are of similar magnitude, we may note the following patterns:

- The Spearman correlation coefficient is always larger than the Pearson correlation coefficient. This feature appears to be an artefact of binary response.
- Kendall's Tau and the Gamma statistic are usually of similar magnitude with the latter being slightly greater ¹⁶.
- The Somers' *D* statistic Y/X is smaller than the statistic X/Y ¹⁵.
- The Somers' *D* X/Y is always higher than the Gamma statistics ¹⁵.
- The Pearson chi-square statistic indicates a significant association between the rows and columns formed by the variables under investigation.

¹⁶ As previously explained in 3.1.1 a).

We note that these results are similar to those obtained for the ordinal analysis and thus explanatory comments made in 3.1.1 a) also apply to this section.

Based on these statistical results, we may conclude that the first-year incoming students with the higher faculty point scores are more likely to fulfil the Faculty's second academic year re-admission criteria in the first calendar year of their study.

Essentially the conclusions reached for the ordinal and binary response analyses point to the similar conclusions i.e. the first year incoming student with *higher* faculty point-score is more likely to pass a *higher* number of first year courses than the first year incoming student with *lower* faculty point score in the first calendar year of his or her study.

What has been established so far?

So far the measures of association have indicated the positive nature and statistical significance of the relationship between the variables under investigation. However they have specified ¹⁷ which particular faculty point scores may be associated with which specific numbers of credits. We may only say that the higher (*lower*) faculty point scores are associated with the higher (*lower*) number of credits, respectively. Also we have not been able to incorporate other explanatory variables ¹⁸ and thus to investigate their collective associations with the performance variable.

The analyses indicate that search for a cut-off value, as a minimum for admission and prospective readmission to a second year, as a quick and dirty indicator, may be useful. But the analyses do not indicate whether a cut-off would be an optimal strategy not the criteria by which a cut-off should be selected. This issue of cut-off may be further clarified in the FIRM analyses.

¹⁷ This type of insight could, to some extent, be established by a careful inspection of the specific contingency tables.

¹⁸ We may, however, investigate the other relationships by computing the measures of association between each of the other variable of interest and the performance variable, individually.

3.1.2 GENERALISED LINEAR MODELLING TECHNIQUES

Out of the family of the Generalised Linear Modelling Techniques (GLIM), initially the simplest Variance Components Method (or the Analysis of Variance (ANOVA)) was used.

The analysis of variance is a convenient and rapid calculation procedure for separating the variability present in a collection of observations into components or sources. It provides a structure for experimental data in terms of mathematical model. It is always connected with a mathematical model and with the type of experimental or data design employed. Here it is applied to observational data.

Extensions of the procedure may be employed on the different types of variables (both discrete or categorical and continuous). This facility is especially useful in this analysis where different types of variables have been isolated in the data set.

The ANOVA models allow to introduce more than one categorical explanatory variable into the analyses into the analysis, and to perform analysis that are similar to multiple regression on artificial or dummy variables. The ANCOVA family of extensions allows for the introduction of explanatory variables that are notionally continuous on a numerical scale, as a covariate. Introduction of appropriate explanatory variables is signalled by decreasing variance and increasing F-statistic.

Again under this method, separate analyses were performed for the ordinal and binary definitions of the performance variable.

Since the results in the measures of association section had indicated the presence of statistically significant correlation between the faculty point score and the number of credits, the faculty point score (*facpoint*) was entered as a covariate¹⁹ in both analyses. The other explanatory variables used in the model²⁰ were:

- home language (*language*)
- educational department (*education*)
- colour
- gender

For each time-period considered, the results are displayed in two tables:

- ANOVA table with the relevant statistics, for Type I sums of squares
- table with the mean number of credits within the different student groupings arising from the distinct categories²¹ of the explanatory variables

¹⁹ A covariate is a continuous variable that is likely to be correlated with the dependent variable of interest. Its inclusion in the analysis may significantly reduce the error variance.

²⁰ The following model is fitted:

$$y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + \psi_l + C \cdot x_{ijkl} + e_{ijkl} \text{ where}$$

y_{ijkl} -denotes the number of credits obtained (credits)

μ -represents an overall mean response for a zero faculty point score

x_{ijkl} -denotes the faculty point score

C -is the coefficient associated with the covariate variable (the pooled within-groups slope)

α_i - represents the effect of the i^{th} level of the home language variable

β_j - represents the effect of the j^{th} level of the educational department variable

γ_k - represents the effect of the k^{th} level of the colour variable

ψ_l - represents the effect of the l^{th} level of the gender variable

e_{ijkl} -is the random error and $e_{ijkl} \sim NID(0, \sigma^2)$

The value ranges of the four subscripts i, j, k, l depend on the number of distinct categories present in the particular data set to which the model is fitted. Therefore, some subscript range values differ between the time periods considered here. The subscripts i, j, k, l simply conveys the inclusion of all relevant individuals from all subsets.

The above model is a multi-way or multi-factor between-groups design since there are four between-groups factors and covariate. It may also be referred to as ANCOVA since it contains a covariate. No interaction terms are present in this model, as the primary interest had concentrated on the apparent effects of the respective explanatory variables on the performance variable and not their mutual associations in contributions to the response variable. Use of interaction terms will rapidly increase the size of the ANCOVA model and hence give rise to difficulties in interpretation.

²¹ A complete list of the categories in the explanatory variables is available in chapter 1 and in TABLE 2.1.1 from chapter 2. Some categories of these variables are not represented in particular years in the study. In these years there were no students with these particular characteristics e.g. no new educational department codes until 1997 (the first year in which the new departments were formed).

b) Ordinal Response Analysis

Calendar year: 1990

ANOVA RESULTS: CREDITS IN 1990: CREDITO							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1	230.04	213.00	0.96	239.03	0.0000
{1}EDUCATION	Fixed	9	8.60	213.00	0.96	8.94	0.0000
{2}LANGUAGE	Fixed	9	1.44	213.00	0.96	1.50	0.1497
{3}COLOUR	Fixed	2	1.01	213.00	0.96	1.05	0.3510
{4}GENDER	Fixed	1	0.10	213.00	0.96	0.10	0.7496

DESCRIPTIVE STATISTICS: CREDITS IN 1990: CREDITO							
	LEVEL OF FACTOR	N	MEAN CREDITO	STD.DEV CREDITO	STD. ERROR CREDITO	-95% CNF. LMT	+95% CNF LMT
TOTAL		236	2.47	1.50	0.10	2.28	2.66
{1}EDUCATION	CO	57	1.98	1.34	0.18	1.63	2.34
{1}EDUCATION	JM	16	2.41	1.57	0.39	1.57	3.24
{1}EDUCATION	CA	81	3.20	1.14	0.13	2.94	3.45
{1}EDUCATION	NE	10	1.70	1.48	0.47	0.64	2.76
{1}EDUCATION	IN	5	3.60	0.89	0.40	2.49	4.71
{1}EDUCATION	TR	21	3.24	1.07	0.23	2.75	3.72
{1}EDUCATION	ET	28	0.52	0.29	0.05	0.41	0.63
{1}EDUCATION	NA	12	3.83	0.25	0.07	3.68	3.99
{1}EDUCATION	OF	3	2.83	2.47	1.42	-3.29	8.96
{1}EDUCATION	TK	3	0.33	0.29	0.17	-0.38	1.05
{2}LANGUAGE	EA	11	1.59	1.41	0.43	0.64	2.54
{2}LANGUAGE	EN	176	2.97	1.28	0.10	2.78	3.16
{2}LANGUAGE	AF	10	1.90	1.43	0.45	0.88	2.92
{2}LANGUAGE	ZU	8	0.50	0.00	0.00	0.50	0.50
{2}LANGUAGE	TW	9	0.61	0.42	0.14	0.29	0.93
{2}LANGUAGE	NS	3	1.50	1.73	1.00	-2.80	5.80
{2}LANGUAGE	XH	13	0.46	0.32	0.09	0.27	0.66
{2}LANGUAGE	SS	4	0.63	0.25	0.13	0.23	1.02
{2}LANGUAGE	SW	1	0.50				
{2}LANGUAGE	VE	1	0.50				
{3}COLOUR	C	57	1.96	1.37	0.18	1.59	2.32
{3}COLOUR	W	127	3.26	1.11	0.10	3.07	3.45
{3}COLOUR	I	13	2.65	1.45	0.40	1.78	3.53
{3}COLOUR	B	39	0.60	0.55	0.09	0.42	0.78
{4}GENDER	M	145	2.36	1.55	0.13	2.10	2.61
{4}GENDER	F	91	2.66	1.40	0.15	2.37	2.95

Calendar year: 1991

ANOVA RESULTS: CREDITS IN 1991: CREDIT1							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1	106.82	243.00	1.06	100.77	0.0000
(1)EDUCATION	Fixed	9	8.10	243.00	1.06	7.64	0.0000
(2)LANGUAGE	Fixed	10	1.07	243.00	1.06	1.01	0.4358
(3)COLOUR	Fixed	2	1.03	243.00	1.06	0.97	0.3805
(4)GENDER	Fixed	1	0.40	243.00	1.06	0.38	0.5379

DESCRIPTIVE STATISTICS CREDITS IN 1991: CREDIT1							
	LEVEL OF FACTOR	N	MEAN CREDIT1	STD.DEV. CREDIT1	STD. ERROR CREDIT1	-.95% CNF. LMT	+95% CNF. LMT
TOTAL		267	2.70	1.30	0.08	2.54	2.86
(1)EDUCATION	CO	86	2.06	1.22	0.13	1.80	2.32
(1)EDUCATION	JM	7	3.43	0.73	0.28	2.75	4.11
(1)EDUCATION	CA	85	3.47	0.89	0.10	3.28	3.66
(1)EDUCATION	NE	14	2.18	1.54	0.41	1.29	3.07
(1)EDUCATION	IN	12	2.71	1.44	0.41	1.80	3.62
(1)EDUCATION	TR	19	3.39	1.19	0.27	2.82	3.97
(1)EDUCATION	ET	24	1.60	0.69	0.14	1.31	1.90
(1)EDUCATION	NA	15	3.30	1.10	0.28	2.69	3.91
(1)EDUCATION	OF	2	1.75	1.77	1.25	-14.13	17.63
(1)EDUCATION	TK	3	1.83	1.04	0.60	-0.75	4.42
(2)LANGUAGE	EA	38	2.47	1.33	0.21	2.04	2.91
(2)LANGUAGE	EN	186	2.90	1.28	0.09	2.72	3.09
(2)LANGUAGE	AF	7	3.00	1.26	0.48	1.84	4.16
(2)LANGUAGE	ZU	5	1.50	1.22	0.55	-0.02	3.02
(2)LANGUAGE	TW	10	1.85	0.53	0.17	1.47	2.23
(2)LANGUAGE	NS	1	0.00				
(2)LANGUAGE	XH	12	1.67	0.83	0.24	1.14	2.20
(2)LANGUAGE	SS	3	2.17	1.76	1.01	-2.20	6.53
(2)LANGUAGE	SW	1	1.50				
(2)LANGUAGE	VE	3	3.17	0.76	0.44	1.27	5.06
(2)LANGUAGE	EA	1	2.00				
(3)COLOUR	C	88	2.11	1.21	0.13	1.85	2.37
(3)COLOUR	W	128	3.40	1.01	0.09	3.22	3.58
(3)COLOUR	I	18	2.44	1.49	0.35	1.70	3.19
(3)COLOUR	B	33	1.70	0.91	0.16	1.37	2.02
(4)GENDER	M	159	2.60	1.34	0.11	2.39	2.81
(4)GENDER	F	108	2.85	1.24	0.12	2.61	3.08

Calendar year: 1992

ANOVA RESULTS: CREDITS IN 1992:CREDIT2							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1	136.10	219.00	0.99	136.93	0.0000
{1}EDUCATION	Fixed	8	10.27	219.00	0.99	10.33	0.0000
{2}LANGUAGE	Fixed	9	0.56	219.00	0.99	0.56	0.8300
{3}COLOUR	Fixed	3	6.49	219.00	0.99	6.53	0.0003
{4}GENDER	Fixed	1	3.29	219.00	0.99	3.31	0.0702

DESCRIPTIVE STATISTICS CREDITS IN 1992: CREDIT2							
	LEVEL OF FACTOR	N	MEAN CREDIT2	STD.DEV. CREDIT2	STD. ERROR CREDIT2	-.95% CNF. LMT	+.95% CNF. LMT
TOTAL		242	2.58	1.39	0.09	2.40	2.76
{1}EDUCATION	CO	54	2.13	1.39	0.19	1.75	2.51
{1}EDUCATION	JM	8	3.19	1.00	0.35	2.35	4.02
{1}EDUCATION	CA	89	3.39	0.96	0.10	3.19	3.59
{1}EDUCATION	NE	13	2.15	1.51	0.42	1.24	3.06
{1}EDUCATION	IN	8	1.75	1.56	0.55	0.45	3.05
{1}EDUCATION	TR	11	3.09	1.50	0.45	2.09	4.10
{1}EDUCATION	ET	41	1.28	0.75	0.12	1.04	1.52
{1}EDUCATION	NA	13	3.73	0.60	0.17	3.37	4.09
{1}EDUCATION	TK	5	1.10	0.22	0.10	0.82	1.38
{2}LANGUAGE	EA	6	2.25	1.04	0.42	1.16	3.34
{2}LANGUAGE	EN	173	2.95	1.35	0.10	2.75	3.15
{2}LANGUAGE	AF	10	2.40	1.02	0.32	1.67	3.13
{2}LANGUAGE	ZU	9	1.44	0.92	0.31	0.74	2.15
{2}LANGUAGE	TW	9	0.89	0.49	0.16	0.52	1.26
{2}LANGUAGE	NS	5	1.60	1.29	0.58	-0.01	3.21
{2}LANGUAGE	XH	20	1.65	0.83	0.19	1.26	2.04
{2}LANGUAGE	SS	7	1.71	0.99	0.38	0.79	2.63
{2}LANGUAGE	VE	1	0.50				
{2}LANGUAGE	EA	2	1.00	0.71	0.50	-5.35	7.35
{3}COLOUR	C	59	2.15	1.40	0.18	1.79	2.52
{3}COLOUR	W	114	3.46	0.93	0.09	3.29	3.64
{3}COLOUR	I	15	1.60	1.43	0.37	0.81	2.39
{3}COLOUR	B	54	1.45	0.88	0.12	1.21	1.69
{4}GENDER	M	160	2.50	1.39	0.11	2.28	2.71
{4}GENDER	F	82	2.74	1.38	0.15	2.44	3.05

Calendar year: 1993

ANOVA RESULTS: CREDITS IN 1993: CREDIT3							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1	112.71	217.00	1.15	97.93	0.0000
{1}EDUCATION	Fixed	9	9.83	217.00	1.15	8.54	0.0000
{2}LANGUAGE	Fixed	12	1.13	217.00	1.15	0.98	0.4683
{3}COLOUR	Fixed	2	7.94	217.00	1.15	6.90	0.0012
{4}GENDER	Fixed	1	0.64	217.00	1.15	0.56	0.4567

DESCRIPTIVE STATISTICS CREDITS IN 1993: CREDIT3							
	LEVEL OF FACTOR	N	MEAN CREDIT3	STD.DEV. CREDIT3	STD. ERROR CREDIT3	-.95% CNF. LMT	+.95% CNF. LMT
TOTAL		243	2.57	1.41	0.09	2.39	2.75
{1}EDUCATION	CO	38	2.13	1.46	0.24	1.65	2.61
{1}EDUCATION	JM	16	2.94	1.38	0.34	2.20	3.67
{1}EDUCATION	CA	75	3.43	0.98	0.11	3.21	3.66
{1}EDUCATION	NE	15	1.73	1.44	0.37	0.94	2.53
{1}EDUCATION	IN	16	2.41	1.59	0.40	1.56	3.26
{1}EDUCATION	TR	18	3.25	1.27	0.30	2.62	3.88
{1}EDUCATION	ET	45	1.52	0.92	0.14	1.25	1.80
{1}EDUCATION	NA	9	3.39	1.05	0.35	2.58	4.20
{1}EDUCATION	OF	3	2.00	1.73	1.00	-2.30	6.30
{1}EDUCATION	TK	8	1.44	0.86	0.31	0.72	2.16
{2}LANGUAGE	EA	4	1.75	2.02	1.01	-1.47	4.97
{2}LANGUAGE	EN	169	2.93	1.35	0.10	2.73	3.14
{2}LANGUAGE	AF	10	2.55	1.40	0.44	1.55	3.55
{2}LANGUAGE	ZU	14	1.46	0.77	0.21	1.02	1.91
{2}LANGUAGE	TW	5	1.50	0.61	0.27	0.74	2.26
{2}LANGUAGE	NS	6	1.42	1.50	0.61	-0.15	2.99
{2}LANGUAGE	XH	21	1.50	0.85	0.19	1.11	1.89
{2}LANGUAGE	SS	6	2.17	0.93	0.38	1.19	3.14
{2}LANGUAGE	SW	1	1.00				
{2}LANGUAGE	VE	3	0.83	1.04	0.60	-1.75	3.42
{2}LANGUAGE	EA	2	4.00	0.00	0.00	4.00	4.00
{2}LANGUAGE	EN	1	0.50				
{2}LANGUAGE	AF	1	4.00				
{3}COLOUR	C	36	2.32	1.40	0.23	1.84	2.79
{3}COLOUR	W	119	3.34	1.08	0.10	3.14	3.54
{3}COLOUR	I	30	1.83	1.55	0.28	1.25	2.41
{3}COLOUR	B	58	1.53	0.96	0.13	1.28	1.79
{4}GENDER	M	145	2.54	1.45	0.12	2.30	2.78
{4}GENDER	F	98	2.62	1.35	0.14	2.35	2.89

Calendar year: 1994

ANOVA RESULTS: CREDITS IN 1994 CREDIT4							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1	132.26	278.00	1.21	109.56	0.0000
{1}EDUCATION	Fixed	9	12.11	278.00	1.21	10.04	0.0000
{2}LANGUAGE	Fixed	11	0.85	278.00	1.21	0.70	0.7371
{3}COLOUR	Fixed	3	1.73	278.00	1.21	1.43	0.2332
{4}GENDER	Fixed	1	0.10	278.00	1.21	0.08	0.7729

DESCRIPTIVE STATISTICS CREDITS IN 1994: CREDIT4							
	LEVEL OF FACTOR	N	MEAN CREDIT4	STD.DEV. CREDIT4	STD. ERROR CREDIT4	-.95% CNF. LMT	+.95% CNF. LMT
TOTAL		304	2.62	1.40	0.08	2.46	2.78
{1}EDUCATION	CO	73	2.12	1.26	0.15	1.82	2.41
{1}EDUCATION	CA	94	3.36	1.10	0.11	3.14	3.59
{1}EDUCATION	NE	13	1.73	1.54	0.43	0.80	2.66
{1}EDUCATION	IN	19	2.34	1.33	0.31	1.70	2.99
{1}EDUCATION	TR	25	3.28	1.32	0.26	2.74	3.82
{1}EDUCATION	ET	45	1.38	1.02	0.15	1.07	1.69
{1}EDUCATION	NA	18	3.17	1.26	0.30	2.54	3.79
{1}EDUCATION	OF	1	2.50				
{1}EDUCATION	TK	1	2.00				
{1}EDUCATION	IE	15	3.60	0.66	0.17	3.23	3.97
{2}LANGUAGE	EA	9	2.22	1.35	0.45	1.19	3.26
{2}LANGUAGE	AF	225	2.88	1.34	0.09	2.71	3.06
{2}LANGUAGE	ZU	14	2.39	1.33	0.36	1.62	3.16
{2}LANGUAGE	TW	6	1.42	1.39	0.57	-0.05	2.88
{2}LANGUAGE	NS	5	1.30	1.30	0.58	-0.32	2.92
{2}LANGUAGE	XH	4	2.13	0.63	0.31	1.12	3.13
{2}LANGUAGE	SS	17	1.21	1.38	0.34	0.50	1.92
{2}LANGUAGE	SW	9	1.50	0.87	0.29	0.83	2.17
{2}LANGUAGE	VE	7	2.36	1.11	0.42	1.33	3.38
{2}LANGUAGE	CH	4	3.13	1.44	0.72	0.84	5.41
{2}LANGUAGE	EA	3	2.17	0.76	0.44	0.27	4.06
{2}LANGUAGE	EN	1	1.50				
{3}COLOUR	C	77	2.14	1.29	0.15	1.85	2.44
{3}COLOUR	W	147	3.30	1.16	0.10	3.11	3.49
{3}COLOUR	I	26	2.35	1.41	0.28	1.78	2.91
{3}COLOUR	B	54	1.59	1.17	0.16	1.27	1.91
{4}GENDER	M	173	2.48	1.41	0.11	2.27	2.69
{4}GENDER	F	131	2.81	1.37	0.12	2.57	3.04

Calendar year: 1995

ANOVA RESULTS: CREDITS IN 1995:CREDITS							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1	162.42	310.00	0.90	180.31	0.0000
{1}EDUCATION	Fixed	9	12.58	310.00	0.90	13.96	0.0000
{2}LANGUAGE	Fixed	11	0.86	310.00	0.90	0.96	0.4864
{3}COLOUR	Fixed	3	1.12	310.00	0.90	1.24	0.2940
{4}GENDER	Fixed	1	0.12	310.00	0.90	0.14	0.7114

DESCRIPTIVE STATISTICS: CREDITS IN 1995:CREDITS							
	LEVEL OF FACTOR	N	MEAN CREDITS	STD.DEV. CREDITS	STD. ERROR CREDITS	-.95% CNF. LMT	+.95% CNF. LMT
TOTAL		336	2.64	1.30	0.07	2.50	2.78
{1}EDUCATION	CO	53	2.01	1.34	0.18	1.64	2.38
{1}EDUCATION	CA	128	3.29	1.02	0.09	3.11	3.46
{1}EDUCATION	NE	5	3.20	1.04	0.46	1.91	4.49
{1}EDUCATION	IN	17	1.15	1.18	0.29	0.64	1.76
{1}EDUCATION	TR	28	3.05	1.21	0.23	2.68	3.62
{1}EDUCATION	ET	47	1.46	0.74	0.11	1.24	1.68
{1}EDUCATION	NA	22	2.95	1.02	0.22	2.50	3.41
{1}EDUCATION	OF	5	3.10	1.24	0.56	1.55	4.65
{1}EDUCATION	TK	5	1.90	0.22	0.10	1.62	2.18
{1}EDUCATION	IE	26	3.08	1.03	0.20	2.66	3.49
{2}LANGUAGE	EA	5	2.10	1.95	0.87	-0.32	4.52
{2}LANGUAGE	AF	257	2.89	1.27	0.08	2.74	3.06
{2}LANGUAGE	ZU	6	1.68	0.86	0.35	0.68	2.49
{2}LANGUAGE	TW	8	2.00	1.20	0.42	1.00	3.00
{2}LANGUAGE	NS	10	1.70	1.14	0.36	0.89	2.51
{2}LANGUAGE	XH	4	2.13	0.63	0.31	1.12	3.13
{2}LANGUAGE	SS	28	1.61	0.46	0.09	1.43	1.78
{2}LANGUAGE	SW	7	1.93	1.43	0.54	0.61	3.25
{2}LANGUAGE	VE	1	1.00				
{2}LANGUAGE	CH	4	1.25	0.65	0.32	0.22	2.28
{2}LANGUAGE	EA	4	3.63	0.48	0.24	2.86	4.39
{2}LANGUAGE	EN	2	1.25	1.06	0.75	-8.28	10.78
{3}COLOUR	C	54	2.07	1.39	0.19	1.70	2.45
{3}COLOUR	W	182	3.27	1.00	0.07	3.12	3.42
{3}COLOUR	I	34	1.97	1.45	0.25	1.47	2.48
{3}COLOUR	B	66	1.70	0.88	0.11	1.48	1.91
{4}GENDER	M	189	2.67	1.36	0.10	2.37	2.76
{4}GENDER	F	147	2.73	1.22	0.10	2.63	2.93

Calendar year: 1996

ANOVA RESULTS: CREDITS IN 1996: CREDIT6							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1	211.49	278.00	1.08	195.51	0.0000
(1)EDUCATION	Fixed	9	12.51	278.00	1.08	11.56	0.0000
(2)LANGUAGE	Fixed	10	0.28	278.00	1.08	0.26	0.9888
(3)COLOUR	Fixed	3	2.40	278.00	1.08	2.22	0.0862
(4)GENDER	Fixed	1	2.09	278.00	1.08	1.93	0.1656

DESCRIPTIVE STATISTICS CREDITS IN 1996: CREDIT6							
	LEVEL OF FACTOR	N	MEAN CREDIT6	STD.DEV. CREDIT6	STD. ERROR CREDIT6	-.95% CNF. LMT	+.95% CNF. LMT
TOTAL		303	2.24	1.45	0.08	2.07	2.40
(1)EDUCATION	CO	50	1.59	1.26	0.18	1.23	1.95
(1)EDUCATION	CA	97	3.13	1.31	0.13	2.87	3.39
(1)EDUCATION	NE	7	1.29	1.58	0.60	-0.17	2.74
(1)EDUCATION	IN	15	1.27	1.59	0.41	0.39	2.15
(1)EDUCATION	TR	28	3.16	1.08	0.20	2.74	3.58
(1)EDUCATION	ET	70	1.19	0.74	0.09	1.02	1.37
(1)EDUCATION	NA	14	3.18	1.12	0.30	2.53	3.83
(1)EDUCATION	OF	4	3.00	1.15	0.58	1.16	4.84
(1)EDUCATION	TK	9	1.50	0.43	0.14	1.17	1.83
(1)EDUCATION	IE	9	2.72	1.25	0.42	1.76	3.69
(2)LANGUAGE	EA	3	2.00	1.50	0.87	-1.73	5.73
(2)LANGUAGE	AF	194	2.66	1.52	0.11	2.44	2.87
(2)LANGUAGE	ZU	9	1.94	1.26	0.42	0.98	2.91
(2)LANGUAGE	TW	16	1.06	0.83	0.21	0.62	1.51
(2)LANGUAGE	NS	12	1.58	1.02	0.29	0.94	2.23
(2)LANGUAGE	XH	12	1.54	0.78	0.23	1.04	2.04
(2)LANGUAGE	SS	29	1.36	0.77	0.14	1.07	1.65
(2)LANGUAGE	SW	9	1.22	0.79	0.26	0.61	1.83
(2)LANGUAGE	VE	8	1.06	0.68	0.24	0.50	1.63
(2)LANGUAGE	CH	9	2.50	1.09	0.36	1.66	3.34
(2)LANGUAGE	EA	2	1.25	0.35	0.25	-1.93	4.43
(3)COLOUR	C	48	1.93	1.36	0.20	1.53	2.32
(3)COLOUR	W	138	3.13	1.26	0.11	2.92	3.34
(3)COLOUR	I	27	1.37	1.48	0.29	0.78	1.96
(3)COLOUR	B	90	1.29	0.82	0.09	1.12	1.46
(4)GENDER	M	166	2.10	1.44	0.11	1.88	2.32
(4)GENDER	F	137	2.40	1.46	0.12	2.16	2.65

Calendar year: 1997

ANOVA RESULTS: CREDITS IN 1997: CREDIT7							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1	213.78	268.00	0.97	221.31	0.0000
{1}EDUCATION	Fixed	17	2.13	268.00	0.97	2.21	0.0045
{2}LANGUAGE	Fixed	11	1.20	268.00	0.97	1.24	0.2593
{3}COLOUR	Fixed	3	2.62	268.00	0.97	2.72	0.0452
{4}GENDER	Fixed	1	0.13	268.00	0.97	0.14	0.7129

DESCRIPTIVE STATISTICS CREDITS IN 1997: CREDIT 7							
	LEVEL OF FACTOR	N	MEAN CREDIT7	STD.DEV. CREDIT7	STD. ERROR CREDIT7	-.95% CNF. LMT	+.95% CNF. LMT
TOTAL		302	2.42	1.33	0.08	2.27	2.57
{1}EDUCATION	CO	3	1.83	1.04	0.60	-0.76	4.42
{1}EDUCATION	CA	8	3.19	0.65	0.23	2.64	3.73
{1}EDUCATION	NE	1	0.50				
{1}EDUCATION	IN	1	0.00				
{1}EDUCATION	TR	9	3.33	0.83	0.28	2.70	3.97
{1}EDUCATION	ET	27	1.54	0.76	0.15	1.24	1.84
{1}EDUCATION	NA	1	1.50				
{1}EDUCATION	OF	2	3.00	3.54	2.50	-28.77	34.77
{1}EDUCATION	TK	3	1.33	1.15	0.67	-1.54	4.20
{1}EDUCATION	IE	26	3.12	1.28	0.25	2.60	3.63
{1}EDUCATION	WC	124	2.67	1.31	0.12	2.44	2.91
{1}EDUCATION	EC	15	2.07	1.35	0.35	1.32	2.81
{1}EDUCATION	KN	30	2.12	1.46	0.27	1.57	2.66
{1}EDUCATION	GT	24	2.56	1.23	0.25	2.04	3.08
{1}EDUCATION	NW	10	1.75	0.98	0.31	1.05	2.45
{1}EDUCATION	NC	5	2.20	0.67	0.30	1.37	3.03
{1}EDUCATION	NP	10	1.15	0.63	0.20	0.70	1.60
{1}EDUCATION	MP	3	2.67	1.53	0.88	-1.13	6.46
{2}LANGUAGE	EA	4	2.63	0.95	0.47	1.12	4.13
{2}LANGUAGE	AF	199	2.82	1.27	0.09	2.64	3.00
{2}LANGUAGE	ZU	10	2.20	1.40	0.44	1.20	3.20
{2}LANGUAGE	TW	15	1.27	0.86	0.22	0.79	1.74
{2}LANGUAGE	NS	14	1.43	0.94	0.25	0.89	1.97
{2}LANGUAGE	XH	10	1.35	0.71	0.22	0.84	1.86
{2}LANGUAGE	SS	25	1.32	0.96	0.19	0.93	1.71
{2}LANGUAGE	SW	7	1.79	0.39	0.15	1.42	2.15
{2}LANGUAGE	VE	3	2.17	1.04	0.60	-0.42	4.75
{2}LANGUAGE	CH	2	1.00	0.00	0.00	1.00	1.00
{2}LANGUAGE	TG	6	3.50	1.22	0.50	2.21	4.79
{2}LANGUAGE	ND	7	1.36	0.56	0.21	0.84	1.87
{3}COLOUR	C	50	2.07	1.25	0.18	1.71	2.43
{3}COLOUR	W	140	3.16	1.09	0.09	2.98	3.34
{3}COLOUR	I	25	2.44	1.53	0.31	1.81	3.07
{3}COLOUR	B	87	1.43	0.87	0.09	1.25	1.62
{4}GENDER	M	173	2.37	1.33	0.10	2.17	2.57
{4}GENDER	F	129	2.48	1.32	0.12	2.25	2.71

Calendar years 1990-1997 combined

ANOVA RESULTS: CREDITS IN 1990-1997 COMBINED: CREDIT							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1	1323.03	2197.00	1.05	1257.77	0.0000
{1}EDUCATION	Fixed	18	33.72	2197.00	1.05	32.06	0.0000
{2}LANGUAGE	Fixed	12	1.76	2197.00	1.05	1.67	0.0669
{3}COLOUR	Fixed	3	10.60	2197.00	1.05	10.07	0.0000
{4}GENDER	Fixed	1	0.79	2197.00	1.05	0.75	0.3875

DESCRIPTIVE STATISTICS: CREDITS IN 1990-1997 COMBINED: CREDIT							
	LEVEL OF FACTOR	N	MEAN CREDIT	STD.DEV. CREDIT	STD. ERROR CREDIT.	-.95% CNF. LMT	+.95% CNF. LMT
TOTAL		2233	2.53	1.39	0.03	2.47	2.59
{1}EDUCATION	CO	414	2.01	1.31	0.06	1.88	2.14
{1}EDUCATION	JM	47	2.87	1.34	0.20	2.48	3.26
{1}EDUCATION	CA	657	3.32	1.07	0.04	3.23	3.40
{1}EDUCATION	NE	78	1.92	1.49	0.17	1.58	2.25
{1}EDUCATION	IN	93	2.00	1.54	0.16	1.68	2.32
{1}EDUCATION	TR	159	3.21	1.18	0.09	3.03	3.40
{1}EDUCATION	ET	327	1.31	0.83	0.05	1.22	1.40
{1}EDUCATION	NA	104	3.29	1.03	0.10	3.09	3.49
{1}EDUCATION	OF	20	2.70	1.60	0.36	1.95	3.45
{1}EDUCATION	TK	37	1.42	0.71	0.12	1.18	1.66
{1}EDUCATION	IE	76	3.15	1.10	0.13	2.90	3.40
{1}EDUCATION	WC	124	2.67	1.31	0.12	2.44	2.91
{1}EDUCATION	EC	15	2.07	1.35	0.35	1.32	2.81
{1}EDUCATION	KN	30	2.12	1.46	0.27	1.67	2.66
{1}EDUCATION	GT	24	2.56	1.23	0.25	2.04	3.08
{1}EDUCATION	NW	10	1.75	0.98	0.31	1.05	2.45
{1}EDUCATION	NC	5	2.20	0.67	0.30	1.37	3.03
{1}EDUCATION	NP	10	1.15	0.63	0.20	0.70	1.60
{1}EDUCATION	MP	3	2.67	1.53	0.88	-1.13	6.46
{2}LANGUAGE	EA	80	2.24	1.37	0.15	1.93	2.54
{2}LANGUAGE	EN	1579	2.87	1.33	0.03	2.81	2.94
{2}LANGUAGE	AF	76	2.26	1.28	0.15	1.97	2.55
{2}LANGUAGE	ZU	81	1.30	0.95	0.11	1.09	1.51
{2}LANGUAGE	TW	74	1.38	0.91	0.11	1.17	1.59
{2}LANGUAGE	NS	45	1.56	0.99	0.15	1.26	1.85
{2}LANGUAGE	XH	165	1.38	0.88	0.07	1.25	1.52
{2}LANGUAGE	SS	52	1.63	0.99	0.14	1.35	1.90
{2}LANGUAGE	SW	6	1.50	1.00	0.41	0.45	2.55
{2}LANGUAGE	VE	27	1.37	0.97	0.19	0.99	1.75
{2}LANGUAGE	CH	28	3.14	1.09	0.21	2.72	3.56
{2}LANGUAGE	TG	18	1.42	0.69	0.16	1.07	1.76
{2}LANGUAGE	ND	2	2.75	1.77	1.25	-13.13	18.83
{3}COLOUR	C	469	2.09	1.32	0.06	1.97	2.21
{3}COLOUR	W	1095	3.28	1.09	0.03	3.22	3.35
{3}COLOUR	I	188	2.04	1.51	0.11	1.82	2.26
{3}COLOUR	B	481	1.43	0.93	0.04	1.34	1.51
{4}GENDER	M	1310	2.44	1.41	0.04	2.36	2.52
{4}GENDER	F	923	2.65	1.34	0.04	2.57	2.74

In the preceding nine ANOVA tables the following variables have emerged as statistically significant:

- faculty point score ²² (in all the time periods considered)
- educational department (in all the time periods considered)
- colour (only in the following years 1992,1993,1997 and the combined data set of all years)

The above factors are sources of variations in the performance variable.

The significant differences in the mean number of credits emerge only between the student groupings in the distinct categories of the variables e.g. in 1992 among the categories for the colour variable or the categories of the educational department variable and not amongst the categories of the language variable or the gender variable. Thus, each of the listed factors taken alone (or along with the faculty point scores) contribute to the fact that the incoming first year students obtain the different numbers of credits in the first calendar year of their studies.

Retrospectively an admission policy over the period 1990 to 1997 might usefully have incorporated predictor performance adjustments for educational departments and for colour. Such adjustments might have been used to indicate refusal an admission, alternative admission options or regular admission.

²² This result is not surprising. The covariate is expected to be significantly correlated with the response variable and was chosen because of the results of 3.1.1.

Upon examination of the descriptive statistics tables we note the following:

For variables that emerge as statistically significant:

- Within the *educational department* variable the departments

Amongst the old educational departments

Departments with the highest mean numbers of credits are:

- Cape Educational Department (CA)
- Transvaal Education Department (TR)
- Joint Matriculation Board (JM)
- Natal Educational Department (NA)
- Department of Education and Culture: House of Delegates (IN)

Department with the lowest mean numbers of credits are:

- Department of Education and Training (ET)
- Transkei Education Department (TK)

Amongst the new educational departments

Departments with the highest mean numbers of credits are:

- Western Cape Educational Department (WC)
- Gauteng Educational Department (GT)
- Mpumalanga Educational Department (MP)

Departments with the lowest mean numbers of credits are:

- Northern Province Educational Department (NP)
- North West Educational Department (NW)

We note that making any comparisons between the old and new educational departments based on these averages maybe at this stage rather misleading. There is only one year (i.e. the year 1997) which partially incorporates the set of the new educational departments, and persistence of the old codes may indicate weaker students from previous cohorts.

- Within the *colour* variable, White is the category with the highest mean number of credits (it is also the largest group of students in all the time periods). The second highest mean of the number of credits is held by either the category Indian or Coloured (changing in different time periods). The category Black has the lowest mean number of credits.

For variables that do not emerge as statistically significant:

- Within the *home language* variable the languages with the highest mean number of credits are Chinese, English, and Afrikaans. The means for all the African languages are lower than in the other categories. Within that group, however, Ndebele, Southern and Northern Sotho emerge in the combined years analysis as the languages with the highest means (however, the respective groups are rather small).
- Within the gender variable, females obtain on average a higher number of credits than males, consistently in all the time periods considered.

The confidence intervals for large samples are narrow, indicating low variability of the calculated means. However as the sample size decreases the confidence intervals become wider. Sometimes the confidence intervals include even negative values. This feature is observed in particular for very small categories where the standard deviations are very large. The estimates with high standard errors ought to be treated with caution.

What has been established so far?

Besides confirming the significant relationship between the faculty point score and the number of credits obtained, we have been able to determine other factors (the educational department and the colour) that retrospectively appear to further explain the different levels of the performance variable in the various student groupings.

b) Binary Response Analysis

Calendar year 1990

ANOVA RESULTS: BINARY CREDITS IN 1990: BCREDITO							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1.00	18.58	213.00	0.11	173.75	0.0000
{1}EDUCATION	Fixed	9.00	0.91	213.00	0.11	8.52	0.0000
{2}LANGUAGE	Fixed	9.00	0.25	213.00	0.11	2.31	0.0171
{3}COLOUR	Fixed	2.00	0.48	213.00	0.11	4.51	0.0120
{4}GENDER	Fixed	1.00	0.12	213.00	0.11	1.15	0.2843

DESCRIPTIVE STATISTICS: BINARY CREDITS IN 1990: BCREDITO							
	LEVEL OF FACTOR	N	MEAN BCREDITO	STD.DEV. BCREDITO	STD. ERROR BCREDITO	-.95% CNF. LMT	+.95% CNF. LMT
TOTAL		236.00	0.66	0.47	0.03	0.60	0.72
{1}EDUCATION	CO	57.00	0.56	0.50	0.07	0.43	0.69
{1}EDUCATION	JM	16.00	0.63	0.50	0.13	0.36	0.89
{1}EDUCATION	CA	81.00	0.88	0.33	0.04	0.80	0.95
{1}EDUCATION	NE	10.00	0.50	0.53	0.17	0.12	0.88
{1}EDUCATION	IN	5.00	1.00	0.00	0.00	1.00	1.00
{1}EDUCATION	TR	21.00	0.90	0.30	0.07	0.77	1.04
{1}EDUCATION	ET	28.00	0.00	0.00	0.00	0.00	0.00
{1}EDUCATION	NA	12.00	1.00	0.00	0.00	1.00	1.00
{1}EDUCATION	OF	3.00	0.67	0.58	0.33	-0.77	2.10
{1}EDUCATION	TK	3.00	0.00	0.00	0.00	0.00	0.00
{2}LANGUAGE	EA	11.00	0.45	0.52	0.16	0.10	0.81
{2}LANGUAGE	EN	176.00	0.82	0.38	0.03	0.77	0.88
{2}LANGUAGE	AF	10.00	0.50	0.53	0.17	0.12	0.88
{2}LANGUAGE	ZU	8.00	0.00	0.00	0.00	0.00	0.00
{2}LANGUAGE	TW	9.00	0.00	0.00	0.00	0.00	0.00
{2}LANGUAGE	NS	3.00	0.33	0.58	0.33	-1.10	1.77
{2}LANGUAGE	XH	13.00	0.00	0.00	0.00	0.00	0.00
{2}LANGUAGE	SS	4.00	0.00	0.00	0.00	0.00	0.00
{2}LANGUAGE	SW	1.00	0.00				
{2}LANGUAGE	VE	1.00	0.00				
{3}COLOUR	C	57.00	0.53	0.50	0.07	0.39	0.66
{3}COLOUR	W	127.00	0.90	0.30	0.03	0.84	0.95
{3}COLOUR	I	13.00	0.85	0.38	0.10	0.62	1.07
{3}COLOUR	B	39.00	0.03	0.16	0.03	-0.03	0.08
{4}GENDER	M	145.00	0.61	0.49	0.04	0.53	0.69
{4}GENDER	F	91.00	0.74	0.44	0.05	0.64	0.83

Calendar year 1991

ANOVA RESULTS: BINARY CREDITS IN 1991: BCREDIT1							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1.00	7.02	243.00	0.15	47.61	0.0000
(1)EDUCATION	Fixed	9.00	0.56	243.00	0.15	3.77	0.0002
(2)LANGUAGE	Fixed	10.00	0.17	243.00	0.15	1.13	0.3420
(3)COLOUR	Fixed	2.00	0.01	243.00	0.15	0.09	0.9178
(4)GENDER	Fixed	1.00	0.13	243.00	0.15	0.89	0.3456

DESCRIPTIVE STATISTICS: BINARY CREDITS IN 1991: BCREDIT1							
	LEVEL OF FACTOR	N	MEAN BCREDIT1	STD.DEV. BCREDIT1	STD. ERROR BCREDIT1	-.95% CNF. LMT	+.95% CNF. LMT
TOTAL		267.00	0.75	0.43	0.03	0.70	0.80
(1)EDUCATION	CO	86.00	0.58	0.50	0.05	0.48	0.69
(1)EDUCATION	JM	7.00	1.00	0.00	0.00	1.00	1.00
(1)EDUCATION	CA	85.00	0.94	0.24	0.03	0.89	0.99
(1)EDUCATION	NE	14.00	0.57	0.51	0.14	0.27	0.87
(1)EDUCATION	IN	12.00	0.83	0.39	0.11	0.59	1.08
(1)EDUCATION	TR	19.00	0.89	0.32	0.07	0.74	1.05
(1)EDUCATION	ET	24.00	0.54	0.51	0.10	0.33	0.76
(1)EDUCATION	NA	15.00	0.93	0.26	0.07	0.79	1.08
(1)EDUCATION	OF	2.00	0.50	0.71	0.50	-6.85	6.85
(1)EDUCATION	TK	3.00	0.33	0.58	0.33	-1.10	1.77
(2)LANGUAGE	EA	38.00	0.63	0.49	0.08	0.47	0.79
(2)LANGUAGE	EN	186.00	0.82	0.39	0.03	0.76	0.87
(2)LANGUAGE	AF	7.00	0.71	0.49	0.18	0.26	1.17
(2)LANGUAGE	ZU	5.00	0.60	0.55	0.24	-0.08	1.28
(2)LANGUAGE	TW	10.00	0.60	0.52	0.16	0.23	0.97
(2)LANGUAGE	NS	1.00	0.00				
(2)LANGUAGE	XH	12.00	0.42	0.51	0.15	0.09	0.74
(2)LANGUAGE	SS	3.00	0.67	0.58	0.33	-0.77	2.10
(2)LANGUAGE	SW	1.00	0.00				
(2)LANGUAGE	VE	3.00	1.00	0.00	0.00	1.00	1.00
(2)LANGUAGE	EA	1.00	1.00				
(3)COLOUR	C	88.00	0.59	0.49	0.05	0.49	0.70
(3)COLOUR	W	128.00	0.92	0.27	0.02	0.87	0.97
(3)COLOUR	I	18.00	0.78	0.43	0.10	0.57	0.99
(3)COLOUR	B	33.00	0.52	0.51	0.09	0.34	0.70
(4)GENDER	M	159.00	0.72	0.45	0.04	0.65	0.79
(4)GENDER	F	108.00	0.81	0.40	0.04	0.73	0.88

Calendar year 1992

ANOVA RESULTS: BINARY CREDITS IN 1992: BCREDIT2							
	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1.00	10.73	219.00	0.13	81.02	0.0000
{1}EDUCATION	Fixed	8.00	1.02	219.00	0.13	7.73	0.0000
{2}LANGUAGE	Fixed	9.00	0.23	219.00	0.13	1.71	0.0887
{3}COLOUR	Fixed	3.00	0.73	219.00	0.13	5.49	0.0012
{4}GENDER	Fixed	1.00	0.02	219.00	0.13	0.13	0.7228

DESCRIPTIVE STATISTICS: BINARY CREDITS IN 1992: BCREDIT2							
	LEVEL OF FACTOR	N	MEAN BCREDIT2	STD.DEV. BCREDIT2	STD. ERROR BCREDIT2	-.95% CNF. LMT	+.95% CNF. LMT
TOTAL		242.00	0.69	0.47	0.03	0.63	0.74
{1}EDUCATION	CO	54.00	0.61	0.49	0.07	0.48	0.76
{1}EDUCATION	JM	8.00	1.00	0.00	0.00	1.00	1.00
{1}EDUCATION	CA	89.00	0.90	0.30	0.03	0.84	0.96
{1}EDUCATION	NE	13.00	0.69	0.48	0.13	0.40	0.98
{1}EDUCATION	IN	8.00	0.38	0.52	0.18	-0.06	0.81
{1}EDUCATION	TR	11.00	0.82	0.40	0.12	0.55	1.09
{1}EDUCATION	ET	41.00	0.27	0.45	0.07	0.13	0.41
{1}EDUCATION	NA	13.00	1.00	0.00	0.00	1.00	1.00
{1}EDUCATION	TK	5.00	0.00	0.00	0.00	0.00	0.00
{2}LANGUAGE	EA	6.00	0.83	0.41	0.17	0.40	1.26
{2}LANGUAGE	EN	173.00	0.80	0.40	0.03	0.74	0.86
{2}LANGUAGE	AF	10.00	0.60	0.52	0.16	0.23	0.97
{2}LANGUAGE	ZU	9.00	0.66	0.53	0.18	0.15	0.96
{2}LANGUAGE	TW	9.00	0.00	0.00	0.00	0.00	0.00
{2}LANGUAGE	NS	5.00	0.40	0.55	0.24	-0.28	1.08
{2}LANGUAGE	XH	20.00	0.30	0.47	0.11	0.08	0.52
{2}LANGUAGE	SS	7.00	0.67	0.53	0.20	0.08	1.07
{2}LANGUAGE	VE	1.00	0.00				
{2}LANGUAGE	EA	2.00	0.00	0.00	0.00	0.00	0.00
{3}COLOUR	C	59.00	0.61	0.49	0.06	0.48	0.74
{3}COLOUR	W	114.00	0.93	0.26	0.02	0.88	0.98
{3}COLOUR	I	15.00	0.40	0.51	0.13	0.12	0.68
{3}COLOUR	B	54.00	0.33	0.48	0.06	0.20	0.46
{4}GENDER	M	160.00	0.66	0.47	0.04	0.59	0.74
{4}GENDER	F	82.00	0.73	0.45	0.05	0.63	0.83

Calendar year 1993

ANOVA RESULTS: BINARY CREDITS IN 1993: BCREDIT3							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1.00	5.76	217.00	0.16	35.91	0.0000
{1}EDUCATION	Fixed	9.00	0.55	217.00	0.16	3.45	0.0005
{2}LANGUAGE	Fixed	12.00	0.12	217.00	0.16	0.77	0.6843
{3}COLOUR	Fixed	2.00	0.76	217.00	0.16	4.77	0.0094
{4}GENDER	Fixed	1.00	0.00	217.00	0.16	0.00	0.9605

DESCRIPTIVE STATISTICS: BINARY CREDITS IN 1993: BCREDIT3							
	LEVEL OF FACTOR	N	MEAN BCREDIT3	STD.DEV. BCREDIT3	STD. ERROR BCREDIT3	-.95% CNF. LMT	+.95% CNF. LMT
TOTAL		243.00	0.72	0.45		0.67	0.78
{1}EDUCATION	CO	38.00	0.63	0.49	0.08	0.47	0.79
{1}EDUCATION	JM	16.00	0.81	0.40	0.10	0.60	1.03
{1}EDUCATION	CA	75.00	0.93	0.25	0.03	0.88	0.99
{1}EDUCATION	NE	15.00	0.47	0.52	0.13	0.18	0.75
{1}EDUCATION	IN	16.00	0.69	0.48	0.12	0.43	0.94
{1}EDUCATION	TR	18.00	0.83	0.38	0.09	0.64	1.02
{1}EDUCATION	ET	45.00	0.49	0.51	0.08	0.34	0.64
{1}EDUCATION	NA	9.00	0.89	0.33	0.11	0.63	1.16
{1}EDUCATION	OF	3.00	0.67	0.58	0.33	-0.77	2.10
{1}EDUCATION	TK	8.00	0.50	0.53	0.19	0.05	0.95
{2}LANGUAGE	EA	4.00	0.50	0.58	0.29	-0.42	1.42
{2}LANGUAGE	EN	169.00	0.80	0.40	0.03	0.74	0.86
{2}LANGUAGE	AF	10.00	0.80	0.42	0.13	0.50	1.10
{2}LANGUAGE	ZU	14.00	0.57	0.51	0.14	0.27	0.87
{2}LANGUAGE	TW	5.00	0.40	0.55	0.24	-0.28	1.08
{2}LANGUAGE	NS	6.00	0.33	0.52	0.21	-0.21	0.88
{2}LANGUAGE	XH	21.00	0.52	0.51	0.11	0.29	0.76
{2}LANGUAGE	SS	6.00	0.67	0.52	0.21	0.12	1.21
{2}LANGUAGE	SW	1.00	0.00				
{2}LANGUAGE	VE	3.00	0.33	0.58	0.33	-1.10	1.77
{2}LANGUAGE	EA	2.00	1.00	0.00	0.00	1.00	1.00
{2}LANGUAGE	EN	1.00	0.00				
{2}LANGUAGE	AF	1.00	1.00				
{3}COLOUR	C	36.00	0.69	0.47	0.08	0.54	0.85
{3}COLOUR	W	119.00	0.90	0.30	0.03	0.84	0.95
{3}COLOUR	I	30.00	0.50	0.51	0.09	0.31	0.69
{3}COLOUR	B	58.00	0.50	0.50	0.07	0.37	0.63
{4}GENDER	M	145.00	0.71	0.46	0.04	0.64	0.79
{4}GENDER	F	98.00	0.74	0.44	0.04	0.66	0.83

Calendar year 1994

ANOVA RESULTS: BINARY CREDITS IN 1994: BCREDIT4							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1.00	7.76	278.00	0.16	49.52	0.0000
(1)EDUCATION	Fixed	9.00	0.74	278.00	0.16	4.73	0.0000
(2)LANGUAGE	Fixed	11.00	0.30	278.00	0.16	1.95	0.0340
(3)COLOUR	Fixed	3.00	0.11	278.00	0.16	0.69	0.5615
(4)GENDER	Fixed	1.00	0.01	278.00	0.16	0.04	0.8443

DESCRIPTIVE STATISTICS: BINARY CREDITS IN 1994: BCREDIT4							
	LEVEL OF FACTOR	N	MEAN BCREDIT4	STD.DEV. BCREDIT4	STD. ERROR BCREDIT4	-.95% CNF. LMT	+.95% CNF. LMT
TOTAL		304.00	0.72	0.45	0.03	0.67	0.77
(1)EDUCATION	CO	73.00	0.59	0.50	0.06	0.47	0.70
(1)EDUCATION	CA	94.00	0.89	0.31	0.03	0.83	0.96
(1)EDUCATION	NE	13.00	0.46	0.52	0.14	0.15	0.78
(1)EDUCATION	IN	19.00	0.68	0.48	0.11	0.45	0.91
(1)EDUCATION	TR	25.00	0.84	0.37	0.07	0.69	0.99
(1)EDUCATION	ET	45.00	0.42	0.50	0.07	0.27	0.57
(1)EDUCATION	NA	18.00	0.83	0.38	0.09	0.64	1.02
(1)EDUCATION	OF	1.00	1.00				
(1)EDUCATION	TK	1.00	1.00				
(1)EDUCATION	IE	15.00	1.00	0.00	0.00	1.00	1.00
(2)LANGUAGE	EA	9.00	0.67	0.50	0.17	0.28	1.05
(2)LANGUAGE	AF	225.00	0.77	0.42	0.03	0.72	0.83
(2)LANGUAGE	ZU	14.00	0.71	0.47	0.13	0.44	0.98
(2)LANGUAGE	TW	6.00	0.17	0.41	0.17	-0.26	0.60
(2)LANGUAGE	NS	5.00	0.40	0.55	0.24	-0.28	1.08
(2)LANGUAGE	XH	4.00	0.75	0.50	0.25	-0.06	1.55
(2)LANGUAGE	SS	17.00	0.29	0.47	0.11	0.05	0.54
(2)LANGUAGE	SW	9.00	0.67	0.50	0.17	0.28	1.05
(2)LANGUAGE	VE	7.00	0.86	0.38	0.14	0.51	1.21
(2)LANGUAGE	CH	4.00	0.76	0.50	0.25	-0.06	1.55
(2)LANGUAGE	EA	3.00	0.67	0.58	0.33	-0.77	2.10
(2)LANGUAGE	EN	1.00	0.00				
(3)COLOUR	C	77.00	0.61	0.49	0.06	0.50	0.72
(3)COLOUR	W	147.00	0.86	0.34	0.03	0.81	0.92
(3)COLOUR	I	26.00	0.65	0.49	0.10	0.46	0.85
(3)COLOUR	B	54.00	0.50	0.50	0.07	0.36	0.64
(4)GENDER	M	173.00	0.69	0.46	0.04	0.62	0.76
(4)GENDER	F	131.00	0.75	0.44	0.04	0.67	0.82

Calendar year 1995

ANOVA RESULTS: BINARY CREDITS IN 1995: BCREDITS							
	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1.00	8.74	310.00	0.14	64.07	0.0000
{1}EDUCATION	Fixed	9.00	1.17	310.00	0.14	8.58	0.0000
{2}LANGUAGE	Fixed	11.00	0.23	310.00	0.14	1.67	0.0803
{3}COLOUR	Fixed	3.00	0.11	310.00	0.14	0.82	0.4834
{4}GENDER	Fixed	1.00	0.06	310.00	0.14	0.45	0.5022

DESCRIPTIVE STATISTICS BINARY CREDITS IN 1995: BCREDITS							
	LEVEL OF FACTOR	N	MEAN BCREDITS	STD.DEV. BCREDITS	STD. ERROR BCREDITS	-.95% CNF. LMT	+.95% CNF. LMT
TOTAL		336.00	0.74	0.44	0.02	0.69	0.79
{1}EDUCATION	CO	53.00	0.60	0.49	0.07	0.47	0.74
{1}EDUCATION	CA	128.00	0.91	0.28	0.02	0.86	0.96
{1}EDUCATION	NE	5.00	0.80	0.45	0.20	0.24	1.36
{1}EDUCATION	IN	17.00	0.35	0.49	0.12	0.10	0.61
{1}EDUCATION	TR	28.00	0.86	0.36	0.07	0.72	1.00
{1}EDUCATION	ET	47.00	0.34	0.48	0.07	0.20	0.48
{1}EDUCATION	NA	22.00	0.86	0.35	0.07	0.71	1.02
{1}EDUCATION	OF	5.00	0.80	0.45	0.20	0.24	1.36
{1}EDUCATION	TK	5.00	0.80	0.45	0.20	0.24	1.36
{1}EDUCATION	IE	26.00	0.88	0.33	0.06	0.75	1.02
{2}LANGUAGE	EA	5.00	0.60	0.55	0.24	-0.08	1.28
{2}LANGUAGE	AF	257.00	0.81	0.39	0.02	0.77	0.86
{2}LANGUAGE	ZU	6.00	0.33	0.52	0.21	-0.21	0.88
{2}LANGUAGE	TW	8.00	0.75	0.46	0.16	0.36	1.14
{2}LANGUAGE	NS	10.00	0.40	0.52	0.16	0.03	0.77
{2}LANGUAGE	XH	4.00	0.75	0.50	0.25	-0.05	1.55
{2}LANGUAGE	SS	28.00	0.43	0.50	0.10	0.23	0.62
{2}LANGUAGE	SW	7.00	0.57	0.53	0.20	0.08	1.07
{2}LANGUAGE	VE	1.00	0.00				
{2}LANGUAGE	CH	4.00	0.25	0.50	0.25	-0.55	1.05
{2}LANGUAGE	EA	4.00	1.00	0.00	0.00	1.00	1.00
{2}LANGUAGE	EN	2.00	0.50	0.71	0.50	-5.85	6.85
{3}COLOUR	C	54.00	0.61	0.49	0.07	0.48	0.75
{3}COLOUR	W	182.00	0.91	0.29	0.02	0.86	0.95
{3}COLOUR	I	34.00	0.56	0.50	0.09	0.38	0.73
{3}COLOUR	B	66.00	0.48	0.50	0.06	0.36	0.61
{4}GENDER	M	189.00	0.70	0.46	0.03	0.64	0.77
{4}GENDER	F	147.00	0.79	0.41	0.03	0.72	0.86

Calendar year 1996

ANOVA RESULTS: BINARY CREDITS IN 1996: BCREDIT6							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1.00	14.69	278.00	0.17	89.05	0.0000
{1}EDUCATION	Fixed	9.00	1.13	278.00	0.17	6.84	0.0000
{2}LANGUAGE	Fixed	10.00	0.18	278.00	0.17	1.11	0.3509
{3}COLOUR	Fixed	3.00	0.31	278.00	0.17	1.85	0.1375
{4}GENDER	Fixed	1.00	0.12	278.00	0.17	0.70	0.4031

DESCRIPTIVE STATISTICS: BINARY CREDITS IN 1996: BCREDIT6							
	LEVEL OF FACTOR	N	MEAN BCREDIT6	STD.DEV. BCREDIT6	STD. ERROR BCREDIT6	-.95% CNF. LMT	+.95% CNF. LMT
TOTAL		303.00	0.58	0.49	0.03	0.53	0.64
{1}EDUCATION	CO	50.00	0.42	0.50	0.07	0.28	0.56
{1}EDUCATION	CA	97.00	0.82	0.38	0.04	0.75	0.90
{1}EDUCATION	NE	7.00	0.29	0.49	0.18	-0.17	0.74
{1}EDUCATION	IN	15.00	0.33	0.49	0.13	0.06	0.60
{1}EDUCATION	TR	28.00	0.86	0.36	0.07	0.72	1.00
{1}EDUCATION	ET	70.00	0.27	0.45	0.05	0.16	0.38
{1}EDUCATION	NA	14.00	0.86	0.36	0.10	0.65	1.07
{1}EDUCATION	OF	4.00	1.00	0.00	0.00	1.00	1.00
{1}EDUCATION	TK	9.00	0.33	0.50	0.17	-0.06	0.72
{1}EDUCATION	IE	9.00	0.78	0.44	0.15	0.44	1.12
{2}LANGUAGE	EA	3.00	0.67	0.58	0.33	-0.77	2.10
{2}LANGUAGE	AF	194.00	0.71	0.45	0.03	0.65	0.78
{2}LANGUAGE	ZU	9.00	0.44	0.53	0.18	0.04	0.85
{2}LANGUAGE	TW	16.00	0.19	0.40	0.10	-0.03	0.40
{2}LANGUAGE	NS	12.00	0.50	0.52	0.15	0.17	0.83
{2}LANGUAGE	XH	12.00	0.50	0.52	0.15	0.17	0.83
{2}LANGUAGE	SS	29.00	0.24	0.44	0.08	0.08	0.41
{2}LANGUAGE	SW	9.00	0.44	0.53	0.18	0.04	0.85
{2}LANGUAGE	VE	8.00	0.13	0.35	0.13	-0.17	0.42
{2}LANGUAGE	CH	9.00	0.67	0.50	0.17	0.28	1.05
{2}LANGUAGE	EA	2.00	0.00	0.00	0.00	0.00	0.00
{3}COLOUR	C	48.00	0.52	0.50	0.07	0.37	0.67
{3}COLOUR	W	138.00	0.84	0.37	0.03	0.78	0.90
{3}COLOUR	I	27.00	0.33	0.48	0.09	0.14	0.52
{3}COLOUR	B	90.00	0.30	0.46	0.05	0.20	0.40
{4}GENDER	M	166.00	0.55	0.50	0.04	0.47	0.62
{4}GENDER	F	137.00	0.63	0.49	0.04	0.55	0.71

Calendar year 1997

ANOVA RESULTS: BINARY CREDITS IN 1997: BCREDIT7							
	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1.00	10.12	268.00	0.17	58.34	0.0000
(1)EDUCATION	Fixed	17.00	0.36	268.00	0.17	2.08	0.0079
(2)LANGUAGE	Fixed	11.00	0.18	268.00	0.17	1.04	0.4135
(3)COLOUR	Fixed	3.00	0.23	268.00	0.17	1.34	0.2610
(4)GENDER	Fixed	1.00	0.06	268.00	0.17	0.37	0.5410

DESCRIPTIVE STATISTICS: BINARY CREDITS IN 1997: BCREDIT7							
	LEVEL OF FACTOR	N	MEAN BCREDIT7	STD.DEV. BCREDIT7	STD. ERROR BCREDIT7	-.95% CNF. LMT	+.95% CNF. LMT
TOTAL		302.00	0.68	0.47	0.03	0.63	0.73
(1)EDUCATION	CO	3.00	0.33	0.58	0.33	-1.10	1.77
(1)EDUCATION	CA	8.00	1.00	0.00	0.00	1.00	1.00
(1)EDUCATION	NE	1.00	0.00				
(1)EDUCATION	IN	1.00	0.00				
(1)EDUCATION	TR	9.00	0.89	0.33	0.11	0.63	1.15
(1)EDUCATION	ET	27.00	0.52	0.51	0.10	0.32	0.72
(1)EDUCATION	NA	1.00	0.00				
(1)EDUCATION	OF	2.00	0.50	0.71	0.50	-5.85	6.85
(1)EDUCATION	TK	3.00	0.67	0.58	0.33	-0.77	2.10
(1)EDUCATION	IE	26.00	0.81	0.40	0.08	0.65	0.97
(1)EDUCATION	WC	124.00	0.76	0.43	0.04	0.68	0.83
(1)EDUCATION	EC	15.00	0.53	0.52	0.13	0.25	0.82
(1)EDUCATION	KN	30.00	0.53	0.51	0.09	0.34	0.72
(1)EDUCATION	GT	24.00	0.79	0.41	0.08	0.62	0.97
(1)EDUCATION	NW	10.00	0.70	0.48	0.15	0.35	1.05
(1)EDUCATION	NC	5.00	0.60	0.55	0.24	-0.08	1.28
(1)EDUCATION	NP	10.00	0.20	0.42	0.13	-0.10	0.50
(1)EDUCATION	MP	3.00	0.67	0.58	0.33	-0.77	2.10
(2)LANGUAGE	EA	4.00	1.00	0.00	0.00	1.00	1.00
(2)LANGUAGE	AF	199.00	0.77	0.42	0.03	0.72	0.83
(2)LANGUAGE	ZU	10.00	0.70	0.48	0.15	0.35	1.05
(2)LANGUAGE	TVV	15.00	0.33	0.49	0.13	0.06	0.60
(2)LANGUAGE	NS	14.00	0.50	0.52	0.14	0.20	0.80
(2)LANGUAGE	XH	10.00	0.40	0.52	0.16	0.03	0.77
(2)LANGUAGE	SS	25.00	0.44	0.51	0.10	0.23	0.65
(2)LANGUAGE	SW	7.00	0.71	0.49	0.18	0.26	1.17
(2)LANGUAGE	VE	3.00	0.67	0.58	0.33	-0.77	2.10
(2)LANGUAGE	CH	2.00	0.00	0.00	0.00	0.00	0.00
(2)LANGUAGE	TG	6.00	0.83	0.41	0.17	0.40	1.26
(2)LANGUAGE	ND	7.00	0.29	0.49	0.18	-0.17	0.74
(3)COLOUR	C	50.00	0.60	0.49	0.07	0.46	0.74
(3)COLOUR	W	140.00	0.86	0.35	0.03	0.80	0.92
(3)COLOUR	I	25.00	0.68	0.48	0.10	0.48	0.88
(3)COLOUR	B	87.00	0.45	0.50	0.05	0.34	0.55
(4)GENDER	M	173.00	0.66	0.48	0.04	0.59	0.73
(4)GENDER	F	129.00	0.71	0.45	0.04	0.63	0.79

Calendar years 1990-1997 combined

ANOVA RESULTS: BCREDITS IN 1990-1997 COMBINED: BCREDIT							
	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1.00	83.59	2197.00	0.15	542.07	0.0000
(1)EDUCATION	Fixed	18.00	2.58	2197.00	0.15	16.75	0.0000
(2)LANGUAGE	Fixed	12.00	0.30	2197.00	0.15	1.92	0.0282
(3)COLOUR	Fixed	3.00	0.56	2197.00	0.15	3.65	0.0121
(4)GENDER	Fixed	1.00	0.37	2197.00	0.15	2.40	0.1213

DESCRIPTIVE STATISTICS BCREDITS IN 1990-1997 COMBINED: BCREDIT							
	LEVEL OF FACTOR	N	MEAN BCREDIT	STD.DEV. BCREDIT	STD. ERROR BCREDIT	-.95% CNF. LMT	+.95% CNF. LMT
TOTAL		2233.00	0.69	0.46	0.01	0.67	0.71
(1)EDUCATION	CO	414.00	0.57	0.50	0.02	0.52	0.62
(1)EDUCATION	JM	47.00	0.81	0.40	0.06	0.69	0.93
(1)EDUCATION	CA	657.00	0.90	0.30	0.01	0.87	0.92
(1)EDUCATION	NE	78.00	0.53	0.50	0.06	0.41	0.64
(1)EDUCATION	IN	93.00	0.57	0.50	0.05	0.47	0.67
(1)EDUCATION	TR	159.00	0.86	0.35	0.03	0.81	0.92
(1)EDUCATION	ET	327.00	0.36	0.48	0.03	0.30	0.40
(1)EDUCATION	NA	104.00	0.89	0.31	0.03	0.83	0.96
(1)EDUCATION	OF	20.00	0.75	0.44	0.10	0.54	0.96
(1)EDUCATION	TK	37.00	0.41	0.50	0.08	0.24	0.67
(1)EDUCATION	IE	76.00	0.87	0.34	0.04	0.79	0.96
(1)EDUCATION	WC	124.00	0.76	0.43	0.04	0.68	0.83
(1)EDUCATION	EC	15.00	0.53	0.52	0.13	0.25	0.82
(1)EDUCATION	KN	30.00	0.53	0.51	0.09	0.34	0.72
(1)EDUCATION	GT	24.00	0.79	0.41	0.08	0.62	0.97
(1)EDUCATION	NW	10.00	0.70	0.48	0.15	0.36	1.06
(1)EDUCATION	NC	5.00	0.60	0.55	0.24	-0.08	1.28
(1)EDUCATION	NP	10.00	0.20	0.42	0.13	-0.10	0.60
(1)EDUCATION	MP	3.00	0.67	0.58	0.33	-0.77	2.10
(2)LANGUAGE	EA	80.00	0.64	0.48	0.05	0.53	0.76
(2)LANGUAGE	EN	1579.00	0.79	0.41	0.01	0.77	0.81
(2)LANGUAGE	AF	76.00	0.62	0.49	0.06	0.51	0.73
(2)LANGUAGE	ZU	81.00	0.38	0.49	0.05	0.27	0.49
(2)LANGUAGE	TW	74.00	0.36	0.48	0.06	0.25	0.48
(2)LANGUAGE	NS	45.00	0.47	0.50	0.08	0.32	0.62
(2)LANGUAGE	XH	165.00	0.35	0.48	0.04	0.27	0.42
(2)LANGUAGE	SS	52.00	0.56	0.50	0.07	0.42	0.70
(2)LANGUAGE	SW	6.00	0.33	0.52	0.21	-0.21	0.88
(2)LANGUAGE	VE	27.00	0.33	0.48	0.09	0.14	0.52
(2)LANGUAGE	CH	28.00	0.82	0.39	0.07	0.67	0.97
(2)LANGUAGE	TG	18.00	0.33	0.49	0.11	0.09	0.57
(2)LANGUAGE	ND	2.00	0.60	0.71	0.50	-5.85	6.85
(3)COLOUR	C	469.00	0.59	0.49	0.02	0.55	0.64
(3)COLOUR	W	1095.00	0.89	0.31	0.01	0.87	0.91
(3)COLOUR	I	188.00	0.57	0.50	0.04	0.50	0.65
(3)COLOUR	B	481.00	0.40	0.49	0.02	0.35	0.44
(4)GENDER	M	1310.00	0.66	0.47	0.01	0.64	0.69
(4)GENDER	F	923.00	0.74	0.44	0.01	0.71	0.76

ANOVA is not usually applied to binary response data. However it is applied here as an exploratory device to focus upon possible indicators of prospective readmission.

Linear discriminant analysis is equivalent to a linear regression of the group indicator (understood as a response) on the continuous measures if ratio of one-way ANOVA of a binary response is a known monotonic function of the conventional Pearson χ^2 for a $2 \times k$ contingency table.

In the binary response analysis the following explanatory variables are found to be significant:

- faculty point score ²³ (in all the time periods considered)
- educational department (in all the time periods considered)
- home language (only in the following years 1994 and all the years combined data set)
- colour (only in the following years 1992, 1993 and all the years combined data set)

The lists of the significant variables and the years in which these variables are significant are the same for the ordinal and binary analyses. This aspect of the data suggests that a shift in focus from the number of credits obtained to the simple issue of (prospective) re-admission to the Faculty on the basis of performance will not require a change in the associated predictors. The faculty point score will dominate with minor adjustments arising for the other factors in a linear model.

²³ However this result was expected and for that reason the faculty point score was entered in the analysis as a covariate.

Due to the way in which the binary response variable had been coded ²⁴, the descriptive statistics tables represent the proportion of students who obtained at least 2 credits within the particular groupings of the categorical variables.

As expected, the categories of the particular variable with the highest proportion of students obtaining at least 2 credits are essentially those with the highest mean number of credits as summarised in the ordinal response analysis. The additional insights gained from the descriptive statistics tables of the binary response analysis are:

- In the time periods considered the proportion of the incoming first year students fulfilling the Faculty's re-admission criteria into the second year is well above 0.60 (1996 is the only exception with the proportion passing at 0.58).
- Only in individual years is any proportion actually 1.00. This feature will generally apply only for very small groups (e.g. in 1991 for Chinese or in 1994 for the Orange Free State Education Department admissions to U.C.T.)
- Also only in individual years is any proportion actually 0.00. Again this feature applies only for small groups (e.g. in 1992 for Venda or in 1997 for the DEC: House of Delegates)
- It is apparent that the proportions of students fulfilling the re-admission requirements from the non-White groupings (expressed either by the colour or the home language or the educational department variables) are lower than the proportions of students from the predominately White groupings.

Again the confidence intervals for large samples are narrow, indicating low variability of the calculated proportions. However as the sample size decreases the confidence intervals become wider. Sometimes the confidence intervals include even negative values or values greater than 1.00. This feature is observed in particular for very small categories where the standard deviations are very large. These estimates should be treated with caution.

²⁴ The variable was assigned 1 if a student obtained at least 2 credits or assigned 0 if a student obtained fewer than 2 credits.

3.1.3 COMPARISON OF REGRESSIONS BY GENDER AND RACE

In this section we have regressed ²⁵ the actual number for credits obtained for the students against the faculty point scores in the four race groups and for race by gender.

The reader may also note that the graphs have been displayed to illustrate the plausible analyses offered in section 2.2 in chapter 2.

However as each plotted data point represents a frequency of bivariate data greater than zero, the effect of any bivariate repetitions in the data is not explicit in the graphs. Better graphical presentations could use dots with size proportional to frequencies.

In examining the graphs whose size is fixed, we must note the origin and scale in interpreting any graphical similarities in contrasts. The regression equations and the correlation coefficients will assist in those interpretations.

²⁵ The following simple regression model has been fitted to each of the specified student groups:

$$y_i = \mu + \beta \cdot x_i$$

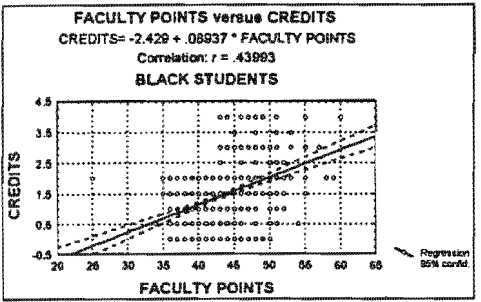
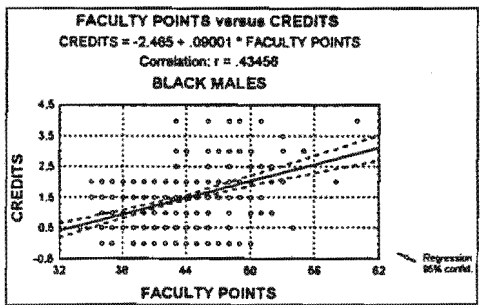
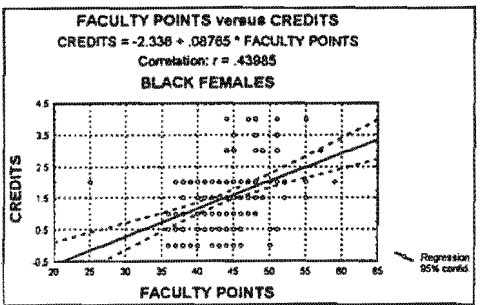
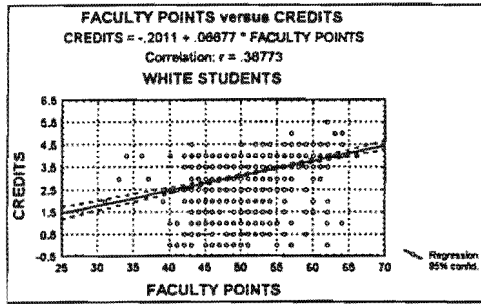
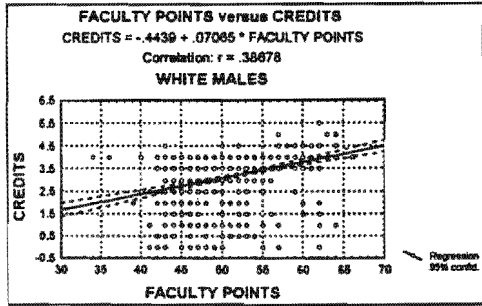
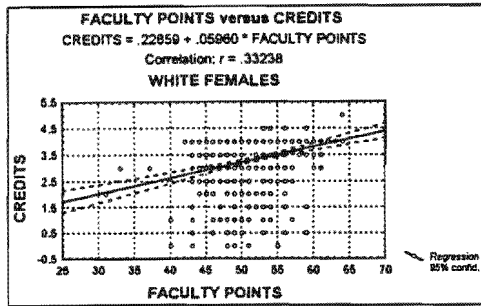
where

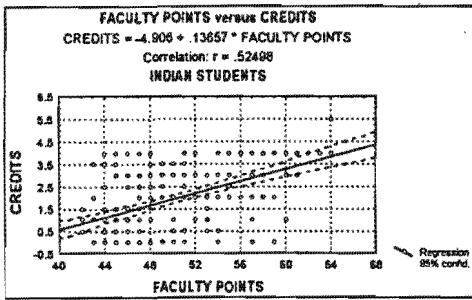
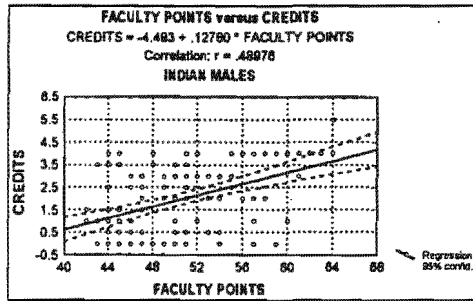
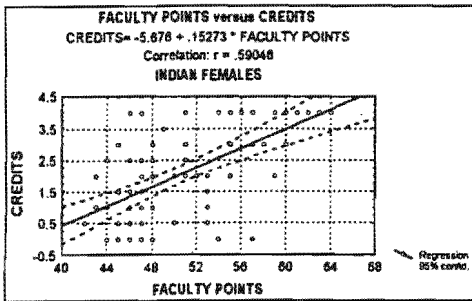
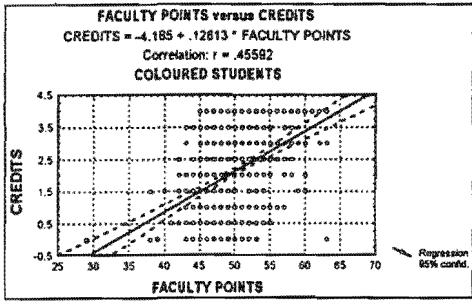
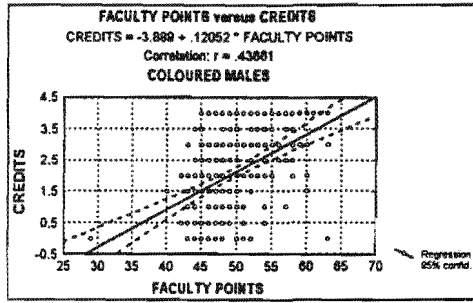
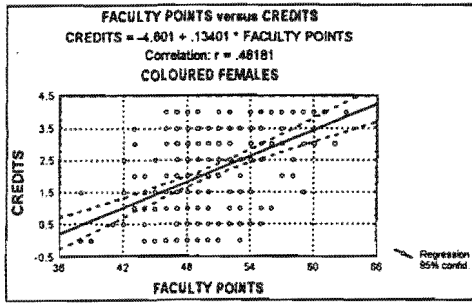
μ -the intercept term

y_i -the actual number of credits obtained by i^{th} student

x_i -the actual faculty point score of the i^{th} student

β -the regression coefficient of the faculty point score variable





The gradients of the derived regression lines increase across the race groups in the following order White, Black, Coloured and Indian. This finding suggests that as we move in the specified order across the four race groups, the increase in the number of credits obtained is more distinctly and sharply related to the increase in the faculty point score.

We observe further that the regression line for those White students with the faculty point scores of 48 and more flatter with fitted values between 2.5 and 4. Equivalently most of these qualifiers at least fulfil the Faculty's minimum readmission criteria into the second year.

The growth in the magnitudes of the gradients ²⁶ also suggests that the faculty point score may be a better predictor of academic performance for the non-White than for the White students.

We also note that the intercepts for most regression lines are negative. They decrease in exactly the order in which the gradients increase. This artefact arises from the restricted range of faculty point scores, and the minimum score being far from zero.

We also observe that the gradients and the intercepts of the regression lines within each of the race groups are almost identical for both female and male students, thus suggesting no statistically significant differences between females and males with the same faculty point score with respect to the average number of credits obtained. This result also confirms previous findings that the gender variable does not appear to contrast academic performance between students at first year level.

Finally we note that these regression methods are simply exploratory because the model fitting takes no account directly of the restricted ranges of the numbers of credits (0 to 4).

²⁶ When the derived regression lines are standardised, the gradients become equal to the Pearson correlation coefficients (or β 's as defined for the simple regression model). The coefficients have been calculated for each of the lines.

3.2 ASSOCIATION BETWEEN MATRIC POINT SCORE AND PERFORMANCE IN FIRST ACADEMIC YEAR AT UCT

3.2.1 MEASURES OF ASSOCIATION

a) Ordinal Response Analysis

The bivariate data sets were formed by the *matric point-scores* of the incoming first-year Science students admitted into the Faculty and the numbers of credits obtained by the students in their first calendar year of study.

A contingency table view of the data is implicit in some measures of association presented.

MATRIC POINT SCORE (X) and CREDITS IN FIRST YEAR (ordinal) (Y)

YEAR : 1990, valid cases: 236		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.66	< 0.05
Spearman correlation coefficient	0.66 (t=13.55)	< 0.05
Kendall's Tau	0.51 (Z=11.77)	< 0.05
Somers' D Statistics	X/Y=0.54 Y/X=0.39	< 0.05 < 0.05
Gamma Statistic	0.57 (Z=11.77)	< 0.05
Pearson Chi-square Statistic	135.78	0.0152

YEAR : 1991, valid cases: 267		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.52	< 0.05
Spearman correlation coefficient	0.55 (t=10.69)	< 0.05
Kendall's Tau	0.41 (Z=10.22)	< 0.05
Somers' D Statistics	X/Y=0.51 Y/X=0.39	< 0.05 < 0.05
Gamma Statistic	0.47 (Z=10.22)	< 0.05
Pearson Chi-square Statistic	171.16	< 0.05

YEAR : 1992, valid cases: 253		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.56	< 0.05
Spearman correlation coefficient	0.59 (t=11.73)	< 0.05
Kendall's Tau	0.45 (t=10.68)	< 0.05
Somers' D Statistics	X/Y=0.57 Y/X=0.42	< 0.05 < 0.05
Gamma Statistic	0.51 (Z=10.63)	< 0.05
Pearson Chi-square Statistic	142.71	< 0.05

YEAR : 1993, valid cases: 243		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.50	< 0.05
Spearman correlation coefficient	0.51 (t=9.38)	< 0.05
Kendall's Tau	0.39 (Z=9.06)	< 0.05
Somers' D Statistics	X/Y=0.47 Y/X=0.36	< 0.05 < 0.05
Gamma Statistic	0.44 (Z=9.38)	< 0.05
Pearson Chi-square Statistic	125.23	0.0015

YEAR : 1994, valid cases: 307		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.48	< 0.05
Spearman correlation coefficient	0.49 (t=9.78)	< 0.05
Kendall's Tau	0.38 (Z=9.83)	< 0.05
Somers' D Statistics	X/Y=0.46 Y/X=0.35	< 0.05 < 0.05
Gamma Statistic	0.43 (Z=9.83)	< 0.05
Pearson Chi-square Statistic	154.71	< 0.05

YEAR : 1995, valid cases: 338		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.54	< 0.05
Spearman correlation coefficient	0.55 (t=11.98)	< 0.05
Kendall's Tau	0.42 (Z=11.42)	< 0.05
Somers' D Statistics	X/Y=0.54 Y/X=0.41	< 0.05 < 0.05
Gamma Statistic	0.47 (Z=11.42)	< 0.05
Pearson Chi-square Statistic	336.61	< 0.05

YEAR : 1996, valid cases: 303		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.58	< 0.05
Spearman correlation coefficient	0.58 (t=12.41)	< 0.05
Kendall's Tau	0.44 (Z=11.54)	< 0.05
Somers' D Statistics	X/Y=0.53 Y/X=0.43	< 0.05 < 0.05
Gamma Statistic	0.49 (Z=11.54)	< 0.05
Pearson Chi-square Statistic	220.91	< 0.05

YEAR : 1997, valid cases: 302		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.62	< 0.05
Spearman correlation coefficient	0.63 (t=13.91)	< 0.05
Kendall's Tau	0.48 (Z=12.54)	< 0.05
Somers' D Statistics	X/Y=0.59 Y/X=0.47	< 0.05 < 0.05
Gamma Statistic	0.61 (Z=12.54)	< 0.05
Pearson Chi-square Statistic	242.67	< 0.05

YEARS: 1990-1997 valid cases: 2249		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.56	<0.05
Spearman correlation coefficient	0.58 (t=33.56)	<0.05
Kendall's Tau	0.44 (Z=31.32)	<0.05
Somers' D Statistics	X/Y=0.53 Y/X=0.41	<0.05 <0.05
Gamma Statistic	0.49 (Z=31.32)	<0.05
Pearson Chi-square statistic	936.14	<0.05

Upon examination of the measures we note that similar patterns to those observed for in section 3.1.1 a) for the *faculty point score* emerge in this section. Virtually all the comments from section 3.1.1 a) could be repeated here. We may add that the measures presented in this section are only slightly higher and in many instances the magnitudes are even equal. These remarks apply for all the individual calendar years and the eight years combined data set. The slightly higher figures for the measures in this section initially suggest that an analysis comparing the explanatory effects of the faculty and matric point scores may be of interest.

Based on these statistical results, we may conclude that the first-year students admitted into the Faculty of Science with the higher (lower) *matric point scores* are more likely to pass the higher (lower) number of the full-year first-year courses or the equivalent in the first calendar year of their study, respectively.

b) Binary Response Analysis ²⁷

The bivariate data sets were formed by the actual *matric point-scores* and the performance variable with two categories: category 1 : incoming first year student obtained 2 or more credits and category 2: incoming first year student obtained fewer than 2 credits in the first calendar year of his or her study.

The results are displayed in the following nine tables below.

MATRIC POINT SCORE (X) and CREDITS IN FIRST YEAR (binary) (Y)

YEAR : 1990, valid cases: 236		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.60	<0.05
Spearman correlation coefficient	0.61 (t=11.96)	<0.05
Kendall's Tau	0.52 (Z=11.84)	<0.05
Somers' D Statistics	X/Y=0.75 Y/X=0.36	<0.05
Gamma Statistic	0.78 (Z=11.84)	<0.05
Pearson Chi-square Statistic	101.45	<0.05

YEAR : 1991, valid cases: 267		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.39	<0.05
Spearman correlation coefficient	0.42 (t=7.49)	<0.05
Kendall's Tau	0.35 (Z=8.61)	<0.05
Somers' D Statistics	X/Y=0.56 Y/X=0.22	<0.05
Gamma Statistic	0.59 (Z=8.61)	<0.05
Pearson Chi-square Statistic	65.31	<0.05

YEAR : 1992, valid cases: 253		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.46	<0.05
Spearman correlation coefficient	0.48 (t=8.58)	<0.05
Kendall's Tau	0.40 (Z=9.49)	<0.05
Somers' D Statistics	X/Y=0.60 Y/X=0.27	<0.05
Gamma Statistic	0.63 (Z=9.49)	<0.05
Pearson Chi-square Statistic	76.95	<0.05

²⁷ Note again that the binary analysis was performed here to address the utility the matric point score in relation to the current re-admission criteria in the Faculty of Science at UCT after first year (similar to analysis in 3.1.1 b).

YEAR : 1993, valid cases: 243		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.36	<0.05
Spearman correlation coefficient	0.35 (t=5.80)	<0.05
Kendall's Tau	0.29 (Z=6.85)	<0.05
Somers' D Statistics	X/Y=0.45 Y/X=0.19	<0.05 <0.05
Gamma Statistic	0.48 (Z=6.85)	<0.05
Pearson Chi-square Statistic	59.23	<0.05

YEAR : 1994, valid cases: 307		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.36	<0.05
Spearman correlation coefficient	0.36 (t=6.66)	<0.05
Kendall's Tau	0.30 (Z=7.85)	<0.05
Somers' D Statistics	X/Y=0.45 Y/X=0.20	<0.05 <0.05
Gamma Statistic	0.48 (Z=7.85)	<0.05
Pearson Chi-square Statistic	62.05	<0.05

YEAR : 1995, valid cases: 338		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.38	<0.05
Spearman correlation coefficient	0.37 (t=7.36)	<0.05
Kendall's Tau	0.31 (Z=8.60)	<0.05
Somers' D Statistics	X/Y=0.49 Y/X=0.20	<0.05 <0.05
Gamma Statistic	0.52 (Z=8.60)	<0.05
Pearson Chi-square Statistic	60.94	<0.05

YEAR : 1996, valid cases: 303		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.46	<0.05
Spearman correlation coefficient	0.48 (t=9.55)	<0.05
Kendall's Tau	0.40 (Z=10.05)	<0.05
Somers' D Statistics	X/Y=0.56 Y/X=0.29	<0.05 <0.05
Gamma Statistic	0.59 (Z=10.50)	<0.05
Pearson Chi-square Statistic	87.64	<0.05

YEAR : 1997, valid cases: 302 MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.38	<0.05
Spearman correlation coefficient	0.38 (t=7.23)	<0.05
Kendall's Tau	0.32 (Z=7.23)	<0.05
Somers' D Statistics	X/Y=0.47 Y/X=0.22	<0.05 <0.05
Gamma Statistic	0.50 (Z=7.23)	<0.05
Pearson Chi-square Statistic	68.90	<0.05

YEARS: 1990-1997 valid cases: 2249 MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.43	<0.05
Spearman correlation coefficient	0.43 (t=22.88)	<0.05
Kendall's Tau	0.36 (Z=25.96)	<0.05
Somers' D Statistics	X/Y=0.54 Y/X=0.24	<0.05 <0.05
Gamma Statistic	0.57 (Z=25.96)	<0.05
Pearson Chi-square statistic	465.90	<0.05

Again upon examination of the measures we observe similar patterns to those from section 3.1.1 b) for the faculty point score. The comments from section 3.1.1 b) apply here. Again the magnitudes of the measures of association presented in this section are lower than those produced in the ordinal response analysis in 3.2.1 a). However the *matric score point* measures are either slightly higher or of equal magnitudes to those for the binary response analysis with the *faculty point score*.

Based on these statistical results, we may conclude that the incoming first-year students with the higher (lower) *matric point scores* are more likely (less likely) to fulfil the Faculty's second academic year re-admission criteria in the first calendar year of their study, respectively. Further evidence for matric points as a possibly superior predictor of performance, at least over this retrospective period, has emerged.

What has been established so far?

So far the measures of association have indicated the positive nature and statistical significance of the relationship between the variables under investigation. However they have specified ²⁸ which particular matric point scores may be associated with which specific numbers of credits. We may only say that the higher (*lower*) matric point scores are associated with the higher (*lower*) number of credits, respectively. Also we have not been able to incorporate other variables ²⁹ and thus to investigate their associations with the performance variable.

²⁸ This type of insight could, to some extent, be established by a careful inspection of the specific contingency tables.

²⁹ We may, however, investigate the other relationships by computing the measures of association between each of the other variable of interest and the performance variable, individually.

3.2.2 GENERALISED LINEAR MODELLING TECHNIQUES

Essentially the structure and the manner in which the results are presented in this section are the same as for the section 3.1.2. The difference is the *matric point score* replacing the faculty point score as a covariate.

In this section we display only the ANOVA tables. The univariate descriptive statistics tables relating to the additional categorical variables in the model are exactly the same whether the analyses are performed with the faculty or matric point scores, and hence been presented in 3.1.2. This feature arises because the same response variable (academic performance) and the same set of categorical variables were used in both investigations.

The ANOVA results are displayed on the following pages.

a) Ordinal Response Analysis

Calendar year 1990

ANOVA:RESULTS: CREDITS IN 1990: CREDIT0							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	229.76	213.00	0.99	231.37	0.0000
{1}EDUCATION	Fixed	9	8.00	213.00	0.99	8.05	0.0000
{2}LANGUAGE	Fixed	9	1.31	213.00	0.99	1.31	0.2309
{3}COLOUR	Fixed	2	1.29	213.00	0.99	1.30	0.2756
{4}GENDER	Fixed	1	0.00	213.00	0.99	0.00	0.9480

Calendar year 1991

ANOVA RESULTS: CREDITS IN 1991: CREDIT1							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	120.48	243.00	1.04	115.87	0.0000
{1}EDUCATION	Fixed	9	7.24	243.00	1.04	6.96	0.0000
{2}LANGUAGE	Fixed	10	0.98	243.00	1.04	0.94	0.4983
{3}COLOUR	Fixed	2	1.08	243.00	1.04	1.04	0.3548
{4}GENDER	Fixed	1	0.25	243.00	1.04	0.24	0.6242

Calendar year 1992

ANOVA RESULTS: CREDITS IN 1992: CREDIT2							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	149.96	219.00	0.99	151.76	0.0000
{1}EDUCATION	Fixed	8	9.38	219.00	0.99	9.49	0.0000
{2}LANGUAGE	Fixed	9	0.56	219.00	0.99	0.57	0.8201
{3}COLOUR	Fixed	3	5.16	219.00	0.99	5.22	0.0017
{4}GENDER	Fixed	1	1.72	219.00	0.99	1.74	0.1879

Calendar year 1993

ANOVA RESULTS: CREDITS IN 1993: CREDIT3							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	121.25	217.00	1.14	106.32	0.0000
{1}EDUCATION	Fixed	9	9.07	217.00	1.14	7.96	0.0000
{2}LANGUAGE	Fixed	12	1.20	217.00	1.14	1.06	0.3995
{3}COLOUR	Fixed	2	7.66	217.00	1.14	6.71	0.0015
{4}GENDER	Fixed	1	0.84	217.00	1.14	0.73	0.3924

Calendar year 1994

ANOVA RESULTS: CREDITS IN 1994: CREDIT4							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	138.80	278.00	1.19	116.29	0.0000
{1}EDUCATION	Fixed	9	11.75	278.00	1.19	9.84	0.0000
{2}LANGUAGE	Fixed	11	0.89	278.00	1.19	0.75	0.6932
{3}COLOUR	Fixed	3	1.79	278.00	1.19	1.50	0.2151
{4}GENDER	Fixed	1	0.00	278.00	1.19	0.00	0.9769

Calendar year 1995

ANOVA RESULTS: CREDITS IN 1995: CREDITS							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	164.08	310.00	0.91	180.61	0.0000
{1}EDUCATION	Fixed	9	12.06	310.00	0.91	13.28	0.0000
{2}LANGUAGE	Fixed	11	0.86	310.00	0.91	0.95	0.4944
{3}COLOUR	Fixed	3	1.30	310.00	0.91	1.44	0.2324
{4}GENDER	Fixed	1	0.18	310.00	0.91	0.20	0.6565

Calendar year 1996

ANOVA RESULTS: CREDITS IN 1996: CREDIT6							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	215.29	278.00	1.07	201.84	0.0000
{1}EDUCATION	Fixed	9	12.44	278.00	1.07	11.66	0.0000
{2}LANGUAGE	Fixed	10	0.45	278.00	1.07	0.42	0.9374
{3}COLOUR	Fixed	3	2.69	278.00	1.07	2.52	0.0581
{4}GENDER	Fixed	1	0.58	278.00	1.07	0.55	0.4605

Calendar year 1997

ANOVA RESULTS: CREDITS IN 1997: CREDIT7							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	200.98	268.00	0.99	202.94	0.0000
{1}EDUCATION	Fixed	17	2.26	268.00	0.99	2.28	0.0032
{2}LANGUAGE	Fixed	11	1.63	268.00	0.99	1.64	0.0868
{3}COLOUR	Fixed	3	2.46	268.00	0.99	2.48	0.0614
{4}GENDER	Fixed	1	0.09	268.00	0.99	0.09	0.7655

Calendar years 1990-1997 combined

ANOVA RESULTS: CREDITS IN 1990-1997 COMBINED: CREDIT							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	1353.77	2197.00	1.05	1288.04	0.0000
{1}EDUCATION	Fixed	18	32.14	2197.00	1.05	30.58	0.0000
{2}LANGUAGE	Fixed	12	1.77	2197.00	1.05	1.68	0.0647
{3}COLOUR	Fixed	3	10.69	2197.00	1.05	10.17	0.0000
{4}GENDER	Fixed	1	0.00	2197.00	1.05	0.00	0.9565

In the nine ANOVA tables the following explanatory variables have emerged as statistically significant:

- matric point score³⁰ (in all the time periods considered)
- educational department (in all the time periods considered)
- colour (only in the following years 1992, 1993 and all the years combined data set)

These variables appear to be sources of variability in the performance variate. They contribute to an explanation of the differences in the number of credits obtained by the students.

The significant differences in the mean number of credits obtained by the incoming first year students are only observed between the student groupings in the distinct categories of these variables (e.g. in 1993 between the categories of the educational department variable or the categories of the colour variable and not between the categories of either the home language variable or the gender variable).

Intrinsically these factors emerge as explanatory whether or not the matric point score variable is used along with them. Again the matric point score is the dominating predictor. In all years prior to 1997 it has a higher F-statistic than the corresponding value for the faculty point score. This feature adds to the noted desirability of a comparative analysis, but also alerts us to a possible anomaly or change-point in the years of departmental change (i.e. the year 1997)

What has been established so far?

Besides confirming the significant relationship between the matric point score and the number of credits obtained, we have been able to determine other factors (the educational department and the colour) that may additionally explain the different levels of the performance variable across the various student groupings.

³⁰ However this result was expected and for that reason the matric point score was entered as a covariate.

b) Binary Response Analysis

Calendar year 1990

ANOVA RESULTS: BINARY CREDITS IN 1990: BCREDIT							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	18.89	213.00	0.11	173.67	0.0000
{1}EDUCATION	Fixed	9	0.85	213.00	0.11	7.84	0.0000
{2}LANGUAGE	Fixed	9	0.23	213.00	0.11	2.09	0.0316
{3}COLOUR	Fixed	2	0.50	213.00	0.11	4.63	0.0107
{4}GENDER	Fixed	1	0.09	213.00	0.11	0.79	0.3744

Calendar year 1991

ANOVA RESULTS: BINARY CREDITS IN 1991: BCREDIT1							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	7.66	243.00	0.15	52.11	0.0000
{1}EDUCATION	Fixed	9	0.50	243.00	0.15	3.42	0.0006
{2}LANGUAGE	Fixed	10	0.16	243.00	0.15	1.11	0.3569
{3}COLOUR	Fixed	2	0.01	243.00	0.15	0.07	0.9285
{4}GENDER	Fixed	1	0.11	243.00	0.15	0.74	0.3905

Calendar year 1992

ANOVA RESULTS: BINARY CREDITS IN 1992: BCREDIT2							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	11.61	219.00	0.13	87.38	0.0000
{1}EDUCATION	Fixed	8	0.96	219.00	0.13	7.23	0.0000
{2}LANGUAGE	Fixed	9	0.21	219.00	0.13	1.60	0.1159
{3}COLOUR	Fixed	3	0.61	219.00	0.13	4.62	0.0037
{4}GENDER	Fixed	1	0.00	219.00	0.13	0.00	0.9711

Calendar year 1993

ANOVA RESULTS: BINARY CREDITS IN 1993: BCREDIT3							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	6.24	217.00	0.16	39.13	0.0000
{1}EDUCATION	Fixed	9	0.51	217.00	0.16	3.20	0.0012
{2}LANGUAGE	Fixed	12	0.13	217.00	0.16	0.83	0.6218
{3}COLOUR	Fixed	2	0.74	217.00	0.16	4.66	0.0105
{4}GENDER	Fixed	1	0.00	217.00	0.16	0.00	0.9806

Calendar year 1994

ANOVA RESULTS: BINARY CREDITS IN 1994: BCREDIT4							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	8.45	278.00	0.16	54.41	0.0000
{1}EDUCATION	Fixed	9	0.71	278.00	0.16	4.59	0.0000
{2}LANGUAGE	Fixed	11	0.29	278.00	0.16	1.90	0.0396
{3}COLOUR	Fixed	3	0.12	278.00	0.16	0.75	0.5219
{4}GENDER	Fixed	1	0.02	278.00	0.16	0.15	0.7034

Calendar year 1995

ANOVA RESULTS: BINARY CREDITS IN 1995: BCREDITS							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	9.11	310.00	0.14	66.94	0.0000
(1)EDUCATION	Fixed	9	1.15	310.00	0.14	8.43	0.0000
(2)LANGUAGE	Fixed	11	0.22	310.00	0.14	1.64	0.0865
(3)COLOUR	Fixed	3	0.12	310.00	0.14	0.88	0.4500
(4)GENDER	Fixed	1	0.05	310.00	0.14	0.38	0.5397

Calendar year 1996

ANOVA RESULTS: BINARY CREDITS IN 1996: BCREDIT6							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	15.66	278.00	0.16	96.57	0.0000
(1)EDUCATION	Fixed	9	1.10	278.00	0.16	6.80	0.0000
(2)LANGUAGE	Fixed	10	0.19	278.00	0.16	1.17	0.3082
(3)COLOUR	Fixed	3	0.33	278.00	0.16	2.01	0.1121
(4)GENDER	Fixed	1	0.03	278.00	0.16	0.18	0.6704

Calendar year 1997

ANOVA RESULTS: BINARY CREDITS IN 1997: BCREDIT7							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	9.32	268.00	0.17	53.26	0.0000
(1)EDUCATION	Fixed	17	0.37	268.00	0.17	2.10	0.0073
(2)LANGUAGE	Fixed	11	0.21	268.00	0.17	1.18	0.3043
(3)COLOUR	Fixed	3	0.23	268.00	0.17	1.31	0.2709
(4)GENDER	Fixed	1	0.07	268.00	0.17	0.42	0.5196

Calendar years 1990-1997 combined

ANOVA RESULTS: BINARY CREDITS IN 1990-1997 COMBINED: BCREDIT							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	86.77	2197.00	0.15	564.06	0.0000
(1)EDUCATION	Fixed	18	2.46	2197.00	0.15	15.99	0.0000
(2)LANGUAGE	Fixed	12	0.30	2197.00	0.15	1.97	0.0232
(3)COLOUR	Fixed	3	0.54	2197.00	0.15	3.53	0.0142
(4)GENDER	Fixed	1	0.17	2197.00	0.15	1.07	0.3003

In the nine ANOVA tables the following explanatory variables emerge as significant:

- matric point score ³¹ (in all the time periods considered)
- educational department (in all the time periods considered, except for 1997)
- home language (only in the following years 1990,1994, and all the years combined data set)
- colour (only in the following four years 1990, 1992,1993 and all the years combined data set)

Again the lists of the significant variables and the years in which they are significant are the same for the ordinal and binary response analyses, except for the home language variable. We may again say that this feature of the data suggests that a shift in focus from the number of credits obtained to the simple issue of (prospective) re-admission to the Faculty on the basis of performance will not require a change in the associated predictors. The matric point score will apparently dominate minor adjustments arising for the other factors in a linear model.

We again observe that in all years prior to 1997 the matric point score variable has a higher F-statistic than the corresponding value for the faculty point score. This feature adds to the noted desirability of a comparative analysis, but also alerts us to a possible anomaly or change-point in the years of departmental change (i.e. the year 1997).

It is possible to fit a linear model with both the faculty and matric point scores as explanatory variables. The corresponding F-statistics may be examined. If one is close to zero the corresponding predictor adds little given the presence of the other. If both F values are large, then there is a case for incorporating both of the predictors in the model.

Similar statements would apply to using the matric point score plus scores for the individual matric subjects that are weighted by the Faculty of Science.

These considerations lead naturally to the application of FIRM.

³¹ This result was expected and hence the matric point score was entered in the analysis as a covariate.

3.3 INFERENCE-BASED RECURSIVE MODELLING: FIRM

A comparative analysis has suggested that the matric point scores may (*post-factum*) have been a better predictor than the faculty point scores. The ANOVA methods have investigated the linear relationships between the number of credits and the two types of scores, individually. In FIRM, on the other hand, at each step a variable that appears to be most significant in splitting the data, is chosen from the set of predictor variables. Both the faculty and matric point scores may be present.

Since FIRM is designed to produce interpretable valid results for large data structures, analyses were performed only for the eight calendar years combined data set. The ordinal and binary response analyses were performed in FIRM 's variable-specific programmes i.e. CONFIRM and CATFIRM respectively.

A list of the predictor variables used in both analyses, together with their codes ³² in FIRM, is given below:

- faculty point score (*facpoint*) (specified as real/ordinal predictor)
- matric point score (*matpoint*) (specified as real/ordinal predictor)
- home language (*language*)
- educational department ³³ (*education*)
- gender (*gender*)
- colour (*colour*)
- matric HG maths (*math*)
- matric SG maths (*mats*)
- matric HG physical science (*psch*)
- matric SG physical science (*pscs*)
- matric HG biology (*bioh*)
- matric SG biology (*bios*)
- matric HG English (*egmh*)
- matric HG Afrikaans (*aghm*)
- matric HG geography (*geoh*)
- calendar year of first year registration (*year*)

³² The listing may be used to improve readability of the dendrogram outputs of FIRM.

³³ Since there are 20 different educational department codes present in the entire data set, the educational department variable was used only in CONFIRM as the programme allows 20 categories per predictor. CATFIRM allows only 16 distinct categories per predictor. It is also permissible to use additional criteria to amalgamate categories and reduce the number entering FIRM to below the maximum number.

Separate dendrograms, produced for the ordinal and binary response analyses, are displayed and discussed on the following pages.

In CONFIRM, the following information, in the specified order, is given for each node:

- number of students in the node
- mean number of credits obtained by students grouped in the node
- standard deviation of mean number of credits in the node
- maximum number of credits obtained by students grouped in the node
- minimum number of credits obtained by students grouped in the node

In CATFIRM, the following information, in the specified order, is given for each node:

- proportion of students obtaining fewer than 2 credits in the node
- proportion of students obtaining more than 2 credits in the node
- number of students grouped in the node

As an extension to the description of the most important features of dendrograms presented in this chapter, the reader may refer to those statistics for additional information.

However it should be pointed out that a dendrogram is only a graphical representation and summary of the results obtained in FIRM. A summary file ³⁴ containing various concise statistics of the different stages in the recursive modelling procedure accompanies each dendrogram. As examples of this feature the summary statistics files for the ordinal and binary response analyses performed on the entire data set with all listed predictors (i.e. the analyses in 3.3.1 a) and in 3.2.2 a)) are included in Addenda 1 and 2 with annotations, respectively. The reader may wish to examine these sections to become familiar with the additional aspects of the FIRM output. Despite the fact that these summary files are presented only for two analyses in the thesis, the comments with respect to other dendrograms have been made after examining the relevant summary files.

³⁴ Additionally to the summary file there are also other output files containing the details of the other calculations performed i.e. in CONFIRM, the split file and in CATFIRM the split file and the split rule table file.

3.3.1 ORDINAL RESPONSE ANALYSIS

a) Entire data set and all listed predictors

The analysis was performed on the entire data set. All listed predictors were used.

Examining the dendrogram produced in this analysis we note:

at first level analyses (node 1)

- Out of all the predictor variables in the data set, the *matric point score* variable (*ordinal variable*) splits the entire data most significantly and into seven groupings³⁵ of the matric point scores (*nodes 2 to 8*).

at second level analyses (nodes 2, 3, 4, 5, 6, and 7)

- For nodes 4, 5, 6,7 and 8, corresponding to the matric point score of at least 34 and greater, the mean number of credits is at least 2.
- For nodes 2 and 3, corresponding to the matric point score less than 34, the mean number credits obtained is less than 2 credits
- The five matric point score groupings (i.e. expect for the node 2 and 5) can be further significantly split into further groups based on the variable *colour*.
- Each of these five nodes is further split into either 2 or 3 subsidiary nodes.

³⁵ Digits 0 to 9 indicate the particular matric point scores obtained by students and coded in FIRM (FIRM separates the values of any real/ordinal predictor into 11 distinct data-determined categories). The primary cut-off points chosen of these categories before any admissible amalgamations are as follows:

MATRIC POINT SCORE
0 ≤ 31.00
1 ≤ 32.00
2 ≤ 34.00
3 ≤ 35.00
4 ≤ 36.00
5 ≤ 37.00
6 ≤ 39.00
7 ≤ 40.00
8 ≤ 44.00
9 > 44.00
"?" is for missing information

- The core of most of these subsidiary nodes is formed either by the category White (*nodes 13,15,20 and 23*) or the categories Black, Coloured and Indian combined (*nodes 12 and 14*) or the categories Coloured and Black or Indian combined (*nodes 19 and 22*) or the categories Indian and Black by themselves (*nodes 18 and 21*), with the category Missing Information attached to some of the mentioned nodes (*nodes 13,14,20 and 21*).
- Within each of the subgroups, the average number of credits obtained is much higher for the White category nodes (greater than 2) than for the Black, Coloured and Indian categories nodes (usually smaller than 2); this is particularly apparent for the nodes descendant from nodes 3 and 4 but it is also a feature of the remaining nodes.
- This output suggests that within the particular student groups formed by the matric point score categories the Black, Coloured or Indian students consistently obtain fewer credits than the White students.
- For node 2, corresponding to the matric point scores of less than 31, the *faculty point scores*³⁶ emerges as significant performance predictor. Although it seems that students with the faculty point scores between 44 and 48 (i.e. the scores for maths and the better of either biology or physical science doubled), obtain on average 2.05 credits, most students in node 2 get fewer than 2 credits.

³⁶ Digits 0 to 9 indicate the particular faculty point scores obtained by students and coded in FIRM (FIRM separates the values of any real/ordinal predictor into 11 distinct data-determined categories). The primary cut-off points chosen of these categories before any admissible amalgamations are as follows:

FACULTY POINT SCORE
0 ≤ 42.00
1 ≤ 44.00
2 ≤ 45.00
3 ≤ 47.00
4 ≤ 48.00
5 ≤ 50.00
6 ≤ 51.00
7 ≤ 54.00
8 ≤ 57.00
9 > 57.00
"?" is for missing information

- For node 5, corresponding to the matric point scores of 36 and 37, the *educational department*³⁷ splits the node into the following two groups, i.e. those from predominately White and non-White departments, with the latter exhibiting substantially lower mean number of credits (i.e. 3.31 in contrast to 1.86 respectively).
- Another interesting features of this dendrogram is that in all the nodes the minimum number of credits obtained is 0, suggesting that an incoming first year student with any matric point score (high or low) may fail all courses in the first calendar year of his or her study.

The CONFIRM analysis has chosen the matric point scores above the faculty point scores, and then finds a further explanatory value for the faculty point scores only for the students with the matric point score of 31 and less, within the aggregated data set for all years. Additionally, the colour variable emerges most often as having additional predictive value with respect to the number of credits obtained.

³⁷ The abbreviations of the educational departments were coded in FIRM as follows:

code	department
a	CO
b	JM
c	CA
d	NE
e	IN
f	TR
g	ET
h	NA
i	OF
j	TK
k	IE
l	WC
m	EC
n	KN
o	GT
p	NW
q	NC
r	NP
s	MP

For full names of the departments refer to chapter 1.

b) Entire data set and all listed predictors except matric point score

To suppress the predictive ability of the matric point scores and thus to investigate other possible alternative performance predictors, the analysis was performed again on the entire data set excluding the matric point score variable from the set of possible predictors.

Examining the dendrogram produced in this analysis we note:

at first level analyses (node 1)

- Out of all the predictor variables in the data set, the *colour* variable splits entire data most significantly into the following three groups with descending mean numbers of credits as follows: White, Coloured and Indian combined (and Missing Information) combined and Black.

at second level analyses (nodes 2, 3, 4 and 5)

- Each of the three groupings can be further significantly split into subsidiary nodes according to the *faculty point scores*. Within each of these subsidiary nodes the higher faculty point scores are associated with the higher mean number of credits obtained.
- For Black students with the faculty point score of 42 and less, there is a significant difference between those admitted in 1990 and those in all the other years.
- For Black students with the faculty point scores of 44 and 45, the *educational department* may well explain the differences in the mean number of credits obtained. Those who were in the privileged departments exhibit higher mean number of credits (of 2.37) than those from the disadvantaged departments. It ought to be pointed out, however, that this privileged departments group is rather small and contains only 23 students.
- For Black students with faculty point scores of at least 47, there are no further significant predictors.
- For Coloured and Indian students in addition to the faculty point score variable, there are no significant predictors of performance.
- For White students with faculty point scores between of 48 and less, there are no other significant predictors.
- For White students with faculty point scores between 50 and 54 , the differences in performance may attributed to the *educational department* attended.
- For White students with the faculty point score of 57 and more, the *matric higher grade physical science* symbol, where available, may be used to determine performance further with symbols A to C corresponding to the mean number of credits of more than 3.5.

In the absence of the matric point scores, the CONFIRM analysis has chosen the colour variable above all other predictors. Only then the faculty point score emerges as having additional predictive value with respect to the number of credits obtained. This aspect of the analysis suggests that the matric point score is better than the faculty point score in explaining the differences in the level of credits obtained by students after the first calendar year of their studies.

Dendrogram of CONFIRM: number of credits (ordinal response analysis)
 entire data set
 matrix point score excluded

split var
 P val (Bonf P)
 levels
 Node 1
 N
 Xbar
 SD
 Min
 Max

1
 2249
 2.5282
 1.3890
 0.0000
 5.5000

ccolour
 7.2E-178: (1.0E-176)

b		ic		w	
2	3	4	5	6	7
481	673	1095			
1.4252	2.0869	3.2840			
0.9319	1.3826	1.0874			
0.0000	0.0000	0.0000			
4.0000	5.5000	5.5000			

facpoint
 3.24E-30: (4.21E-29)

facpoint
 1.37E-31: (1.92E-30)

facpoint
 2.96E-22: (3.84E-21)

0	5	12	6	7	8	9	10	11	12	13	14	15	89
242	135	70	34	242	34	242	158	169	104	395	368	332	
1.0702	1.5037	1.9571	2.5441	1.4277	1.8956	1.4911	2.4911	3.2548	3.3071	2.8316	3.3071	3.7967	
0.6932	0.8584	1.0206	1.0543	1.2094	1.2770	1.2730	1.0950	1.0579	1.0579	1.2007	1.0579	0.7157	
0.0000	0.0000	0.0000	0.5000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
2.0000	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000	4.0000	4.5000	4.5000	4.5000	5.5000	

education
 6.35E-03: (0.0825)

education
 0.0180: (0.2334)

education
 5.23E-03: (0.0627)

16	17	18	19	20	21	22	23	24
4637215	4637215	4637215	4637215	4637215	4637215	4637215	4637215	4637215
261	216	112	23	32	336	336	3	111
0.4808	1.1412	1.3259	2.3696	2.3906	3.3943	3.3943	2.0000	3.6622
0.1721	0.6865	0.7343	0.9073	1.3722	0.9463	0.9463	1.8028	0.7235
0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.5000	0.5000
1.0000	2.0000	3.0000	4.0000	4.0000	4.5000	4.5000	4.0000	4.5000

c) Data of 1997 excluded and all listed predictors

Up to 1996 the data included only 10 distinct educational department. From 1997 this number was increased by another 10 new departments. In the previous analyses it has appeared evident that the variable has too many departments for CONFIRM to make any significant splits. Thus in order to investigate whether there were not any significant difference between the departments prior to 1997, a new analysis was performed on the entire data set excluding 1997. All listed predictors were used.

Examining the dendrogram produced in this analysis we note:

at first level analyses (node 1)

- Out of all the predictor variables in the data set, the *educational department* variable splits the entire data most significantly and into the four groupings³⁸ of the *educational departments*.
- Again the racially-based departments have been grouped together e.g. the Department of Education and Training and the Transkei Educational Department (*node 2*) or the Cape Education Department Cape, the Natal Education Department , the Transvaal Education Department and IE (*node 5*)
- The mean numbers of credits in the Black departments groupings are lower than those in the predominately White departments.

³⁸ The abbreviations of the educational departments were coded in FIRM as follows:

Code	department
A	CO
B	JM
C	CA
D	NE
E	IN
F	TR
G	ET
H	NA
I	OF
J	TK
K	IE

For full names of the departments refer to chapter 1.

at second level analyses (nodes 2,3,4 and 5)

- The three department groupings (nodes 2,3 and 4) can be further significantly split into further nodes according to the *faculty point scores*³⁹.
- Within the grouping of departments ET and TK (node 2), for those students with the *faculty point score* of 42 and less, there appears to be a significant difference between those admitted in 1990 in comparison to those in all the other years.
- For students from CA, TR, NA and IE departments (node 5) with the *matric point scores* emerges as a further performance predictor.
- Also, for students in node 5 with the matric point score of 44 and more, the *matric higher-grade maths* symbol may be used to predict performance further. Although the *matric higher-grade English* emerges as the next significant predictor, it does not appear to add any further explanation as the one of the subsidiary nodes (node 28) contains only one student.
- Also, for students in node 5 with the matric point score of 31 and less, the *faculty point score* emerges as a further significant predictor.

Careful examination of the descendent nodes (nodes 22, 23 and 24) reveals the anomaly some faculty and matric point scores. A matric point score of 31 may possibly result in the category of faculty point score of 48 and less. It cannot feasibly result in the faculty point scores of 50 or more. The handling of the data cases with anomalous point scores has been discussed already in the predicament section at the beginning of this chapter.

³⁹ Digits 0 to 9 indicate the particular faculty and matric point scores obtained by students and coded in FIRM (FIRM separates the values of any real/ordinal predictor into 11 distinct data-determined categories). The primary cut-off points chosen of these categories before any admissible amalgamations are as follows:

FACULTY POINT SCORE	MATRIC POINT SCORE
0 ≤ 42.00	0 ≤ 31.00
1 ≤ 44.00	1 ≤ 33.00
2 ≤ 46.00	2 ≤ 34.00
3 ≤ 47.00	3 ≤ 35.00
4 ≤ 48.00	4 ≤ 36.00
5 ≤ 50.00	5 ≤ 38.00
6 ≤ 51.00	6 ≤ 39.00
7 ≤ 54.00	7 ≤ 40.00
8 ≤ 58.00	8 ≤ 44.00
9 > 58.00	9 > 44.00
"?" is for missing information	"?" is for missing information

We note that these cut-off point are slightly different than those derived for the entire data set.

In the absence of the 1997 data, the CONFIRM analysis has chosen the educational department above all other predictors. Only then the faculty and matric point score variables have been chosen as having additional predictive value with respect to the number of credits obtained with the former emerging for the traditionally non-White educational departments.

d) Data of 1997 excluded and all listed predictors except for matric point score

In order to eliminate the anomalous data cases in the predicament specified at the beginning of this chapter and in the previous section 3.1.1 c), the analysis was repeated on a reduced data set excluding the 1997 data and suppressing the matric point score in the list of predictors, on the assumption that the matric point score was more likely to be the cause of anomalies than the faculty point score.

The dendrogram produced in this analysis has a similar structure of the dendrogram to that of dendrogram produced of section 3.1.1 c). The major difference occurs at node 5 where the matric point score is replaced by the faculty point score. Additionally two other predictors have emerged for the particular groupings of the faculty point score, i.e. the matric higher-grade Afrikaans and the matric standard-grade maths.

We may also note that when the matric point score is suppressed, adjacent faculty point score are amalgamated into categories⁴⁰ at stage 2 over all four nodes (labelled 2 to 5 in the graph). This amalgamation is slightly different when the matric point score is not suppressed as the faculty point score is entered and categorised at stage 2 over only 3 of the four nodes. This categorisation is transferred and unchanged by node 5 in stage 3.

⁴⁰ Digits 0 to 9 indicate the particular faculty point scores obtained by students and coded in FIRM (FIRM separates the values of any real/ordinal predictor into 11 distinct data-determined categories). The primary cut-off points chosen of these categories before any admissible amalgamations are as follows:

FACULTY POINT SCORE
0 ≤ 42.00
1 ≤ 44.00
2 ≤ 46.00
3 ≤ 47.00
4 ≤ 49.00
5 ≤ 50.00
6 ≤ 52.00
7 ≤ 55.00
8 ≤ 59.00
9 > 59.00
"?" is for missing information

We may note that these cut-off points are slightly different from those derived for the analysis in 3.1.1 c).

3.3.2 BINARY RESPONSE ANALYSIS

All the analyses performed in 3.3.1 (except for 3.1.1 d)) were repeated for the binary performance variable.

a) Entire data set and all listed predictors

Examining the dendrogram produced in this analysis we note:

at first level analysis

- The *colour* variable splits the entire data most significantly and into three nodes with the decreasing proportion of those who obtained at least 2 credits as follows: White, Coloured and Indian combined, and Black.

at second level analyses (nodes 2, 3, 4, 5 and 6)

- The White category and the Coloured and Indian categories combined may be further significantly split according to the *matric point scores*.
- For White students with any matric point score the proportion of those who obtained at least 2 credits is always substantially higher than the proportion of those with fewer than 2 credits.
- For the Coloured and Indian categories combined, the proportion of those who obtained at least 2 credits becomes lower than the complementary proportion of those who obtained fewer than 2 credits, for students with the matric point scores of 35 and less.
- The Black category may be further split according to the *faculty point scores*. The proportion of those who obtained at least 2 credits is lower than the complementary proportion of those who obtained fewer than 2 credits, for students with the faculty point score of 34 and less.
- There are no further significant predictors.

The CATFIRM analysis has chosen again the colour variable above either the matric and faculty point scores variables, within the aggregated data set for all years. Only then the two point scores have emerged as having additional predictive performance ability.

Despite the fact that the most significant predictors for the ordinal and binary response analyses are different, the observed results are essentially similar. The incoming first year students with the higher matric point scores or the higher faculty point score are more likely to get 2 or more credits in the first calendar year of their study than students with the lower matric point scores. Irrespective of the actual point scores the White students appear far more likely to fulfil the Faculty's re-admission criteria into a second year than either the Coloured, Indian or Black students.

b) Entire data set and all listed predictors except matric point score

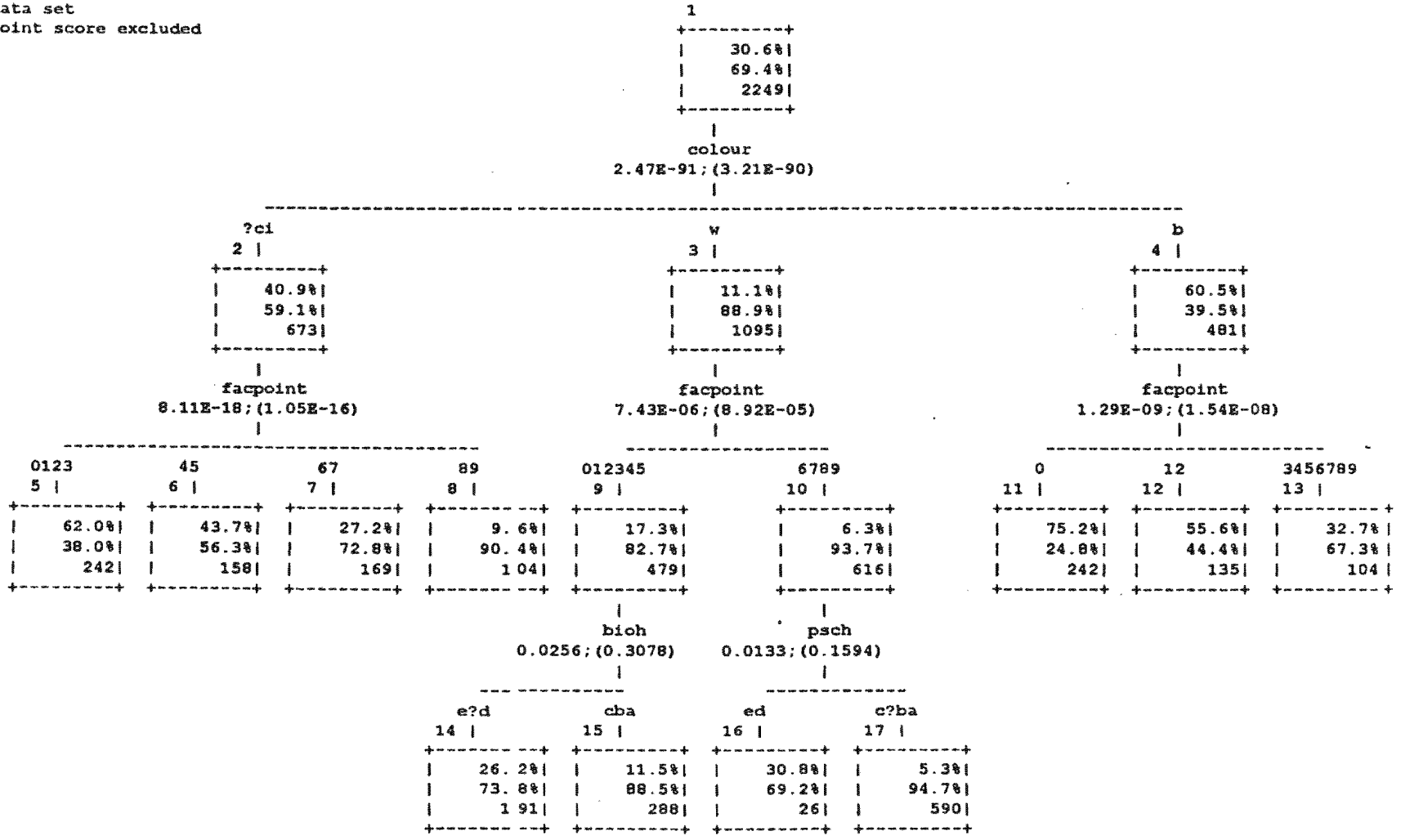
The overall structure and the splitting variables of the binary analysis dendrogram are essentially similar to the ordinal analysis dendrogram, except for the following features:

- For the Black, Coloured and Indian students there are no other significant predictors except for the faculty point scores.
- For the White students with faculty point scores of 51 and more, *the matric higher-grade biology* symbol emerges as significant performance predictor. And for the White students with the faculty point scores of 50 and less, *the matric higher-grade physical science* symbol emerges as significant performance predictor.

Thus again, in the absence of the matric point scores, the CATFIRM analysis has chosen the colour variable above all other predictors. The faculty point score has emerged as having additional predictive ability in distinguishing between students more and less likely to fulfil the Faculty's re-admission criteria into the second year.

Dendrogram of CATFIRM: number of credits (binary response analysis)
 |
 split var entire data set
 p val (Bonf P) matric point score excluded

 |
 levels
 Node |
 +-----+
 |% less than 2 |
 |% at least 2 |
 |_N_ |
 +-----+



c) Data of 1997 excluded and all listed predictors ⁴¹

The overall structure and the splitting variables of the binary analysis dendrogram are similar to the ordinal analysis dendrogram, except for the following features:

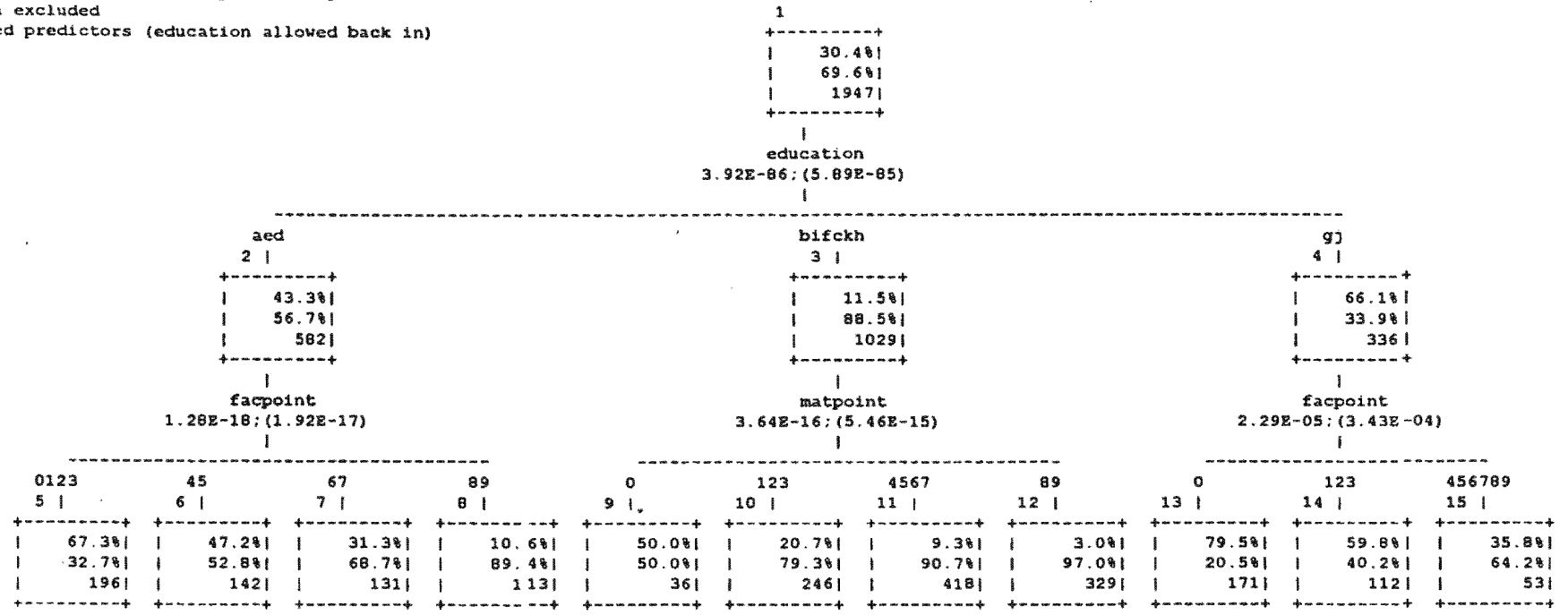
- The educational department variable splits the entire set most significantly and only into three educational department groupings. However the historically White and non-White departments have formed separate nodes.
- Besides the matric and faculty point scores, no other significant predictors of performance have been found.

Thus, in the absence of the 1997 data, the CATFIRM analysis has chosen again the educational department above all other predictors. Thereafter either the matric and faculty point score variables have emerged in different educational departments as having additional predictive ability in distinguishing between students more and less likely to fulfil the Faculty's re-admission criteria into the second year.

⁴¹ The educational department variable has been allowed back into the set of predictors as up to 1996 there were only 10 distinct educational departments

Dendrogram of CATFIRM: number of credits (binary response analysis)
 1997 data excluded
 all listed predictors (education allowed back in)

split var
 P val (Bonf P)
 levels
 Node |
 +-----+
 |% less than 2 |
 |% at least 2 |
 |_N_ |
 +-----+



3.3.3 ADDITIONAL ANALYSES FOR BLACK, COLOURED AND INDIAN STUDENTS COMBINED SET

To investigate further any additional or more significant predictors of performance amongst the disadvantaged students, some analyses were repeated on the data set containing only the Black, Coloured and Indian students i.e. the ordinal and binary response analyses on all the years combined data set and the ordinal response analysis on the 1997 excluded data set. Where appropriate all listed predictors were used.

Examining these three dendrograms we note that the faculty point score variable has emerged as the most significant predictor of performance. At the first level analyses, the major and significant difference in the dendrograms relates to the particular groupings of the faculty point scores. There are fewer point score groupings in the binary response analysis than in the two ordinal analyses. However the cut-off points⁴² for the groupings formed differ slightly between the two ordinal response analyses.

Additionally in the ordinal response analyses the calendar year of first year registration and the educational department variables have been found significant. For students with the faculty point scores of 42 and less there were differences in the mean numbers of credits amongst the particular calendar years (1990 and 1993, 1994, 1996 and 1991, 1992, 1995, 1997). For students with the faculty point scores between 44 and 46 (in the 1997 excluded set) and 43 and 48 (in all the years combined) the number of credits were higher for those from the privileged educational departments.

These analyses have indicated that the faculty point score is a more efficient predictor in distinguishing between the potentially successful and unsuccessful for the set of the disadvantaged students. This finding confirms the results in section 3.1.3 on the regression plots.

⁴² The table below includes the cut-off points of the faculty point scores and the educational department codes as expressed in FIRM for all the years combined and the 1997 excluded data sets.

ALL YEARS COMBINED DATA SET					1997 EXCLUDED DATA SET		
FACULTY POINT SCORE	EDUCATIONAL DEPARTMENT				FACULTY POINT SCORE	EDUCATIONAL DEPARTMENT	
0 ≤ 39.00	ET	a	KN	L	0 ≤ 39.00	ET	a
1 ≤ 42.00	TR	b	NW	M	1 ≤ 42.00	TR	b
2 ≤ 43.00	TK	c	WC	N	2 ≤ 44.00	TK	c
3 ≤ 45.00	JM	d	NP	o	3 ≤ 45.00	JM	d
4 ≤ 46.00	NE	e	EC	p	4 ≤ 46.00	NE	e
5 ≤ 48.00	NA	f	GT	q	5 ≤ 48.00	NA	f
6 ≤ 49.00	CA	g	MP	r	6 ≤ 50.00	CA	g
7 ≤ 51.00	CO	h	NC	s	7 ≤ 51.00	CO	h
8 ≤ 54.00	IE	i			8 ≤ 56.00	IE	i
9 > 54.00	IN	j			9 > 56.00	IN	j
"?" missing info.	OF	k			"?" missing info.	OF	k

Dendrogram of CONFIRM: number of credits (ordinal response analysis)

Black, Coloured and Indian students
all listed predictors

split var
P val (Bonf P)
levels
Node
N
Xbar
SD
Min
Max

1				
1138				
1.8010				
1.2490				
0.0000				
5.5000				
facpoint 1.79E-70; (2.69E-69)				
01	2345	67	8	9
2	3	4	5	6
257	456	179	113	133
1.0681	1.5581	2.0559	2.4292	3.1729
0.6989	1.1121	1.2861	1.2190	1.0784
0.0000	0.0000	0.0000	0.0000	0.0000
2.5000	4.0000	4.0000	4.0000	5.5000
year 1.54E-03; (0.0231)		education 3.48E-05; (5.22E-04)		
0	463	7251	eojhalfcnqsm	gtpidr
7	8	9	10	11
28	103	126	403	53
0.4821	1.0000	1.2540	1.4293	2.5377
0.1657	0.6788	0.7092	1.0387	1.1720
0.0000	0.0000	0.0000	0.0000	0.0000
1.0000	2.0000	2.5000	4.0000	4.0000

Dendrogram of CATFIRM: number of credits (binary response analysis)

Black, Coloured, Indian
all listed predictors

split var
P val (Bonf P)
levels
Node
% less than 2
% at least 2
N

1			
49.4%			
50.6%			
1138			
facpoint 1.83E-36; (2.56E-35)			
01	2345	678	9
2	3	4	5
75.5%	55.7%	34.6%	9.8%
24.5%	44.3%	65.4%	90.2%
257	456	292	133

Dendrogram of CONFIRM: number of credits (ordinal response analysis)
 Black, Coloured, Indian students
 1997 data excluded
 all listed predictors

split var
 P val (Bonf P)

levels

Node |
 +-----+
 |N |
 |Xbar |
 |SD |
 |Min |
 |Max |
 +-----+

1
 +-----+
 | 976 |
 | 1.8038 |
 | 1.2607 |
 | 0.0000 |
 | 4.0000 |
 +-----+

facpoint
 5.74E-61; (8.62E-60)

01			234		56	78	9
2			3		4	5	6
+-----+			+-----+		+-----+	+-----+	+-----+
201			263		243	184	85
1.0348			1.4487		1.7860	2.5136	3.2353
0.6755			1.0914		1.2127	1.2319	1.0567
0.0000			0.0000		0.0000	0.0000	0.0000
2.0000			4.0000		4.0000	4.0000	4.0000
+-----+			+-----+		+-----+	+-----+	+-----+
year			education				
4.41E-04; (6.61E-03)			6.86E-05; (1.03E-03)				
+-----+			+-----+		+-----+	+-----+	+-----+
0	463	251	fejzac	gbid			
7	8	9	10	11			
+-----+	+-----+	+-----+	+-----+	+-----+			
28	103	70	228	35			
0.4821	1.0000	1.3071	1.2939	2.4571			
0.1657	0.6788	0.6555	1.0105	1.0736			
0.0000	0.0000	0.0000	0.0000	0.0000			
1.0000	2.0000	2.0000	4.0000	4.0000			
+-----+	+-----+	+-----+	+-----+	+-----+			

3.4 TECHNICAL COMMENTS ON METHODS USED

Variance Components

For each of the variance components analyses only five distinct explanatory variables were used. This restriction was chosen to ensure that the variance within the data was split only between few factors. The chosen variables were of priority importance in the study and were expected to be significantly related to the response variable.

A possibility, however, remains to perform other variance components analyses where new variables, such as individual matric subjects, together with the current variables could be examined.

FIRM

The significance levels for the splitting and merging stages were stringent, 0.9% and 1%, respectively. Higher significance levels could possibly result in additional variables to split the data significantly further. The chosen levels were taken as per an introductory example in FIRM's manual. They were also different to the levels applied in the educational data example discussed in Appendix C. The respective levels used there were 5% and 6%. We performed our analyses with stricter levels to extract predictors with very small p-values.

We should also apply caution to the interpretations and the implications of some of the results. The White students form the largest and best represented group in the data set. The numbers in other categories of the colour variable are much smaller. The majority of the Black students are enrolled in the SFP Programme and by design of the Programme attend fewer courses and thus can only obtain fewer number of credits.

CHAPTER 4: ASSOCIATIONS BETWEEN POINT SCORE AND FIRST YEAR PERFORMANCE

In the first section of this chapter we re-examine the statistical relationships within categories of the biographical ¹ explanatory variables i.e. colour, educational department and home language, individually for the faculty and matric point scores against the academic performance of the incoming first year students. These issues have been investigated with the log-linear models.

In the second section the focus of the investigation slightly changes as we inspect the chances of passing/failing the first year for across different student groups. These issues have been addressed by fitting the logistic regression model.

¹ Gender was excluded, as it has not emerged as significant in any of the analyses in chapter 3.

4.1. RELATIONSHIPS WITHIN CATEGORIES BETWEEN POINT SCORES AND ACADEMIC PERFORMANCE

As specified ² in Table 2.3 from chapter 2, the data was summarised as frequencies in multi-way contingency tables with rows and columns formed by the categories of the specified variables. Log-linear models ³ were fitted to those tables. The log-linear modelling technique works optimally for contingency tables with few empty cells. Thus to reduce the total number of cells in the tables, both types of point scores were divided into categories. The cut-off points ⁴ for these categories are presented below.

FACULTY POINT SCORE	MATRIC POINT SCORE
Category 1: point scores of 47 and less	Category 1: point scores of 35 and less
Category 2: point scores between 48 and 56 inclusive	Category 2: point scores between 36 and 42 inclusive
Category 3: point scores of 57 and greater	Category 3: point scores of 43 and greater

The ordinal and binary definitions of the performance variable, as described in chapter 3, were used in the current investigation.

² Refer to section 7.1 in Table 2.3 in chapter 2.

³ The discussion of the log-linear modelling technique is given in section B6 of Appendix B.

⁴ Since the faculty point score of 48 is the minimum score required for admission into the Faculty, we may attach the following meanings to the newly formed groupings of the faculty point scores:

- category 1: students conditionally admitted into the Faculty
- category 2: students with average matric results admitted into the Faculty
- category 3: students with good matric results admitted into the Faculty

This choice is just one of many possible groupings of point scores. Other groupings may be investigated.

Since the admission criteria in the Faculty are specified only in terms of the faculty point scores, the cut-off points for the categories of the matric points scores were chosen as near equivalents on the following basis. The faculty point 48 may be interpreted as the score for six higher-grade subjects for which a student obtained C symbol ($6 \times 6 + 2 \times 6(\text{doubling of points}) = 48$). Thus, corresponding matric point score is 36. Similarly for the faculty point score of 56, corresponding matric point score is 42. Thus, the scores of 36 and 42 may form the cut-off points for the middle category of the matric point scores.

The cut-off points of the remaining two categories follow accordingly.

4.1.1 RELATIONSHIPS WITHIN CATEGORIES BETWEEN FACULTY POINT SCORE AND ACADEMIC PERFORMANCE

a) ORDINAL RESPONSE ANALYSES

The tables below summarise the interactions terms found to be statistically significant in the log-linear models⁵ fitted to the three contingency tables formed individually for the variables colour, educational department and home language against the faculty point score and the actual number of credits obtained.

VARIABLES IN THE MODELS	SIGNIFICANT INTERACTION TERMS	STATISTICS FOR THE FINAL MODEL FITTED
faculty point score colour number of credits	number of credits * faculty point score number of credits * colour	chi-square = 279.20 df = 36 p<0.05

VARIABLES IN THE MODELS	SIGNIFICANT INTERACTION TERMS	STATISTICS FOR THE FINAL MODEL FITTED
faculty point score educational department number of credits	number of credits * faculty point score number of credits * educational department	chi-square = 364.17 df = 216 p<0.05

VARIABLES IN THE MODELS	SIGNIFICANT INTERACTION TERMS	STATISTICS FOR THE FINAL MODEL FITTED
faculty point score home language number of credits	number of credits * faculty point score number of credits * home language	chi-square = 272.43 df = 114 p<0.05

⁵ Three log-linear models have been fitted with the following general structure:

$$\ln(m_{ijk}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$$

where

- $m_{ijk} = E(y_{ijk})$: expected ijk^{th} cell frequency (expected number of students in the ijk^{th} cell)
- μ is an overall mean effect (overall base for the logarithm of cell frequencies)
- λ_{level}^{factor} represents the main effect of the particular level of a specified factor as a background condition
(faculty point score and number of credits in all the models and colour or educational department or home language)
- $\lambda_{level, level}^{factor, factor}$ represents the first order interaction term between particular levels of any two specified factors
(either faculty point score and number of credits in all the models and either colour and number of credits or educational department and number of credits or home language and number of credits)

We are interested in the first order interaction terms only involving the number of credits obtained by students and any of the other variables, as indicators of some explanatory effect.

- λ_{ijk}^{ABC} represents the second order interaction term between the i^{th} level of factor A , the j^{th} level of factor B and the k^{th} level of factor C
(either faculty point score, number of credits and one of the other three variables in each of the models)

The significant interaction terms ⁶ may be interpreted in the following way:

- interdependencies between the categories of the faculty point score and the actual number of credits obtained e.g. the students in the higher faculty point score category is more likely to obtain higher than lower number of credits and the students in the lower faculty point score category are more likely to obtain lower than higher number of credits.
- interdependencies between the different colour groups and the actual number of credits obtained e.g. the White students are more likely to obtain higher than lower number of credits and the Black students are more likely to obtain lower than higher number of credits.
- interdependencies between the different educational departments and the number of credits obtained e.g. the students from the privileged educational departments are more likely to obtain higher than lower number of credits and the students from the disadvantaged educational departments are more likely to obtain lower than higher number of credits .
- interdependencies between the different home languages and the number of credits obtained i.e. the students speaking any of the African languages at home are more likely to lower than higher credits and than students speaking any of the "White" languages are more likely to obtain higher than lower number of credits.

These findings confirm the results of the ordinal response analyses of chapter 3.

We may also note that in the second order interaction terms (i.e. interaction between all three variables) have not been found significant in any of the models fitted. This feature indicates that the analyses is not sensitive to the types of indicators of explanatory variables which FIRM has explored. Nevertheless the method may be sub-optimal in that it ignores an ordinality of categories and treats them as nominal

⁶ We draw the specific conclusions relating to the models fitted from the actual structure of the contingency tables formed by cross-classifying the chosen variables.

b) BINARY RESPONSE ANALYSIS

The tables below summarise the interactions terms found to be statistically significant in the log-linear models fitted to the three contingency tables formed individually for the variables colour, educational department and home language against the faculty point score and the binary definition of performance.

VARIABLES IN THE MODELS	SIGNIFICANT INTERACTION TERMS	STATISTICS FOR THE FINAL MODEL FITTED
faculty point score colour performance	performance * faculty point score performance* colour	chi-square = 356.60 df = 12 p<0.05

VARIABLES IN THE MODELS	SIGNIFICANT INTERACTION TERMS	STATISTICS FOR THE FINAL MODEL FITTED
faculty point score educational department performance	performance * faculty point score performance * educational department	chi-square = 371.57 df = 72 p<0.05

VARIABLES IN THE MODELS	SIGNIFICANT INTERACTION TERMS	STATISTICS FOR THE FINAL MODEL FITTED
faculty point score home language performance	performance * faculty point score performance * home language	chi-square = 355.37 df = 48 p<0.05

The significant interaction terms may be interpreted in the following way:

- interdependencies between the categories of the faculty point score and the categories of the performance variable e.g. the students in the higher (*lower*) faculty point score category are more likely (*less likely*) to obtain 2 or more credits than less than 2 credits, respectively.
- interdependencies between the different colour groups and the categories of the performance variable e.g. the White students are more likely to obtain 2 or more credits than less than 2 credits and the Black students are more likely to obtain less than 2 credits than 2 or more credits .
- interdependencies between the different educational departments and the categories of the performance variable e.g. the students in the privileged educational departments are more likely to obtain 2 or more credits than less than 2 credits and the students in the disadvantaged educational departments are more likely to obtain less than 2 than 2 or more credits .

- interdependencies between the different home languages and the categories of performance variable i.e. the students speaking any of the African languages at home are more likely to obtain less than 2 credits than 2 or more credits and the students speaking any of English at home are more likely to obtain 2 or more credits than less than 2 credits .

Again these findings confirm the results on the binary response analyses of the chapter 3.

We may also note that in the second order interaction terms (i.e. interaction between all three variables) have not been found significant in any of the models fitted. This feature indicates that log linear modelling for these multi-way contingency tables is not as sensitive to interactions or to the ordinal structure of categories as FIRM.

4.1.2 RELATIONSHIPS WITHIN CATEGORIES BETWEEN MATRIC POINT SCORE AND ACADEMIC PERFORMANCE

The near equivalence of the matric point score categories and the faculty point score categories induces near equivalence of the corresponding multi-way tables so that the output and the interpretation effectively mimic the corresponding elements of section 4.1.1.

a) ORDINAL RESPONSE ANALYSES

The tables below summarise the interaction terms found to be statistically significant in the log-linear models ⁷ fitted to the three contingency tables formed individually for the variables colour, educational department and home language against the matric point score and the ordinal definition of performance.

VARIABLES IN THE MODELS	SIGNIFICANT INTERACTION TERMS	STATISTICS FOR THE FINAL MODEL FITTED
matric point score colour number of credits	number of credits* matric point score number of credits * colour	chi-square = 230.19 df = 36 p<0.05

VARIABLES IN THE MODELS	SIGNIFICANT INTERACTION TERMS	STATISTICS FOR THE FINAL MODEL FITTED
matric point score educational department number of credits	number of credits* matric point score number of credits * educational department	chi-square = 350.07 df = 216 p<0.05

VARIABLES IN THE MODELS	SIGNIFICANT INTERACTION TERMS	STATISTICS FOR THE FINAL MODEL FITTED
matric point score home language number of credits	number of credits* matric point score number of credits * home language	chi-square = 266.56 df = 144 p<0.05

The additional significant interaction terms may be interpreted in the following way:

- interdependencies between the categories of the matric point score and the actual number of credits obtained i.e. the incoming first year students in the higher faculty point score category obtain the higher number of credits.

⁷ The structure of the models fitted is similar to those described in section 4.1. The matric point score has replaced the faculty point score.

b) BINARY RESPONSE ANALYSES

The tables below summarise the interactions terms found to be statistically significant in the log-linear models fitted to the three contingency tables formed individually for the variables colour, educational department and home language against the matric point score and the binary definition of performance.

VARIABLES IN THE MODELS	SIGNIFICANT INTERACTION TERMS	STATISTICS FOR THE FINAL MODEL FITTED
matric point score colour performance	performance * matric point score performance* colour	chi-square = 370.89 df = 12 p< 0.05

VARIABLES IN THE MODELS	SIGNIFICANT INTERACTION TERMS	STATISTICS FOR THE FINAL MODEL FITTED
matric point score educational department performance	performance * matric point score performance* educational department	chi-square = 382.01 df = 72 p< 0.05

VARIABLES IN THE MODELS	SIGNIFICANT INTERACTION TERMS	STATISTICS FOR THE FINAL MODEL FITTED
matric point score home language performance	performance * matric point score performance* home language	chi-square = 337.45 df = 48 p< 0.05

The additional significant interaction terms may be interpreted in the following way

- interdependencies between the categories of the matric point score and the categories of the performance variable i.e. the incoming first year students in the higher faculty point score category are more likely to obtain 2 or more credits.

4.2 ODDS ON PASSING FIRST ACADEMIC YEAR

The results of previous sections have indicated that the major contrasts in the level of performance in first year have emerged amongst the four race groups. This feature has been expressed by the different number of credits obtained by students either in the distinct categories of the colour or the educational department variables or by lower faculty point scores (i.e. sub-standard matric results).

In order to compare the chances of passing versus failing first year amongst the race groups, the logistic regression model⁸ has been fitted to the data. The estimates of the parameters are displayed in the table on the following page.

⁸ The model fitted has the following structure:

$$\ln\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \mu + \alpha_i + (C + \beta_i) \cdot x_j$$

where

π_{ij} - proportion of students passing first year in the i^{th} category of the colour variable with the j^{th} faculty point score

$1 - \pi_{ij}$ - proportion of students failing first year in the i^{th} category of the colour variable with the j^{th} faculty point score

μ - overall mean proportion of students passing first year

x_j - the j^{th} faculty point score

α_i - effect of the i^{th} level of the colour variable

C - effect of the faculty point score variable

β_i - interaction term between the faculty point score variable and the i^{th} level of the colour variable

The colour variable has been chosen over the educational department variable since it contains fewer and better represented categories.

The interaction term between the colour and the faculty point score variables was fitted as the incoming first year students from the four race groups were admitted into the Faculty with different levels of the faculty point scores.

MODEL: LOGISTIC REGRESSION (LOGIT) N OF O'S:688 I'S:1561								
Final loss: 1059.7487378 Chi²(7)=650.34 p=0.0000								
PARAMETER	CONSTANT	EFFECT OF COLOUR			FACULTY POINT SCORE	INTERACTION BETWEEN FACULTY POINT SCORE AND LEVELS OF COLOUR		
		BLACK	COLOURED	INDIAN		BLACK	COLOURED	INDIAN
ESTIMATE	-3.3148	-3.7066	-5.7526	-6.5952	0.1058	0.0463	0.0856	0.0977
Standard Err.	0.8961	1.3709	1.6171	2.1072	0.0181	0.0299	0.0329	0.0425
t(2241)	-3.6991	-2.7038	-3.5573	-3.1299	5.8376	1.5509	2.6015	2.2987
P-LEVEL	0.0002	0.0069	0.0004	0.0018	0.0000	0.1211	0.0093	0.0216
-95%CL	-5.0721	-6.3950	-8.9239	-10.7275	0.0703	-0.0123	0.0211	0.0144
+95%CL	-1.5575	-1.0182	-2.5814	-2.4630	0.1413	0.1049	0.1502	0.1811
Wald chi-square	13.6833	7.3104	12.6542	9.7960	34.0780	2.4054	6.7680	5.2839
p-level	0.0002	0.0069	0.0004	0.0018	0.0000	0.1209	0.0093	0.0215
ODDS RATIO	0.0363	0.0246	0.0032	0.0014	1.1116	1.0474	1.0894	1.1027
-95%CL	0.0063	0.0017	0.0001	0.0000	1.0728	0.9878	1.0213	1.0145
+95%CL	0.2107	0.3612	0.0757	0.0852	1.1518	1.1106	1.1621	1.1985

The reference categories are for:

- the effect of colour: the white category
- the interaction between the faculty point score and the levels of colour: the interaction between the faculty point score and the white category

Thus we may calculate the odds of passing to failing for the four race groups at the particular ⁹ faculty point scores e.g. 43, 48 and 50. The odds ¹⁰ are displayed in the table below.

FACULTY POINT SCORE	WHITE	BLACK	COLOURED	INDIAN
43	3.45 (2.55;4.68)	0.62	0.44	0.31
48	5.86 (4.27;7.27)	1.32	1.13	0.87
50	7.24 (5.87;8.93)	1.79	1.66	1.31

⁹ The faculty point score of 48 is the minimum point score required for admission into the Faculty. The border faculty point scores were examined, as the question of real concern is whether or not the students with low faculty point scores have a reasonable chance of passing the first year.

¹⁰ Odds are defined as the ratio of the probability of passing to the probability of failing. In terms of the specified model the odds may be expressed as:

$$\left(\frac{\pi_{ij}}{1 - \pi_{ij}} \right) = e^{\mu + \alpha_i + (\gamma + \beta_i) x_j}$$

The odds, displayed in the table, were calculated according to the expression above, substituting the estimates of the respective parameters.

The odds on passing first year increase with an increasing faculty point scores across all the four race groups. At the faculty point score of 48 (i.e. the minimum admission requirement), the odds in favour of passing are greater than 1 across all the race groups. At the faculty point score of 43 (i.e. sub-standard matric results), the odds in favour of passing are less than 1 across all the race groups, except for the White group.

The confidence intervals for the odds of passing first year for White students at 43, 48 and 50 are presented in the table. Confidence intervals may be calculated for the other three groups, but are omitted here because the essential known features of interests have already been confirmed.

CHAPTER 5: PREDICTORS OF PERFORMANCE IN SPECIFIC FIRST YEAR UNIVERSITY COURSES

In this chapter we search for the likely predictors ¹ of academic performance in the selected ² university courses offered at first year level in the Faculty of Science . For each of the investigated courses, we first discuss the statistical relationships as summarised by the measures of association and secondly examine dendrograms produced in FIRM ³.

Ordinal and binary response definitions of academic performance variable have been used. The ordinal response variable was defined as the actual percentage obtained by an incoming first year student for a university course at first year level. The categories of the binary response variable were defined as follows:

- an incoming first year student obtained 50% or more for a first year level university course
- an incoming first year student obtained less than 50% for a first year level university course

The measures of association were calculated for bivariate data formed individually for the faculty point score, the matric point score and the specific matric subjects (i.e. either biology or maths or physical science) against the ordinal response variable separately for each of the analysed courses. The tables of summary results were constructed for each of the individual calendar years and the combined years data set. In addition, the measures were calculated for the two types of point scores and the matric subjects of interest against the binary response variable for the combined years data set only.

¹ Analyses in this chapter follow section 3.1 in Table 2.3 in chapter 2.

² The courses analysed in this chapter are the most popular and best-attended courses at first year level in the Faculty of Science.

³ None of the GLIM techniques was used in this investigation as the additional insights obtained from these techniques were of secondary importance in the project. The volume of the output to be produced and the additional discussions to be written comprised also an important consideration and constraint on the extent of analysis.

In FIRM, the ordinal and binary response definitions of the performance variable were used. FIRM analyses were performed on the combined years data set only. A list of the investigated predictors (with their codes in FIRM) is displayed below.

- faculty point score (*facpoint*)
- matric point score (*matpoint*)
- ✓• home language (*language*) ✓
- ✓• gender (*gender*)
- colour (*colour*)
- educational department ⁴ (*education*)
- ✓• matric HG maths (*math*)
- ✓• matric SG maths (*mats*)
- matric HG physical science (*psch*)
- matric SG physical science (*pscs*)
- matric HG biology (*bioh*)
- matric SG biology (*bios*)
- matric HG English (*egmh*)
- matric HG Afrikaans (*aghm*)
- matric HG geography (*geoh*)
- matric African language ⁵ (*african*)
- calendar year of first year registration (*year*)

Finally the results for the measures of association and FIRM were grouped in the three sections according to the particular matric subject of interest i.e. biology, maths and physical science.

A fourth section has been especially devoted to three courses offered in the Science Foundation Programme. Each of these four sections ends off with a short summary of the most important findings. A fifth section at the end of this chapter contains additional comments and reservations on the performed analyses.

⁴ The educational department variable was used only in the ordinal response analyses. As investigated in chapter 3 the department variable, when applied to all years-combined data set has more categories than CATFIRM can meaningfully deal with.

⁵ The field was used only in the analyses of the SFP courses. It was formed by combining marks for all the African languages completed at matric level.

Additionally, in CONFIRM, the following information, in the specified order, is available for each node:

- number of students in the node
- mean percentage mark obtained by students grouped in the node
- standard deviation of mean percentage mark in the node
- maximum percentage mark obtained by students grouped in the node
- minimum percentage mark obtained by students grouped in the node

In CATFIRM, the following information, in the specified order, is available for each node:

- proportion of students obtaining less than 50% for the course in the node
- proportion of students obtaining 50% or more for the course in the node
- number of students grouped in the node

As a complement to the descriptions of the important features of dendrograms presented, the reader may refer to those statistics for additional information and insight.

The analyses in this chapter focus upon the Science Faculty students within the various courses. Several of the courses have very large numbers of participants from other faculties, and by design of this study, those students are not included in the analyses.

Technical aspects of the individual courses analyses

Before we begin the presentation and discussion of the results, we note some technical aspects of the analyses. Namely, the sum of the numbers of the students who completed a specific university course in each of the relevant calendar years (i.e. the sum of students in the individual courses fields) is not necessarily equal to the total number of students who completed the course in any year as represented by the combined years field for a specific course.

This feature has appeared firstly as a result of the only feasible and quick method available, under the time and financial constraints of the project to create the combined years fields for courses, and secondly, the way in which the specific analyses were performed.

The combined years field for a particular university course was formed by aligning all the fields of that course next to one another in *Statistica* and taking the minimum in each of the created rows. Thus, each of the combined years field reports the marks for the particular course obtained at the presumed first attempt for students who completed the courses in any calendar year, not necessarily in the calendar year of their first year registration. The later attempts at a course were not considered.

In each of the individual years analyses only students who completed the specific course for the first time but in precisely the calendar year of their first year registration were included. Thus, those students who completed the course at first attempt but in a year after their first year registration were not considered in the individual years' analyses.

In both the individual and combined years fields those students who had zeros as their course marks were excluded from the analyses. The "zeros" indicated that they either were absent from the final examination or were not able to write it for academic or personal reasons. This anomaly of "zero" marks was inherited from the original ADP source files.

We may also note that the numbers of students are greater in the contingency tables formed for both the faculty and matric point scores against the course marks than in the contingency tables formed for the individual matric subjects against the course marks. This difference is attributed to the fact both types of point scores are present for all students whereas the marks or symbols for specific matric subjects may not be available for all students, as some may not have taken a particular subject in matric.

5.1 BIOLOGY RELATED UNIVERSITY COURSES

In this section we examine the following three courses, BIO100F, BOT102S and ZOO103S offered at first year level by the Zoology and Botany Departments in the Faculty of Science at UCT.

5.1.1 FIRST YEAR BIOLOGY COURSE: BIO100F

BIO100F has been offered in the Faculty since 1993. Thus, the measures of association were determined separately for the years 1993 to 1997 and for the five years combined field.

The results are displayed in the tables below.

FACULTY POINT SCORE (X) AND BIO100F (ordinal) (Y)

YEAR : 1993, valid cases: 107		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.57	<0.05
Spearman correlation coefficient	0.61	<0.05
Kendall's Tau	0.44	<0.05
Somers' D Statistics	X/Y=0.43 Y/X=0.45	<0.05 <0.05
Gamma Statistic	0.46	<0.05

YEAR : 1994, valid cases: 161		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.42	<0.05
Spearman correlation coefficient	0.42	<0.05
Kendall's Tau	0.30	<0.05
Somers' D Statistics	X/Y=0.29 Y/X=0.30	<0.05 <0.05
Gamma Statistic	0.31	<0.05

YEAR : 1995, valid cases: 157		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.45	<0.05
Spearman correlation coefficient	0.40	<0.05
Kendall's Tau	0.29	<0.05
Somers' D Statistics	X/Y=0.28 Y/X=0.29	<0.05 <0.05
Gamma Statistic	0.30	<0.05

YEAR : 1996, valid cases: 96		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.46	<0.05
Spearman correlation coefficient	0.43	<0.05
Kendall's Tau	0.31	<0.05
Somers' D Statistics	X/Y=0.31 Y/X=0.32	<0.05 <0.05
Gamma Statistic	0.32	<0.05

YEAR : 1997, valid cases: 100		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.55	<0.05
Spearman correlation coefficient	0.53	<0.05
Kendall's Tau	0.39	<0.05
Somers' D Statistics	X/Y=0.39 Y/X=0.40	<0.05 <0.05
Gamma Statistic	0.41	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1993-1997 valid cases: 682		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.50	<0.05
Spearman correlation coefficient	0.49	<0.05
Kendall's Tau	0.35	<0.05
Somers' D Statistics	X/Y=0.35 Y/X=0.36	<0.05 <0.05
Gamma Statistic	0.37	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1993-1997 valid cases: 682		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.34	<0.05
Spearman correlation coefficient	0.34	<0.05
Kendall's Tau	0.28	<0.05
Somers' D Statistics	X/Y=0.46 Y/X=0.18	<0.05 <0.05
Gamma Statistic	0.48	<0.05

MATRIC POINT SCORE (X) AND BIO100F (ordinal) (Y)

YEAR : 1993, valid cases: 107		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.63	<0.05
Spearman correlation coefficient	0.64	<0.05
Kendall's Tau	0.47	<0.05
Somers' D Statistics	X/Y=0.46 Y/X=0.49	<0.05 <0.05
Gamma Statistic	0.50	<0.05

YEAR : 1994, valid cases: 161		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.45	<0.05
Spearman correlation coefficient	0.44	<0.05
Kendall's Tau	0.32	<0.05
Somers' D Statistics	X/Y=0.31 Y/X=0.33	<0.05 <0.05
Gamma Statistic	0.44	<0.05

YEAR : 1995, valid cases: 157		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.43	<0.05
Spearman correlation coefficient	0.39	<0.05
Kendall's Tau	0.29	<0.05
Somers' D Statistics	X/Y=0.28 Y/X=0.30	<0.05 <0.05
Gamma Statistic	0.30	<0.05

YEAR : 1996, valid cases: 96		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.47	<0.05
Spearman correlation coefficient	0.49	<0.05
Kendall's Tau	0.38	<0.05
Somers' D Statistics	X/Y=0.32 Y/X=0.45	<0.05 <0.05
Gamma Statistic	0.38	<0.05

YEAR : 1997, valid cases: 100		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.60	<0.05
Spearman correlation coefficient	0.57	<0.05
Kendall's Tau	0.42	<0.05
Somers' D Statistics	X/Y=0.41 Y/X=0.43	<0.05 <0.05
Gamma Statistic	0.44	<0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1993-1997 valid cases: 682		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.54	<0.05
Spearman correlation coefficient	0.51	<0.05
Kendall's Tau	0.37	<0.05
Somers' D Statistics	X/Y=0.36 Y/X=0.38	<0.05 <0.05
Gamma Statistic	0.39	<0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1993-1997 valid cases: 682		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.36	<0.05
Spearman correlation coefficient	0.34	<0.05
Kendall's Tau	0.30	<0.05
Somers' D Statistics	X/Y=0.48 Y/X=0.19	<0.05 <0.05
Gamma Statistic	0.52	<0.05

MATRIC HIGHER-GRADE BIOLOGY (X) AND BIO100F (ordinal) (Y)

YEAR : 1993, valid cases: 96		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.68	<0.05
Spearman correlation coefficient	0.69	<0.05
Kendall's Tau	0.55	<0.05
Somers' D Statistics	X/Y=0.48 Y/X=0.64	<0.05 <0.05
Gamma Statistic	0.65	<0.05

YEAR : 1994, valid cases: 145		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.34	<0.05
Spearman correlation coefficient	0.34	<0.05
Kendall's Tau	0.27	<0.05
Somers' D Statistics	X/Y=0.22 Y/X=0.32	<0.05 <0.05
Gamma Statistic	0.33	<0.05

YEAR : 1995, valid cases: 141		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.28	<0.05
Spearman correlation coefficient	0.26	<0.05
Kendall's Tau	0.20	<0.05
Somers' D Statistics	X/Y=0.17 Y/X=0.23	<0.05 <0.05
Gamma Statistic	0.24	<0.05

YEAR : 1996, valid cases: 89		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.56	<0.05
Spearman correlation coefficient	0.49	<0.05
Kendall's Tau	0.38	<0.05
Somers' D Statistics	X/Y=0.34 Y/X=0.45	<0.05 <0.05
Gamma Statistic	0.47	<0.05

YEAR : 1997, valid cases: 90		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.55	<0.05
Spearman correlation coefficient	0.55	<0.05
Kendall's Tau	0.44	<0.05
Somers' D Statistics	X/Y=0.38 Y/X=0.51	<0.05 <0.05
Gamma Statistic	0.51	<0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1993-1997 valid cases: 616		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.44	<0.05
Spearman correlation coefficient	0.44	<0.05
Kendall's Tau	0.34	<0.05
Somers' D Statistics	X/Y=0.29 Y/X=0.40	<0.05 <0.05
Gamma Statistic	0.41	<0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1993-1997 valid cases: 616		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.29	<0.05
Spearman correlation coefficient	0.28	<0.05
Kendall's Tau	0.26	<0.05
Somers' D Statistics	X/Y=0.37 Y/X=0.18	<0.05 <0.05
Gamma Statistic	0.49	<0.05

In the preceding tables we observe:

- All the measures are *statistically significant*.

Based on these observations, we may conclude that incoming first year student with either the higher (*lower*) faculty point score or the higher (*lower*) matric point score or the higher (*lower*) matric biology symbol are more likely (*less likely*) to achieve the higher (*lower*) percentage for BIO100F, respectively.

- The magnitudes of the measures are usually highest for associations involving the matric point scores, the lowest for associations involving the matric biology. This feature suggests that the matric point score, of the three investigated predictors, may be best over the 1993-97 period in predicting the actual percentage obtained by the students for BIO100F.

Examining the dendrograms we note:

- For the ordinal response analysis (CONFIRM output) the *matric point score* appears to be the best predictor of performance for BIO100F.

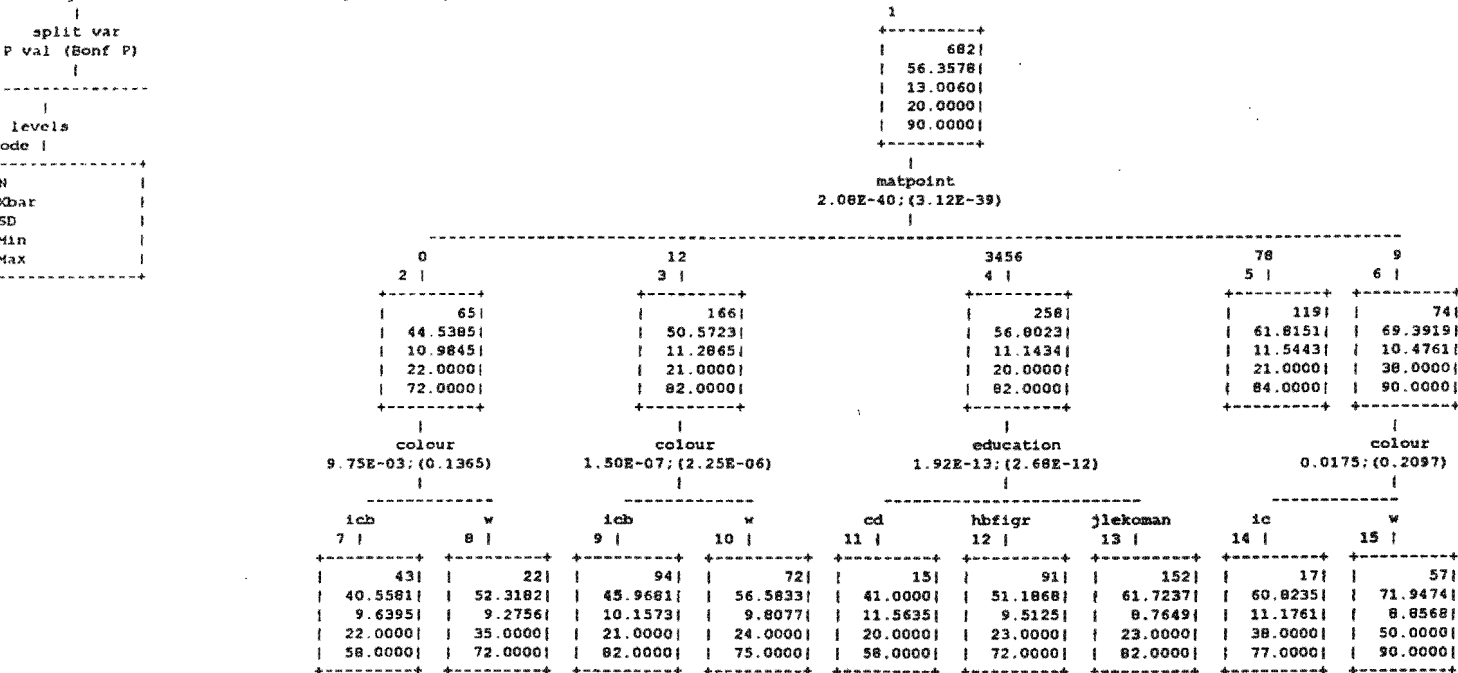
For most groupings of the matric point scores, the significant differences in the average mark emerge between the White category and the Black, Coloured and Indian combined categories. Only for the matric point scores between 36 and 39, the marked differences are present amongst the students from the formerly privileged and the students from the disadvantaged educational departments with the letter group, however, performing sufficiently worst.

- In contrast, for the binary response analysis (CATFIRM output) the *colour* variable appears to split the entire data most significantly and into two groups, namely the White category and the Black, Coloured and Indian categories combined.

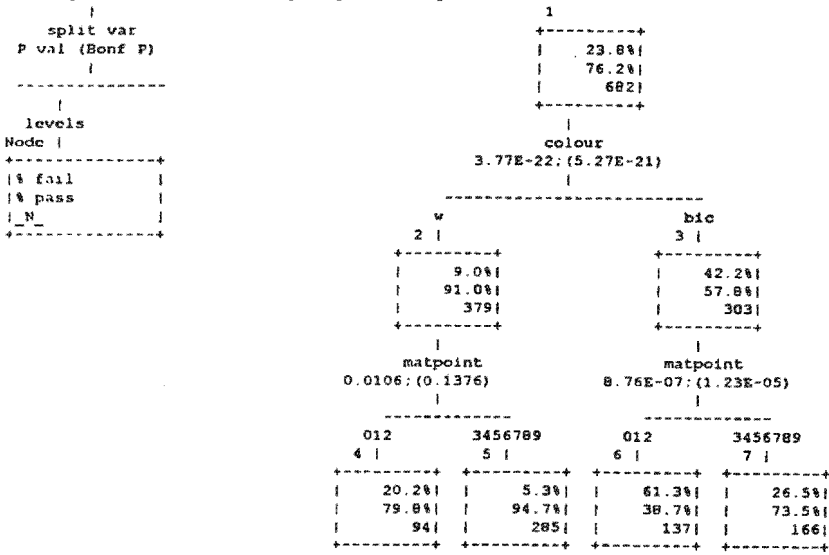
Only then the *matric point score* emerges as the second best predictor in identifying the students groupings with significantly different proportions of those passing the biology course. The proportions of the White students passing in either of the two groupings of the matric point scores are substantially higher than the proportions of the Black, Coloured or Indian students passing.

The analyses in FIRM have confirmed that the matric point score is a significant determinant of academic performance in BIO100F.

Dendrogram of COBFIRM: ordinal response analysis: bio100f



Dendrogram of CATFIRM: binary response analysis: bio100f



5.1.2 FIRST YEAR BOTANY COURSE: BOT102S

The database includes records for BOT102S in the years from 1990 to 1995. Thus, the relevant measures were calculated for each of these six individual years and then the combined years field.

The results are displayed in the tables below.

FACULTY POINT SCORE (X) AND BOT102S (ordinal) (Y)

YEAR : 1990, valid cases: 27		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.00	>0.05
Spearman correlation coefficient	0.04	>0.05
Kendall's Tau	0.04	>0.05
Somers' D Statistics	X/Y=0.04 Y/X=0.05	>0.05 >0.05
Gamma Statistic	0.05	>0.05

YEAR : 1991, valid cases: 33		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.49	<0.05
Spearman correlation coefficient	0.37	<0.05
Kendall's Tau	0.26	<0.05
Somers' D Statistics	X/Y=0.25 Y/X=0.26	<0.05 <0.05
Gamma Statistic	0.37	<0.05

YEAR : 1992, valid cases: 36		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.42	<0.05
Spearman correlation coefficient	0.46	<0.05
Kendall's Tau	0.34	<0.05
Somers' D Statistics	X/Y=0.33 Y/X=0.34	<0.05 <0.05
Gamma Statistic	0.35	<0.05

YEAR : 1993, valid cases: 27		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.40	>0.05
Spearman correlation coefficient	0.44	>0.05
Kendall's Tau	0.31	>0.05
Somers' D Statistics	X/Y=0.31 Y/X=0.31	>0.05 >0.05
Gamma Statistic	0.33	>0.05

YEAR : 1994, valid cases: 27		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.53	<0.05
Spearman correlation coefficient	0.55	<0.05
Kendall's Tau	0.43	<0.05
Somers' D Statistics	X/Y=0.43 Y/X=0.44	<0.05 <0.05
Gamma Statistic	0.46	<0.05

YEAR : 1995, valid cases: 48		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.46	<0.05
Spearman correlation coefficient	0.49	<0.05
Kendall's Tau	0.35	<0.05
Somers' D Statistics	X/Y=0.34 Y/X=0.36	<0.05 <0.05
Gamma Statistic	0.36	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1995 valid cases: 243		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.35	<0.05
Spearman correlation coefficient	0.46	<0.05
Kendall's Tau	0.34	<0.05
Somers' D Statistics	X/Y=0.33 Y/X=0.34	<0.05 <0.05
Gamma Statistic	0.35	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1995 valid cases: 243		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.15	<0.05
Spearman correlation coefficient	0.14	<0.05
Kendall's Tau	0.12	<0.05
Somers' D Statistics	X/Y=0.27 Y/X=0.05	<0.05 >0.05
Gamma Statistic	0.29	<0.05

MATRIC POINT SCORE (X) AND BOT102S (ordinal) (Y)

YEAR : 1990, valid cases: 27		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.15	<0.05
Spearman correlation coefficient	0.17	<0.05
Kendall's Tau	0.13	<0.05
Somers' D Statistics	X/Y=0.12	<0.05
	Y/X=0.13	<0.05
Gamma Statistic	0.14	<0.05

YEAR : 1991, valid cases: 33		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.53	<0.05
Spearman correlation coefficient	0.42	<0.05
Kendall's Tau	0.29	<0.05
Somers' D Statistics	X/Y=0.28	<0.05
	Y/X=0.30	<0.05
Gamma Statistic	0.31	<0.05

YEAR : 1992, valid cases: 36		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.49	<0.05
Spearman correlation coefficient	0.56	<0.05
Kendall's Tau	0.40	<0.05
Somers' D Statistics	X/Y=0.39	<0.05
	Y/X=0.42	<0.05
Gamma Statistic	0.56	<0.05

YEAR : 1993, valid cases: 27		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.43	<0.05
Spearman correlation coefficient	0.48	<0.05
Kendall's Tau	0.35	<0.05
Somers' D Statistics	X/Y=0.34	<0.05
	Y/X=0.36	<0.05
Gamma Statistic	0.37	<0.05

YEAR : 1994, valid cases: 27		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.51	<0.05
Spearman correlation coefficient	0.51	<0.05
Kendall's Tau	0.38	<0.05
Somers' D Statistics	X/Y=0.37	<0.05
	Y/X=0.39	<0.05
Gamma Statistic	0.41	<0.05

YEAR : 1995, valid cases: 48		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.52	<0.05
Spearman correlation coefficient	0.57	<0.05
Kendall's Tau	0.43	<0.05
Somers' D Statistics	X/Y=0.42	<0.05
	Y/X=0.44	<0.05
Gamma Statistic	0.45	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1995 valid cases: 243		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.49	<0.05
Spearman correlation coefficient	0.52	<0.05
Kendall's Tau	0.37	<0.05
Somers' D Statistics	X/Y=0.37	<0.05
	Y/X=0.40	<0.05
Gamma Statistic	0.41	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1995 valid cases: 243		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.20	<0.05
Spearman correlation coefficient	0.20	<0.05
Kendall's Tau	0.17	<0.05
Somers' D Statistics	X/Y=0.39	<0.05
	Y/X=0.07	>0.05
Gamma Statistic	0.42	<0.05

MATRIC HIGHER-GRADE BIOLOGY (X) AND BOT102S (ordinal) (Y)

YEAR : 1990, valid cases: 27		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	-0.18	>0.05
Spearman correlation coefficient	-0.07	>0.05
Kendall's Tau	-0.05	>0.05
Somers' D Statistics	X/Y=-0.05 Y/X=-0.06	>0.05 >0.05
Gamma Statistic	-0.06	>0.05

YEAR : 1991, valid cases: 30		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.16	>0.05
Spearman correlation coefficient	0.11	>0.05
Kendall's Tau	0.09	>0.05
Somers' D Statistics	X/Y=0.07 Y/X=0.12	>0.05 >0.05
Gamma Statistic	0.12	>0.05

YEAR : 1992, valid cases: 34		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.34	>0.05
Spearman correlation coefficient	0.16	>0.05
Kendall's Tau	0.11	>0.05
Somers' D Statistics	X/Y=0.09 Y/X=0.13	>0.05 >0.05
Gamma Statistic	0.13	>0.05

YEAR : 1993, valid cases: 26		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.30	<0.05
Spearman correlation coefficient	0.32	>0.05
Kendall's Tau	0.25	>0.05
Somers' D Statistics	X/Y=0.21 Y/X=0.30	>0.05 >0.05
Gamma Statistic	0.32	>0.05

YEAR : 1994, valid cases: 23		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.50	<0.05
Spearman correlation coefficient	0.57	<0.05
Kendall's Tau	0.45	<0.05
Somers' D Statistics	X/Y=0.38 Y/X=0.54	<0.05 <0.05
Gamma Statistic	0.57	<0.05

YEAR : 1995, valid cases: 41		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.09	>0.05
Spearman correlation coefficient	0.12	>0.05
Kendall's Tau	0.08	>0.05
Somers' D Statistics	X/Y=0.07 Y/X=0.10	>0.05 >0.05
Gamma Statistic	0.10	>0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1995 valid cases: 221		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.19	<0.05
Spearman correlation coefficient	0.23	<0.05
Kendall's Tau	0.18	<0.05
Somers' D Statistics	X/Y=0.15 Y/X=0.21	<0.05 <0.05
Gamma Statistic	0.21	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1995 valid cases: 221		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	-0.02	>0.05
Spearman correlation coefficient	-0.03	>0.05
Kendall's Tau	-0.03	>0.05
Somers' D Statistics	X/Y=-0.06 Y/X=-0.01	>0.05 >0.05
Gamma Statistic	-0.08	>0.05

In the proceeding tables we note:

- The measures emerge as *statistically significant* mostly for the relationships involving the faculty or the matric point scores. There does not seem to be any significant association between the actual matric biology symbol and the performance in the botany course.

Thus, we may conclude that the incoming first year student with either the higher (*lower*) faculty point score or the higher (*lower*) matric point score are more likely (*less likely*) to achieve the higher (*lower*) percentage for BOT102S, respectively, over the period of analysis.

- The magnitudes of the measures are usually higher for the relationships involving the matric point score than for the faculty point score. This feature once again suggests that the matric point score, out of the three investigated variables, may better in explaining variation in marks obtained by students for BOT102S.

Examining the dendrograms we note:

- Whether the analysis is performed with the ordinal or the binary definitions of the academic performance, the *colour* variable appears to be the best predictor for BOT102F.
- In both analyses the White category obtains the highest mean percentage or has the highest pass rate in comparison to any other colour categories.
- In the ordinal response analysis the *matric point score* is found to be a further significant predictor of performance for the White category only. The higher matric point scores are associated with the higher mean percentage. Additionally within this White student grouping, those with the matric point scores of 42 and more, those students who obtained A for the *higher-grade matric geography* have a higher percentage mark (74.95%) than those who obtained B (68.10%).

Referring to the results obtained for the measures of association, the CONFIRM analysis has found a predictive value for the matric point scores but only for the performance of the White students.

Dendrogram of CONFIRM: ordinal response analysis: bot102s

	1		
split var	+-----+		
P val (Bonf P)	243		
	62.5021		
	10.6427		
	22.0000		
	84.0000		
	+-----+		
levels			
Node	+-----+		
	colour		
	6.41E-19; (9.61E-18)		
	+-----+		
	b	c1?	w
	2	3	4
	+-----+		
	18	61	164
	48.3333	55.6066	66.6220
	9.2546	10.8294	7.6826
	31.0000	22.0000	34.0000
	62.0000	77.0000	84.0000
	+-----+		
	matpoint		
	2.55E-09; (3.31E-08)		
	+-----+		
	012	34567	89
	5	6	7
	+-----+		
	44	88	32
	60.6818	67.3409	72.8125
	7.6547	6.2989	5.1584
	34.0000	53.0000	63.0000
	73.0000	84.0000	83.0000
	+-----+		
	geoh		
	0.0384; (0.4607)		
	+-----+		
	b	?a	
	8	9	
	+-----+		
	10	22	
	68.1000	74.9545	
	2.8067	4.5301	
	63.0000	66.0000	
	72.0000	83.0000	
	+-----+		

Dendrogram of CATFIRM: binary response analysis: bot102s

	1	
split var	+-----+	
P val (Bonf P)	9.5%	
	90.5%	
	243	
	+-----+	
levels		
Node	+-----+	
	colour	
	3.00E-06; (4.20E-05)	
	+-----+	
	?w	bci
	2	3
	+-----+	
	1.8%	25.6%
	98.2%	74.4%
	165	78
	+-----+	

5.1.3 FIRST YEAR ZOOLOGY COURSE: ZOO103S

The database includes the records of ZOO103S in the years 1990-1995. Thus, the measures were determined for the individual six years and the combined years field.

The results are presented in the tables below.

FACULTY POINT SCORE (X) AND ZOO103S (ordinal) (Y)

YEAR : 1990, valid cases: 69		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.54	<0.05
Spearman correlation coefficient	0.48	<0.05
Kendall's Tau	0.36	<0.05
Somers' D Statistics	X/Y=0.35	<0.05
	Y/X=0.36	<0.05
Gamma Statistic	0.37	<0.05

YEAR : 1991, valid cases: 81		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.30	<0.05
Spearman correlation coefficient	0.22	<0.05
Kendall's Tau	0.15	<0.05
Somers' D Statistics	X/Y=0.15	<0.05
	Y/X=0.15	<0.05
Gamma Statistic	0.16	<0.05

YEAR : 1992, valid cases: 96		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.28	<0.05
Spearman correlation coefficient	0.25	<0.05
Kendall's Tau	0.18	<0.05
Somers' D Statistics	X/Y=0.18	<0.05
	Y/X=0.18	<0.05
Gamma Statistic	0.19	<0.05

YEAR : 1993, valid cases: 80		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.37	<0.05
Spearman correlation coefficient	0.43	<0.05
Kendall's Tau	0.30	<0.05
Somers' D Statistics	X/Y=0.30	<0.05
	Y/X=0.30	<0.05
Gamma Statistic	0.31	<0.05

YEAR : 1994, valid cases: 124		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.34	<0.05
Spearman correlation coefficient	0.33	<0.05
Kendall's Tau	0.23	<0.05
Somers' D Statistics	X/Y=0.23 Y/X=0.24	<0.05
Gamma Statistic	0.24	<0.05

YEAR : 1995, valid cases: 134		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.39	<0.05
Spearman correlation coefficient	0.36	<0.05
Kendall's Tau	0.26	<0.05
Somers' D Statistics	X/Y=0.25 Y/X=0.26	<0.05
Gamma Statistic	0.27	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1995 valid cases: 642		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.41	<0.05
Spearman correlation coefficient	0.38	<0.05
Kendall's Tau	0.28	<0.05
Somers' D Statistics	X/Y=0.27 Y/X=0.28	<0.05
Gamma Statistic	0.29	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1995 valid cases: 642		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.14	<0.05
Spearman correlation coefficient	0.14	<0.05
Kendall's Tau	0.11	<0.05
Somers' D Statistics	X/Y=0.28 Y/X=0.05	<0.05
Gamma Statistic	0.30	<0.05

MATRIC POINT SCORE (X) AND ZOO103S (ordinal) (Y)

YEAR : 1990, valid cases: 69		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.56	<0.05
Spearman correlation coefficient	0.50	<0.05
Kendall's Tau	0.37	<0.05
Somers' D Statistics	X/Y=0.36	<0.05
	Y/X=0.38	<0.05
Gamma Statistic	0.39	<0.05

YEAR : 1991, valid cases: 81		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.46	<0.05
Spearman correlation coefficient	0.42	<0.05
Kendall's Tau	0.30	<0.05
Somers' D Statistics	X/Y=0.30	<0.05
	Y/X=0.31	<0.05
Gamma Statistic	0.32	<0.05

YEAR : 1992, valid cases: 96		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.41	<0.05
Spearman correlation coefficient	0.34	<0.05
Kendall's Tau	0.26	<0.05
Somers' D Statistics	X/Y=0.25	<0.05
	Y/X=0.26	<0.05
Gamma Statistic	0.27	<0.05

YEAR : 1993, valid cases: 80		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.44	<0.05
Spearman correlation coefficient	0.46	<0.05
Kendall's Tau	0.33	<0.05
Somers' D Statistics	X/Y=0.32	<0.05
	Y/X=0.34	<0.05
Gamma Statistic	0.35	<0.05

YEAR : 1994, valid cases: 124		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.38	<0.05
Spearman correlation coefficient	0.36	<0.05
Kendall's Tau	0.26	<0.05
Somers' D Statistics	X/Y=0.25	<0.05
	Y/X=0.26	<0.05
Gamma Statistic	0.27	<0.05

YEAR : 1995, valid cases: 134		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.39	<0.05
Spearman correlation coefficient	0.38	<0.05
Kendall's Tau	0.28	<0.05
Somers' D Statistics	X/Y=0.27 Y/X=0.29	<0.05 <0.05
Gamma Statistic	0.30	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1995 valid cases: 642		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.47	<0.05
Spearman correlation coefficient	0.44	<0.05
Kendall's Tau	0.32	<0.05
Somers' D Statistics	X/Y=0.31 Y/X=0.33	<0.05 <0.05
Gamma Statistic	0.34	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1995 valid cases: 642		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.18	<0.05
Spearman correlation coefficient	0.16	<0.05
Kendall's Tau	0.26	<0.05
Somers' D Statistics	X/Y=0.34 Y/X=0.06	<0.05 >0.05
Gamma Statistic	0.37	<0.05

MATRIC HIGHER-GRADE BIOLOGY (X) AND ZOO103S (ordinal) (Y)

YEAR : 1990, valid cases: 63		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.20	>0.05
Spearman correlation coefficient	0.18	>0.05
Kendall's Tau	0.14	>0.05
Somers' D Statistics	X/Y=0.12 Y/X=0.16	>0.05 >0.05
Gamma Statistic	0.17	>0.05

YEAR : 1991, valid cases: 72		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.31	<0.05
Spearman correlation coefficient	0.24	<0.05
Kendall's Tau	0.19	<0.05
Somers' D Statistics	X/Y=0.17 Y/X=0.23	<0.05 <0.05
Gamma Statistic	0.24	<0.05

YEAR : 1992, valid cases: 88		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.24	<0.05
Spearman correlation coefficient	0.25	<0.05
Kendall's Tau	0.20	<0.05
Somers' D Statistics	X/Y=0.17 Y/X=0.23	<0.05 <0.05
Gamma Statistic	0.24	<0.05

YEAR : 1993, valid cases: 76		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.55	<0.05
Spearman correlation coefficient	0.53	<0.05
Kendall's Tau	0.41	<0.05
Somers' D Statistics	X/Y=0.36 Y/X=0.48	<0.05 <0.05
Gamma Statistic	0.50	<0.05

YEAR : 1994, valid cases: 111		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.27	<0.05
Spearman correlation coefficient	0.29	<0.05
Kendall's Tau	0.22	<0.05
Somers' D Statistics	X/Y=0.19 Y/X=0.27	<0.05 <0.05
Gamma Statistic	0.27	<0.05

YEAR : 1995, valid cases: 120		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.17	>0.05
Spearman correlation coefficient	0.18	>0.05
Kendall's Tau	0.14	>0.05
Somers' D Statistics	X/Y=0.12 Y/X=0.17	>0.05 >0.05
Gamma Statistic	0.17	>0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1995 valid cases: 582		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.35	<0.05
Spearman correlation coefficient	0.33	<0.05
Kendall's Tau	0.26	<0.05
Somers' D Statistics	X/Y=0.22 Y/X=0.29	<0.05 <0.05
Gamma Statistic	0.31	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1995 valid cases: 582		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.17	<0.05
Spearman correlation coefficient	0.14	<0.05
Kendall's Tau	0.13	<0.05
Somers' D Statistics	X/Y=0.28 Y/X=0.06	<0.05 >0.05
Gamma Statistic	0.36	<0.05

Examining the dendrograms we note:

- In CONFIRM and CATFIRM outputs, the *colour* variable splits the data set most significantly and into groupings with worsening levels of performance in ZOO103S: in the ordinal response analysis into three groups, namely the White category and the Black category and the Coloured and Indian combined categories, and in the binary response analysis into two groups, namely the White and the non-White categories.
- In both dendrograms the White category, in comparison to the other colour groupings, is either the group with the highest mean percentage mark or the one with the highest pass rate.
- Additionally in the ordinal response analyses in the combined group of the Coloured and Indian students, the mean mark for those who completed the course in 1990 is substantially lower than those who completed the courses in any later year.

Somewhat contrary to the expectations created by the results in the measures of association, the FIRM analyses have found no predictive value for the matric point scores in relation to the performance in ZOO103S.

Dendrogram of CONFIRM: ordinal response analysis: zool03s ; alldata.txt

split var		1		
P val (Bonf P)		642		
		1.9159		
		0.2778		
		1.0000		
		2.0000		
levels		colour		
Node		4.25E-12; (6.37E-11)		
		Year		
		0.0190; (0.2844)		
		0 42531		
		5 6		
		b ?ci w		
		2 3 4		
		55 214 373		
		1.7091 1.8505 1.9839		
		0.4584 0.3574 0.1260		
		1.0000 1.0000 1.0000		
		2.0000 2.0000 2.0000		
		28 186		
		1.5714 1.8925		
		0.5040 0.3106		
		1.0000 1.0000		
		2.0000 2.0000		

Dendrogram of CATFIRM: bzool03s ; alldata.txt

split var		1	
P val (Bonf P)		8.4%	
		91.6%	
		642	
levels		colour	
Node		2.07E-10; (2.90E-09)	
		?cib w	
		2 3	
		17.8% 1.6%	
		82.2% 98.4%	
		269 373	

5.2 MATHEMATICS RELATED UNIVERSITY COURSES

5.2.1 FIRST YEAR MATHS COURSES

Following the amalgamation of the Departments of Mathematics and Applied Mathematics in 1995, codes with respect to subjects offered by both departments were changed. Thus, MAM100W is essentially a continuation of MTH105W and MAM104F is a continuation of MTH101F with slight syllabus changes.

a) MATHS COURSES (full year courses): MTH105W and MAM100W

a1) MTH105W

For MTH105W the measures of association were calculated for the years 1990-1994 and the combined years field.

The results are displayed in the tables below.

FACULTY POINT SCORE (X) AND MTH105W (ordinal) (Y)

YEAR : 1990, valid cases: 93		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.67	<0.05
Spearman correlation coefficient	0.68	< 0.05
Kendall's Tau	0.50	< 0.05
Somers' D Statistics	X/Y=0.49 Y/X=0.50	< 0.05 < 0.05
Gamma Statistic	0.51	< 0.05

YEAR : 1991, valid cases: 98		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.61	<0.05
Spearman correlation coefficient	0.61	< 0.05
Kendall's Tau	0.44	< 0.05
Somers' D Statistics	X/Y=0.44 Y/X=0.45	< 0.05 < 0.05
Gamma Statistic	0.46	< 0.05

YEAR : 1992, valid cases: 73		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.49	<0.05
Spearman correlation coefficient	0.44	< 0.05
Kendall's Tau	0.31	< 0.05
Somers' D Statistics	X/Y=0.31 Y/X=0.32	< 0.05 < 0.05
Gamma Statistic	0.33	< 0.05

YEAR : 1993, valid cases: 70		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.47	<0.05
Spearman correlation coefficient	0.50	< 0.05
Kendall's Tau	0.34	< 0.05
Somers' D Statistics	X/Y=0.34 Y/X=0.34	< 0.05 < 0.05
Gamma Statistic	0.35	< 0.05

YEAR : 1994, valid cases: 97		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.52	<0.05
Spearman correlation coefficient	0.50	< 0.05
Kendall's Tau	0.37	< 0.05
Somers' D Statistics	X/Y=0.37 Y/X=0.38	< 0.05 < 0.05
Gamma Statistic	0.39	< 0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1995 valid cases: 448		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.59	<0.05
Spearman correlation coefficient	0.59	< 0.05
Kendall's Tau	0.42	< 0.05
Somers' D Statistics	X/Y=0.41 Y/X=0.43	< 0.05 < 0.05
Gamma Statistic	0.44	< 0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1995 valid cases: 448		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.38	<0.05
Spearman correlation coefficient	0.37	< 0.05
Kendall's Tau	0.31	< 0.05
Somers' D Statistics	X/Y=0.52 Y/X=0.19	< 0.05 < 0.05
Gamma Statistic	0.54	< 0.05

MATRIC POINT SCORE (X) AND MTH105W (ordinal) (Y)

YEAR : 1990, valid cases: 93		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.64	<0.05
Spearman correlation coefficient	0.68	< 0.05
Kendall's Tau	0.49	< 0.05
Somers' D Statistics	X/Y=0.49 Y/X=0.50	< 0.05 < 0.05
Gamma Statistic	0.51	< 0.05

YEAR : 1991, valid cases: 98		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.60	<0.05
Spearman correlation coefficient	0.61	< 0.05
Kendall's Tau	0.44	< 0.05
Somers' D Statistics	X/Y=0.43 Y/X=0.45	< 0.05 < 0.05
Gamma Statistic	0.46	< 0.05

YEAR : 1992, valid cases: 73		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.42	<0.05
Spearman correlation coefficient	0.43	< 0.05
Kendall's Tau	0.31	< 0.05
Somers' D Statistics	X/Y=0.30 Y/X=0.32	< 0.05 < 0.05
Gamma Statistic	0.33	< 0.05

YEAR : 1993, valid cases: 70		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.48	<0.05
Spearman correlation coefficient	0.50	< 0.05
Kendall's Tau	0.35	< 0.05
Somers' D Statistics	X/Y=0.35 Y/X=0.36	< 0.05 < 0.05
Gamma Statistic	0.36	< 0.05

YEAR : 1994, valid cases: 97		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.51	,0.05
Spearman correlation coefficient	0.48	< 0.05
Kendall's Tau	0.38	< 0.05
Somers' D Statistics	X/Y=0.35 Y/X=0.37	< 0.05 < 0.05
Gamma Statistic	0.38	< 0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1995 valid cases: 448		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.57	<0.05
Spearman correlation coefficient	0.58	< 0.05
Kendal's Tau	0.42	< 0.05
Somers' D Statistics	X/Y=0.41 Y/X=0.43	< 0.05 < 0.05
Gamma Statistic	0.44	< 0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1995 valid cases: 448		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.39	<0.05
Spearman correlation coefficient	0.37	< 0.05
Kendal's Tau	0.31	< 0.05
Somers' D Statistics	X/Y=0.52 Y/X=0.19	< 0.05 < 0.05
Gamma Statistic	0.55	< 0.05

MATRIC HIGHER-GRADE MATHS (X) AND MTH105W (Y)

YEAR : 1990, valid cases: 93		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.54	<0.05
Spearman correlation coefficient	0.57	< 0.05
Kendall's Tau	0.46	< 0.05
Somers' D Statistics	X/Y=0.38	< 0.05
	Y/X=0.57	< 0.05
Gamma Statistic	0.58	< 0.05

YEAR : 1991, valid cases: 98		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.62	<0.05
Spearman correlation coefficient	0.62	< 0.05
Kendall's Tau	0.49	< 0.05
Somers' D Statistics	X/Y=0.43	< 0.05
	Y/X=0.56	< 0.05
Gamma Statistic	0.57	< 0.05

YEAR : 1992, valid cases: 73		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.36	<0.05
Spearman correlation coefficient	0.36	< 0.05
Kendall's Tau	0.27	< 0.05
Somers' D Statistics	X/Y=0.24	< 0.05
	Y/X=0.31	< 0.05
Gamma Statistic	0.32	< 0.05

YEAR : 1993, valid cases: 70		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.43	<0.05
Spearman correlation coefficient	0.49	< 0.05
Kendall's Tau	0.37	< 0.05
Somers' D Statistics	X/Y=0.32	< 0.05
	Y/X=0.43	< 0.05
Gamma Statistic	0.44	< 0.05

YEAR : 1994, valid cases: 97		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.47	<0.05
Spearman correlation coefficient	0.47	< 0.05
Kendall's Tau	0.37	< 0.05
Somers' D Statistics	X/Y=0.31	< 0.05
	Y/X=0.43	< 0.05
Gamma Statistic	0.44	< 0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1995 valid cases: 447		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.53	<0.05
Spearman correlation coefficient	0.54	<0.05
Kendall's Tau	0.43	<0.05
Somers' D Statistics	X/Y=0.37 Y/X=0.49	<0.05 <0.05
Gamma Statistic	0.50	<0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1995 valid cases: 447		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.35	<0.05
Spearman correlation coefficient	0.34	<0.05
Kendall's Tau	0.31	<0.05
Somers' D Statistics	X/Y=0.45 Y/X=0.21	<0.05 <0.05
Gamma Statistic	0.57	<0.05

In the preceding tables we note:

- All the measures are *statistically significant*.

Therefore, we may conclude that the incoming first year students with either the higher (*lower*) faculty point score or the higher (*lower*) matric point score or the higher (*lower*) matric higher- grade maths symbol are more likely (*less likely*) to achieve the higher (*lower*) percentage for MTH105W, respectively.

- The measures are of comparable magnitudes with the measures involving the matric point score being only marginally but not meaningfully higher.

Examining the dendrograms we note:

- In both the ordinal and binary response analyses (CONFIRM and CATFIRM outputs) the *faculty point score* is chosen as the most significant predictor of academic performance in MTH105W.
- Additionally, in the ordinal response analyses, for the students with the faculty point scores between 47 and 51, those from the disadvantaged educational departments obtained on average a lower percentage mark (44.30%) than those from the privileged educational departments (56.20%).
- On the other hand, in the binary response analysis, for the students with the faculty point scores of between 47 and 55, the proportion of the White students passing the subject is substantially higher (83.6%) than the proportions of the Black, Coloured and Indian students combined (55.9%).

Dendrogram of CONFIRM: ordinal response analysis: mth105w

split var	1
P val (Bonf P)	448
levels	57.9063 15.9370 13.0000 100.0000
Node	facpoint 2.68E-37; (4.02E-36)
N	0 123 45 67 89
Xbar	2 3 4 5 6
SD	48 133 103 87 77
Min	42.7917 51.1955 55.9223 65.0000 73.5584
Max	11.4668 12.4988 13.6207 13.0277 12.8805
	13.0000 20.0000 25.0000 35.0000 25.0000
	63.0000 74.0000 88.0000 100.0000 100.0000
	education 1.95E-04; (2.73E-03)
	diaf hecqb
	7 8
	56 77
	44.3036 56.2078
	11.6836 10.5872
	20.0000 24.0000
	69.0000 74.0000

Dendrogram of CATFIRM: binary response analysis: mth105w

split var	1
P val (Bonf P)	22.3% 77.7% 448
levels	facpoint 2.35E-12; (3.29E-11)
Node	0 12345 6789
% fail	2 3 4
% pass	56.3% 28.4% 3.7%
N	43.8% 71.6% 96.3%
	48 236 164
	colour 2.55E-03; (0.0357)
	?w bci
	5 6
	16.4% 44.1%
	83.6% 55.9%
	134 102

a2) MAM100W

For MAM100W the measures of association were calculated for the years 1995-1997 and the combined years field.

The results are displayed in the tables below.

FACULTY POINT SCORE (X) and MAM100W (ordinal) (Y)

YEAR : 1995, valid cases: 117		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.58	<0.05
Spearman correlation coefficient	0.56	< 0.05
Kendall's Tau	0.40	< 0.05
Somers' D Statistics	X/Y=0.38 Y/X=0.41	< 0.05 < 0.05
Gamma Statistic	0.42	< 0.05

YEAR : 1996, valid cases: 93		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.66	<0.05
Spearman correlation coefficient	0.70	< 0.05
Kendall's Tau	0.53	< 0.05
Somers' D Statistics	X/Y=0.52 Y/X=0.54	< 0.05 < 0.05
Gamma Statistic	0.55	< 0.05

YEAR : 1997, valid cases: 95		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.56	<0.05
Spearman correlation coefficient	0.59	< 0.05
Kendall's Tau	0.44	< 0.05
Somers' D Statistics	X/Y=0.44 Y/X=0.45	< 0.05 < 0.05
Gamma Statistic	0.45	< 0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1995-1997 valid cases: 357		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.60	<0.05
Spearman correlation coefficient	0.61	< 0.05
Kendall's Tau	0.45	< 0.05
Somers' D Statistics	X/Y=0.45	< 0.05
	Y/X=0.46	< 0.05
Gamma Statistic	0.47	< 0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1995-1997 valid cases: 357		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.44	<0.05
Spearman correlation coefficient	0.43	< 0.05
Kendall's Tau	0.36	< 0.05
Somers' D Statistics	X/Y=0.53	< 0.05
	Y/X=0.24	< 0.05
Gamma Statistic	0.55	< 0.05

MATRIC POINT SCORE (X) AND MAM100W (ordinal) (Y)

YEAR : 1995, valid cases: 117		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.57	<0.05
Spearman correlation coefficient	0.53	< 0.05
Kendall's Tau	0.38	< 0.05
Somers' Statistics	X/Y=0.37	< 0.05
	Y/X=0.39	< 0.05
Gamma Statistic	0.39	< 0.05

YEAR : 1996, valid cases: 93		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.62	<0.05
Spearman correlation coefficient	0.66	< 0.05
Kendall's Tau	0.48	< 0.05
Somers' D Statistics	X/Y=0.47	< 0.05
	Y/X=0.49	< 0.05
Gamma Statistic	0.49	< 0.05

YEAR : 1997, valid cases: 95		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.56	<0.05
Spearman correlation coefficient	0.59	< 0.05
Kendall's Tau	0.44	< 0.05
Somers' D Statistics	X/Y=0.43	< 0.05
	Y/X=0.45	< 0.05
Gamma Statistic	0.46	< 0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1995-1997 valid cases: 357		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.59	<0.05
Spearman correlation coefficient	0.59	< 0.05
Kendall's Tau	0.43	< 0.05
Somers' D Statistics	X/Y=0.43 Y/X=0.44	< 0.05 < 0.05
Gamma Statistic	0.45	< 0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1995-1997 valid cases: 357		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.44	<0.05
Spearman correlation coefficient	0.40	< 0.05
Kendall's Tau	0.34	< 0.05
Somers' D Statistics	X/Y=0.50 Y/X=0.23	< 0.05 < 0.05
Gamma Statistic	0.53	< 0.05

MATRIC HIGHER-GRADE MATH (X) AND MAM100W (ordinal) (Y)

YEAR : 1996, valid cases: 117		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.54	<0.05
Spearman correlation coefficient	0.54	< 0.05
Kendall's Tau	0.43	< 0.05
Somers' D Statistics	X/Y=0.37 Y/X=0.50	< 0.05 < 0.05
Gamma Statistic	0.51	< 0.05

YEAR : 1996, valid cases: 93		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.58	<0.05
Spearman correlation coefficient	0.63	< 0.05
Kendall's Tau	0.51	< 0.05
Somers' D Statistics	X/Y=0.43 Y/X=0.61	< 0.05 < 0.05
Gamma Statistic	0.62	< 0.05

YEAR : 1997, valid cases: 95		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.56	<0.05
Spearman correlation coefficient	0.57	< 0.05
Kendall's Tau	0.46	< 0.05
Somers' D Statistics	X/Y=0.38 Y/X=0.55	< 0.05 < 0.05
Gamma Statistic	0.55	< 0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1995-1997 valid cases: 355		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.56	<0.05
Spearman correlation coefficient	0.58	< 0.05
Kendall's Tau	0.46	< 0.05
Somers' D Statistics	X/Y=0.40	< 0.05
	Y/X=0.53	< 0.05
Gamma Statistic	0.54	< 0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1995-1997 valid cases: 355		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.43	<0.05
Spearman correlation coefficient	0.42	< 0.05
Kendall's Tau	0.39	< 0.05
Somers' D Statistics	X/Y=0.51	< 0.05
	Y/X=0.30	< 0.05
Gamma Statistic	0.63	< 0.05

In the preceding tables we note:

- Similarly to MTH105W, all the measures are *statistically significant*.

Therefore, we may conclude that the incoming first year students with either the higher (*lower*) faculty point score or the higher (*lower*) matric point score or the higher (*lower*) matric higher- grade maths symbol are more likely (*less likely*) to achieve the higher (*lower*) percentage for MAM100W, respectively.

- The measures are of comparable magnitudes with those involving the faculty point score being marginally but not meaningfully higher.

Examining the dendrograms we note:

- In the ordinal response analysis (CONFIRM output) the *faculty point score* seems to be the best performance predictor.

Within the group of students with the faculty point scores between 57 and 60, females obtain lower percentage mark (54.39%) than males (64.98%). Additionally, for that group of female students, the *matric higher-grade English* symbol may act as further performance predictor with the symbols A, B and C corresponding to the higher mark (57.62%) than the symbols D and E (38.71%).

Within the group of students with the faculty point score of 62, the students from the privileged educational departments obtained higher marks (74.54%) than the students from the disadvantaged departments (45.17%).

- In the binary response analyses (CATFIRM output) the *matric higher-grade physical science* appears to predict the proportions of students passing or failing MAM100W significantly well.

It appears that different factors should be used to predict performance in the two maths courses, MTH105W and MAM100W. This observation may be attributed to a changed profile of students admitted into the course after amalgamation of the two departments.

b) MATHS COURSES (half-year courses): MTH101F and MAM104F

b1) MTH101F

For MTH101F the measures of association were calculated for the years 1990-1994 and the combined years field.

The results are displayed in the tables below.

FACULTY POINT SCORE (X) AND MTH101F (ordinal) (Y)

YEAR : 1990, valid cases: 68		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.42	<0.05
Spearman correlation coefficient	0.40	<0.05
Kendall's Tau	0.30	<0.05
Somers' D Statistics	X/Y=0.30	<0.05
	Y/X=0.30	<0.05
Gamma Statistic	0.32	<0.05

YEAR : 1991, valid cases: 74		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.21	>0.05
Spearman correlation coefficient	0.16	>0.05
Kendall's Tau	0.12	>0.05
Somers' D Statistics	X/Y=0.12	>0.05
	Y/X=0.13	>0.05
Gamma Statistic	0.13	>0.05

YEAR : 1992, valid cases: 88		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.14	>0.05
Spearman correlation coefficient	0.17	>0.05
Kendall's Tau	0.11	>0.05
Somers' D Statistics	X/Y=0.11	>0.05
	Y/X=0.12	>0.05
Gamma Statistic	0.12	>0.05

YEAR : 1993, valid cases: 83		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.36	<0.05
Spearman correlation coefficient	0.46	<0.05
Kendall's Tau	0.32	<0.05
Somers' D Statistics	X/Y=0.31	<0.05
	Y/X=0.33	<0.05
Gamma Statistic	0.33	<0.05

YEAR : 1994, valid cases: 115		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.45	<0.05
Spearman correlation coefficient	0.45	<0.05
Kendall's Tau	0.32	<0.05
Somers' D Statistics	X/Y=0.31 Y/X=0.33	<0.05 <0.05
Gamma Statistic	0.34	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1995 valid cases: 453		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.32	<0.05
Spearman correlation coefficient	0.34	<0.05
Kendall's Tau	0.24	<0.05
Somers' D Statistics	X/Y=0.23 Y/X=0.25	<0.05 <0.05
Gamma Statistic	0.25	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1996 valid cases: 453		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.16	<0.05
Spearman correlation coefficient	0.17	<0.05
Kendall's Tau	0.15	<0.05
Somers' D Statistics	X/Y=0.26 Y/X=0.08	<0.05 <0.05
Gamma Statistic	0.28	<0.05

MATRIC POINT SCORE (X) AND MTH101F (ordinal) (Y)

YEAR : 1990, valid cases: 68		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.42	<0.05
Spearman correlation coefficient	0.35	<0.05
Kendall's Tau	0.26	<0.05
Somers' D Statistics	X/Y=0.25 Y/X=0.26	<0.05 <0.05
Gamma Statistic	0.28	<0.05

YEAR : 1991, valid cases: 74		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.35	<0.05
Spearman correlation coefficient	0.29	<0.05
Kendall's Tau	0.21	<0.05
Somers' D Statistics	X/Y=0.20 Y/X=0.22	<0.05 <0.05
Gamma Statistic	0.22	<0.05

YEAR : 1992, valid cases: 88		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.23	<0.05
Spearman correlation coefficient	0.19	<0.05
Kendall's Tau	0.13	<0.05
Somers' D Statistics	X/Y=0.13 Y/X=0.13	<0.05 <0.05
Gamma Statistic	0.14	<0.05

YEAR : 1993, valid cases: 83		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.42	<0.05
Spearman correlation coefficient	0.46	<0.05
Kendall's Tau	0.33	<0.05
Somers' D Statistics	X/Y=0.31 Y/X=0.34	<0.05 <0.05
Gamma Statistic	0.35	<0.05

YEAR : 1994, valid cases: 115		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.45	<0.05
Spearman correlation coefficient	0.45	<0.05
Kendall's Tau	0.32	<0.05
Somers' D Statistics	X/Y=0.31 Y/X=0.33	<0.05 <0.05
Gamma Statistic	0.34	<0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1995 valid cases: 453		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.36	<0.05
Spearman correlation coefficient	0.36	<0.05
Kendall's Tau	0.26	<0.05
Somers' D Statistics	X/Y=0.25 Y/X=0.26	<0.05 <0.05
Gamma Statistic	0.27	<0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1995 valid cases: 453		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.21	<0.05
Spearman correlation coefficient	0.20	<0.05
Kendall's Tau	0.17	<0.05
Somers' D Statistics	X/Y=0.31 Y/X=0.10	<0.05 <0.05
Gamma Statistic	0.34	<0.05

MATRIC HIGHER-GRADE MATHS (X) AND MTH101F (ordinal) (Y)

YEAR : 1990, valid cases: 64		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.37	<0.05
Spearman correlation coefficient	0.44	<0.05
Kendall's Tau	0.34	<0.05
Somers' D Statistics	X/Y=0.30 Y/X=0.39	<0.05 <0.05
Gamma Statistic	0.41	<0.05

YEAR : 1991, valid cases: 71		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.54	<0.05
Spearman correlation coefficient	0.52	<0.05
Kendall's Tau	0.43	<0.05
Somers' D Statistics	X/Y=0.38 Y/X=0.48	<0.05 <0.05
Gamma Statistic	0.50	<0.05

YEAR : 1992, valid cases: 79		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.30	<0.05
Spearman correlation coefficient	0.35	<0.05
Kendall's Tau	0.27	<0.05
Somers' D Statistics	X/Y=0.23 Y/X=0.32	<0.05 <0.05
Gamma Statistic	0.33	<0.05

YEAR : 1993, valid cases: 78		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.33	<0.05
Spearman correlation coefficient	0.36	<0.05
Kendall's Tau	0.27	<0.05
Somers' D Statistics	X/Y=0.23 Y/X=0.32	<0.05 <0.05
Gamma Statistic	0.32	<0.05

YEAR : 1994, valid cases: 101		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.39	<0.05
Spearman correlation coefficient	0.37	<0.05
Kendall's Tau	0.29	<0.05
Somers' D Statistics	X/Y=0.25 Y/X=0.33	<0.05 <0.05
Gamma Statistic	0.34	<0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1995 valid cases: 417		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.36	<0.05
Spearman correlation coefficient	0.39	<0.05
Kendall's Tau	0.30	<0.05
Somers' D Statistics	X/Y=0.26 Y/X=0.35	<0.05 <0.05
Gamma Statistic	0.36	<0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1995 valid cases: 417		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.17	<0.05
Spearman correlation coefficient	0.16	<0.05
Kendall's Tau	0.14	<0.05
Somers' D Statistics	X/Y=0.24 Y/X=0.09	<0.05 <0.05
Gamma Statistic	0.33	<0.05

In the preceding tables we observe:

- In all the years, except for 1990 and 1991 for the faculty point score, the measures have been found to be *statistically significant*.

Therefore, we may say that the incoming first year student with either the higher (*lower*) faculty point score or the higher (*lower*) matric point score or the higher (*lower*) matric maths symbol are more likely (*less likely*) to achieve the higher (*lower*) percentage for MTH101F, respectively.

- The measures are of comparable magnitudes. Out of the investigated factors, none appears to be more significantly correlated with the response than the others. There may however be other factors, not examined here, which are more closely related with the response.

Examining the dendrograms we note:

- In CONFIRM the *matric point score* emerges as the best predictor of the actual percentage obtained for MTH101F.

For the students with the matric point scores of 36 and less, the White and the Black students have a higher mean percentage (56.89%) than the Coloured and the Indian students (48.00%). Further within the combined White and Black student group, those registered in the years 1990 and 1991 have a lower percentage mark (51.98%) than those registered in 1992, 1993 and 1994 (59.33%).

Also for the students with the matric point scores between 37 and 39 those registered in 1993 have a higher mean percentage (67.81%) than those registered in the earlier years (58.21%).

- In CATFIRM the *colour* variable splits most significantly the entire data set and into two groups, namely the White and Black categories combined and the Missing Information (node 2) and the Coloured and Indian categories combined (node 3). The former group exhibits a higher passing rate (91.0%) than the latter (66.7%).

Dendrogram of CONFIRM: ordinal response analysis: mth101f

split var	1		
P val (Bonf P)	453		
	56.7903		
	12.7558		
	16.0000		
	89.0000		
levels	matpoint		
Node	2.16E-12; (3.24E-11)		
	012345	678	9
	2	3	4
	284	133	36
	53.5775	60.0902	69.9444
	12.1375	11.6108	9.9282
	16.0000	29.0000	50.0000
	84.0000	89.0000	89.0000
	colour	year	
	1.15E-06; (1.73E-05)	0.0324; (0.4542)	
	1c	bw?	0142
	5	6	7
	106	178	107
	48.0094	56.8933	58.2150
	12.8155	10.4159	11.1601
	16.0000	20.0000	29.0000
	80.0000	84.0000	85.0000
	year		
	8.44E-03; (0.1266)		
	01	342	
	9	10	
	59	119	
	51.9831	59.3277	
	9.4695	10.0314	
	20.0000	30.0000	
	67.0000	84.0000	

Dendrogram of CATFIRM: binary response analysis: mth101f

split var	1	
P val (Bonf P)	17.9%	
	82.1%	
	453	
levels	colour	
Node	1.00E-07; (1.40E-06)	
	?wb	ci
	2	3
	9.0%	33.3%
	91.0%	66.7%
	288	165

b2) MAM104F

For MAM104F the measures of association were calculated for the years 1995-1997 and the combined years field.

The results are presented in the tables below.

FACULTY POINT SCORE (X) AND MAM104F (ordinal) (Y)

YEAR : 1995, valid cases: 128		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.41	<0.05
Spearman correlation coefficient	0.34	<0.05
Kendall's Tau	0.25	<0.05
Somers' D Statistics	X/Y=0.27 Y/X=0.25	<0.05 <0.05
Gamma Statistic	0.27	<0.05

YEAR : 1996, valid cases: 73		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.48	<0.05
Spearman correlation coefficient	0.48	<0.05
Kendall's Tau	0.35	<0.05
Somers' D Statistics	X/Y=0.34 Y/X=0.36	<0.05 <0.05
Gamma Statistic	0.37	<0.05

YEAR : 1997, valid cases: 74		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.60	<0.05
Spearman correlation coefficient	0.60	<0.05
Kendall's Tau	0.45	<0.05
Somers' D Statistics	X/Y=0.44 Y/X=0.45	<0.05 <0.05
Gamma Statistic	0.46	<0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1995-1997 valid cases: 306		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.47	<0.05
Spearman correlation coefficient	0.43	<0.05
Kendall's Tau	0.31	<0.05
Somers' D Statistics	X/Y=0.31 Y/X=0.32	<0.05 <0.05
Gamma Statistic	0.33	<0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1995-1997 valid cases: 306		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.30	<0.05
Spearman correlation coefficient	0.30	<0.05
Kendall's Tau	0.25	<0.05
Somers' D Statistics	X/Y=0.40 Y/X=0.16	<0.05 <0.05
Gamma Statistic	0.42	<0.05

MATRIC POINT SCORE (X) AND MAM104F (ordinal) (Y)

YEAR : 1995, valid cases: 128		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.38	<0.05
Spearman correlation coefficient	0.31	<0.05
Kendall's Tau	0.23	<0.05
Somers' D Statistics	X/Y=0.23 Y/X=0.24	<0.05 <0.05
Gamma Statistic	0.25	<0.05

YEAR : 1996, valid cases: 73		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.47	<0.05
Spearman correlation coefficient	0.51	<0.05
Kendall's Tau	0.37	<0.05
Somers' D Statistics	X/Y=0.36 Y/X=0.39	<0.05 <0.05
Gamma Statistic	0.40	<0.05

YEAR : 1997, valid cases: 74		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.55	<0.05
Spearman correlation coefficient	0.55	<0.05
Kendall's Tau	0.40	<0.05
Somers' D Statistics	X/Y=0.39 Y/X=0.41	<0.05 <0.05
Gamma Statistic	0.42	<0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1995-1997 valid cases: 306		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.43	<0.05
Spearman correlation coefficient	0.40	<0.05
Kendall's Tau	0.29	<0.05
Somers' D Statistics	X/Y=0.29 Y/X=0.30	<0.05 <0.05
Gamma Statistic	0.31	<0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1995-1997 valid cases: 306		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.30	<0.05
Spearman correlation coefficient	0.28	<0.05
Kendall's Tau	0.24	<0.05
Somers' D Statistics	X/Y=0.38 Y/X=0.15	<0.05 <0.05
Gamma Statistic	0.41	<0.05

MATRIC HIGHER-GRADE MATHS (X) AND MAM104F (ordinal) (Y)

YEAR : 1995, valid cases: 119		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.51	<0.05
Spearman correlation coefficient	0.49	<0.05
Kendall's Tau	0.40	<0.05
Somers' D Statistics	X/Y=0.35	<0.05
	Y/X=0.45	<0.05
Gamma Statistic	0.49	<0.05

YEAR : 1996, valid cases: 62		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.30	<0.05
Spearman correlation coefficient	0.29	<0.05
Kendall's Tau	0.23	<0.05
Somers' D Statistics	X/Y=0.20	<0.05
	Y/X=0.26	<0.05
Gamma Statistic	0.29	<0.05

YEAR : 1997, valid cases: 69		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.56	<0.05
Spearman correlation coefficient	0.60	<0.05
Kendall's Tau	0.48	<0.05
Somers' D Statistics	X/Y=0.43	<0.05
	Y/X=0.53	<0.05
Gamma Statistic	0.55	<0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1995-1997 valid cases: 274		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.47	<0.05
Spearman correlation coefficient	0.45	<0.05
Kendall's Tau	0.36	<0.05
Somers' D Statistics	X/Y=0.31 Y/X=0.40	<0.05 <0.05
Gamma Statistic	0.42	<0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1995-1997 valid cases: 274		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.30	<0.05
Spearman correlation coefficient	0.30	<0.05
Kendall's Tau	0.27	<0.05
Somers' D Statistics	X/Y=0.41 Y/X=0.18	<0.05 <0.05
Gamma Statistic	0.54	<0.05

In the preceding tables we observe:

- All the measures are *statistically significant*.

Therefore, we may conclude that the incoming first year student with either the higher (*lower*) faculty point score or the higher (*lower*) matric point score or the higher (*lower*) matric higher-grade maths symbol are more likely (*less likely*) to achieve higher (*lower*) percentage for the first year half-course MAM104F, respectively.

- The magnitudes of the measures are of comparable magnitudes. There is no evidence suggesting that any of the investigated variables may be more highly correlated with the academic performance in MAM104F than the others.

Examining the dendrograms we observe:

- In both the ordinal and binary response analyses (CONFIRM and CATFIRM outputs) the *matric higher-grade maths* emerges as the best performance predictor.

It appears that the students with the higher matric maths symbols are more likely to obtain the higher actual percentage marks for MAM104F, and that more of these students pass than fail the course.

- Additionally, in the ordinal response analysis, for the students with A for the matric maths on higher grade, the symbol for the *matric higher-grade physical science* emerges as a further determinant of the actual percentage mark. Thus, the students with A for the matric physical science have higher mean percentage marks (82.55%) than those with B or C for the matric physical science (63.00%).

On the other hand, for the students with C or D for the matric maths on higher grade, the mean percentage marks are lower for those from the predominately disadvantaged educational departments (46.14%) than from the privileged educational departments (56.76%).

Comparison between the two half year courses may suggest that MAM104F (course after the departmental amalgamation) was better structured and adapted to the needs of the combined set of students than MTH101F. Performance in MAM104F may be directly associated with the academic achievement in the matric higher-grade maths.

Dendrogram of CONFIRM: ordinal response analysis: mam104f

split var	1			
P val (Bonf P)	306			
	53.2418			
	14.9844			
	10.0000			
levels	93.0000			
Node	1			
N	math			
Xbar	1.23E-15; (1.73E-14)			
SD				
Min				
Max				
	?fe	dc	b	a
	2	3	4	5
	63	173	53	17
	42.2857	52.9538	60.7547	73.3529
	14.2188	12.9914	11.5341	12.6389
	10.0000	16.0000	21.0000	51.0000
	70.0000	86.0000	82.0000	93.0000
		education		psch
		0.0177; (0.2474)		0.0396; (0.4357)
	dfnicbrk	eghiam	cb	a
	6	7	8	9
	62	111	8	9
	46.1452	56.7568	63.0000	82.5556
	13.4492	11.0850	6.1179	9.1803
	16.0000	21.0000	51.0000	63.0000
	72.0000	86.0000	72.0000	93.0000

Dendrogram of CATFIRM: binary response analysis: mam104f

split var	1		
P val (Bonf P)	25.2%		
	74.8%		
	306		
levels			
Node	math		
% fail	8.73E-08; (1.13E-06)		
% pass			
N			
	?fe	d	cba
	2	3	4
	55.6%	26.3%	11.5%
	44.4%	73.7%	88.5%
	63	95	148

5.2.2 FIRST YEAR STATISTICS COURSES: SMS100W and STA100S

a) SMS100W

For SMS100W the measures of association were calculated for the years 1990-1997 and for the combined years field.

The results are displayed in the tables below.

FACULTY POINT SCORE (X) AND SMS100W (ordinal) (Y)

YEAR : 1990, valid cases: 38		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.78	<0.05
Spearman correlation coefficient	0.82	<0.05
Kendall's Tau	0.63	<0.05
Somers' D Statistics	X/Y=0.63 Y/X=0.64	<0.05 <0.05
Gamma Statistic	0.65	<0.05

YEAR : 1991, valid cases: 42		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.75	<0.05
Spearman correlation coefficient	0.70	<0.05
Kendall's Tau	0.53	<0.05
Somers' D Statistics	X/Y=0.53 Y/X=0.54	<0.05 <0.05
Gamma Statistic	0.55	<0.05

YEAR : 1992, valid cases: 27		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.67	<0.05
Spearman correlation coefficient	0.73	<0.05
Kendall's Tau	0.55	<0.05
Somers' D Statistics	X/Y=0.54 Y/X=0.56	<0.05 <0.05
Gamma Statistic	0.57	<0.05

YEAR : 1993, valid cases: 29		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.57	<0.05
Spearman correlation coefficient	0.54	<0.05
Kendall's Tau	0.42	<0.05
Somers' D Statistics	X/Y=0.41 Y/X=0.43	<0.05 <0.05
Gamma Statistic	0.43	<0.05

YEAR : 1994, valid cases: 28		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.66	<0.05
Spearman correlation coefficient	0.60	<0.05
Kendall's Tau	0.46	<0.05
Somers' D Statistics	X/Y=0.46 Y/X=0.45	<0.05 <0.05
Gamma Statistic	0.52	<0.05

YEAR : 1995, valid cases: 35		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.73	<0.05
Spearman correlation coefficient	0.74	<0.05
Kendall's Tau	0.58	<0.05
Somers' D Statistics	X/Y=0.57 Y/X=0.59	<0.05 <0.05
Gamma Statistic	0.59	<0.05

YEAR : 1996, valid cases:		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.55	<0.05
Spearman correlation coefficient	0.55	<0.05
Kendall's Tau	0.41	<0.05
Somers' D Statistics	X/Y=0.40 Y/X=0.42	<0.05 <0.05
Gamma Statistic	0.43	<0.05

YEAR : 1997, valid cases: 16		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.31	<0.05
Spearman correlation coefficient	0.28	<0.05
Kendall's Tau	0.20	<0.05
Somers' D Statistics	X/Y=0.20 Y/X=0.20	<0.05 <0.05
Gamma Statistic	0.21	<0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 257		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.62	<0.05
Spearman correlation coefficient	0.63	<0.05
Kendall's Tau	0.46	<0.05
Somers' D Statistics	X/Y=0.46 Y/X=0.47	<0.05 <0.05
Gamma Statistic	0.48	<0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 257		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.37	<0.05
Spearman correlation coefficient	0.38	<0.05
Kendall's Tau	0.32	<0.05
Somers' D Statistics	X/Y=0.50 Y/X=0.21	<0.05 <0.05
Gamma Statistic	0.52	<0.05

MATRIC POINT SCORE (X) AND SMS100W (ordinal) (Y)

YEAR : 1990, valid cases: 38		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.83	<0.05
Spearman correlation coefficient	0.84	<0.05
Kendall's Tau	0.67	<0.05
Somers' D Statistics	X/Y=0.66 Y/X=0.69	<0.05 <0.05
Gamma Statistic	0.70	<0.05

YEAR : 1991, valid cases: 42		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.75	<0.05
Spearman correlation coefficient	0.72	<0.05
Kendall's Tau	0.56	<0.05
Somers' D Statistics	X/Y=0.54 Y/X=0.57	<0.05 <0.05
Gamma Statistic	0.58	<0.05

YEAR : 1992, valid cases: 27		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.68	<0.05
Spearman correlation coefficient	0.72	<0.05
Kendall's Tau	0.54	<0.05
Somers' D Statistics	X/Y=0.53 Y/X=0.55	<0.05 <0.05
Gamma Statistic	0.56	<0.05

YEAR : 1993, valid cases: 29		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.52	<0.05
Spearman correlation coefficient	0.51	<0.05
Kendall's Tau	0.37	<0.05
Somers' D Statistics	X/Y=0.36 Y/X=0.38	<0.05 <0.05
Gamma Statistic	0.38	<0.05

YEAR : 1994, valid cases: 28		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.69	<0.05
Spearman correlation coefficient	0.62	<0.05
Kendall's Tau	0.49	<0.05
Somers' D Statistics	X/Y=0.48 Y/X=0.50	<0.05 <0.05
Gamma Statistic	0.52	<0.05

YEAR : 1996, valid cases: 35		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.72	<0.05
Spearman correlation coefficient	0.76	<0.05
Kendall's Tau	0.59	<0.05
Somers' D Statistics	X/Y=0.58 Y/X=0.61	<0.05 <0.05
Gamma Statistic	0.62	<0.05

YEAR : 1996, valid cases: 29		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.48	<0.05
Spearman correlation coefficient	0.50	<0.05
Kendall's Tau	0.37	<0.05
Somers' D Statistics	X/Y=0.36 Y/X=0.37	<0.05 <0.05
Gamma Statistic	0.38	<0.05

YEAR : 1997, valid cases: 16		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.34	<0.05
Spearman correlation coefficient	0.41	<0.05
Kendall's Tau	0.33	<0.05
Somers' D Statistics	X/Y=0.32 Y/X=0.34	<0.05 <0.05
Gamma Statistic	0.35	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 257		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.62	<0.05
Spearman correlation coefficient	0.63	<0.05
Kendall's Tau	0.47	<0.05
Somers' D Statistics	X/Y=0.46 Y/X=0.48	<0.05 <0.05
Gamma Statistic	0.49	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 257		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.37	<0.05
Spearman correlation coefficient	0.39	<0.05
Kendall's Tau	0.33	<0.05
Somers' D Statistics	X/Y=0.51 Y/X=0.21	<0.05 <0.05
Gamma Statistic	0.54	<0.05

MATRIC HIGHER-GRADE MATHS (X) AND SMS100W (ordinal) (Y)

YEAR : 1990, valid cases: 38		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.72	<0.05
Spearman correlation coefficient	0.73	<0.05
Kendall's Tau	0.59	<0.05
Somers' D Statistics	X/Y=0.50 Y/X=0.70	<0.05 <0.05
Gamma Statistic	0.71	<0.05

YEAR : 1991, valid cases: 42		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.56	<0.05
Spearman correlation coefficient	0.57	<0.05
Kendall's Tau	0.45	<0.05
Somers' D Statistics	X/Y=0.38 Y/X=0.52	<0.05 <0.05
Gamma Statistic	0.53	<0.05

YEAR : 1992, valid cases: 27		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.48	<0.05
Spearman correlation coefficient	0.56	<0.05
Kendall's Tau	0.42	<0.05
Somers' D Statistics	X/Y=0.34 Y/X=0.51	<0.05 <0.05
Gamma Statistic	0.53	<0.05

YEAR : 1993, valid cases: 28		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.73	<0.05
Spearman correlation coefficient	0.73	<0.05
Kendall's Tau	0.61	<0.05
Somers' D Statistics	X/Y=0.54 Y/X=0.68	<0.05 <0.05
Gamma Statistic	0.69	<0.05

YEAR : 1994, valid cases: 28		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.50	<0.05
Spearman correlation coefficient	0.52	<0.05
Kendall's Tau	0.42	<0.05
Somers' D Statistics	X/Y=0.36 Y/X=0.50	<0.05 <0.05
Gamma Statistic	0.52	<0.05

YEAR : 1995, valid cases: 35		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.63	<0.05
Spearman correlation coefficient	0.61	<0.05
Kendall's Tau	0.49	<0.05
Somers' D Statistics	X/Y=0.42 Y/X=0.56	<0.05 <0.05
Gamma Statistic	0.57	<0.05

YEAR : 1996, valid cases: 29		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.45	<0.05
Spearman correlation coefficient	0.48	<0.05
Kendall's Tau	0.38	<0.05
Somers' D Statistics	X/Y=0.34 Y/X=0.43	<0.05 <0.05
Gamma Statistic	0.44	<0.05

YEAR : 1997, valid cases: 16		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.31	<0.05
Spearman correlation coefficient	0.29	<0.05
Kendall's Tau	0.19	<0.05
Somers' D Statistics	X/Y=0.17 Y/X=0.22	<0.05 <0.05
Gamma Statistic	0.23	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 253		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.57	<0.05
Spearman correlation coefficient	0.57	<0.05
Kendall's Tau	0.45	<0.05
Somers' D Statistics	X/Y=0.39 Y/X=0.52	<0.05 <0.05
Gamma Statistic	0.53	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 253		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.37	<0.05
Spearman correlation coefficient	0.37	<0.05
Kendall's Tau	0.34	<0.05
Somers' D Statistics	X/Y=0.46 Y/X=0.25	<0.05 <0.05
Gamma Statistic	0.58	<0.05

In the preceding tables we observe:

- All the measures are *statistically significant*.

Therefore, we may conclude that the incoming first year students with either the higher (*lower*) faculty point score or the higher (*lower*) matric point score or the higher (*lower*) matric higher-grade maths symbol are more likely (*less likely*) to achieve the higher (*lower*) percentage marks for SMS100W, respectively.

- Generally, the magnitudes of the measures are comparable magnitudes with those involving the matric higher-grade maths being marginally but not meaningfully lower than those involving both types of the point scores.

Examining the dendrograms we observe:

- In ordinal response analysis, the *faculty point score* has been selected as the best predictor of the percentage marks obtained by the students for SMS100W.

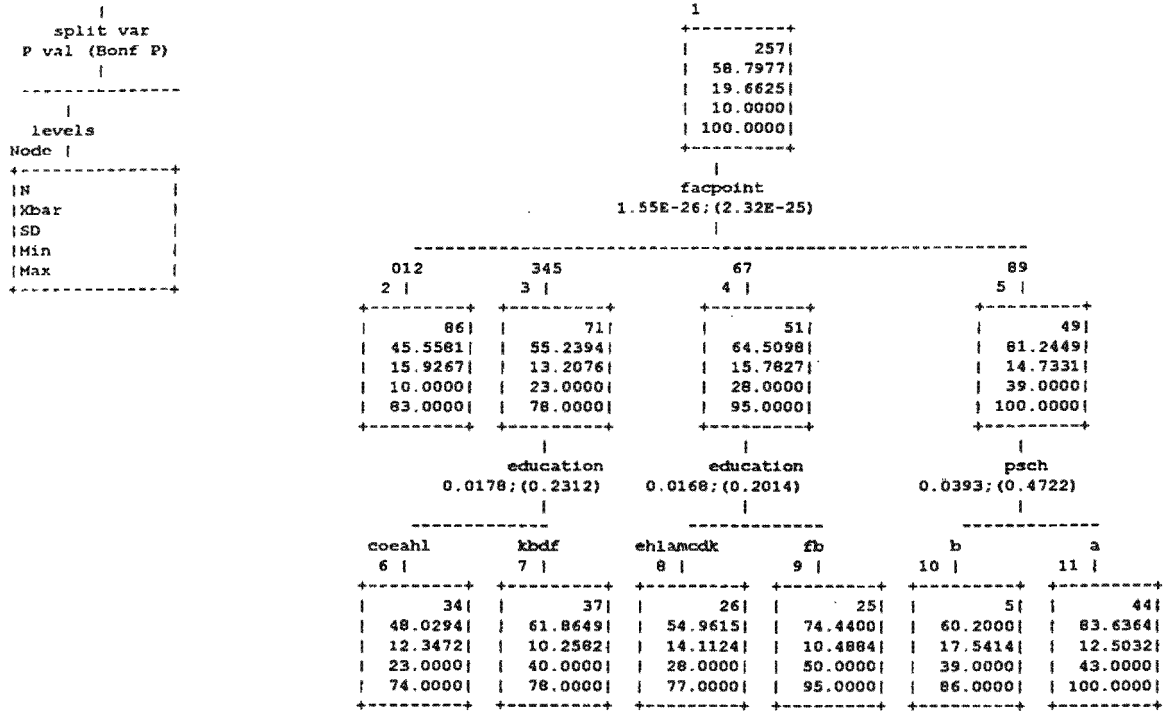
Additionally, for the students with the faculty point scores between 51 and 57 inclusive (nodes 3 and 4), the students from the historically disadvantaged educational departments obtain on average lower marks (48.03% in node 3 and 61.86% in node 4) than those from the privileged departments (54.96% in node 3 and 74.44% in node 4). We observe again that the students with higher faculty point scores obtained higher percentages than those with lower faculty point scores.

Further, for the students with the faculty point scores of 62 and more, those with A for the *matric higher-grade physical science* obtained a higher percentage (83.63%) than those with B for the matric subject (60.20%).

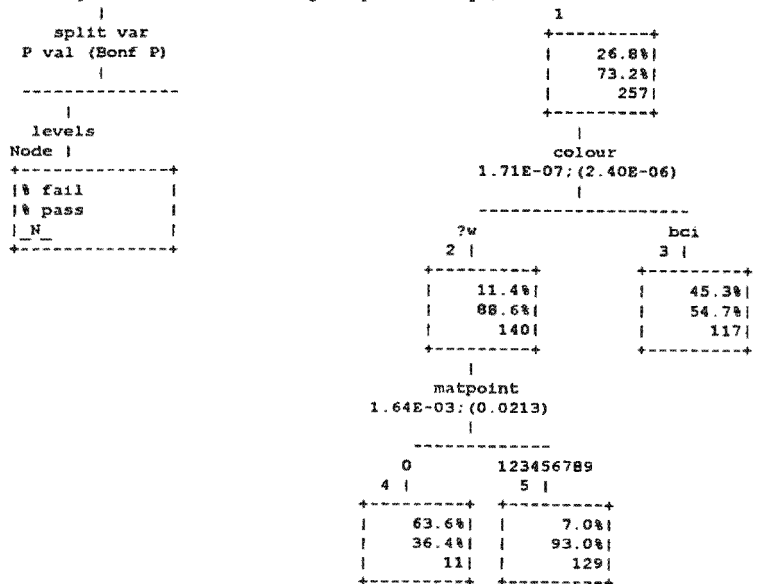
- In contrast, in the binary response analysis, the *colour* variable splits the data most significantly and into the White and the non-White student groups. The proportion of the White students passing SMS100W is higher than that of the non-White students.

The *matric point score* may be further use to predict pass or failure for the White students. The proportion of students with the matric point score of 33 and less who fail is substantially higher (63.6%) than the proportion of the students with the matric point score of 35 and more who fail (7.0%).

Dendrogram of CONFIRM: ordinal response analysis: sms100w



Dendrogram of CATFIRM: binary response analysis: sms100w



b) STA100S

For STA100S the measures of association were calculated for each of the years 1990-1997 and the combined years field.

The results are presented in the tables below.

FACULTY POINT SCORE (X) AND STA100S (ordinal) (Y)

YEAR : 1990, valid cases: 49		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.71	<0.05
Spearman correlation coefficient	0.71	<0.05
Kendall's Tau	0.52	<0.05
Somers' D Statistics	X/Y=0.52 Y/X=0.52	<0.05
Gamma Statistic	0.53	<0.05

YEAR : 1991, valid cases: 52		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.20	<0.05
Spearman correlation coefficient	0.17	<0.05
Kendall's Tau	0.12	<0.05
Somers' D Statistics	X/Y=0.12 Y/X=0.12	<0.05
Gamma Statistic	0.13	<0.05

YEAR : 1992, valid cases: 54		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.54	<0.05
Spearman correlation coefficient	0.46	<0.05
Kendall's Tau	0.36	<0.05
Somers' D Statistics	X/Y=0.35 Y/X=0.36	<0.05
Gamma Statistic	0.37	<0.05

YEAR : 1993, valid cases: 39		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.28	<0.05
Spearman correlation coefficient	0.41	<0.05
Kendall's Tau	0.28	<0.05
Somers' D Statistics	X/Y=0.28 Y/X=0.28	<0.05
Gamma Statistic	0.29	<0.05

YEAR : 1994, valid cases: 62		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.30	<0.05
Spearman correlation coefficient	0.31	<0.05
Kendall's Tau	0.22	<0.05
Somers' D Statistics	X/Y=0.21 Y/X=0.22	<0.05 <0.05
Gamma Statistic	0.23	<0.05

YEAR : 1995, valid cases: 76		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.36	<0.05
Spearman correlation coefficient	0.38	<0.05
Kendall's Tau	0.27	<0.05
Somers' D Statistics	X/Y=0.26 Y/X=0.27	<0.05 <0.05
Gamma Statistic	0.28	<0.05

YEAR : 1996 valid cases: 59		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.68	<0.05
Spearman correlation coefficient	0.66	<0.05
Kendall's Tau	0.49	<0.05
Somers' D Statistics	X/Y=0.49 Y/X=0.50	<0.05 <0.05
Gamma Statistic	0.51	<0.05

YEAR : 1997, valid cases: 84		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.62	<0.05
Spearman correlation coefficient	0.60	<0.05
Kendall's Tau	0.44	<0.05
Somers' D Statistics	X/Y=0.44 Y/X=0.45	<0.05 <0.05
Gamma Statistic	0.46	<0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 631		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.53	<0.05
Spearman correlation coefficient	0.51	<0.05
Kendall's Tau	0.37	<0.05
Somers' D Statistics	X/Y=0.36 Y/X=0.37	<0.05 <0.05
Gamma Statistic	0.38	<0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 631		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.38	<0.05
Spearman correlation coefficient	0.37	<0.05
Kendall's Tau	0.31	<0.05
Somers' D Statistics	X/Y=0.50 Y/X=0.19	<0.05 <0.05
Gamma Statistic	0.52	<0.05

MATRIC POINT SCORE (X) AND STA100S (ordinal) (Y)

YEAR : 1990, valid cases: 49		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.67	<0.05
Spearman correlation coefficient	0.65	<0.05
Kendall's Tau	0.48	<0.05
Somers' D Statistics	X/Y=0.46 Y/X=0.49	<0.05 <0.05
Gamma Statistic	0.50	<0.05

YEAR : 1991, valid cases: 52		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.45	<0.05
Spearman correlation coefficient	0.41	<0.05
Kendall's Tau	0.30	<0.05
Somers' D Statistics	X/Y=0.29 Y/X=0.31	<0.05 <0.05
Gamma Statistic	0.32	<0.05

YEAR : 1992, valid cases: 51		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.53	<0.05
Spearman correlation coefficient	0.42	<0.05
Kendall's Tau	0.32	<0.05
Somers' D Statistics	X/Y=0.31 Y/X=0.33	<0.05 <0.05
Gamma Statistic	0.34	<0.05

YEAR : 1993 valid cases: 39		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.27	<0.05
Spearman correlation coefficient	0.33	<0.05
Kendall's Tau	0.23	<0.05
Somers' D Statistics	X/Y=0.22 Y/X=0.24	<0.05 <0.05
Gamma Statistic	0.24	<0.05

YEAR : 1994 valid cases: 62		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.33	<0.05
Spearman correlation coefficient	0.36	<0.05
Kendall's Tau	0.25	<0.05
Somers' D Statistics	X/Y=0.24 Y/X=0.25	<0.05 <0.05
Gamma Statistic	0.26	<0.05

YEAR : 1995, valid cases: 76		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.36	<0.05
Spearman correlation coefficient	0.37	<0.05
Kendall's Tau	0.26	<0.05
Somers' D Statistics	X/Y=0.25 Y/X=0.26	<0.05 <0.05
Gamma Statistic	0.27	<0.05

YEAR : 1996, valid cases: 59		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.69	<0.05
Spearman correlation coefficient	0.68	<0.05
Kendall's Tau	0.52	<0.05
Somers' D Statistics	X/Y=0.51 Y/X=0.53	<0.05 <0.05
Gamma Statistic	0.54	<0.05

YEAR : 1997, valid cases: 84		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient		
Spearman correlation coefficient	0.64	<0.05
Kendall's Tau	0.45	<0.05
Somers' D Statistics	X/Y=0.44 Y/X=0.45	<0.05 <0.05
Gamma Statistic	0.46	<0.05

ORDINAL RESPONSE ANALYSIS		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 631		
Pearson correlation coefficient		
Spearman correlation coefficient	0.52	<0.05
Kendall's Tau	0.38	<0.05
Somers' D Statistics	X/Y=0.37 Y/X=0.39	<0.05 <0.05
Gamma Statistic	0.40	<0.05

BINARY RESPONSE ANALYSIS		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 631		
Pearson correlation coefficient	0.40	<0.05
Spearman correlation coefficient	0.39	<0.05
Kendall's Tau	0.32	<0.05
Somers' D Statistics	X/Y=0.52 Y/X=0.20	<0.05 <0.05
Gamma Statistic	0.56	<0.05

MATRIC HIGHER-GRADE MATHS (X) AND STA100S (ordinal) (Y)

YEAR : 1990, valid cases: 49		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.59	<0.05
Spearman correlation coefficient	0.60	<0.05
Kendall's Tau	0.48	<0.05
Somers' D Statistics	X/Y=0.42 Y/X=0.54	<0.05 <0.05
Gamma Statistic	0.55	<0.05

YEAR : 1991, valid cases: 52		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.38	<0.05
Spearman correlation coefficient	0.40	<0.05
Kendall's Tau	0.31	<0.05
Somers' D Statistics	X/Y=0.28 Y/X=0.34	<0.05 <0.05
Gamma Statistic	0.36	<0.05

YEAR : 1992, valid cases: 46		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.37	<0.05
Spearman correlation coefficient	0.35	<0.05
Kendall's Tau	0.28	<0.05
Somers' D Statistics	X/Y=0.24 Y/X=0.33	<0.05 <0.05
Gamma Statistic	0.34	<0.05

YEAR : 1993, valid cases: 38		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.23	<0.05
Spearman correlation coefficient	0.28	<0.05
Kendall's Tau	0.20	<0.05
Somers' D Statistics	X/Y=0.18 Y/X=0.23	<0.05 <0.05
Gamma Statistic	0.24	<0.05

YEAR : 1994, valid cases: 59		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.18	<0.05
Spearman correlation coefficient	0.23	<0.05
Kendall's Tau	0.18	<0.05
Somers' D Statistics	X/Y=0.16 Y/X=0.20	<0.05 <0.05
Gamma Statistic	0.21	<0.05

YEAR : 1995, valid cases: 74		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.35	<0.05
Spearman correlation coefficient	0.38	<0.05
Kendall's Tau	0.29	<0.05
Somers' D Statistics	X/Y=0.26 Y/X=0.33	<0.05 <0.05
Gamma Statistic	0.34	<0.05

YEAR : 1996, valid cases: 56		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.49	<0.05
Spearman correlation coefficient	0.52	<0.05
Kendall's Tau	0.39	<0.05
Somers' D Statistics	X/Y=0.35 Y/X=0.44	<0.05 <0.05
Gamma Statistic	0.45	<0.05

YEAR : 1997, valid cases: 84		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.59	<0.05
Spearman correlation coefficient	0.59	<0.05
Kendall's Tau	0.46	<0.05
Somers' D Statistics	X/Y=0.41 Y/X=0.52	<0.05 <0.05
Gamma Statistic	0.54	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 609		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.42	<0.05
Spearman correlation coefficient	0.41	<0.05
Kendall's Tau	0.32	<0.05
Somers' D Statistics	X/Y=0.28 Y/X=0.35	<0.05 <0.05
Gamma Statistic	0.36	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 609		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.28	<0.05
Spearman correlation coefficient	0.29	<0.05
Kendall's Tau	0.26	<0.05
Somers' D Statistics	X/Y=0.38 Y/X=0.18	<0.05 <0.05
Gamma Statistic	0.48	<0.05

In the preceding tables we see that:

- All the measures have been found to be *statistically significant*.

We may conclude that the incoming first year students with either the higher (*lower*) faculty point score or the higher (*lower*) matric point score or the higher (*lower*) matric higher-grade maths symbol are more likely (*less likely*) to achieve the higher (*lower*) percentage marks for STA100S, respectively.

- Across the different calendar years, the measures are of comparable magnitudes with those involving the matric point score being marginally but not importantly larger.

Examining the dendrograms we observe:

- In both the ordinal and binary response analyses, the *matric point score* splits the data most significantly and into five groups of point scores in CONFIRM and four groups of point scores in CATFIRM. In CONFIRM the point scores form two separate nodes as opposed to CATFIRM where these point scores are combined in one node. It emerges clearly that the students with the higher point scores obtained the higher percentages and that a substantially higher proportion of these students pass the course.
- Additionally in the ordinal response analysis, for the students with the matric point score of 43 and more those with A for the matric higher-grade physical science obtained a higher average percentage (76.87%) than those with C or B for the matric subject (60.93%).

Despite the fact that the statistics syllabi ⁶ for SMS100W and STA100S are essentially the same, it does not appear that a similar hierarchy of factors contribute to academic achievement in these courses.

⁶ In addition the syllabus for SMS100W includes the applied maths and programming modules.

Dendrogram of CONFIRM: ordinal response analysis: stal00s

split var	1				
P val (Bonf P)	631				
	56.0491				
	15.5691				
	11.0000				
	96.0000				
levels					
Node					
	matpoint				
	2.00E-40; (3.01E-39)				
	0	123	4567	8	9
	2	3	4	5	6
	65	202	248	46	70
	42.8462	50.3812	57.5282	64.7826	73.6857
	13.1413	13.9115	12.8988	11.2880	13.3423
	20.0000	11.0000	12.0000	36.0000	39.0000
	71.0000	83.0000	87.0000	96.0000	96.0000
	psch				
	0.0122; (0.1343)				
	cb ?a				
	7				8
	14				56
	60.9286				76.8750
	9.7465				12.2104
	39.0000				50.0000
	74.0000				96.0000

Dendrogram of CATFIRM: binary response analysis: stal00s

split var	1			
P val (Bonf P)	23.8%			
	76.2%			
	631			
levels				
Node				
	matpoint			
	2.78E-18; (3.89E-17)			
	0	123	4567	89
	2	3	4	5
	63.1%	32.7%	16.5%	1.7%
	36.9%	67.3%	83.5%	98.3%
	65	202	248	116
% fail				
% pass				
N				

5.2.3 FIRST YEAR COMPUTER SCIENCE COURSES: CSC105W

For CSS105W the measures of association were calculated for each of the years 1990-1997 and the combined years field.

The results are given in the tables below.

FACULTY POINT SCORE (X) AND CSC105W (ordinal) (Y)

YEAR : 1990, valid cases: 51		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.62	<0.05
Spearman correlation coefficient	0.63	<0.05
Kendall's Tau	0.45	<0.05
Somers' D Statistics	X/Y=0.45	<0.05
	Y/X=0.46	<0.05
Gamma Statistic	0.47	<0.05

YEAR : 1991, valid cases: 56		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.62	<0.05
Spearman correlation coefficient	0.59	<0.05
Kendall's Tau	0.41	<0.05
Somers' D Statistics	X/Y=0.41	<0.05
	Y/X=0.42	<0.05
Gamma Statistic	0.43	<0.05

YEAR : 1992, valid cases: 39		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.61	<0.05
Spearman correlation coefficient	0.56	<0.05
Kendall's Tau	0.40	<0.05
Somers' D Statistics	X/Y=0.39	<0.05
	Y/X=0.40	<0.05
Gamma Statistic	0.41	<0.05

YEAR : 1993, valid cases: 34		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.46	<0.05
Spearman correlation coefficient	0.48	<0.05
Kendall's Tau	0.34	<0.05
Somers' D Statistics	X/Y=0.34	<0.05
	Y/X=0.34	<0.05
Gamma Statistic	0.35	<0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 420		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.54	<0.05
Spearman correlation coefficient	0.53	<0.05
Kendall's Tau	0.38	<0.05
Somers' D Statistics	X/Y=0.38 Y/X=0.39	<0.05 <0.05
Gamma Statistic	0.39	<0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 420		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.31	<0.05
Spearman correlation coefficient	0.32	<0.05
Kendall's Tau	0.27	<0.05
Somers' D Statistics	X/Y=0.48 Y/X=0.15	<0.05 <0.05
Gamma Statistic	0.50	<0.05

YEAR : 1995, valid cases: 57		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.48	<0.05
Spearman correlation coefficient	0.44	<0.05
Kendall's Tau	0.32	<0.05
Somers' D Statistics	X/Y=0.32 Y/X=0.33	<0.05 <0.05
Gamma Statistic	0.34	<0.05

YEAR : 1996, valid cases: 53		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.62	<0.05
Spearman correlation coefficient	0.65	<0.05
Kendall's Tau	0.48	<0.05
Somers' D Statistics	X/Y=0.48 Y/X=0.49	<0.05 <0.05
Gamma Statistic	0.50	<0.05

YEAR : 1997, valid cases: 43		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.38	<0.05
Spearman correlation coefficient	0.37	<0.05
Kendall's Tau	0.27	<0.05
Somers' D Statistics	X/Y=0.26 Y/X=0.28	<0.05 <0.05
Gamma Statistic	0.28	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 420		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.55	<0.05
Spearman correlation coefficient	0.53	<0.05
Kendall's Tau	0.38	<0.05
Somers' D Statistics	X/Y=0.38 Y/X=0.39	<0.05 <0.05
Gamma Statistic	0.40	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 420		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.32	<0.05
Spearman correlation coefficient	0.33	<0.05
Kendall's Tau	0.28	<0.05
Somers' D Statistics	X/Y=0.49 Y/X=0.15	<0.05 <0.05
Gamma Statistic	0.52	<0.05

MATRIC HIGHER-GRADE MATHS AND CSC105W (ordinal) (Y)

YEAR : 1990, valid cases: 51		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.57	<0.05
Spearman correlation coefficient	0.62	<0.05
Kendall's Tau	0.50	<0.05
Somers' D Statistics	X/Y=0.43 Y/X=0.58	<0.05 <0.05
Gamma Statistic	0.59	<0.05

YEAR : 1991, valid cases: 56		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.47	<0.05
Spearman correlation coefficient	0.47	<0.05
Kendall's Tau	0.36	<0.05
Somers' D Statistics	X/Y=0.32 Y/X=0.41	<0.05 <0.05
Gamma Statistic	0.42	<0.05

YEAR : 1992, valid cases: 38		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.40	<0.05
Spearman correlation coefficient	0.43	<0.05
Kendall's Tau	0.32	<0.05
Somers' D Statistics	X/Y=0.28 Y/X=0.37	<0.05 <0.05
Gamma Statistic	0.38	<0.05

YEAR : 1993, valid cases: 33		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.39	<0.05
Spearman correlation coefficient	0.46	<0.05
Kendall's Tau	0.36	<0.05
Somers' D Statistics	X/Y=0.30 Y/X=0.43	<0.05 <0.05
Gamma Statistic	0.45	<0.05

YEAR : 1994, valid cases: 48		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.28	<0.05
Spearman correlation coefficient	0.25	<0.05
Kendall's Tau	0.19	<0.05
Somers' D Statistics	X/Y=0.17 Y/X=0.22	<0.05 <0.05
Gamma Statistic	0.23	<0.05

YEAR : 1995, valid cases: 57		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.32	<0.05
Spearman correlation coefficient	0.33	<0.05
Kendall's Tau	0.26	<0.05
Somers' D Statistics	X/Y=0.23 Y/X=0.30	<0.05
Gamma Statistic	0.31	<0.05

YEAR : 1996, valid cases: 52		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.42	<0.05
Spearman correlation coefficient	0.43	<0.05
Kendall's Tau	0.33	<0.05
Somers' D Statistics	X/Y=0.29 Y/X=0.38	<0.05
Gamma Statistic	0.39	<0.05

YEAR : 1997, valid cases: 43		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.30	<0.05
Spearman correlation coefficient	0.27	<0.05
Kendall's Tau	0.22	<0.05
Somers' D Statistics	X/Y=0.18 Y/X=0.27	<0.05
Gamma Statistic	0.27	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 413		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.41	<0.05
Spearman correlation coefficient	0.41	<0.05
Kendall's Tau	0.31	<0.05
Somers' D Statistics	X/Y=0.27 Y/X=0.36	<0.05
Gamma Statistic	0.37	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 413		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.27	<0.05
Spearman correlation coefficient	0.27	<0.05
Kendall's Tau	0.25	<0.05
Somers' D Statistics	X/Y=0.46 Y/X=0.25	<0.05
Gamma Statistic	0.58	<0.05

In the preceding tables we observe:

- All the measures are *statistically significant*.

Therefore, we may conclude that the incoming first year student with either the higher (*lower*) faculty point score or the higher (*lower*) matric point score or the higher (*lower*) matric higher-grade maths symbol are more likely (*less likely*) to achieve the higher (*lower*) percentage for the first year computer science CSC105W course, respectively.

- The measures are of comparable magnitudes. No evidence emerges suggesting that any of the investigated variables may be more highly correlated with the academic performance in CSC105W than the others.

Examining the dendrograms we observe:

- In the ordinal response analysis (CONFIRM output) the *matric point score* emerges as the best performance predictor of the actual percentage obtained for CSC105W.

For the students with the matric point scores of 46 and more, the *matric higher-grade physical science* may serve as a further performance predictor where students with A exhibit a higher mean percentage mark (77.76%) than those with B and C.

For the students with the matric point score between 42 and 44, the White students obtain on average a higher mark (70.38%) than either the Black, the Coloured or the Indian students (55.29%).

For the students with the matric point scores between 35 and 40, the students from the privileged educational departments obtained higher percentages (61.57%) than those from the disadvantaged departments (49.94%).

- On the other hand, in the binary response analysis (CATFIRM output) the *faculty point score* splits the data most significantly and into two groups of the faculty point scores.

For the students with the faculty point scores of 51 and less, a higher proportion of male (73.5%) than female students (41.7%) passed CSC105W.

Dendrogram of CONFIRM: ordinal response analysis: csc105w

split var	1							
P val (Bonf P)	420							
	61.8929							
	14.5060							
	20.0000							
	93.0000							
levels								
Node								
	matpoint							
	1.59E-28; (2.38E-27)							
	0	12345	67		89			
	2	3	4		5			
	68	185	92		75			
	50.9412	58.2378	65.9565		75.8533			
	11.2182	12.4227	13.0299		11.1060			
	23.0000	20.0000	38.0000		46.0000			
	75.0000	88.0000	92.0000		93.0000			
	education				colour		psch	
	3.15E-04; (4.73E-03)				8.21E-05; (1.15E-03)		2.41E-03; (0.0289)	
	gad	fhceknbjl	cib	w?	cb	?a		
	6	7	8	9	10	11		
	53	132	27	65	8	67		
	49.9434	61.5682	55.2963	70.3846	59.8750	77.7612		
	9.5221	11.9062	7.7401	12.2151	10.5484	9.5839		
	30.0000	20.0000	38.0000	41.0000	46.0000	48.0000		
	73.0000	88.0000	75.0000	92.0000	80.0000	93.0000		

Dendrogram of CATFIRM: binary response analysis: csc105w

split var	1	
P val (Bonf P)	17.9%	
	82.1%	
	420	
levels		
Node		
	facpoint	
	3.10E-09; (4.35E-08)	
	0123	456789
	2	3
	32.8%	6.7%
	67.2%	93.3%
	180	240
	gender	
	0.0242; (0.3393)	
	?m	f
	4	5
	26.4%	58.3%
	73.6%	41.7%
	144	36

5.3. PHYSICAL SCIENCE RELATED SUBJECTS

In this section we discuss the two physics courses, PHY104W and PHY100W and the two chemistry courses, CEM100W and CEM101W offered at first year level by the Departments of Physics and Chemistry in the Faculty of Science at UCT, respectively.

5.3.1 CHEMISTRY FIRST YEAR COURSES: CEM100W and CEM101W

a) CEM100W

For CEM100W the measures of association were calculated for the years 1990-1997 and the combined years field.

The results are displayed in the tables below.

FACULTY POINT SCORE (X) AND CEM100W (ordinal) (Y)

YEAR : 1990, valid cases: 39		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.71	<0.05
Spearman correlation coefficient	0.73	<0.05
Kendall's Tau	0.55	<0.05
Somers' D Statistics	X/Y=0.54 Y/X=0.55	<0.05 <0.05
Gamma Statistic	0.56	<0.05

YEAR : 1991, valid cases: 17		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.45	>0.05
Spearman correlation coefficient	0.34	>0.05
Kendall's Tau	0.29	>0.05
Somers' D Statistics	X/Y=0.29 Y/X=0.29	>0.05 >0.05
Gamma Statistic	0.30	>0.05

YEAR : 1992, valid cases: 14		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.44	>0.05
Spearman correlation coefficient	0.14	>0.05
Kendall's Tau	0.09	>0.05
Somers' D Statistics	X/Y=0.09 Y/X=0.09	>0.05 >0.05
Gamma Statistic	0.09	>0.05

YEAR : 1993, valid cases: 16		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.78	<0.05
Spearman correlation coefficient	0.76	<0.05
Kendall's Tau	0.60	<0.05
Somers' D Statistics	X/Y=0.60 Y/X=0.58	<0.05 <0.05
Gamma Statistic	0.62	<0.05

YEAR : 1994, valid cases: 25		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.23	>0.05
Spearman correlation coefficient	0.25	>0.05
Kendall's Tau	0.20	>0.05
Somers' D Statistics	X/Y=0.20 Y/X=0.21	>0.05 >0.05
Gamma Statistic	0.21	>0.05

YEAR : 1995, valid cases: 24		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.59	<0.05
Spearman correlation coefficient	0.54	<0.05
Kendall's Tau	0.42	<0.05
Somers' D Statistics	X/Y=0.41 Y/X=0.44	<0.05 <0.05
Gamma Statistic	0.44	<0.05

YEAR : 1996, valid cases: 30		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.70	<0.05
Spearman correlation coefficient	0.67	<0.05
Kendall's Tau	0.51	<0.05
Somers' D Statistics	X/Y=0.51 Y/X=0.51	<0.05 <0.05
Gamma Statistic	0.54	<0.05

YEAR : 1997, valid cases: 24		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.55	<0.05
Spearman correlation coefficient	0.50	<0.05
Kendall's Tau	0.39	<0.05
Somers' D Statistics	X/Y=0.39 Y/X=0.40	<0.05 <0.05
Gamma Statistic	0.41	<0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 230		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.60	<0.05
Spearman correlation coefficient	0.59	<0.05
Kendall's Tau	0.43	<0.05
Somers' D Statistics	X/Y=0.43 Y/X=0.43	<0.05 <0.05
Gamma Statistic	0.45	<0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 230		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.48	<0.05
Spearman correlation coefficient	0.48	<0.05
Kendall's Tau	0.40	<0.05
Somers' D Statistics	X/Y=0.59 Y/X=0.27	<0.05 <0.05
Gamma Statistic	0.61	<0.05

MATRIC POINT SCORE (X) AND CEM100W (ordinal) (Y)

YEAR : 1990, valid cases:		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.69	<0.05
Spearman correlation coefficient	0.71	<0.05
Kendall's Tau	0.53	<0.05
Somers' D Statistics	X/Y=0.53 Y/X=0.54	<0.05 <0.05
Gamma Statistic	0.55	<0.05

YEAR : 1991, valid cases: 17		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.42	>0.05
Spearman correlation coefficient	0.29	>0.05
Kendall's Tau	0.22	>0.05
Somers' D Statistics	X/Y=0.22 Y/X=0.23	>0.05 >0.05
Gamma Statistic	0.23	>0.05

YEAR : 1992, valid cases: 14		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.38	>0.05
Spearman correlation coefficient	0.14	>0.05
Kendall's Tau	0.10	>0.05
Somers' D Statistics	X/Y=0.10 Y/X=0.11	>0.05 >0.05
Gamma Statistic	0.11	>0.05

YEAR : 1993, valid cases: 16		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.76	<0.05
Spearman correlation coefficient	0.78	<0.05
Kendall's Tau	0.60	<0.05
Somers' D Statistics	X/Y=0.60 Y/X=0.60	<0.05 <0.05
Gamma Statistic	0.63	<0.05

YEAR : 1994, valid cases: 25		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.30	>0.05
Spearman correlation coefficient	0.28	>0.05
Kendall's Tau	0.23	>0.05
Somers' D Statistics	X/Y=0.23 Y/X=0.23	>0.05 >0.05
Gamma Statistic	0.23	>0.05

YEAR : 1995, valid cases: 24		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.64	<0.05
Spearman correlation coefficient	0.61	<0.05
Kendall's Tau	0.49	<0.05
Somers' D Statistics	X/Y=0.47 Y/X=0.52	<0.05 <0.05
Gamma Statistic	0.53	<0.05

YEAR : 1996, valid cases: 30		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.67	<0.05
Spearman correlation coefficient	0.61	<0.05
Kendall's Tau	0.45	<0.05
Somers' D Statistics	X/Y=0.45 Y/X=0.44	<0.05 <0.05
Gamma Statistic	0.47	<0.05

YEAR : 1997, valid cases: 24		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.46	<0.05
Spearman correlation coefficient	0.46	<0.05
Kendall's Tau	0.37	<0.05
Somers' D Statistics	X/Y=0.36 Y/X=0.37	<0.05 <0.05
Gamma Statistic	0.38	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 230		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.58	<0.05
Spearman correlation coefficient	0.57	<0.05
Kendall's Tau	0.42	<0.05
Somers' D Statistics	X/Y=0.42 Y/X=0.43	<0.05 <0.05
Gamma Statistic	0.44	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 230		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.47	<0.05
Spearman correlation coefficient	0.48	<0.05
Kendall's Tau	0.40	<0.05
Somers' D Statistics	X/Y=0.59 Y/X=0.27	<0.05 <0.05
Gamma Statistic	0.61	<0.05

MATRIC HIGHER-GRADE PHYSICAL SCIENCE (X) AND CEM100W (ordinal) (Y)

YEAR : 1990, valid cases: 39		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.75	<0.05
Spearman correlation coefficient	0.75	<0.05
Kendall's Tau	0.60	<0.05
Somers' D Statistics	X/Y=0.53 Y/X=0.68	<0.05 <0.05
Gamma Statistic	0.69	<0.05

YEAR : 1991, valid cases: 17		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.08	>0.05
Spearman correlation coefficient	0.05	>0.05
Kendall's Tau	0.05	>0.05
Somers' D Statistics	X/Y=0.04 Y/X=0.06	>0.05 >0.05
Gamma Statistic	0.05	>0.05

YEAR : 1992, valid cases: 14		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.61	<0.05
Spearman correlation coefficient	0.57	<0.05
Kendall's Tau	0.44	<0.05
Somers' D Statistics	X/Y=0.40 Y/X=0.50	<0.05 <0.05
Gamma Statistic	0.52	<0.05

YEAR : 1993, valid cases: 16		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.50	<0.05
Spearman correlation coefficient	0.51	<0.05
Kendall's Tau	0.42	<0.05
Somers' D Statistics	X/Y=0.37 Y/X=0.48	<0.05 <0.05
Gamma Statistic	0.51	<0.05

YEAR : 1994, valid cases: 24		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.05	>0.05
Spearman correlation coefficient	0.05	>0.05
Kendall's Tau	0.04	>0.05
Somers' D Statistics	X/Y=0.03 Y/X=0.04	>0.05 >0.05
Gamma Statistic	0.05	>0.05

YEAR : 1996, valid cases: 17		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.17	>0.05
Spearman correlation coefficient	0.13	>0.05
Kendall's Tau	0.09	>0.05
Somers' D Statistics	X/Y=0.08 Y/X=0.12	>0.05 >0.05
Gamma Statistic	0.12	>0.05

YEAR : 1996, valid cases: 23		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.47	<0.05
Spearman correlation coefficient	0.58	<0.05
Kendall's Tau	0.45	<0.05
Somers' D Statistics	X/Y=0.41 Y/X=0.49	<0.05 <0.05
Gamma Statistic	0.50	<0.05

YEAR : 1997, valid cases: 23		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.47	<0.05
Spearman correlation coefficient	0.40	<0.05
Kendall's Tau	0.32	<0.05
Somers' D Statistics	X/Y=0.27 Y/X=0.37	<0.05 <0.05
Gamma Statistic	0.37	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 225		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.52	<0.05
Spearman correlation coefficient	0.51	<0.05
Kendall's Tau	0.39	<0.05
Somers' D Statistics	X/Y=0.35 Y/X=0.44	<0.05 <0.05
Gamma Statistic	0.45	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 225		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.39	<0.05
Spearman correlation coefficient	0.39	<0.05
Kendall's Tau	0.35	<0.05
Somers' D Statistics	X/Y=0.48 Y/X=0.26	<0.05 <0.05
Gamma Statistic	0.59	<0.05

In the preceding tables we observe:

- Most of the measures have emerged as *statistically significant* (except for the years 1991, 1992 and 1994 for associations involving both types of the point scores and for 1991, 1994 and 1995 for associations involving the matric higher-grade physical science; the numbers of students in the respective contingency tables are usually very small).

We may conclude that the incoming first year student with either the higher (*lower*) faculty point score or the higher (*lower*) matric point score or the higher (*lower*) matric higher-grade physical science symbol are more likely (*less likely*) to achieve the higher (*lower*) percentage for CEM100W , respectively.

- The measures involving both types of the point scores are of comparable magnitudes. The measures involving the matric physical science are usually of the smallest magnitudes.

Examining the dendrograms we note:

- In the ordinal response analysis, the *matric higher-grade maths* emerges as the best predictor of the actual percentage obtained for CEM100W. The students with the higher symbols for the matric maths obtained on average the higher percentage marks for the course.
- In the binary response analysis, the *faculty point score* emerges as the best predictor of passing/ failing rate for CEM100W. The higher proportions of students passing the courses correspond with the student groupings with the higher faculty point scores.
- There are no other significant predictors in both analyses.

Dendrogram of CONFIRM: ordinal response analysis: cem100w

split var	1
P val (Bonf P)	230
	54.7652
	13.9955
	24.0000
	93.0000
levels	
Node	
	math
	7.75E-19; (1.16E-17)
	ped cb a
	2 3 4
	64 102 64
	45.4688 52.7941 67.2031
	11.2094 10.2168 12.9519
	24.0000 32.0000 36.0000
	71.0000 83.0000 93.0000

Dendrogram of CATFIRM: binary response analysis: cem100w

split var	1
P val (Bonf P)	31.7%
	68.3%
	230
levels	
Node	
	facpoint
	4.17E-10; (5.83E-09)
	0123 4 5 6789
	2 3 4 5
	56.6% 4.5% 47.6% 6.8%
	43.4% 95.5% 52.4% 93.2%
	99 22 21 88

b) CEM101W

For CEM101W the measures of association were calculated for the years 1990-1997 and the combined years field.

The results are displayed in the tables below.

FACULTY POINT SCORE (X) AND CEM101W (ordinal) (Y)

YEAR : 1990, valid cases: 43		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.68	<0.05
Spearman correlation coefficient	0.63	<0.05
Kendall's Tau	0.47	<0.05
Somers' D Statistics	X/Y=0.47 Y/X=0.48	<0.05 <0.05
Gamma Statistic	0.49	<0.05

YEAR : 1991, valid cases: 56		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.26	<0.05
Spearman correlation coefficient	0.22	<0.05
Kendall's Tau	0.16	<0.05
Somers' D Statistics	X/Y=0.16 Y/X=0.16	<0.05 <0.05
Gamma Statistic	0.17	<0.05

YEAR : 1992, valid cases: 55		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.36	<0.05
Spearman correlation coefficient	0.31	<0.05
Kendall's Tau	0.23	<0.05
Somers' D Statistics	X/Y=0.22 Y/X=0.23	<0.05 <0.05
Gamma Statistic	0.24	<0.05

YEAR : 1993, valid cases: 50		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.38	<0.05
Spearman correlation coefficient	0.39	<0.05
Kendall's Tau	0.28	<0.05
Somers' D Statistics	X/Y=0.27 Y/X=0.28	<0.05 <0.05
Gamma Statistic	0.29	<0.05

YEAR : 1994, valid cases: 74		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.50	<0.05
Spearman correlation coefficient	0.52	<0.05
Kendall's Tau	0.39	<0.05
Somers' D Statistics	X/Y=0.39 Y/X=0.39	<0.05 <0.05
Gamma Statistic	0.41	<0.05

YEAR : 1995, valid cases: 91		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.23	<0.05
Spearman correlation coefficient	0.19	>0.05
Kendall's Tau	0.14	>0.05
Somers' D Statistics	X/Y=0.13 Y/X=0.14	>0.05 >0.05
Gamma Statistic	0.14	>0.05

YEAR : 1996, valid cases: 47		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.64	<0.05
Spearman correlation coefficient	0.55	<0.05
Kendall's Tau	0.41	<0.05
Somers' D Statistics	X/Y=0.40 Y/X=0.42	<0.05 <0.05
Gamma Statistic	0.43	<0.05

YEAR : 1997, valid cases: 66		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.74	<0.05
Spearman correlation coefficient	0.67	<0.05
Kendall's Tau	0.52	<0.05
Somers' D Statistics	X/Y=0.51 Y/X=0.52	<0.05 <0.05
Gamma Statistic	0.54	<0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 557		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.45	<0.05
Spearman correlation coefficient	0.41	<0.05
Kendall's Tau	0.30	<0.05
Somers' D Statistics	X/Y=0.29 Y/X=0.30	<0.05 <0.05
Gamma Statistic	0.31	<0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 557		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.31	<0.05
Spearman correlation coefficient	0.31	<0.05
Kendall's Tau	0.26	<0.05
Somers' D Statistics	X/Y=0.37 Y/X=0.18	<0.05 <0.05
Gamma Statistic	0.39	<0.05

MATRIC POINT SCORE (X) AND CEM101W (ordinal) (Y)

YEAR : 1990, valid cases: 43		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.66	<0.05
Spearman correlation coefficient	0.64	<0.05
Kendall's Tau	0.48	<0.05
Somers' D Statistics	X/Y=0.46 Y/X=0.49	<0.05 <0.05
Gamma Statistic	0.50	<0.05

YEAR : 1991, valid cases: 56		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.47	<0.05
Spearman correlation coefficient	0.41	<0.05
Kendall's Tau	0.29	<0.05
Somers' D Statistics	X/Y=0.28 Y/X=0.29	<0.05 <0.05
Gamma Statistic	0.31	<0.05

YEAR : 1992, valid cases: 55		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.34	<0.05
Spearman correlation coefficient	0.27	<0.05
Kendall's Tau	0.20	<0.05
Somers' D Statistics	X/Y=0.19 Y/X=0.20	<0.05 <0.05
Gamma Statistic	0.21	<0.05

YEAR : 1993, valid cases: 50		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.44	<0.05
Spearman correlation coefficient	0.41	<0.05
Kendall's Tau	0.29	<0.05
Somers' D Statistics	X/Y=0.29 Y/X=0.30	<0.05 <0.05
Gamma Statistic	0.31	<0.05

YEAR : 1994, valid cases: 74		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.49	<0.05
Spearman correlation coefficient	0.50	<0.05
Kendall's Tau	0.38	<0.05
Somers' D Statistics	X/Y=0.37 Y/X=0.39	<0.05 <0.05
Gamma Statistic	0.40	<0.05

YEAR : 1995, valid cases: 91		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.21	<0.05
Spearman correlation coefficient	0.19	>0.05
Kendall's Tau	0.14	>0.05
Somers' D Statistics	X/Y=0.14 Y/X=0.15	>0.05 >0.05
Gamma Statistic	0.15	>0.05

YEAR : 1996, valid cases: 47		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.60	<0.05
Spearman correlation coefficient	0.50	<0.05
Kendall's Tau	0.38	<0.05
Somers' D Statistics	X/Y=0.37 Y/X=0.39	<0.05 <0.05
Gamma Statistic	0.40	<0.05

YEAR : 1997, valid cases: 66		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient		
Spearman correlation coefficient	0.68	<0.05
Kendall's Tau	0.52	
Somers' D Statistics	X/Y=0.51 Y/X=0.53	
Gamma Statistic	0.55	

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 557		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.46	<0.05
Spearman correlation coefficient	0.43	<0.05
Kendall's Tau	0.31	<0.05
Somers' D Statistics	X/Y=0.30 Y/X=0.32	<0.05 <0.05
Gamma Statistic	0.33	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 557		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.32	<0.05
Spearman correlation coefficient	0.31	<0.05
Kendall's Tau	0.27	<0.05
Somers' D Statistics	X/Y=0.38 Y/X=0.19	<0.05 <0.05
Gamma Statistic	0.41	<0.05

MATRIC HIGHER-GRADE PHYSICAL SCIENCE (X) AND CEM101W (ordinal) (Y)

YEAR : 1990, valid cases: 53		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.53	<0.05
Spearman correlation coefficient	0.48	<0.05
Kendall's Tau	0.37	<0.05
Somers' D Statistics	X/Y=0.33 Y/X=0.41	<0.05 <0.05
Gamma Statistic	0.43	<0.05

YEAR : 1991, valid cases: 54		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.31	<0.05
Spearman correlation coefficient	0.31	<0.05
Kendall's Tau	0.23	<0.05
Somers' D Statistics	X/Y=0.20 Y/X=0.27	<0.05 <0.05
Gamma Statistic	0.28	<0.05

YEAR : 1992, valid cases: 54		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.53	<0.05
Spearman correlation coefficient	0.52	<0.05
Kendall's Tau	0.39	<0.05
Somers' D Statistics	X/Y=0.34 Y/X=0.44	<0.05 <0.05
Gamma Statistic	0.45	<0.05

YEAR : 1993, valid cases: 49		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.24	<0.05
Spearman correlation coefficient	0.24	<0.05
Kendall's Tau	0.18	<0.05
Somers' D Statistics	X/Y=0.15 Y/X=0.21	<0.05 <0.05
Gamma Statistic	0.22	<0.05

YEAR : 1994, valid cases: 71		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.48	<0.05
Spearman correlation coefficient	0.53	<0.05
Kendall's Tau	0.41	<0.05
Somers' D Statistics	X/Y=0.36 Y/X=0.47	<0.05 <0.05
Gamma Statistic	0.46	<0.05

YEAR : 1995, valid cases: 90		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.19	>0.05
Spearman correlation coefficient	0.24	<0.05
Kendall's Tau	0.19	<0.05
Somers' D Statistics	X/Y=0.16 Y/X=0.22	<0.05 <0.05
Gamma Statistic	0.24	<0.05

YEAR : 1996, valid cases: 47		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.63	<0.05
Spearman correlation coefficient	0.52	<0.05
Kendall's Tau	0.42	<0.05
Somers' D Statistics	X/Y=0.35 Y/X=0.50	<0.05 <0.05
Gamma Statistic	0.51	<0.05

YEAR : 1997, valid cases: 65		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.61	<0.05
Spearman correlation coefficient	0.58	<0.05
Kendall's Tau	0.47	<0.05
Somers' D Statistics	X/Y=0.42 Y/X=0.52	<0.05 <0.05
Gamma Statistic	0.54	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 536		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.43	<0.05
Spearman correlation coefficient	0.41	<0.05
Kendall's Tau	0.32	<0.05
Somers' D Statistics	X/Y=0.27 Y/X=0.36	<0.05 <0.05
Gamma Statistic	0.37	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 536		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.32	<0.05
Spearman correlation coefficient	0.32	<0.05
Kendall's Tau	0.29	<0.05
Somers' D Statistics	X/Y=0.38 Y/X=0.23	<0.05 <0.05
Gamma Statistic	0.51	<0.05

In the preceding tables we observe:

- In most of the years (except for 1995 for associations involving both types of point scores) the measures have been found to be *statistically significant*.

We may conclude that the incoming first year student with either the higher (*lower*) faculty point score or the higher (*lower*) matric point score or the higher (*lower*) matric higher-grade physical science symbol are more likely (*less likely*) to achieve higher (*lower*) percentage for CEM101W, respectively.

- The magnitudes of the measures are comparable. It does not seem that any of the investigated factors is a better determinant of the actual percentage marks for CEM101W than the others.

Examining the dendrograms we note:

- In the ordinal response analysis (CONFIRM output), the *matric point score* is the best predictor of the actual percentage obtained for CEM101W.

For those students with the matric point scores of 36 and less there are significant differences between the White and non-White students, with the former obtaining on average higher percentage (52.00% and 44.49% respectively).

For those students with the matric point scores 37 and 41 there are significant differences between those who registered for first year in different calendar years. The students registered in 1995 exhibiting the lowest mean percentage (47.34% in contrast with 53.44% and 59.08%).

For those students with the matric point score of 43, those with A for the matric higher-grade maths obtained higher mark (69.15%) than those with any lower symbol (59.13%).

- In the binary response analysis (CATFIRM output), the *colour* variable emerges as the best predictor for passing or failing the chemistry course.

For the White category, the *matric higher-grade physical science* appears to act as a further discriminant of passing or failing the course. The proportions of students passing the course are higher for the symbols A and B (91.0%) than for the symbols C, D and E (71.4%).

For the combined Black, Coloured, Indian and Missing Information categories, the *faculty point score* may play that role with the higher (*lower*) point scores exhibiting higher (*lower*) pass rates for CEM101W, respectively.

Despite the fact that CEM100W and CEM101W are effectively exactly the same courses (the split had been made essentially for administrative reasons), the different factors emerge as significant performance predictors. We find, however, in the larger sets (i.e. the set for CEM101W) evidence, at least within the White students grouping, that passing or failing the chemistry course at first year level may be strongly associated with the actual symbol achieved in matric for physical science.

Dendrogram of CONFIRM: ordinal response analysis: cem101w

split var	1						
P val (Bonf P)	557						
	53.0180						
	12.1670						
	22.0000						
	94.0000						
levels							
Node							
N	matpoint						
Xbar	1.59E-26; (2.39E-25)						
SD							
Min							
Max							
	0123	4567	89				
	2	3	4				
	225	248	84				
	47.7956	53.7863	64.7381				
	9.8308	11.3233	11.5999				
	26.0000	22.0000	32.0000				
	78.0000	86.0000	94.0000				
	colour	year	math				
	5.80E-06; (8.70E-05)	5.58E-03; (0.0837)	0.0297; (0.3855)				
	cbl	w7	5	20176	34	7dcb	a
	5	6	7	8	9	10	11
	126	99	46	137	65	37	47
	44.4921	52.0000	47.3478	53.4380	59.0769	59.1351	69.1489
	8.4737	9.8706	10.4631	10.6397	10.9006	9.9671	10.9504
	26.0000	27.0000	22.0000	24.0000	35.0000	34.0000	32.0000
	70.0000	78.0000	75.0000	79.0000	86.0000	82.0000	94.0000

Dendrogram of CATFIRM: binary response analysis: cem101w

split var	1			
P val (Bonf P)	34.5%			
	65.5%			
	557			
levels				
Node				
% fail	colour			
% pass	2.61E-13; (3.65E-12)			
N				
	?cbl	w		
	2	3		
	53.4%	20.4%		
	46.6%	79.6%		
	238	319		
	f_cpoin	psch		
	2.60E-04; (3.64E-03)	0.0133; (0.1726)		
	01234	56789	edc	b?a
	4	5	6	7
	65.6%	31.0%	28.6%	9.0%
	34.4%	69.0%	71.4%	91.0%
	154	84	185	134

c) PHY104W

For PHY104W the measures of association were calculated for the years 1990-1997 and the combined years field.

The results are presented in the tables below.

FACULTY POINT SCORE (X) AND PHY104W (ordinal) (Y)

YEAR : 1990, valid cases: 31		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.60	<0.05
Spearman correlation coefficient	0.56	<0.05
Kendall's Tau	0.41	<0.05
Somers' D Statistics	X/Y=0.40 Y/X=0.42	<0.05
Gamma Statistic	0.42	<0.05

YEAR : 1991, valid cases: 31		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.68	<0.05
Spearman correlation coefficient	0.66	<0.05
Kendall's Tau	0.51	<0.05
Somers' D Statistics	X/Y=0.50 Y/X=0.51	<0.05
Gamma Statistic	0.52	<0.05

YEAR : 1992, valid cases: 39		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.63	<0.05
Spearman correlation coefficient	0.61	<0.05
Kendall's Tau	0.45	<0.05
Somers' D Statistics	X/Y=0.44 Y/X=0.46	<0.05
Gamma Statistic	0.47	<0.05

YEAR : 1993, valid cases: 51		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.57	<0.05
Spearman correlation coefficient	0.52	<0.05
Kendall's Tau	0.38	<0.05
Somers' D Statistics	X/Y=0.38 Y/X=0.38	<0.05
Gamma Statistic	0.39	<0.05

YEAR : 1994, valid cases: 31		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.71	<0.05
Spearman correlation coefficient	0.72	<0.05
Kendall's Tau	0.50	<0.05
Somers' D Statistics	X/Y=0.52 Y/X=0.53	<0.05 <0.05
Gamma Statistic	0.54	<0.05

YEAR : 1995, valid cases: 53		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.63	<0.05
Spearman correlation coefficient	0.62	<0.05
Kendall's Tau	0.46	<0.05
Somers' D Statistics	X/Y=0.46 Y/X=0.47	<0.05 <0.05
Gamma Statistic	0.48	<0.05

YEAR : 1996, valid cases: 43		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.68	<0.05
Spearman correlation coefficient	0.66	<0.05
Kendall's Tau	0.49	<0.05
Somers' D Statistics	X/Y=0.48 Y/X=0.50	<0.05 <0.05
Gamma Statistic	0.51	<0.05

YEAR : 1997, valid cases: 17		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.71	<0.05
Spearman correlation coefficient	0.73	<0.05
Kendall's Tau	0.56	<0.05
Somers' D Statistics	X/Y=0.54 Y/X=0.58	<0.05 <0.05
Gamma Statistic	0.59	<0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 307		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.61	<0.05
Spearman correlation coefficient	0.61	<0.05
Kendall's Tau	0.45	<0.05
Somers' D Statistics	X/Y=0.44 Y/X=0.45	<0.05 <0.05
Gamma Statistic	0.46	<0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 307		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.37	<0.05
Spearman correlation coefficient	0.42	<0.05
Kendall's Tau	0.35	<0.05
Somers' D Statistics	X/Y=0.60 Y/X=0.20	<0.05 <0.05
Gamma Statistic	0.63	<0.05

MATRIC POINT SCORE (X) AND PHY104W (ordinal) (Y)

YEAR : 1990, valid cases: 31		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.55	<0.05
Spearman correlation coefficient	0.49	<0.05
Kendall's Tau	0.37	<0.05
Somers' D Statistics	X/Y=0.36 Y/X=0.38	<0.05 <0.05
Gamma Statistic	0.38	<0.05

YEAR : 1991, valid cases: 31		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.62	<0.05
Spearman correlation coefficient	0.58	<0.05
Kendall's Tau	0.43	<0.05
Somers' D Statistics	X/Y=0.43 Y/X=0.44	<0.05 <0.05
Gamma Statistic	0.45	<0.05

YEAR : 1992, valid cases: 39		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.61	<0.05
Spearman correlation coefficient	0.58	<0.05
Kendall's Tau	0.44	<0.05
Somers' D Statistics	X/Y=0.43 Y/X=0.46	<0.05 <0.05
Gamma Statistic	0.47	<0.05

YEAR : 1993, valid cases: 51		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.55	<0.05
Spearman correlation coefficient	0.51	<0.05
Kendall's Tau	0.37	<0.05
Somers' D Statistics	X/Y=0.37 Y/X=0.38	<0.05 <0.05
Gamma Statistic	0.39	<0.05

YEAR : 1994, valid cases: 31		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.67	<0.05
Spearman correlation coefficient	0.69	<0.05
Kendall's Tau	0.50	<0.05
Somers' D Statistics	X/Y=0.49 Y/X=0.51	<0.05 <0.05
Gamma Statistic	0.52	<0.05

YEAR : 1995, valid cases: 53		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.62	<0.05
Spearman correlation coefficient	0.60	<0.05
Kendall's Tau	0.45	<0.05
Somers' D Statistics	X/Y=0.45 Y/X=0.46	<0.05 <0.05
Gamma Statistic	0.48	<0.05

YEAR : 1996, valid cases: 43		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.61	<0.05
Spearman correlation coefficient	0.64	<0.05
Kendall's Tau	0.46	<0.05
Somers' D Statistics	X/Y=0.45 Y/X=0.47	<0.05 <0.05
Gamma Statistic	0.48	<0.05

YEAR : 1997, valid cases: 17		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.76	<0.05
Spearman correlation coefficient	0.80	<0.05
Kendall's Tau	0.63	<0.05
Somers' D Statistics	X/Y=0.61 Y/X=0.67	<0.05 <0.05
Gamma Statistic	0.80	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 307		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.58	<0.05
Spearman correlation coefficient	0.58	<0.05
Kendall's Tau	0.42	<0.05
Somers' D Statistics	X/Y=0.42 Y/X=0.43	<0.05 <0.05
Gamma Statistic	0.44	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 307		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.42	<0.05
Spearman correlation coefficient	0.41	<0.05
Kendall's Tau	0.34	<0.05
Somers' D Statistics	X/Y=0.59 Y/X=0.20	<0.05 <0.05
Gamma Statistic	0.62	<0.05

MATRIC HIGHER-GRADE PHYSICAL SCIENCE (X) AND PHY104W (ordinal) (Y)

YEAR : 1990, valid cases: 31		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.62	<0.05
Spearman correlation coefficient	0.56	<0.05
Kendall's Tau	0.45	<0.05
Somers' D Statistics	X/Y=0.35 Y/X=0.58	<0.05 <0.05
Gamma Statistic	0.59	<0.05

YEAR : 1991, valid cases: 31		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.68	<0.05
Spearman correlation coefficient	0.67	<0.05
Kendall's Tau	0.56	<0.05
Somers' D Statistics	X/Y=0.45 Y/X=0.69	<0.05 <0.05
Gamma Statistic	0.70	<0.05

YEAR : 1992, valid cases: 39		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.34	<0.05
Spearman correlation coefficient	0.33	<0.05
Kendall's Tau	0.26	<0.05
Somers' D Statistics	X/Y=0.22 Y/X=0.32	<0.05 <0.05
Gamma Statistic	0.32	<0.05

YEAR : 1993, valid cases: 50		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.45	<0.05
Spearman correlation coefficient	0.39	<0.05
Kendall's Tau	0.30	<0.05
Somers' D Statistics	X/Y=0.26 Y/X=0.35	<0.05 <0.05
Gamma Statistic	0.36	<0.05

YEAR : 1994, valid cases: 30		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.57	<0.05
Spearman correlation coefficient	0.58	<0.05
Kendall's Tau	0.45	<0.05
Somers' D Statistics	X/Y=0.39 Y/X=0.51	<0.05 <0.05
Gamma Statistic	0.52	<0.05

YEAR : 1995, valid cases: 53		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.53	<0.05
Spearman correlation coefficient	0.59	<0.05
Kendall's Tau	0.44	<0.05
Somers' D Statistics	X/Y=0.38 Y/X=0.50	<0.05 <0.05
Gamma Statistic	0.51	<0.05

YEAR : 1996, valid cases: 43		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.61	<0.05
Spearman correlation coefficient	0.64	<0.05
Kendall's Tau	0.52	<0.05
Somers' D Statistics	X/Y=0.44 Y/X=0.63	<0.05 <0.05
Gamma Statistic	0.64	<0.05

YEAR : 1997, valid cases: 17		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.51	<0.05
Spearman correlation coefficient	0.53	<0.05
Kendall's Tau	0.44	<0.05
Somers' D Statistics	X/Y=0.29 Y/X=0.68	<0.05 <0.05
Gamma Statistic	0.53	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 305		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.53	<0.05
Spearman correlation coefficient	0.55	<0.05
Kendall's Tau	0.43	<0.05
Somers' D Statistics	X/Y=0.36 Y/X=0.50	<0.05 <0.05
Gamma Statistic	0.52	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 305		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.44	<0.05
Spearman correlation coefficient	0.43	<0.05
Kendall's Tau	0.40	<0.05
Somers' D Statistics	X/Y=0.58 Y/X=0.27	<0.05 <0.05
Gamma Statistic	0.70	<0.05

In the preceding tables we observe:

- All the measures are *statistically significant*.

Therefore, we may conclude that the incoming first year student with either the higher (*lower*) faculty point score or the higher (*lower*) matric point score or the higher (*lower*) matric higher-grade physical science symbol are more likely (*less likely*) to achieve higher (*lower*) percentage PHY104W, respectively.

- There is no specific pattern observed between the magnitudes of the measures. The strength of interrelationships appears to be year specific.

Examining the dendrograms we note:

- In both the ordinal and binary response analyses (CONFIRM and CATFIRM outputs), the *faculty point score* emerges as the best performance predictor for the physics PHY104W course. There are two groupings of the faculty point scores in the binary response analysis in contrast to the four groupings of the faculty point score produced for the ordinal response analysis.
- Additionally in the ordinal response analysis, for the faculty point score groupings of 46 and 52, and 54 and 58, there are significant differences in the mean percentages between students registered in the different calendar years.

For students with the faculty point scores of 60 and 61, those from the disadvantaged educational departments obtained on average a lower percentage (60.74%) than those from the privileged educational departments (75.34%).

d) PHY100W

For PHY100W the measures of association were calculated for the years 1990-1997 and the combined years field.

The results are displayed in the tables below.

FACULTY POINT SCORE (X) AND PHY100W (ordinal) (Y)

YEAR : 1990, valid cases: 86		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.60	<0.05
Spearman correlation coefficient	0.54	<0.05
Kendall's Tau	0.39	<0.05
Somers' D Statistics	X/Y=0.39 Y/X=0.39	<0.05 <0.05
Gamma Statistic	0.40	<0.05

YEAR : 1991, valid cases: 101		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.45	<0.05
Spearman correlation coefficient	0.43	<0.05
Kendall's Tau	0.32	<0.05
Somers' D Statistics	X/Y=0.31 Y/X=0.33	<0.05 <0.05
Gamma Statistic	0.33	<0.05

YEAR : 1992, valid cases: 96		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.29	<0.05
Spearman correlation coefficient	0.22	<0.05
Kendall's Tau	0.16	<0.05
Somers' D Statistics	X/Y=0.16 Y/X=0.17	<0.05 <0.05
Gamma Statistic	0.17	<0.05

YEAR : 1993, valid cases: 79		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.50	<0.05
Spearman correlation coefficient	0.48	<0.05
Kendall's Tau	0.34	<0.05
Somers' D Statistics	X/Y=0.34 Y/X=0.35	<0.05 <0.05
Gamma Statistic	0.36	<0.05

YEAR : 1994, valid cases: 134		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.49	<0.05
Spearman correlation coefficient	0.51	<0.05
Kendall's Tau	0.36	<0.05
Somers' D Statistics	X/Y=0.36 Y/X=0.37	<0.05 <0.05
Gamma Statistic	0.38	<0.05

YEAR : 1995, valid cases: 114		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.40	<0.05
Spearman correlation coefficient	0.39	<0.05
Kendall's Tau	0.28	<0.05
Somers' D Statistics	X/Y=0.27 Y/X=0.28	<0.05 <0.05
Gamma Statistic	0.29	<0.05

YEAR : 1996, valid cases: 70		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.46	<0.05
Spearman correlation coefficient	0.44	<0.05
Kendall's Tau	0.32	<0.05
Somers' D Statistics	X/Y=0.32 Y/X=0.33	<0.05 <0.05
Gamma Statistic	0.34	<0.05

YEAR : 1997, valid cases: 95		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.58	<0.05
Spearman correlation coefficient	0.55	<0.05
Kendall's Tau	0.42	<0.05
Somers' D Statistics	X/Y=0.41 Y/X=0.42	<0.05 <0.05
Gamma Statistic	0.44	<0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 838		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.45	<0.05
Spearman correlation coefficient	0.42	<0.05
Kendall's Tau	0.20	<0.05
Somers' D Statistics	X/Y=0.29 Y/X=0.30	<0.05 <0.05
Gamma Statistic	0.31	<0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 838		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.26	<0.05
Spearman correlation coefficient	0.27	<0.05
Kendall's Tau	0.23	<0.05
Somers' D Statistics	X/Y=0.41 Y/X=0.13	<0.05 <0.05
Gamma Statistic	0.44	<0.05

MATRIC POINT SCORE (X) AND PHY100W (ordinal) (Y)

YEAR : 1990, valid cases: 86		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.57	<0.05
Spearman correlation coefficient	0.50	<0.05
Kendall's Tau	0.37	<0.05
Somers' D Statistics	X/Y=0.36 Y/X=0.38	<0.05 <0.05
Gamma Statistic	0.39	<0.05

YEAR : 1991, valid cases: 101		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.58	<0.05
Spearman correlation coefficient	0.57	<0.05
Kendall's Tau	0.42	<0.05
Somers' D Statistics	X/Y=0.40 Y/X=0.43	<0.05 <0.05
Gamma Statistic	0.44	<0.05

YEAR : 1992, valid cases: 96		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.32	<0.05
Spearman correlation coefficient	0.22	<0.05
Kendall's Tau	0.16	<0.05
Somers' D Statistics	X/Y=0.16 Y/X=0.17	<0.05 <0.05
Gamma Statistic	0.17	<0.05

YEAR : 1993, valid cases: 79		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.56	<0.05
Spearman correlation coefficient	0.51	<0.05
Kendall's Tau	0.37	<0.05
Somers' D Statistics	X/Y=0.36 Y/X=0.38	<0.05 <0.05
Gamma Statistic	0.39	<0.05

YEAR : 1994, valid cases: 130		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.51	<0.05
Spearman correlation coefficient	0.52	<0.05
Kendall's Tau	0.38	<0.05
Somers' D Statistics	X/Y=0.37 Y/X=0.39	<0.05 <0.05
Gamma Statistic	0.41	<0.05

YEAR : 1995, valid cases: 114		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.38	<0.05
Spearman correlation coefficient	0.36	<0.05
Kendall's Tau	0.26	<0.05
Somers' D Statistics	X/Y=0.25 Y/X=0.27	<0.05 <0.05
Gamma Statistic	0.28	<0.05

YEAR : 1996, valid cases: 70		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.42	<0.05
Spearman correlation coefficient	0.44	<0.05
Kendall's Tau	0.32	<0.05
Somers' D Statistics	X/Y=0.32 Y/X=0.33	<0.05 <0.05
Gamma Statistic	0.34	<0.05

YEAR : 1997, valid cases: 95		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.54	<0.05
Spearman correlation coefficient	0.53	<0.05
Kendall's Tau	0.41	<0.05
Somers' D Statistics	X/Y=0.40 Y/X=0.41	<0.05 <0.05
Gamma Statistic	0.43	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 838		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.46	<0.05
Spearman correlation coefficient	0.43	<0.05
Kendall's Tau	0.31	<0.05
Somers' D Statistics	X/Y=0.30 Y/X=0.32	<0.05 <0.05
Gamma Statistic	0.33	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 838		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.30	<0.05
Spearman correlation coefficient	0.29	<0.05
Kendall's Tau	0.24	<0.05
Somers' D Statistics	X/Y=0.43 Y/X=0.14	<0.05 <0.05
Gamma Statistic	0.46	<0.05

MATRIC HIGHER-GRADE PHYSICAL SCIENCE (X) AND PHY100W (ordinal) (Y)

YEAR : 1990, valid cases: 83		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.51	<0.05
Spearman correlation coefficient	0.47	<0.05
Kendall's Tau	0.36	<0.05
Somers' D Statistics	X/Y=0.32 Y/X=0.41	<0.05 <0.05
Gamma Statistic	0.43	<0.05

YEAR : 1991, valid cases: 96		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.37	<0.05
Spearman correlation coefficient	0.37	<0.05
Kendall's Tau	0.27	<0.05
Somers' D Statistics	X/Y=0.24 Y/X=0.32	<0.05 <0.05
Gamma Statistic	0.37	<0.05

YEAR : 1992, valid cases: 93		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.36	<0.05
Spearman correlation coefficient	0.32	<0.05
Kendall's Tau	0.25	<0.05
Somers' D Statistics	X/Y=0.22 Y/X=0.29	<0.05 <0.05
Gamma Statistic	0.29	<0.05

YEAR : 1993, valid cases: 77		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.37	<0.05
Spearman correlation coefficient	0.35	<0.05
Kendall's Tau	0.28	<0.05
Somers' D Statistics	X/Y=0.24 Y/X=0.32	<0.05 <0.05
Gamma Statistic	0.33	<0.05

YEAR : 1994, valid cases: 128		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.43	<0.05
Spearman correlation coefficient	0.47	<0.05
Kendall's Tau	0.38	<0.05
Somers' D Statistics	X/Y=0.32 Y/X=0.42	<0.05 <0.05
Gamma Statistic	0.44	<0.05

YEAR : 1995, valid cases: 112		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.38	<0.05
Spearman correlation coefficient	0.39	<0.05
Kendall's Tau	0.31	<0.05
Somers' D Statistics	X/Y=0.27 Y/X=0.38	<0.05 <0.05
Gamma Statistic	0.39	<0.05

YEAR : 1996, valid cases: 69		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.44	<0.05
Spearman correlation coefficient	0.43	<0.05
Kendall's Tau	0.33	<0.05
Somers' D Statistics	X/Y=0.29 X/Y=0.38	<0.05 <0.05
Gamma Statistic	0.39	<0.05

YEAR : 1997, valid cases: 94		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.58	<0.05
Spearman correlation coefficient	0.55	<0.05
Kendall's Tau	0.43	<0.05
Somers' D Statistics	X/Y=0.37 Y/X=0.49	<0.05 <0.05
Gamma Statistic	0.50	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 812		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.40	<0.05
Spearman correlation coefficient	0.39	<0.05
Kendall's Tau	0.30	<0.05
Somers' D Statistics	X/Y=0.26 Y/X=0.35	<0.05 <0.05
Gamma Statistic	0.36	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 812		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.21	<0.05
Spearman correlation coefficient	0.21	<0.05
Kendall's Tau	0.19	<0.05
Somers' D Statistics	X/Y=0.31 Y/X=0.12	<0.05 <0.05
Gamma Statistic	0.42	<0.05

In the preceding tables we observe:

- All the measures are *statistically significant*.

Therefore, we may conclude that the incoming first year student with either the higher (*lower*) faculty point score or the higher (*lower*) matric point score or the higher (*lower*) matric higher-grade physical science symbol are more likely (*less likely*) to achieve the higher (*lower*) percentage for the first year physics PHY100W course, respectively.

- The magnitudes of the measures are usually lowest for associations involving the matric physical science. The magnitudes of the measures are slightly larger for associations involving the faculty point score than the matric point score. However these differences are only of marginal importance.

Examining the dendrograms we note:

- In the ordinal response analysis (CONFIRM output), the *faculty point score* emerges as the best performance predictor for the actual percentage for PHY100W.

For the students with the faculty point score of 56 and less, there are significant differences in the mean percentages obtained amongst students from the predominately disadvantaged educational departments (node 6) and the privileged educational departments (nodes 7 and 8).

- In the binary response analysis (CATFIRM output), the *matric point score* emerges as the best predictor for passing or failing PHY100W.

For the students with the matric point score of 36 and less, there are significant differences in the proportions of students passing PHY100W between the White and the non-White student groupings with the latter exhibiting a substantially lower pass rate than the former (59.5% in contrast to 84.5%).

Additionally, for the White students the proportion passing is substantially lower for those registered in first year in 1997 (42.9%) than in any of the earlier years (88.7%).

Despite the fact that comparing the two physics courses at first year level would be rather difficult as they follow different syllabi and are aimed at different groups of students, we note that the faculty point scores appear to be the most significant determinant of the actual marks obtained by students for the courses.

No statistical evidence has emerged as to whether the performance in any physics course at first year level might be attributed to the performance in the physical science in matric.

Dendrogram of CONFIRM: ordinal response analysis: phy104w

split var	1			
P val (Bonf P)	307			
	61.7264			
	16.0095			
	23.0000			
	97.0000			
levels	facpoint			
Node	1.32E-26; (1.97E-25)			
N	0123	45	67	89
Xbar	2	3	4	5
SD	122	71	59	55
Min	51.0902	62.5211	68.6610	76.8545
Max	13.0010	11.7167	11.9577	14.1440
	23.0000	35.0000	36.0000	32.0000
	80.0000	87.0000	89.0000	97.0000
	year	year	education	
	0.0131; (0.1960)	0.0165; (0.2140)	0.0121; (0.1453)	
	6540	3712	56	740231
	6	7	8	9
	66	56	18	53
	46.0152	57.0714	51.7778	66.1698
	11.6017	12.0602	9.9384	9.9494
	23.0000	29.0000	35.0000	46.0000
	80.0000	80.0000	74.0000	87.0000
	bhken	aldic		
	10	11		
	27	32		
	60.7407	75.3437		
	9.8160	9.2723		
	36.0000	53.0000		
	75.0000	89.0000		

Dendrogram of CATFIRM: binary response analysis: phy104w

split var	1	
P val (Bonf P)	19.9%	
	80.1%	
	307	
levels	facpoint	
Node	3.06E-11; (4.28E-10)	
% fail	0123	456789
% pass	2	3
N	41.0%	5.9%
	59.0%	94.1%
	122	185

5.4 SCIENCE FOUNDATION PROGRAMME COURSES

5.4.1 SFP BIOLOGY COURSE: BIO001F

BIO001F has been offered in the Faculty since 1993. Thus, the measures of association were calculated for the years from 1993 to 1997 and the combined years field.

The results are presented in the tables below.

FACULTY POINT SCORE (X) AND BIO001F (ordinal) (Y)

YEAR : 1993, valid cases: 29		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.63	<0.05
Spearman correlation coefficient	0.58	<0.05
Kendall's Tau	0.44	<0.05
Somers' D Statistics	X/Y=0.43 Y/X=0.46	<0.05 <0.05
Gamma Statistic	0.48	<0.05

YEAR : 1994, valid cases: 23		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.23	>0.05
Spearman correlation coefficient	0.23	>0.05
Kendall's Tau	0.17	>0.05
Somers' D Statistics	X/Y=0.16 Y/X=0.18	>0.05 >0.05
Gamma Statistic	0.18	>0.05

YEAR : 1995, valid cases: 25		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.17	>0.05
Spearman correlation coefficient	0.20	>0.05
Kendall's Tau	0.15	>0.05
Somers' D Statistics	X/Y=0.14 Y/X=0.15	>0.05 >0.05
Gamma Statistic	0.16	>0.05

YEAR : 1996, valid cases: 53		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.29	<0.05
Spearman correlation coefficient	0.16	>0.05
Kendall's Tau	0.10	>0.05
Somers' D Statistics	X/Y=0.10 Y/X=0.11	>0.05 >0.05
Gamma Statistic	0.11	>0.05

YEAR : 1997, valid cases: 48		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.41	<0.05
Spearman correlation coefficient	0.33	<0.05
Kendall's Tau	0.24	<0.05
Somers' D Statistics	X/Y=0.23 Y/X=0.24	<0.05 <0.05
Gamma Statistic	0.25	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1993-1997 valid cases: 178		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.33	<0.05
Spearman correlation coefficient	0.31	<0.05
Kendall's Tau	0.22	<0.05
Somers' D Statistics	X/Y=0.21 Y/X=0.23	<0.05 <0.05
Gamma Statistic	0.23	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1993-1997 valid cases: 178		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.17	<0.05
Spearman correlation coefficient	0.16	<0.05
Kendall's Tau	0.14	<0.05
Somers' D Statistics	X/Y=0.19 Y/X=0.10	<0.05 <0.05
Gamma Statistic	0.20	<0.05

MATRIC POINT SCORE (X) AND BIO001F (ordinal) (Y)

YEAR : 1993, valid cases: 29		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.60	<0.05
Spearman correlation coefficient	0.56	<0.05
Kendall's Tau	0.40	<0.05
Somers' D Statistics	X/Y=0.39 Y/X=0.41	<0.05 <0.05
Gamma Statistic	0.42	<0.05

YEAR : 1994, valid cases: 23		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.28	>0.05
Spearman correlation coefficient	0.34	>0.05
Kendall's Tau	0.24	>0.05
Somers' D Statistics	X/Y=0.23 Y/X=0.25	>0.05 >0.05
Gamma Statistic	0.27	>0.05

YEAR : 1995, valid cases: 25		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.28	>0.05
Spearman correlation coefficient	0.34	>0.05
Kendall's Tau	0.25	>0.05
Somers' D Statistics	X/Y=0.23 Y/X=0.26	>0.05 >0.05
Gamma Statistic	0.28	>0.05

YEAR : 1996, valid cases: 53		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.36	<0.05
Spearman correlation coefficient	0.24	>0.05
Kendall's Tau	0.17	>0.05
Somers' D Statistics	X/Y=0.16 Y/X=0.18	>0.05 >0.05
Gamma Statistic	0.18	>0.05

YEAR : 1997, valid cases: 48		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.36	<0.05
Spearman correlation coefficient	0.32	<0.05
Kendall's Tau	0.23	<0.05
Somers' D Statistics	X/Y=0.22 Y/X=0.24	<0.05 <0.05
Gamma Statistic	0.24	<0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1993-1997 valid cases: 178		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.34	<0.05
Spearman correlation coefficient	0.34	<0.05
Kendall's Tau	0.24	<0.05
Somers' D Statistics	X/Y=0.23 Y/X=0.25	<0.05 <0.05
Gamma Statistic	0.26	<0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1993-1997 valid cases: 178		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.21	<0.05
Spearman correlation coefficient	0.22	<0.05
Kendall's Tau	0.19	<0.05
Somers' D Statistics	X/Y=0.25 Y/X=0.14	<0.05 <0.05
Gamma Statistic	0.28	<0.05

MATRIC HIGHER-GRADE BIOLOGY AND BIO001F (ordinal) (Y)

YEAR : 1993, valid cases: 28		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.23	>0.05
Spearman correlation coefficient	0.33	>0.05
Kendall's Tau	0.26	>0.05
Somers' D Statistics	X/Y=0.22 Y/X=0.30	>0.05 >0.05
Gamma Statistic	0.32	>0.05

YEAR : 1994, valid cases: 22		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.18	>0.05
Spearman correlation coefficient	0.16	>0.05
Kendall's Tau	0.14	>0.05
Somers' D Statistics	X/Y=0.10 Y/X=0.20	>0.05 >0.05
Gamma Statistic	0.21	>0.05

YEAR : 1995, valid cases: 19		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.22	>0.05
Spearman correlation coefficient	0.23	>0.05
Kendall's Tau	0.19	>0.05
Somers' D Statistics	X/Y=0.15 Y/X=0.24	>0.05 >0.05
Gamma Statistic	0.25	>0.05

YEAR : 1996, valid cases: 52		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.25	>0.05
Spearman correlation coefficient	0.18	>0.05
Kendall's Tau	0.16	>0.05
Somers' D Statistics	X/Y=0.13 Y/X=0.19	>0.05 >0.05
Gamma Statistic	0.20	>0.05

YEAR : 1997, valid cases: 47		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.37	<0.05
Spearman correlation coefficient	0.32	<0.05
Kendall's Tau	0.26	<0.05
Somers' D Statistics	X/Y=0.23 Y/X=0.31	<0.05 <0.05
Gamma Statistic	0.32	<0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1993-1997 valid cases: 168		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.25	<0.05
Spearman correlation coefficient	0.22	<0.05
Kendall's Tau	0.27	<0.05
Somers' D Statistics	X/Y=0.14 Y/X=0.21	<0.05 <0.05
Gamma Statistic	0.22	<0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1993-1997 valid cases: 168		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.08	<0.05
Spearman correlation coefficient	0.09	<0.05
Kendall's Tau	0.08	<0.05
Somers' D Statistics	X/Y=0.10 Y/X=0.07	<0.05 <0.05
Gamma Statistic	0.15	<0.05

In the preceding tables most of the measures in the individual years analyses have been found *not to be statistically significant*. On the other hand, the measures for the combined years analyses have emerged as statistically significant. This feature is attributed to the fact that the sample sizes in the combined analyses have been substantially higher than the sample sizes in the individual years' analyses.

Given the data available, we may say that the performance in BIO001F does not appear to be significantly associated with either the faculty point score or the matric point score or the symbol for the matric higher-grade biology. However, there may exist other factors that are more strongly correlated with the performance variable.

Examining the dendrograms we note:

- In CONFIRM output, the *faculty point score* has emerged as the best predictor of the actual percentage obtained for BIO001F.

For the students with the faculty point scores of 46 and less, the mean percentage mark is 50.82%. For the students with the faculty point scores of more than 46, the mean percentage mark is 62.72%.

- There are no further significant predictors for the ordinal response analysis.
- CATFIRM did not produce a dendrogram since no predictors were found to split significantly the data set with respect to the proportions of students passing or failing BIO001F.

Dendrogram of CONFIRM: ordinal response analysis: bio001f

```
      |
      | split var
      | P val (Bonf P)
      |
      |-----|
      |
      | levels
Node |
+-----+
|N |
|Xbar |
|SD |
|Min |
|Max |
+-----+
      |
      | facpoint
      | 2.49E-04; (3.73E-03)
      |
      |-----|
      | 012345678 9
      | 2 | 3 |
      |-----+ +-----+
      | 160 | | 18 |
      | 50.8250 | | 62.7222 |
      | 9.0939 | | 7.4975 |
      | 24.0000 | | 49.0000 |
      | 77.0000 | | 74.0000 |
      |-----+ +-----+
```

No dendrogram in CATFIRM

5.4.2 SFP CHEMISTRY COURSE: CEM002F

For CEM002F the measures of association were calculated for the years 1990-1997 and the combined years field.

The results are displayed in the tables below.

FACULTY POINT SCORE (X) AND CEM002F (ordinal) (Y)

YEAR : 1990, valid cases: 36		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.20	>0.05
Spearman correlation coefficient	0.14	>0.05
Kendall's Tau	0.10	>0.05
Somers' D Statistics	X/Y=0.10 Y/X=0.10	>0.05 >0.05
Gamma Statistic	0.11	>0.05

YEAR : 1991, valid cases: 28		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.43	<0.05
Spearman correlation coefficient	0.37	<0.05
Kendall's Tau	0.27	<0.05
Somers' D Statistics	X/Y=0.27 Y/X=0.27	<0.05 <0.05
Gamma Statistic	0.28	<0.05

YEAR : 1992, valid cases: 47		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.20	>0.05
Spearman correlation coefficient	0.21	>0.05
Kendall's Tau	0.15	>0.05
Somers' D Statistics	X/Y=0.14 Y/X=0.16	>0.05 >0.05
Gamma Statistic	0.16	>0.05

YEAR : 1993, valid cases: 49		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.26	>0.05
Spearman correlation coefficient	0.30	<0.05
Kendall's Tau	0.21	>0.05
Somers' D Statistics	X/Y=0.20 Y/X=0.22	>0.05 >0.05
Gamma Statistic	0.22	>0.05

YEAR : 1994, valid cases: 38		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.11	>0.05
Spearman correlation coefficient	0.12	>0.05
Kendall's Tau	0.08	>0.05
Somers' D Statistics	X/Y=0.08 Y/X=0.08	>0.05 >0.05
Gamma Statistic	0.12	>0.05

YEAR : 1995, valid cases: 48		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.17	>0.05
Spearman correlation coefficient	0.23	>0.05
Kendall's Tau	0.17	>0.05
Somers' D Statistics	X/Y=0.16 Y/X=0.18	>0.05 >0.05
Gamma Statistic	0.18	>0.05

YEAR : 1996, valid cases: 80		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.35	<0.05
Spearman correlation coefficient	0.30	<0.05
Kendall's Tau	0.22	<0.05
Somers' D Statistics	X/Y=0.22 Y/X=0.23	<0.05 <0.05
Gamma Statistic	0.24	<0.05

YEAR : 1997, valid cases: 79		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.15	>0.05
Spearman correlation coefficient	0.17	<0.05
Kendall's Tau	0.12	>0.05
Somers' D Statistics	X/Y=0.12 Y/X=0.13	>0.05 >0.05
Gamma Statistic	0.13	>0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 405		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.18	<0.05
Spearman correlation coefficient	0.20	<0.05
Kendall's Tau	0.14	<0.05
Somers' D Statistics	X/Y=0.14 Y/X=0.14	<0.05 <0.05
Gamma Statistic	0.15	<0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 405		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.17	<0.05
Spearman correlation coefficient	0.19	<0.05
Kendall's Tau	0.16	<0.05
Somers' D Statistics	X/Y=0.32 Y/X=0.07	<0.05 <0.05
Gamma Statistic	0.35	<0.05

MATRIC POINT SCORE (X) AND CEM002F (ordinal) (Y)

YEAR : 1990, valid cases: 28		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.28	>0.05
Spearman correlation coefficient	0.19	>0.05
Kendall's Tau	0.13	>0.05
Somers' D Statistics	X/Y=0.13 Y/X=0.14	>0.05 >0.05
Gamma Statistic	0.14	>0.05

YEAR : 1991, valid cases: 28		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.38	<0.05
Spearman correlation coefficient	0.32	<0.05
Kendall's Tau	0.22	<0.05
Somers' D Statistics	X/Y=0.21 Y/X=0.23	<0.05 <0.05
Gamma Statistic	0.32	<0.05

YEAR : 1992, valid cases: 47		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.20	>0.05
Spearman correlation coefficient	0.19	<0.05
Kendall's Tau	0.14	>0.05
Somers' D Statistics	X/Y=0.14 Y/X=0.15	>0.05 >0.05
Gamma Statistic	0.15	>0.05

YEAR : 1993, valid cases: 49		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.32	<0.05
Spearman correlation coefficient	0.35	<0.05
Kendall's Tau	0.25	<0.05
Somers' D Statistics	X/Y=0.24 Y/X=0.25	<0.05 <0.05
Gamma Statistic	0.26	<0.05

YEAR : 1994, valid cases: 38		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.15	>0.05
Spearman correlation coefficient	0.16	<0.05
Kendall's Tau	0.12	>0.05
Somers' D Statistics	X/Y=0.12 Y/X=0.12	>0.05 >0.05
Gamma Statistic	0.13	>0.05

YEAR : 1995, valid cases: 48		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.18	>0.05
Spearman correlation coefficient	0.15	>0.05
Kendall's Tau	0.10	>0.05
Somers' D Statistics	X/Y=0.10 Y/X=0.11	>0.05 >0.05
Gamma Statistic	0.15	>0.05

YEAR : 1996, valid cases: 80		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.31	<0.05
Spearman correlation coefficient	0.26	<0.05
Kendall's Tau	0.19	<0.05
Somers' D Statistics	X/Y=0.18 Y/X=0.19	<0.05 <0.05
Gamma Statistic	0.20	<0.05

YEAR : 1997, valid cases: 79		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.06	>0.05
Spearman correlation coefficient	0.10	>0.05
Kendall's Tau	0.07	>0.05
Somers' D Statistics	X/Y=0.07 Y/X=0.07	>0.05 >0.05
Gamma Statistic	0.07	>0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 405		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.15	<0.05
Spearman correlation coefficient	0.16	<0.05
Kendall's Tau	0.11	<0.05
Somers' D Statistics	X/Y=0.11 Y/X=0.12	<0.05 <0.05
Gamma Statistic	0.12	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 405		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.12	<0.05
Spearman correlation coefficient	0.13	<0.05
Kendall's Tau	0.11	<0.05
Somers' D Statistics	X/Y=0.18 Y/X=0.07	<0.05 <0.05
Gamma Statistic	0.20	<0.05

MATRIC HIGHER GRADE PHYSICAL SCIENCE (X) AND CEM002F (ordinal) (Y)

YEAR : 1990, valid cases: 32		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.02	>0.05
Spearman correlation coefficient	-0.12	>0.05
Kendall's Tau	-0.12	>0.05
Somers' D Statistics	X/Y=-0.09 Y/X=-0.14	>0.05 >0.05
Gamma Statistic	-0.16	>0.05

YEAR : 1991, valid cases: 24		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.34	>0.05
Spearman correlation coefficient	0.39	>0.05
Kendall's Tau	0.31	>0.05
Somers' D Statistics	X/Y=0.23 Y/X=0.38	>0.05 >0.05
Gamma Statistic	0.39	>0.05

YEAR : 1992, valid cases: 43		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.12	>0.05
Spearman correlation coefficient	0.06	>0.05
Kendall's Tau	0.05	>0.05
Somers' D Statistics	X/Y=0.04 Y/X=0.06	>0.05 >0.05
Gamma Statistic	0.06	>0.05

YEAR : 1993, valid cases: 43		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.27	>0.05
Spearman correlation coefficient	0.28	>0.05
Kendall's Tau	0.21	>0.05
Somers' D Statistics	X/Y=0.18 Y/X=0.24	>0.05 >0.05
Gamma Statistic	0.25	>0.05

YEAR : 1994, valid cases: 35		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.10	>0.05
Spearman correlation coefficient	0.09	>0.05
Kendall's Tau	0.07	>0.05
Somers' D Statistics	X/Y=0.05 Y/X=0.10	>0.05 >0.05
Gamma Statistic	0.11	>0.05

YEAR : 1996, valid cases: 42		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.04	>0.05
Spearman correlation coefficient	0.06	>0.05
Kendall's Tau	0.04	>0.05
Somers' D Statistics	X/Y=0.03 Y/X=0.05	>0.05 >0.05
Gamma Statistic	0.06	>0.05

YEAR : 1996, valid cases: 64		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.26	>0.05
Spearman correlation coefficient	0.26	>0.05
Kendall's Tau	0.21	>0.05
Somers' D Statistics	X/Y=0.17 Y/X=0.28	>0.05 >0.05
Gamma Statistic	0.28	>0.05

YEAR : 1997, valid cases: 65		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.21	>0.05
Spearman correlation coefficient	0.20	>0.05
Kendall's Tau	0.16	>0.05
Somers' D Statistics	X/Y=0.12 Y/X=0.21	>0.05 >0.05
Gamma Statistic	0.22	>0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 348		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.20	>0.05
Spearman correlation coefficient	0.21	>0.05
Kendall's Tau	0.16	>0.05
Somers' D Statistics	X/Y=0.13 Y/X=0.20	>0.05 >0.05
Gamma Statistic	0.21	>0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 348		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.17	>0.05
Spearman correlation coefficient	-0.03	>0.05
Kendall's Tau	-0.03	>0.05
Somers' D Statistics	X/Y=-0.04 Y/X=-0.02	>0.05 >0.05
Gamma Statistic	-0.07	>0.05

Similarly to BIO001F, the measures for the individual years have been found *not to be statistically significant*. The measures for the combined years have emerged as statistically significant due to the increased sample sizes.

We may similarly conclude that the performance in CEM002F does not seem to be significantly associated with either the faculty point score or the matric point score or the symbol for the matric higher-grade physical science. There may, however, exist other factors that have significant correlation with the performance variable.

Examining the dendrograms we note:

- In the ordinal and the binary response analyses (both in CONFIRM and CATFIRM outputs) the significant differences in either the mean percentage mark or the proportion passing or failing CEM002F have emerged amongst students registered in the different calendar years. It appears that the mean marks and the proportions passing have been decreasing across the years.
- There are no further significant performance predictors.

Dendrogram of CONFIRM: ordinal response analysis: cem002f

split var	1		
P val (Bonf P)	405		
	57.6494		
	13.0809		
	10.0000		
	95.0000		
levels	year		
Node	1.32E-24; (1.98E-23)		
	5467 30 21		
	2	3	4
	245	85	75
	53.0531	58.7647	71.4000
	10.6502	10.6867	12.9427
	10.0000	22.0000	40.0000
	83.0000	81.0000	95.0000

Dendrogram of CATFIRM: binary response analysis: cem002f

split var	1	
P val (Bonf P)	22.5%	
	77.5%	
	405	
levels	year	
Node	0.0203; (0.2836)	
% fail		
% pass		
N	021	34567
	2	3
	6.3%	28.6%
	93.7%	71.4%
	111	294

5.3.3 SFP PHYSICS COURSE: PHY002F

For PHY002F the measures of association were calculated for the years 1990-1997 and the combined years field.

The results are displayed in the tables below.

FACULTY POINT SCORE (X) AND PHY002F (ordinal) (Y)

YEAR : 1990, valid cases: 36		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.20	>0.05
Spearman correlation coefficient	0.29	>0.05
Kendall's Tau	0.24	>0.05
Somers' D Statistics	X/Y=0.23 Y/X=0.26	>0.05 >0.05
Gamma Statistic	0.26	>0.05

YEAR : 1991, valid cases: 28		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.43	<0.05
Spearman correlation coefficient	0.48	<0.05
Kendall's Tau	0.36	<0.05
Somers' D Statistics	X/Y=0.35 Y/X=0.36	<0.05 <0.05
Gamma Statistic	0.37	<0.05

YEAR : 1992, valid cases: 47		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.20	>0.05
Spearman correlation coefficient	0.35	<0.05
Kendall's Tau	0.25	<0.05
Somers' D Statistics	X/Y=0.24 Y/X=0.26	<0.05 <0.05
Gamma Statistic	0.27	<0.05

YEAR : 1993, valid cases: 49		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.26	>0.05
Spearman correlation coefficient	0.32	<0.05
Kendall's Tau	0.23	<0.05
Somers' D Statistics	X/Y=0.23 Y/X=0.24	<0.05 <0.05
Gamma Statistic	0.25	<0.05

YEAR : 1994, valid cases: 38		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.11	>0.05
Spearman correlation coefficient	0.35	<0.05
Kendall's Tau	0.28	<0.05
Somers' D Statistics	X/Y=0.27 Y/X=0.29	<0.05 <0.05
Gamma Statistic	0.30	<0.05

YEAR : 1995, valid cases: 48		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.17	>0.05
Spearman correlation coefficient	0.23	<0.05
Kendall's Tau	0.17	>0.05
Somers' D Statistics	X/Y=0.17 Y/X=0.18	>0.05 >0.05
Gamma Statistic	0.23	>0.05

YEAR : 1996, valid cases: 80		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.35	<0.05
Spearman correlation coefficient	0.24	<0.05
Kendall's Tau	0.16	<0.05
Somers' D Statistics	X/Y=0.16 Y/X=0.17	<0.05 <0.05
Gamma Statistic	0.24	<0.05

YEAR : 1997, valid cases: 79		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.15	>0.05
Spearman correlation coefficient	0.13	>0.05
Kendall's Tau	0.09	>0.05
Somers' D Statistics	X/Y=0.09 Y/X=0.09	>0.05 >0.05
Gamma Statistic	0.09	>0.05

ORDINAL RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 398		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient		<0.05
Spearman correlation coefficient	0.25	<0.05
Kendall's Tau	0.18	<0.05
Somers' D Statistics	X/Y=0.18 Y/X=0.19	<0.05 <0.05
Gamma Statistic	0.19	<0.05

BINARY RESPONSE ANALYSIS	VALUE	P-VALUE
YEARS: 1990-1997 valid cases: 398		
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	0.17	<0.05
Spearman correlation coefficient	0.17	<0.05
Kendall's Tau	0.15	<0.05
Somers' D Statistics	X/Y=0.32 Y/X=0.07	<0.05 <0.05
Gamma Statistic	0.35	<0.05

MATRIC POINT SCORE (X) AND PHY002F (ordinal) (Y)

YEAR : 1990, valid cases: 32		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.31	>0.05
Spearman correlation coefficient	0.29	>0.05
Kendall's Tau	0.24	>0.05
Somers' D Statistics	X/Y=0.23 Y/X=0.26	>0.05 >0.05
Gamma Statistic	0.26	>0.05

YEAR : 1991, valid cases: 27		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.41	>0.05
Spearman correlation coefficient	0.40	<0.05
Kendall's Tau	0.29	<0.05
Somers' D Statistics	X/Y=0.28 Y/X=0.31	<0.05 <0.05
Gamma Statistic	0.31	<0.05

YEAR : 1992, valid cases: 47		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.29	>0.05
Spearman correlation coefficient	0.29	<0.05
Kendall's Tau	0.20	<0.05
Somers' D Statistics	X/Y=0.20 Y/X=0.21	<0.05 <0.05
Gamma Statistic	0.22	<0.05

YEAR : 1993, valid cases: 48		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.26	>0.05
Spearman correlation coefficient	0.31	<0.05
Kendall's Tau	0.20	<0.05
Somers' D Statistics	X/Y=0.22 Y/X=0.23	<0.05 <0.05
Gamma Statistic	0.24	<0.05

YEAR : 1994, valid cases: 38		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.31	>0.05
Spearman correlation coefficient	0.33	<0.05
Kendall's Tau	0.26	<0.05
Somers' D Statistics	X/Y=0.26 Y/X=0.27	<0.05 <0.05
Gamma Statistic	0.28	<0.05

YEAR : 1995, valid cases: 48		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.26	>0.05
Spearman correlation coefficient	0.19	>0.05
Kendall's Tau	0.14	>0.05
Somers' D Statistics	X/Y=0.13 Y/X=0.14	>0.05 >0.05
Gamma Statistic	0.15	>0.05

YEAR : 1996, valid cases: 80		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.22	>0.05
Spearman correlation coefficient	0.17	>0.05
Kendall's Tau	0.12	>0.05
Somers' D Statistics	X/Y=0.11 Y/X=0.12	>0.05 >0.05
Gamma Statistic	0.12	>0.05

YEAR : 1997, valid cases: 78		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.13	>0.05
Spearman correlation coefficient	0.12	>0.05
Kendall's Tau	0.09	>0.05
Somers' D Statistics	X/Y=0.09 Y/X=0.09	>0.05 >0.05
Gamma Statistic	0.09	>0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 398		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.24	<0.05
Spearman correlation coefficient	0.22	<0.05
Kendall's Tau	0.16	<0.05
Somers' D Statistics	X/Y=0.15 Y/X=0.16	<0.05 <0.05
Gamma Statistic	0.17	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 398		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.15	<0.05
Spearman correlation coefficient	0.14	<0.05
Kendall's Tau	0.12	<0.05
Somers' D Statistics	X/Y=0.26 Y/X=0.05	<0.05 <0.05
Gamma Statistic	0.29	<0.05

MATRIC HIGHER-GRADE PHYSICAL SCIENCE (X) AND PHY002F (Y)

YEAR : 1990, valid cases: 30		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.06	>0.05
Spearman correlation coefficient	-0.05	>0.05
Kendall's Tau	-0.05	>0.05
Somers' D Statistics	X/Y=-0.04 Y/X=-0.07	>0.05 >0.05
Gamma Statistic	-0.05	>0.05

YEAR : 1991, valid cases: 24		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.35	>0.05
Spearman correlation coefficient	0.31	>0.05
Kendall's Tau	0.24	>0.05
Somers' D Statistics	X/Y=0.19 Y/X=0.29	>0.05 >0.05
Gamma Statistic	0.30	>0.05

YEAR : 1992, valid cases: 43		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.23	>0.05
Spearman correlation coefficient	0.15	>0.05
Kendall's Tau	0.11	>0.05
Somers' D Statistics	X/Y=0.09 Y/X=0.14	>0.05 >0.05
Gamma Statistic	0.14	>0.05

YEAR : 1993, valid cases: 42		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.38	<0.05
Spearman correlation coefficient	0.42	<0.05
Kendall's Tau	0.32	<0.05
Somers' D Statistics	X/Y=0.28 Y/X=0.38	<0.05 <0.05
Gamma Statistic	0.39	<0.05

YEAR : 1994, valid cases: 35		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.12	>0.05
Spearman correlation coefficient	0.08	>0.05
Kendall's Tau	0.06	>0.05
Somers' D Statistics	X/Y=0.04 Y/X=0.09	>0.05 >0.05
Gamma Statistic	0.09	>0.05

YEAR : 1995, valid cases: 42		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	-0.14	>0.05
Spearman correlation coefficient	-0.15	>0.05
Kendall's Tau	-0.12	>0.05
Somers' D Statistics	X/Y=-0.10 Y/X=-0.16	>0.05 >0.05
Gamma Statistic	-0.16	>0.05

YEAR : 1996, valid cases: 64		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.26	<0.05
Spearman correlation coefficient	0.27	<0.05
Kendall's Tau	0.21	<0.05
Somers' D Statistics	X/Y=0.17 Y/X=0.26	<0.05 <0.05
Gamma Statistic	0.27	<0.05

YEAR : 1997, valid cases: 65		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.36	<0.05
Spearman correlation coefficient	0.34	<0.05
Kendall's Tau	0.28	<0.05
Somers' D Statistics	X/Y=0.22 Y/X=0.38	<0.05 <0.05
Gamma Statistic	0.39	<0.05

ORDINAL RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 345		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.20	<0.05
Spearman correlation coefficient	0.19	<0.05
Kendall's Tau	0.15	<0.05
Somers' D Statistics	X/Y=0.12 Y/X=0.19	<0.05 <0.05
Gamma Statistic	0.19	<0.05

BINARY RESPONSE ANALYSIS		
YEARS: 1990-1997 valid cases: 345		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.16	<0.05
Spearman correlation coefficient	0.07	<0.05
Kendall's Tau	0.07	<0.05
Somers' D Statistics	X/Y=0.13 Y/X=0.03	<0.05 <0.05
Gamma Statistic	0.20	<0.05

In the preceding tables the measures for associations involving either the matric point score or the matric higher-grade physical science have been found *not to be statistically significant* for most of the individual years analyses. The measures involving the faculty point score have emerged as *statistically significant* for all the individual years except for 1990, 1995 and 1997. For the combined years analyses the measures have emerged as *significant* since the respective samples have been substantially increased.

We may say that the higher (*lower*) faculty point scores are associated with the higher (*lower*) percentages for PHY002F. However there may be other factors that also have significant association with the performance variable.

Examining the dendrograms we note:

- In CONFIRM output, the *faculty point score* splits the entire data set most significantly and into the three groupings of the faculty point scores.

Those students with the faculty point scores of 40 and less have the mean percentage mark of 58.49%. Those students with the faculty point scores between 42 and 47 have the mean percentage of 62.66%. Those students with the faculty point score of 47 and more have the mean percentage mark of 69.03%.

There are no further significant predictors for the ordinal response analysis.

- CATFIRM did not produce a dendrogram as no predictors were found to split significantly the data with respect to the proportions of students who pass or fail PHY002F.

Summary of Section 5.4

Finally we note that for all three SFP courses the extra variable added in the FIRM analyses, namely the matric African language, has not provided any additional insights into the students' performance. In fact, in the summary files the variable has emerged usually as the one with the lowest significance levels.

Summarising the analyses of the SFP courses, there is not a single common performance predictor for all of them, except for the faculty point score for the actual percentages for BIO001F and PHY002F.

Despite the fact that less stringent criteria (i.e. 5% for the splitting phase and 6% for the merging phase) than in the analyses of the regular B.Sc. courses were used, no dendrogram was produced either for BIO001F or PHY002F in CATFIRM.

Those features may suggest that when selecting students for the special programme a combination of different factors (perhaps other than purely the academic ones) be considered.

5.5 ADDITIONAL COMMENTS ON THE PERFORMED ANALYSES

As regards the measures of associations, no explicit comments have been made with respect to the differences or similarities between the magnitudes and statistical significance of the measures in the ordinal and binary response analyses calculated for the combined years fields. The corresponding measures in these analyses are of different magnitudes but there is no common patterns observed e.g. the corresponding measures in the two analyses are neither consistently higher nor lower. Nevertheless we may note that the measures in the binary response analyses have emerged as statistically significant for the same relationships as those in the ordinal response analyses. Therefore, we may draw some additional conclusions.

The first year students with the higher (*lower*) values of the particular factors (i.e. either the faculty or the matric point score or the symbol for the specific subject at matric level) are more likely (*less likely*) to have passed than have failed the examined university courses, respectively.

We note that for certain courses the measures of associations calculated for the combined years sets have been found to be statistically significant despite the fact that the measures have not merged significant for some of the individual years analyses. This peculiarity is mainly attributed to the increased sample sizes in the combined years analyses, where the increase of sample size increases the power of the statistical test. The contrast between separate and combined analyses is just an artefact of the sizes of the data set, and the relative size of the overall association. It does not provide us with any extra insights into the possible predictors of the performance in the particular university subjects.

As regard the FIRM analyses, they were originally performed with rather stringent criteria (i.e. the significance levels of 0.9% and 1% for the splitting and merging stages, respectively). Thus we may expect that other significance levels (5% and 6%) could have possibly resulted in additional variables chosen to split the data significantly further. However, examination of the additional output files have indicated that in order to make additional splits on the remaining variables, the significance levels would have to be particularly low.

We need also to point out the cut-off points for the faculty and matric point scores in the presented dendrograms are not necessarily the same for all the courses investigated. Full details of the respective cut-off points may be found in the summary files, not explicitly included in the thesis, but available directly from the author.

CHAPTER 6: RELATIONSHIPS BETWEEN ACADEMIC YEARS

In this chapter we present the statistical results with respect to the mutual interdependencies amongst performance in the different academic years. These relationships were investigated using the measures of association. Only the ordinal definition of performance was examined i.e. the actual number of credits obtained by students in each of the specific academic years.

Once again presentation of the respective results follows the order and manner defined ¹ in Table 2.3 in chapter 2.

¹ Refer sections 5.1 and 5.2 in Table 2.3 in chapter 2.

6.1 RELATIONSHIP BETWEEN ACADEMIC PERFORMANCE IN FIRST AND SECOND YEARS

The rows and columns of the contingency tables were formed by the number of credits earned with respect to university courses completed in the first and in the second years, respectively. The analyses were performed only for the individual calendar years.

The results are displayed in the seven ² tables below.

CREDITS IN FIRST YEAR (X) AND CREDITS IN SECOND YEAR (Y)

YEAR : 1990, valid cases: 280		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.46	<0.05
Spearman correlation coefficient	0.55	<0.05
Kendall's Tau	0.48	<0.05
Somers' D Statistics	X/Y=0.47 Y/X=0.48	<0.05 <0.05
Gamma Statistic	0.61	<0.05
Pearson Chi-square Statistic	73.30	<0.05

YEAR : 1991, valid cases: 303		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.40	<0.05
Spearman correlation coefficient	0.35	<0.05
Kendall's Tau	0.31	<0.05
Somers' D Statistics	X/Y=0.30 Y/X=0.30	<0.05 <0.05
Gamma Statistic	0.39	<0.05
Pearson Chi-square Statistic	80.00	<0.05

YEAR : 1992, valid cases: 301		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.61	<0.05
Spearman correlation coefficient	0.61	<0.05
Kendall's Tau	0.54	<0.05
Somers' D Statistics	X/Y=0.53 Y/X=0.54	<0.05 <0.05
Gamma Statistic	0.69	<0.05
Pearson Chi-square Statistic	98.25	<0.05

YEAR : 1993, valid cases: 292		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.56	<0.05
Spearman correlation coefficient	0.54	<0.05
Kendall's Tau	0.45	<0.05
Somers' D Statistics	X/Y=0.46 Y/X=0.45	<0.05 <0.05
Gamma Statistic	0.58	<0.05
Pearson Chi-square Statistic	91.09	<0.05

² There are only seven tables since the information with respect to second year credits for students registered in first year in 1997 is not available in the current data set.

YEAR : 1994, valid cases: 318		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.57	<0.05
Spearman correlation coefficient	0.60	<0.05
Kendall's Tau	0.52	<0.05
Somers' D Statistics	X/Y=0.54 Y/X=0.51	<0.05 <0.05
Gamma Statistic	0.68	<0.05
Pearson Chi-square Statistic	93.29	<0.05

YEAR : 1995, valid cases: 340		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.52	<0.05
Spearman correlation coefficient	0.53	<0.05
Kendall's Tau	0.47	<0.05
Somers' D Statistics	X/Y=0.48 Y/X=0.46	<0.05 <0.05
Gamma Statistic	0.59	<0.05
Pearson Chi-square Statistic	176.65	<0.05

YEAR : 1996, valid cases: 350		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.64	<0.05
Spearman correlation coefficient	0.69	<0.05
Kendall's Tau	0.59	<0.05
Somers' D Statistics	X/Y=0.61 Y/X=0.58	<0.05 <0.05
Gamma Statistic	0.74	<0.05
Pearson Chi-square Statistic	124.58	<0.05

In each of the seven tables all the measures emerge as significant and are of similar magnitudes. Similar patterns to those described in chapter 3 are observed between the measures determined in this section. The only exception is the relative magnitude of the two Somers' *D* statistics³.

Based on these results, there is a significant relationship between the number of credits obtained in the first year and those obtained second year i.e. a student with a higher (*lower*) number of credits in the first year is more likely to obtain higher (*lower*) number of credits in the second year, respectively.

³ The two statistics are often similar or even equal. Since the two variables under investigation have similar number of distinct categories, there are similar numbers of ties present in both variables.

6.2 RELATIONSHIP BETWEEN ACADEMIC PERFORMANCE IN FIRST AND THIRD YEAR

The rows and columns of the contingency tables were formed by the number of credits earned with respect to university courses completed in the first and in the third years, respectively. The analyses were performed only for the individual calendar years.

The results are displayed in the six ⁴ tables below.

CREDITS IN FIRST YEAR (X) AND CREDITS IN THIRD YEAR (Y)

YEAR : 1990, valid cases: 280		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.45	<0.05
Spearman correlation coefficient	0.50	<0.05
Kendall's Tau	0.42	<0.05
Somers' D Statistics	X/Y=0.43 Y/X=0.41	<0.05 <0.05
Gamma Statistic	0.55	<0.05
Pearson Chi-square Statistic	57.79	<0.05

YEAR : 1991, valid cases: 303		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.37	<0.05
Spearman correlation coefficient	0.36	<0.05
Kendall's Tau	0.30	<0.05
Somers' D Statistics	X/Y=0.32 Y/X=0.29	<0.05 <0.05
Gamma Statistic	0.41	<0.05
Pearson Chi-square Statistic	42.82	<0.05

YEAR : 1992, valid cases: 301		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.59	<0.05
Spearman correlation coefficient	0.61	<0.05
Kendall's Tau	0.52	<0.05
Somers' D Statistics	X/Y=0.54 Y/X=0.50	<0.05 <0.05
Gamma Statistic	0.69	<0.05
Pearson Chi-square Statistic	91.93	<0.05

YEAR : 1993, valid cases: 292		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.56	<0.05
Spearman correlation coefficient	0.62	<0.05
Kendall's Tau	0.53	<0.05
Somers' D Statistics	X/Y=0.57 Y/X=0.49	<0.05 <0.05
Gamma Statistic	0.71	<0.05
Pearson Chi-square Statistic	103.87	<0.05

⁴ There are only six tables since the information with respect to third year credits for students registered in first year in 1996 and 1997 is not available in the current data set.

YEAR : 1994, valid cases: 318		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.53	<0.05
Spearman correlation coefficient	0.54	<0.05
Kendall's Tau	0.44	<0.05
Somers' D Statistics	X/Y=0.46 Y/X=0.43	<0.05 <0.05
Gamma Statistic	0.59	<0.05
Pearson Chi-square Statistic	101.08	<0.05

YEAR : 1995, valid cases: 340		
MEASURE OF ASSOCIATION	VALUE	P-VALUE
Pearson correlation coefficient	0.49	<0.05
Spearman correlation coefficient	0.51	<0.05
Kendall's Tau	0.41	<0.05
Somers' D Statistics	X/Y=0.41 Y/X=0.41	<0.05 <0.05
Gamma Statistic	0.52	<0.05
Pearson Chi-square Statistic	140.29	<0.05

In each of the six tables all the measures emerge as significant and are of similar magnitudes. Similar patterns to those described in 6.1.1 are observed here.

Thus, there is a significant relationship between the number of credits obtained in the first and those earned in the third year i.e. a student with a higher (*lower*) number of credits in the first year is more likely to obtain higher (*lower*) number of credits in the third year, respectively.

Important comments on the interpretation of results in section 6.1. and 6.2

The definitions of the variables used in the two sections may give rise some questions. Although, the definition of the academic performance in the first year is straightforward i.e. it is the equivalent number of first year courses completed by the incoming first year students, the definition of the performance in the second year may seem rather arbitrary. It is defined here as the equivalent number of courses completed by the incoming students in the second calendar year of their studies. For some students (especially those who obtain 2 or fewer credits) the credits may not necessarily be in respect of only second year but also first year courses, as they had to complete some of the curriculum required first year courses in the second calendar year of their studies. A similar comment applies to the number of credits obtained in third year. For some students the third year credits may be in respect of first, second and third year subjects.

All these difficulties of definition originate from the re-admissions rules in the Faculty and the fact that the data set in its current form does not distinguish between the credits obtained in different calendar years but with respect to different year subjects. Thus, it would, perhaps, be more appropriate to treat the presented relationships as the associations between credits obtained in the different calendar rather than academic years of students' studies. It may also be possible to extend the analyses to include all the previous academic history, in multiple regressions. For example third year performance may be modelled on the second, third and matric outcomes.

CHAPTER 7: GRADUATION ISSUES

In this chapter we investigate potential predictors of the duration ¹ of the Bachelor of Science (B.Sc.) degree, amongst those who actually graduate. The University would be interested in the dichotomy between completing and dropping out of a degree, and also in shortening the actual degree durations towards a minimum duration.

The statistical methods and models specified in the preceding chapters were applied to the subset of students for whom the graduation details could be reasonably retrieved ² from the current database.

Ordinal and binary definitions of the performance variable were used. The ordinal response variable was defined as the actual number of years taken to complete the degree. The binary response variable had the following two categories:

- a student took only three years to complete the degree
- a student took more than three years to complete the degree

In contrast to the preceding chapters, the results obtained under the measures of association, GLIM and FIRM are discussed in this chapter separately for the ordinal and binary response analyses and not under each technique for both analyses together. However details of the investigations still follow the manner specified ³ in Table 2.3 from chapter 2.

¹ The minimum number of years required to complete the B.Sc. degree is three. However, students take four, five or in exceptional circumstances up to six years to complete the degree.

Also note that the duration of the B.Sc. degree is yet another definition of the academic performance.

² The actual number of years taken to graduate was not directly available either from the original ADP source files or the designed unified data set file. The graduation details were obtained by manipulating the fields with promotion codes in particular years and the calendar year of first year registration. It was possible only to count the complete full calendar years taken by students to complete the degree.

Within the time frame allocated to the project it was not possible to establish what had happened to those students for whom the graduation indicator was not present. It is believed that they either left the Faculty due to academic or personal reasons or moved to other Faculties at UCT. We have therefore no further information on which to model their experience.

In respect of students who were registered in the first year in 1995, the graduation indicators were available only for those who completed the degree within three years. The remaining students were expected to have graduated in the subsequent years, outside the time frame of the current database. Thus, the data under analysis may be biased in favour of completing the degree within the three-year period.

Also the graduation indicators were not available for students who had registered in the first year in 1996 or 1997, as at most they could have had progressed to the second academic year in the time frame of the current database.

³ Refer to section 1.3 in Table 2.3 in chapter 2.

In CONFIRM, the following information, in the specified order, is given for each node:

- number of students in the node
- mean number of years taken by students grouped in the node to graduate
- standard deviation of mean number of years to graduate in the node
- maximum number of years taken to graduate in the node
- minimum number of years taken to graduate in the node

In CATFIRM, the following information, in the specified order, is given for each node:

- proportion of students who took only three years to graduate in the node
- proportion of students who took more than three years to graduate in the node
- number of students grouped in the node

As a complement to the descriptions of the most important features of dendrograms presented in this chapter, the reader may refer to those statistics for additional information and insight.

7.1 ORDINAL RESPONSE ANALYSIS

7.1.1 MEASURES OF ASSOCIATION

The data were arranged in two contingency tables formed individually for the faculty and for the matric point scores against the actual number of years taken to graduate.

The results are displayed in the following two tables.

FACULTY POINT SCORE vs. NUMBER OF YEARS TAKEN TO GRADUATE (ordinal) valid cases: 849	VALUE	P-VALUE
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	-0.33	<0.05
Spearman correlation coefficient	-0.34 (t=-10.58)	<0.05
Kendall's Tau	-0.28 (Z=-12.08)	<0.05
Somers' D Statistics	X/Y=-0.38 Y/X=-0.20	<0.05 <0.05
Gamma Statistic	-0.39 (Z=-12.08)	<0.05
Pearson Chi-square Statistic	260.24	<0.05

MATRIC POINT SCORE vs. NUMBER OF YEARS TAKEN TO GRADUATE (ordinal) valid cases: 849	VALUE	P-VALUE
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	-0.33	<0.05
Spearman correlation coefficient	-0.34 (t=-10.69)	<0.05
Kendall's Tau	-0.28 (Z=-12.30)	<0.05
Somers' D Statistics	X/Y=-0.38 Y/X=-0.21	<0.05 <0.05
Gamma Statistic	-0.40 (Z=-12.30)	<0.05
Pearson Chi-square Statistic	200.52	<0.05

In both tables all the measures are statistically significant. They are negative and moderately strong i.e. mostly greater 0.30. Similar interrelationships between the particular measures to those from chapter 3 emerge in these analyses. Thus, the explanatory comments made those sections apply also here.

The magnitudes of the measures for both the faculty and matric point scores are virtually identical (only minor differences for the Gamma and the Somer's *D* *Y/X* Statistics). This observation suggests that both types of point scores have similar predictive values in relation to the actual number of years taken to graduate.

The magnitudes of the measures are negative, as we might expect that the higher (*lower*) faculty or matric point scores are associated with the lower (*higher*) actual number of years taken to complete the B.Sc. degree.

7.1.2 GENERALISED LINEAR MODELLING TECHNIQUES

The ANOVA method was applied to identify others factors which taken alone or along with either the faculty or matric point score could further explain the differences in the duration of the B.Sc. degree amongst various student groupings. This analysis has treated the duration as a measure rather than a discrete number and is simply exploratory because the usual ANOVA assumptions are clearly invalid.

The ANOVA results together with the descriptive statistics table are displayed below.

ANOVA RESULTS: GRADUATION STATUS: GSTATUS							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1	52.20	814.00	0.42	124.36	0.0000
{1}EDUCATION	Fixed	10	6.36	814.00	0.42	15.14	0.0000
{2}LANGUAGE	Fixed	12	0.77	814.00	0.42	1.83	0.0405
{3}COLOUR	Fixed	3	1.22	814.00	0.42	2.90	0.0344
{4}GENDER	Fixed	1	0.10	814.00	0.42	0.24	0.6224

ANOVA RESULTS: GRADUATION STATUS: GSTATUS							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	51.99	814.00	0.42	123.55	0.0000
{1}EDUCATION	Fixed	10	6.29	814.00	0.42	14.95	0.0000
{2}LANGUAGE	Fixed	12	0.77	814.00	0.42	1.82	0.0409
{3}COLOUR	Fixed	3	1.22	814.00	0.42	2.89	0.0345
{4}GENDER	Fixed	1	0.06	814.00	0.42	0.14	0.7075

DESCRIPTIVE STATISTICS: GRADUATION STATUS							
	LEVEL OF FACTOR	N	MEAN GSTATUS	STD.DEV GSTATUS	STD. ERROR. GSTATUS	-.95% CNF. LMT	+95% CNF LMT
TOTAL		842	3.49	0.75	0.03	3.44	3.54
{1}EDUCATION	CO	148	3.72	0.82	0.07	3.59	3.86
{1}EDUCATION	JM	28	3.71	0.94	0.18	3.35	4.08
{1}EDUCATION	CA	399	3.31	0.59	0.03	3.25	3.36
{1}EDUCATION	NE	29	3.69	0.66	0.12	3.44	3.94
{1}EDUCATION	IN	24	4.04	1.16	0.24	3.55	4.53
{1}EDUCATION	TR	76	3.26	0.57	0.07	3.13	3.39
{1}EDUCATION	ET	50	4.48	0.79	0.11	4.26	4.70
{1}EDUCATION	NA	53	3.17	0.38	0.05	3.07	3.27
{1}EDUCATION	OF	7	3.29	0.49	0.18	2.83	3.74
{1}EDUCATION	TK	7	4.14	0.69	0.26	3.50	4.78
{1}EDUCATION	IE	21	3.14	0.36	0.08	2.98	3.31
{2}LANGUAGE	EA	37	3.62	0.79	0.13	3.36	3.89
{2}LANGUAGE	EN	701	3.39	0.66	0.02	3.34	3.43
{2}LANGUAGE	AF	27	3.78	1.01	0.19	3.38	4.18
{2}LANGUAGE	ZU	12	4.25	0.97	0.28	3.64	4.86
{2}LANGUAGE	TW	10	4.70	0.67	0.21	4.22	5.18
{2}LANGUAGE	XH	26	4.19	0.80	0.16	3.87	4.52
{2}LANGUAGE	NS	4	3.75	0.50	0.25	2.95	4.55
{2}LANGUAGE	SS	11	4.45	0.93	0.28	3.83	5.08
{2}LANGUAGE	SW	1	6.00				
{2}LANGUAGE	VE	4	4.50	1.29	0.65	2.45	6.55
{2}LANGUAGE	CH	5	3.40	0.89	0.40	2.29	4.51
{2}LANGUAGE	TG	3	4.00	0.00	0.00	4.00	4.00
{2}LANGUAGE	ND	1	3.00				
{3}COLOUR	C	159	3.70	0.81	0.06	3.58	3.83
{3}COLOUR	W	565	3.28	0.57	0.02	3.24	3.33
{3}COLOUR	I	45	3.96	0.98	0.15	3.66	4.25
{3}COLOUR	B	73	4.30	0.86	0.10	4.10	4.50
{4}GENDER	M	476	3.51	0.78	0.04	3.44	3.58
{4}GENDER	F	366	3.45	0.71	0.04	3.38	3.53

In the preceding ANOVA tables all the explanatory variables, except for gender, are statistically significant. Thus, the important differences in the number of years taken to graduate emerge between the respective categories of these variables. However both the faculty and matric point scores have the largest F-statistics and hence are again the dominating effects in the respective models. The educational department variable with the second largest F-statistic also appears to have a strong secondary influence in the model.

Upon examination of the descriptive statistics tables we observe that the students who graduate in the Faculty take on average 3.49 years to complete the degree. The student groupings with the shortest duration of the degree in each of the significant explanatory variables are respectively, from the Natal Educational Department (3.17), speaking English at home (3.39) and White (3.38). These categories refer to some of the formerly (and currently) privileged population groups. The means of the categories representing the disadvantaged student groups are mostly greater than 4. The confidence intervals are narrow for large, well-represented categories and wider for the smaller groups with only few students, as is to be expected.

7.1.3 INFERENCE-BASED RECURSIVE MODELLING

The same set of the predictor variables as in the previous CONFIRM analyses was used in this binary response analysis.

Upon examination of the dendrogram we note that:

at the first analyses level (node 1)

- The *colour* variable splits the data set most significantly and into three nodes with the mean duration increasing in the particular categories in the following manner: White, Coloured and Indian combined and Black.

at the second analyses level (nodes 2,3 and 4)

- For the White students, the differences in the average time until graduation may be attributed further to the particular symbol for the *matric higher- grade English*. Those with A and B complete the degree more quickly (3.16) than those with C, D or E (3.38).
- On the other hand, for the Coloured and Indian students the differences in the average time until graduation may be attributed further to the particular symbol for the *matric higher- grade biology*. Those with A and B complete the degree more quickly (3.60) than those with C or D (4.16).
- For Black students the actual faculty point score may serve as a good predictor of the duration with those with a point score of 45 and less taking a considerably longer time to graduate (4.76) than those with any higher point score (3.80).

Somewhat contrary to the results emerging from the ordinal response analysis in sections 7.1.1 and 7.1.2, the CONFIRM analysis has found the colour variable above both of the point scores to contrast the different student groupings with respect to the average duration of the undergraduate degree.

7.2 BINARY RESPONSE ANALYSIS

7.2.1 MEASURES OF ASSOCIATION

The data were arranged in two contingency tables formed individually for the faculty and for the matric point scores against the graduation performance variable with two categories:

- a student took only three years to graduate
- a student took more than three years to graduate.

The results are displayed in the following two tables below.

FACULTY POINT SCORE vs. NUMBER OF YEARS TAKEN TO GRADUATE (binary) valid cases: 849	VALUE	P-VALUE
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	-0.33	<0.05
Spearman correlation coefficient	-0.33 (t=-10.15)	<0.05
Kendall's Tau	-0.27 (Z=-11.09)	<0.05
Somers' D Statistics	X/Y=-0.40 Y/X=-0.19	<0.05 <0.05
Gamma Statistic	-0.41 (Z=-11.09)	<0.05
Pearson Chi-square Statistic	138.33	<0.05

MATRIC POINT SCORE vs. NUMBER OF YEARS TAKEN TO GRADUATE (binary) valid cases: 849	VALUE	P-VALUE
MEASURE OF ASSOCIATION		
Pearson correlation coefficient	-0.34	<0.05
Spearman correlation coefficient	-0.34 (t=-10.46)	<0.05
Kendall's Tau	-0.28 (Z=-12.42)	<0.05
Somers' D Statistics	X/Y=-0.41 Y/X=-0.20	<0.05 <0.05
Gamma Statistic	-0.43 (Z=-12.42)	<0.05
Pearson Chi-square Statistic	122.66	<0.05

In both tables all the measures are statistically significant. They are negative and moderately strong i.e. mostly greater than 0.30.

As with the ordinal response analyses, the magnitudes of the measures for the faculty and matric point scores are similar.

We may conclude that a student either with the higher faculty or matric point score is more likely to have completed the B.Sc. degree within three years.

Although the measures calculated in the ordinal and binary response analyses are of very similar magnitudes, the binary magnitudes are marginally but not meaningfully higher.

7.2.1 GENERALISED LINEAR MODELLING TECHNIQUES

The ANOVA method was used as exploratory to identify additional factors which taken alone or along with either the faculty or the matric point scores might describe distinct characteristics of those students who took more than three years to graduate in contrast to those students who took only three years to graduate.

The ANOVA results together with the descriptive statistics table are displayed below.

ANOVA RESULTS: BINARY GRADUATION STATUS: BGSTATUS							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
FACPOINT	Fixed	1	20.55	814.00	0.18	113.63	0.0000
{1}EDUCATION	Fixed	10	2.11	814.00	0.18	11.64	0.0000
{2}LANGUAGE	Fixed	12	0.08	814.00	0.18	0.47	0.9341
{3}COLOUR	Fixed	3	1.00	814.00	0.18	5.53	0.0009
{4}GENDER	Fixed	1	0.02	814.00	0.18	0.10	0.7468

ANOVA RESULTS: BINARY GRADUATION STATUS: BGSTATUS							
FACTORS	EFFECT (F/R)	DF EFFECT	MS EFFECT	DF ERROR	MS ERROR	F	P
MATPOINT	Fixed	1	22.29	814.00	0.18	123.84	0.0000
{1}EDUCATION	Fixed	10	2.01	814.00	0.18	11.17	0.0000
{2}LANGUAGE	Fixed	12	0.08	814.00	0.18	0.44	0.9466
{3}COLOUR	Fixed	3	1.00	814.00	0.18	5.54	0.0009
{4}GENDER	Fixed	1	0.01	814.00	0.18	0.04	0.8416

DESCRIPTIVE STATISTICS: BINARY GRADUATION STATUS: BGSTATUS							
	LEVEL OF FACTOR	N	MEAN BGSTATUS	STD.DEV BGSTATUS	STD. ERROR. BGSTATUS	-.95% CNF. LMT	+.95% CNF LMT
Total		842	0.36	0.48	0.02	0.32	0.39
{1}EDUCATION	CO	148	0.53	0.50	0.04	0.45	0.61
{1}EDUCATION	JM	28	0.43	0.50	0.10	0.23	0.62
{1}EDUCATION	CA	399	0.25	0.43	0.02	0.20	0.29
{1}EDUCATION	NE	29	0.59	0.50	0.09	0.40	0.78
{1}EDUCATION	IN	24	0.50	0.51	0.10	0.28	0.72
{1}EDUCATION	TR	76	0.21	0.41	0.05	0.12	0.30
{1}EDUCATION	ET	50	0.92	0.27	0.04	0.84	1.00
{1}EDUCATION	NA	53	0.17	0.38	0.05	0.07	0.27
{1}EDUCATION	OF	7	0.29	0.49	0.18	-0.17	0.74
{1}EDUCATION	TK	7	0.86	0.38	0.14	0.51	1.21
{1}EDUCATION	IE	21	0.14	0.36	0.08	-0.02	0.31
{2}LANGUAGE	EA	37	0.46	0.51	0.08	0.29	0.63
{2}LANGUAGE	EN	701	0.30	0.46	0.02	0.27	0.33
{2}LANGUAGE	AF	27	0.44	0.51	0.10	0.24	0.64
{2}LANGUAGE	ZU	12	0.75	0.45	0.13	0.46	1.04
{2}LANGUAGE	TW	10	1.00	0.00	0.00	1.00	1.00
{2}LANGUAGE	XH	26	0.81	0.40	0.08	0.65	0.97
{2}LANGUAGE	NS	4	0.75	0.50	0.25	-0.05	1.55
{2}LANGUAGE	SS	11	0.82	0.40	0.12	0.55	1.09
{2}LANGUAGE	SW	1	1.00				
{2}LANGUAGE	VE	4	0.75	0.50	0.25	-0.05	1.55
{2}LANGUAGE	CH	5	0.20	0.45	0.20	-0.36	0.76
{2}LANGUAGE	TG	3	1.00	0.00	0.00	1.00	1.00
{2}LANGUAGE	ND	1	0.00				
{3}COLOUR	C	159	0.52	0.50	0.04	0.44	0.59
{3}COLOUR	W	565	0.23	0.42	0.02	0.20	0.26
{3}COLOUR	I	45	0.60	0.50	0.07	0.45	0.75
{3}COLOUR	B	73	0.82	0.39	0.05	0.73	0.91
{4}GENDER	M	476	0.37	0.48	0.02	0.32	0.41
{4}GENDER	F	366	0.34	0.47	0.02	0.29	0.39

In the ANOVA tables, in addition to the faculty or matric point scores (representing the dominating effects in the respective models), the educational department and the colour variables emerge as statistically significant. Thus, the particular attributes of these variables define the specific characteristics of the students in the two graduation categories.

Due to the particular way in which the graduation performance variable was coded ⁴, the third table reports as means the proportions of students in the specific categories of the explanatory variables who took more than three years to complete the degree. Out of all the students who graduated in the Faculty some 36% had completed the undergraduate degree in more than 3 years. The categories with the highest proportions are respectively, from the Education and Training Department (0.92), speaking Southern Sotho at home (0.82) and Black (0.82). The proportions in the categories representing the formerly privileged population groups have very small proportions. The confidence intervals are narrow for large, well-represented categories and wider for the smaller groups with only few students, as generally expected.

Essentially the ANOVA results of both the ordinal and binary response analyses indicate that students from the disadvantaged communities take longer to graduate.

⁴ The variable was assigned 1 if a student took more than three years to graduate and assigned 0 if a student took only three years to graduate.

7.2.2 INFERENCE-BASED RECURSIVE MODELLING

The same set of the predictor variables as in the previous CATFIRM analyses was used in the current ordinal response investigation.

Upon examination of the dendrogram we note that:

at the first analyses level (node 1)

- The *colour* variable splits the data set most significantly and into three nodes with the proportion of students graduating within three years decreasing through the nodes: White, Coloured and Indian combined and Black.

at the second analyses level (nodes 2,3 and 4)

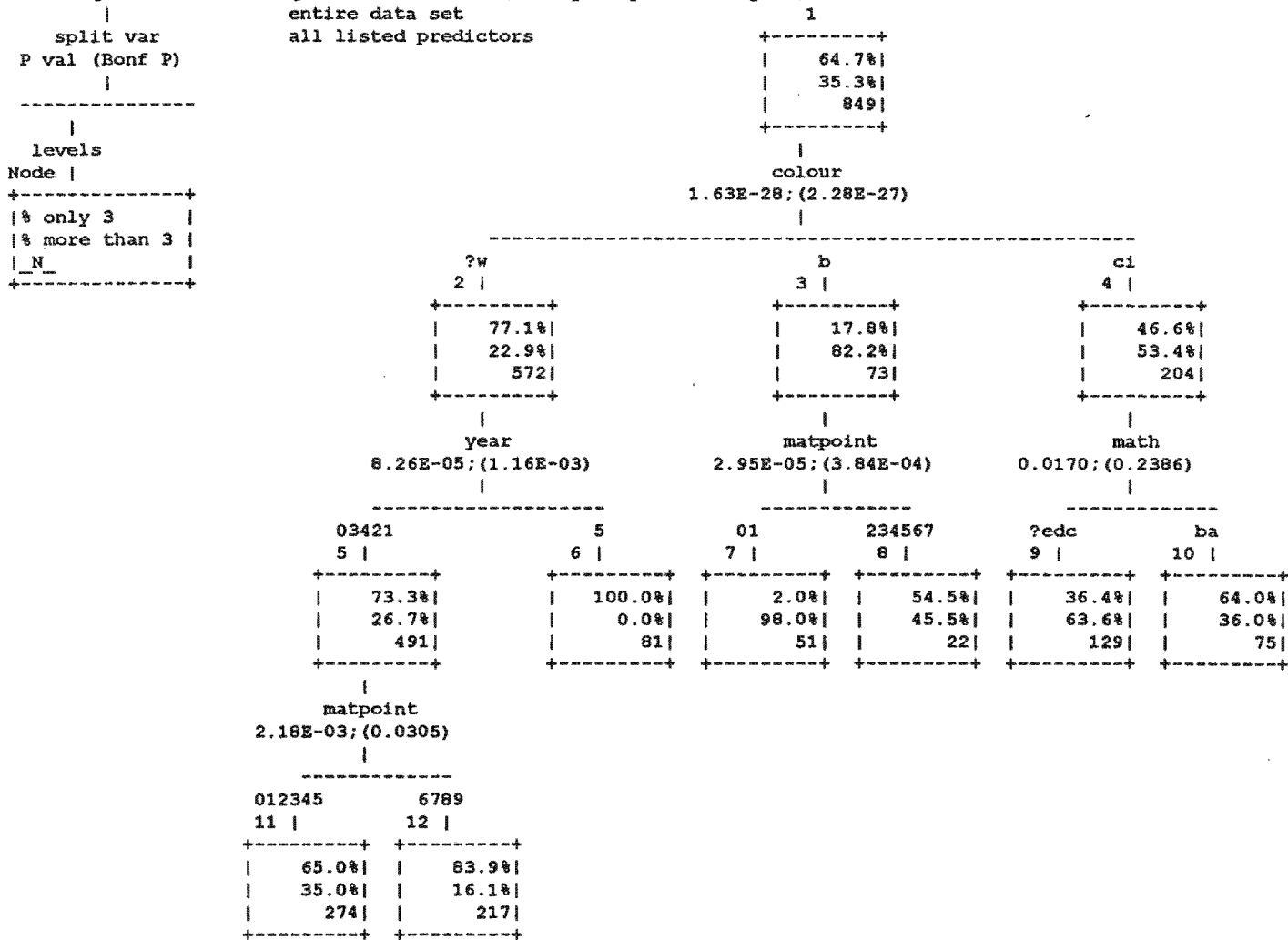
- For the White students there are significant differences between those who were registered in the first year in before 1995 and those registered in 1995. This feature is just an artefact of the database since the graduation details for those students who registered in 1995 and took longer than three years to graduate were outside of the time frame of the current database. Thus, the calendar year split can be omitted.

For those registered before 1995, the proportion completing the degree in three years is substantially higher for those with the *matric point score* of 41 or more (0.839) than those with any lower matric point score (0.65).

- For the Coloured and Indian students the proportion completing the degree in three years is substantially higher for those with A or B on the *matric higher-grade maths* (0.64) than those with C, D or E (0.36)
- For Black students the proportion completing the degree in three years is substantially higher for those with the *matric point score* of 36 or more (0.545) than those with any lower matric point score (0.02)

The CATFIRM analysis has chosen the colour variable above all other predictors. Then it has found further explanatory value for only the matric point score within the White and Black students.

Dendrogram of CATFIRM: graduation status (binary response analysis)



7.2.3 ODDS ON GRADUATING WITHIN MORE THAN THREE YEARS

The results of the previous sections indicated that the major contrast in the number of years taken to graduate arises between the race groups.

In order to contrast the apparent effect of faculty point score on the odds of graduating in three years, within race groups, separate logistic regression models⁵ have been fitted.

The estimates for parameters of the models and the odds calculated at the mean faculty point score and at the faculty point score of 48 for each race group, are displayed and discussed on the following pages.

⁵ The model fitted to each race group has the following structure:

$$\ln\left(\frac{\pi_i}{1-\pi_i}\right) = \alpha + \beta(x_i - \bar{x})$$

where

π_i - proportion of students graduating within more than three year with the i^{th} faculty point score

$1 - \pi_i$ - proportion of students graduating within three years with the i^{th} faculty point score

x_i - the i^{th} faculty point score

\bar{x} - mean faculty point score (for particular race group)

α - intercept term

β - effect of the faculty point score variable

Thus the odds may be expressed as

$$\left(\frac{\pi_i}{1-\pi_i}\right) = e^{\alpha + \beta(x_i - \bar{x})}$$

White Students

WHITE STUDENTS		
MODEL: LOGISTIC REGRESSION (LOGIT) N OF 0'S:435 1'S:130		
Final loss: 294.57981048 chi ² (1)=20.343 p=.00001		
PARAMETER	CONSTANT	FACULTY POINT SCORE
ESTIMATE	-1.2674	-0.0776
Standard Error	0.1051	0.0178
t(563)	-12.0599	-4.3694
P-LEVEL	0.0000	0.0000
-95%CL	-1.4739	-0.1125
+95%CL	-1.0610	-0.0427
Wald's chi-square	145.4403	19.0920
p-level	0.0000	0.0000
ODDS RATIO	0.2816	0.9253
-95%CL	0.2290	0.8936
+95%CL	0.3461	0.9582

FACULTY POINT SCORE	ODDS
52.93 (mean faculty point score)	0.28 (0.23 ; 0.35)
48	12.91 (9.62 ; 17.05)

At the mean faculty point score of 52.93, the odds are lower than unity, indicating that most of the White students, who graduated in the Faculty, took ⁶ only three years to complete the undergraduate degree. However at the cut-off point of 48, the odds are substantially greater than unity, indicating that the (below average) White students who just fulfilled the minimum admission requirement, were more likely to take than the minimum period to complete their degree.

⁶ We note again the results may be biased towards completing the degree within three years as described at the beginning of this chapter.

Black Students

BLACK STUDENTS		
MODEL: LOGISTIC REGRESSION (LOGIT) N OF 0'S:13 1'S:60		
Final loss: 22.126116853 Chi²(1)=24.145 p=.00000		
PARAMETER	CONSTANT	FACULTY POINT SCORE
ESTIMATE	3.2011	-0.4455
Standard Error	0.7266	0.1251
t(71)	4.4058	-3.5622
P-LEVEL	0.0000	0.0007
-95%CL	1.7524	-0.6948
+95%CL	4.6498	-0.1961
Wald's Chi-square	19.4110	12.6895
p-level	0.0000	0.0004
ODDS RATIO	24.5593	0.6405
-95%CL	5.7682	0.4992
+95%CL	104.5661	0.8219

FACULTY POINT SCORE	ODDS
43.13 (mean faculty point score)	24.55 (5.77; 104.57)
48	2.81 (0.83 ;9.50)

At the mean faculty point score, substantially lower than the minimum admission requirement, the odds are greater than unity. This finding indicates that most Black students, who graduated in the Faculty, took more than three years to complete the undergraduate degree. At the cut-off faculty point score of 48, the odds are lower than at the previous point score but are still greater than unity. This result shows that even those Black students, who fulfilled the minimum admission requirement, generally complete the degree over more than three years. We may note that the established confidence intervals for the odds are very wide.

Coloured Students

COLOURED STUDENTS		
MODEL: LOGISTIC REGRESSION (LOGIT) N OF 0'S:77 1'S:82		
Final loss: 105.50544157 Chi ² (1)=9.2527 p=.00235		
PARAMETER	CONSTANT	FACULTY POINT SCORE
ESTIMATE	0.22715	-0.10489
Standard Error	0.17248	0.03569
t(157)	1.31697	-2.93913
P-LEVEL	0.18977	0.00379
-95%CL	-0.11353	-0.17538
+95%CL	0.56783	-0.03440
Wald's Chi-square	1.73441	8.63847
p-level	0.18786	0.00329
ODDS RATIO	1.25502	0.90043
-95%CL	0.89268	0.83914
+95%CL	1.76443	0.96618

FACULTY POINT SCORE	ODDS
49.75	1.25 (0.89; 1.76)
48	1.51 (1.09; 2.08)

At the cut-off faculty point score of 48, the odds are greater than unity, indicating that the Coloured students, who just fulfilled the minimum admission requirement and graduated in the Faculty, took more than three years to complete their degree. The mean faculty point score for the Coloured students is not very much greater than 48, the odds are still greater than unity.

Indian Students

INDIAN STUDENTS		
MODEL: LOGISTIC REGRESSION (LOGIT) N OF 0'S:18 1'S:27		
Final loss: 23.967519906 Chi ² (1)=12.636 p=.00038		
PARAMETER	CONSTANT	FACULTY POINT SCORE
ESTIMATE	1.0871	-0.1905
Standard Error	0.4430	0.0620
t(43)	2.4542	-3.0747
P-LEVEL	0.0182	0.0037
-95%CL	0.1938	-0.3154
+95%CL	1.9804	-0.0655
Wald's Chi-square	6.0233	9.4535
p-level	0.0141	0.0021
ODDS RATIO	2.9657	0.8266
-95%CL	1.2139	0.7295
+95%CL	7.2458	0.9366

FACULTY POINT SCORE	ODDS
50.86	2.96 (1.21; 7.24)
48	5.11 (2.29; 11.36)

At the cut-off faculty point score of 48, the odds are greater than unity. This finding shows that the Indian students, who just fulfilled the minimum admission requirement, took more than three years to graduate. Despite the fact that the mean faculty point score of the Indian students is greater than 48, the odds are still greater than unity, again reflecting the fact that most of the Indian students took more than three years to complete the degree. We again note that the established confidence intervals for the odds are rather wide.

CHAPTER 8: CONCLUSIONS AND RECOMMENDATIONS

Examination of academic performance at a tertiary institution is a non-trivial and decision-intensive task, as it requires inspection of many aspects of an extensive research problem.

Firstly, the issues of primary interest and particular importance must be carefully specified and isolated. Secondly, the various ways of representing and defining the problems at issue must be identified. Thirdly, out of the many factors related to different aspects of an individual's life and partially determining academic performance, only some can be meaningfully explored in statistical analyses. Additionally, there ought to exist a reliable, easily accessible and manageable data source to facilitate the necessary statistical analyses.

Despite these complications, the analyses in the current study have identified, on the retrospective basis, several factors which may have assisted in distinguishing between potentially successful and unsuccessful students in some of the identified definitions of academic performance in the Faculty of Science at UCT.

8.1 SUMMARY OF MOST IMPORTANT FINDINGS

The current findings essentially do not provide any additional or significant insights into the issues of academic performance in the Faculty but rather decisively confirm what has been observed and anticipated all along. The levels of academic performance are distinctly varied across the racial groups. The contrasts have not only been expressed in terms of the colour variable but also the variables synonymous in the South African context with race i.e. the educational department, the home language or even the sub-standard matric results as summarised either by lower matric or lower faculty point scores.

Generally, the performance of the White students is markedly better than that of the non-White students. These contrasts¹ are conventionally attributed to the resource allocation discrepancies and the quality of tuition provided to the White and the non-White sectors of the South African educational system in the past.

Thus acknowledging the differences in the level of performance amongst the four race groups, we briefly summarise the other significant findings as discussed in the result chapters.

Higher point scores (both matric and faculty) are associated with better academic performance in first years. The matric point score has emerged as a more efficient and powerful single predictor than the faculty point score in determining the academic performance of the incoming first year students in the first year of their studies for the entire set of students, which was largely dominated by the White students. Separate analyses have shown, however, the faculty point score may serve as a better single predictor for the non-White students alone.

The educational departments have emerged as not comparable with respect to their apparent predictive value with respect to the number of credits. Prediction patterns² vary across the educational departments. This variation could have arisen from the differences in educational departments such as differences in resources or difference in examination standards.

Also important interdependencies have been indicated amongst the categories of the variables such as educational department or home language and the levels of academic performance for the incoming first year students. The students from the disadvantaged educational departments or speaking at home one of the African languages performed significantly worse than their colleges from the privileged educational departments or speaking at home one of the "White" languages.

No common predictors have been found to predict performance across the investigated university courses offered at first year level. Both of the point scores

¹ Thus we may further say that the problem at stake is not only that of identifying the reliable predictors of success or failure at an academic institution but also important considerations of how the observed gaps could be meaningfully and satisfactorily patched up.

² This finding should also form part of the debate on the standard of examination results across the educational departments.