University of Cape Town

CIV5000W

Research Project

A Spatial-Temporal Geo-spatial Database Model For Transport Justice. Case Study: Cape Town, South Africa

Author:
Brent Kotzee
Supervisor:
Mark Zuidgeest

Student Number:
KTZBRE001
August 15, 2024

Declaration

I, Brent Kotzee, declare that:

1. This report and the work presented in it, is my own.

2. I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.

3. These calculations/report/plans are my own work.

4. I have not allowed and will not allow anyone to copy my work with the intention of passing it off as his or her own work.

Signature: Signed by candidate

Date: 15/08/2024

ETHICS APPLICATION FORM

Please Note:
Any person planning to undertake research in the Faculty of Engineering and the Built Environment (EBE) at the University of Cape Town is required to complete this form **before** collecting or analysing data. The objective of submitting this application *prior* to embarking on research is to ensure that the highest ethical standards in research, conducted under the auspices of the EBE Faculty, are met. Please ensure that you have read, and understood the **EBE Ethics in Research Handbook** (available from the UCT EBE, Research Ethics website) prior to completing this application form: http://www.ebe.uct.ac.za/ebe/research/ethics1

APPLICANT'S DETAILS		
Name of principal researcher, student or external applicant		Brent Kotzee
Department		Civil Engineering
Preferred email address of applicant:		KTZBRE001@myuct.ac.za
If Student	Your Degree: e.g., MSc, PhD, etc.	MSc in Engineering in Civil Engineering
	Credit Value of Research: e.g., 60/120/180/360 etc.	180
	Name of Supervisor (if supervised):	Prof. Mark Zuidgeest
If this is a researchcontract, indicate the source of funding/sponsorship		Masakh'iSizwe Bursary Programme
Project Title		Analysis of Various Modes of Transport to Form A Spatial-Temporal Geospatial Database Model For Transport Justice Analysis

I hereby undertake to carry out my research in such a way that:

- there is no apparent legal objection to the nature or the method of research; and
- the research will not compromise staff or students or the other responsibilities of the University;
- the stated objective will be achieved, and the findings will have a high degree of validity;
- limitations and alternative interpretations will be considered;
- the findings could be subject to peer review and publicly available; and
- I will comply with the conventions of copyright and avoid any practice that would constitute plagiarism.

APPLICATION BY	Full name	Signature	Date
Principal Researcher/ Student/External applicant	Brent Kotzee	Signed by candidate	12/03/2020
SUPPORTED BY	**Full name**	**Signature**	**Date**
Supervisor (where applicable)	MHP Zuidgeest		12/03/2020

APPROVED BY	Full name	Signature	Date
HOD (or delegated nominee) Final authority for all applicants who have answered NO to all questions in Section 1; and for all Undergraduate research (Including Honours).			
Chair: Faculty EIR Committee For applicants other than undergraduate students who have answered YES to any of the questions in Section 1.			

Acknowledgements

Firstly, I would acknowledge my advisor Prof. Mark Zuidgeest for his efforts in aiding the development for the required processes to complete this research project. I would like to thank WhereIsMyTransport for providing the GFTS data of the various transport services used in this research project. I also acknowledge my friends, Keegan Klein, Toufeeq Fisher, Jesse Dreyer, and Chester Isaacs my sister, Amy Kotzee, my father, Franklin Kotzee and my late mother, Gail Kotzee for standing behind me in order for me to complete this research project.

Abstract

This research project develops a database model for the analysis of spatial and temporal differences in potential accessibility and mobility across the various public transport modes in the context of Cape Town, South Africa. This database model provides the foundation on which further transport justice tools are to be created. These transport justice tools aim to assist with redressing shortcomings in transport planning and operations which branched from amongst others, the apartheid land use transportation planning model and the effects it has on people until today.

The ultimate goal of the research was to create a database model that served as the foundation for transport justice tools to use in transport justice analysis. The first tool was a Space-Time Cube model with potential accessibility, potential mobility and time being the three respective axes. This cube represents the three parameters of a particular TAZ region for specific modes of public transportation. This cube can be used to perform analysis on the TAZ regions of interest, e.g. to analyse and compare the levels of potential accessibility, potential mobility across time in the area. The second tool provides a two-dimensional Temporal Aggregation of the Space-Time Cube. This allows us to study particular hours of the day of interest for potential accessibility and mobility of all possible TAZ regions of interest for the City of Cape Town. Income and racial demographics and population data can be used with both of these transport justice tools to assist the analysis.

Having more information about the potential accessibility and mobility, especially temporally will assist the user in making better decisions for planning their use of public transportation systems and operations. The service provider can use this database model to identify disparities in the potential accessibility and mobility of transport services they are providing. The government could benefit from this research by identifying regions that have potential accessibility and mobility deficiencies, possibly due to poor transport planning of areas where previously disadvantaged people were displaced. In addition, the government can identify which regions require improvements in potential accessibility and mobility for example, by looking at population, race, and income data of the areas that fall within this

potential accessibility and mobility deficit.

Using the Temporal Aggregation Python tool created for this project showed Accessibility levels are low at 06h00, then increases at 10h00, then decreases at 16h00 again. This was due to taking operational hours of the opportunities into account in the accessibility calculation as well as how frequently the public transit operates in the peak and off-peak periods of the day. The PMI levels are high at 06h00, then decreases at 10h00, then increases at 16h00 again. This was due to how frequently the public transit operates in the peak and off-peak periods of the day. This indicates that a person would reach destinations much faster when there is an increase in the availability of a particular public transit mode in the off-peak hours.

Further analysis can be done on each mode of transport to analyse the potential accessibility and mobility both spatially and temporally. Deductions have been made about this analysis to determine the fairness of these modes of transport with respect to potential accessibility and mobility.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Racial Segregation In South Africa

The probability that two individuals drawn randomly from the country's groups are not from the same group is known as fractionalization (Alesina & Ferrara, 2005; Patsiurko, Campbell & Hall, 2011). Fractionalization ranges from 0 to 1, where 0 indicates that they belong to the same group and 1 implying that belong to different groups (Alesina & Ferrara, 2005; Patsiurko, Campbell & Hall, 2011). South Africa has a very heterogeneous population amongst other countries with regards to race and ethnicity (Ramutsindela, 1997; Posel, 2001; Rondganger, 2006). It has fractionalization scores with regards to ethnicity, linguistics, and religion of 0.7517, 0.8652, 0.8603 respectively (Fearon, 2003; Alesina, Devleeschauwer, Easterly, Kurlat & Wacziarg, 2003). It makes a place such as South Africa so unique with every race and culture contributing in their own way to the society (Ramutsindela, 1997).

Unfortunately, part of this heterogeneity is rooted in the era of apartheid, when people were segregated based on their racial classification (Van Den Berghe, 1966; Nengwekhulu, 1986). Each racial class was also assigned against a racial hierarchy, with the black population being the lowest and the white population being the highest in the racial hierarchy (Van Den Berghe, 1966; Nengwekhulu, 1986). The Indian and Coloured population groups remained to be ranked above the black population, but still ranked below the white population in the racial hierarchy of apartheid (Van Den Berghe, 1966).

This racial classification segregation ideology was funneled into various aspects of South Africa's society such as healthcare education, income, residential locations, and transportation (Van Den Berghe, 1966; Parnell, 1993). Each of these inequalities are discussed in the next section.

1.2 Inequalities In South Africa

1.2.1 Transportation Inequality

Accessibility is a measure of convenience and the ease of with which opportunities for services and activities can be reached (Pirie, 1979; Thompson et al, 2012). Most employment opportunities and facilities were situated close to urban centers (Mabin, 1992). Since the population of South Africa was racially segregated based on their racial groups, the non-white population had to overcome larger travel distances to access these employment opportunities and facilities as opposed to the white population as they were mainly located within the urban centers due to the group area act of 1950 (Group Areas Act, 1950). The services and infrastructure for transportation were largely developed and established in urban areas as compared to the underdeveloped regions of the displaced non-white population groups which had low accessibility levels (Christie & Collins, 1982). The non-white population had to spend more time and money with regard to transportation to have the same access to opportunities as compared to the white population (Christie & Collins, 1982).

Mobility is the ease of movement that a person can move through space (Cervero, 2013; Martens, 2017). Most of the non-white population could not afford a personal vehicle due to these population groups having low-income levels and were forced to use public transportation services as a cheaper alternative to participate in employment opportunities and facilities (Christie & Collins, 1982; Orcutt, 2019). Since the non-white population regions were underdeveloped concerning transportation services and infrastructure, these factors made it difficult for them to access transportation services and reach their desired destinations, resulting in low potential mobility levels (Christie & Collins, 1982). The black population groups had the worst levels of potential mobility as they were moved toward the urban peripherals (Christie & Collins, 1982). In the context of South Africa, the main public transportation modes were bus, rail, and MiniBus Taxi (MBT). Although most of these public transport modes were affordable to the low-income population, often the user experienced discomfort due to these transport modes being designed to accommodate large volumes of people. This discomfort has a negative effect on a person's quality of life, especially on the population that experiences large travel distances (Pirie, 2009). Further elaboration of transport modes is discussed in chapter 1.4.

1.2.2 Healthcare Inequality

During apartheid, the white population were segregated from the non-white populations within healthcare facilities (Coovadia et al., 2009; Delobelle, 2013). The right to access of healthcare facilities was geared towards the benefit of the white population as they had a larger amount and higher quality standards of these facilities closer to their residential locations (Coovadia et al., 2009; Delobelle, 2013). These same standards of healthcare facilities could not be compared to the rest of the non-white population as often far away from their residential locations and experienced poor standards of these healthcare facilities (Coovadia et al., 2009; Delobelle, 2013). The inequalities in healthcare facilities between racial groups during apartheid contributed towards racial tension between the white population and the non-white population (Delobelle, 2013).

1.2.3 Education Inequality

In 1953 the Bantu Education Act was introduced by the apartheid government to give the black population an inferior quality level of education as compared to the white population education standards (Bantu Education Act, No. 47 of 1953, 1953). This was done to oppress the black population by educating them just enough to only get labour jobs (Van Den Berghe, 1966; Giliomee & Mbenga, 2007; Ross, 2008). This inequality in education kept the black population helped reinforce the distribution of wealth within South Africa as labour jobs fell into the low-income job category (Giliomee & Mbenga, 2007; Ross, 2008).

1.2.4 Income Inequality

The working environment was favoured for the white population to have more professional and most senior positions as most of these jobs were made available only for them (Van Den Berghe, 1966; Parnell, 1993). People were paid according to their racial classification (Van Den Berghe, 1966; Parnell, 1993). Therefore, the white population were paid the most compared to any other racial groups doing the same job (Van Den Berghe, 1966; Nengwekhulu, 1986). This meant that most of the country's wealth was distributed among the white population (Parnell, 1993). The low-income jobs were left for the non-white population, where the black population was paid the least among the entire population groups (Van Den Berghe, 1966; Nengwekhulu, 1986). This made it difficult for the non-white population groups to progress financially in life (Nengwekhulu, 1986; Parnell, 1993)

1.2.5 Residential Racial Segregation

The main outcome of apartheid was to favour the white population in every aspect of living in South Africa's society and eliminate racial unity (van den Berghe, 1966). Enforcing laws of racial segregation was one of the ways the apartheid government caused a racial divide within the country (Mabin, 1992). The Prohibition of Mixed Marriages Act was introduced in 1949, which made it illegal for white people to marry non-white people. Couples that were involved in these mixed-race relationships could face prison time if convicted (Prohibition of Mixed Marriages Act, Act No. 55 of 1949, 1949). The Group Areas Act of 1950 made it possible to reclassify particular locations to only be occupied by singular race groups (Group Areas Act, No. 41 of 1950, 1950, Mabin, 1992).

This reclassification was mainly done in urban areas (Mabin, 1992). The best suitable locations with regards to existing facilities, infrastructure, and employment opportunities were allocated to the white populations irrespective if there were other race groups originally living in these urban locations (Mabin, 1992; Ramutsindela, 1997). The non-white population was forcefully removed from their original residence to accommodate the newly enforced apartheid laws (Mabin, 1992). The black and coloured population were allocated regions that were underdeveloped in addition to these locations being far away from urban centers (Mabin, 1992; Ramutsindela, 1997) The black population was displaced the furthest away from the urban centers, mainly towards the urban peripherals (Mabin, 1992; Ramutsindela, 1997) These locations had low levels of accessibility and potential mobility to opportunities (Ramutsindela, 1997).

These apartheid laws made it difficult for the black population to move closer to locations of employment opportunities and facilities as locations closer to the urban peripherals were declared for the black population only (Mabin, 1992). The property value also had a high correlation with which race lives in a particular area (Mabin, 1992). This meant that white areas were valued more than the non-white populations displaced locations (Mabin, 1992; Ramutsindela, 1997). This created yet another obstacle for the non-white population trying to move closer to the urban centers as most of the non-white population had low-income jobs and could not afford housing closer to the white population (Van Den Berghe, 1966; Parnell, 1993).

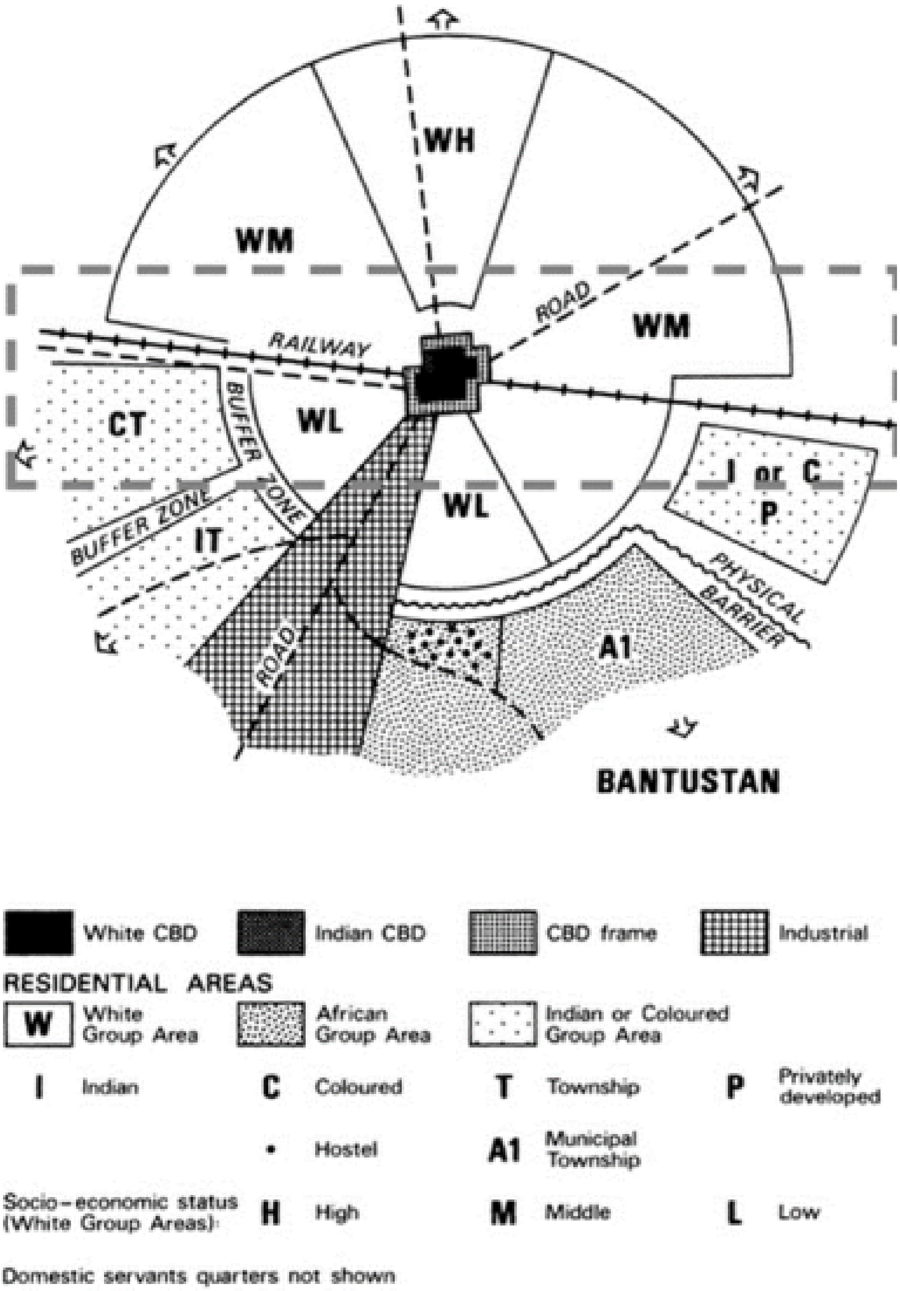


Figure 1.1: Apartheid city spatial structure (Simons, 1989)

The figure 1.1 shows the typical spatial planning structure of cities in South Africa during apartheid. As mentioned earlier the white population was moved the closest to the city center while the black population was moved the furthest away towards the urban peripherals (Mabin, 1992). The large travel distance was not the only thing the non-white population had to overcome due to this spatial structure of apartheid, they also needed to bypass barriers that separated the white population from the non-white population (Simons, 1989). These spatial obstacle barriers were open fields, physical barriers, and railway lines (Simons, 1989). These barriers largely hindered the accessibility and mobility of the non-white population as this spatial planning largely resulted in the non-white population all having to enter the city from the same roads (Simons, 1989).

Due to apartheid policies and education, income, and opportunities were less towards the non-white population. Therefore, income distribution would be different across racial groups. It would indicate that income and residential racial segregation is correlated (Bhorat, van der Westhuizen and Jacobs, 2009). This spatial income distribution with regard to race groups will be highlighted in a further chapter (chapter 4.3.4).

1.3 Indicators of Inequalities

The GINI Coefficient and the Theil Index will assist with highlighting income inequality and racial segregation respectively within South Africa. This information can assist with understanding the effects that the apartheid laws and spatial planning have on the current distribution of the various population groups within South Africa.

1.3.1 GINI Coefficient for South Africa

A measure of income inequality within a population of a country or any population group is known as a GINI coefficient (GINI, 1912). The range of the GINI coefficient is from 0 to 1, where 0 represents perfect income equality and 1 represents perfect inequality (GINI, 1912). The Lorenz curve determines the value of the GINI coefficient, which plots a curve using the cumulative population in question as the independent variable and the cumulative income of the population in question as its dependent variable (GINI, 1912). Both independent and dependent variables are normalised so that both variables range from 0 to 1 (GINI, 1912).

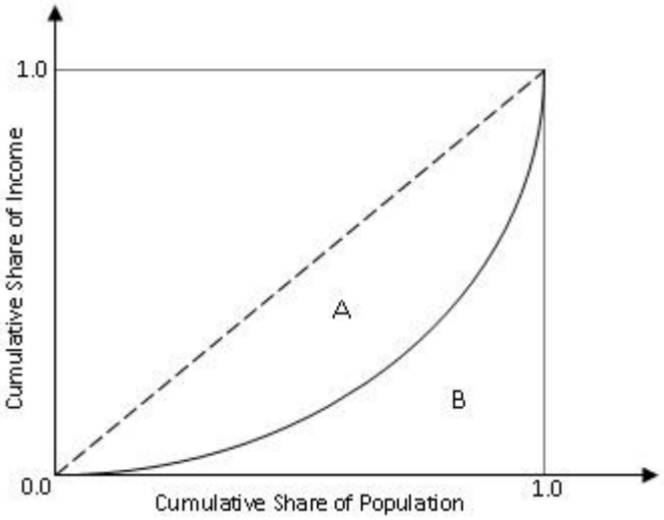


Figure 1.2: A Lorenz Curve Example (Bhorat, van der Westhuizen and Jacobs, 2009:9)

In figure 1.2, the GINI coefficient can be visually described as area A, which is the area below the dashed line, representing the perfect line of equality and above the Lorenz curve (Bhorat, van der Westhuizen and Jacobs, 2009). The further away the curve is from the line of perfect equality the larger the GINI coefficient will be and vice versa (Bhorat, van der Westhuizen and Jacobs, 2009). The higher the GINI coefficient, the higher the inequality of the distribution of income is within the country's population and racial groups (Theil, 1967; Leibbrandt et al., 2001).

By analysing the GINI coefficient we can find the extent of income inequality distribution among the population of South Africa in 1995 and 2005 due to prolonged inequality (Bhorat, van der Westhuizen and Jacobs, 2009).

Table 1.1: GINI coefficient of South Africa's Race Groups for 1995 and 2005 (Bhorat, van der Westhuizen and Jacobs, 2009:10)

Racial Groups	**1995**	**2005**
African	0.56	0.61
Coloured	0.49	0.59
Asian	0.46	0.56
White	0.44	0.51
Across The Population GINI Coefficient	0.64	0.72

Table 1.1 shows the GINI coefficient of the different race groups and the entire population of South Africa between 1995 and 2005 (Bhorat, van der Westhuizen and Jacobs, 2009). The year 1995 describes the income inequality with regard to race as a result of apartheid and 2005 describes the income inequality after apartheid (Bhorat, van der Westhuizen and Jacobs, 2009). As described in Table 1.1 the GINI coefficient values increased between 1995 and 2005 for all race groups and the entire population of South Africa (Bhorat, van der Westhuizen and Jacobs, 2009). This would indicate that the income inequality in South Africa increased after the apartheid regime was removed (Bhorat, van der Westhuizen and Jacobs, 2009).

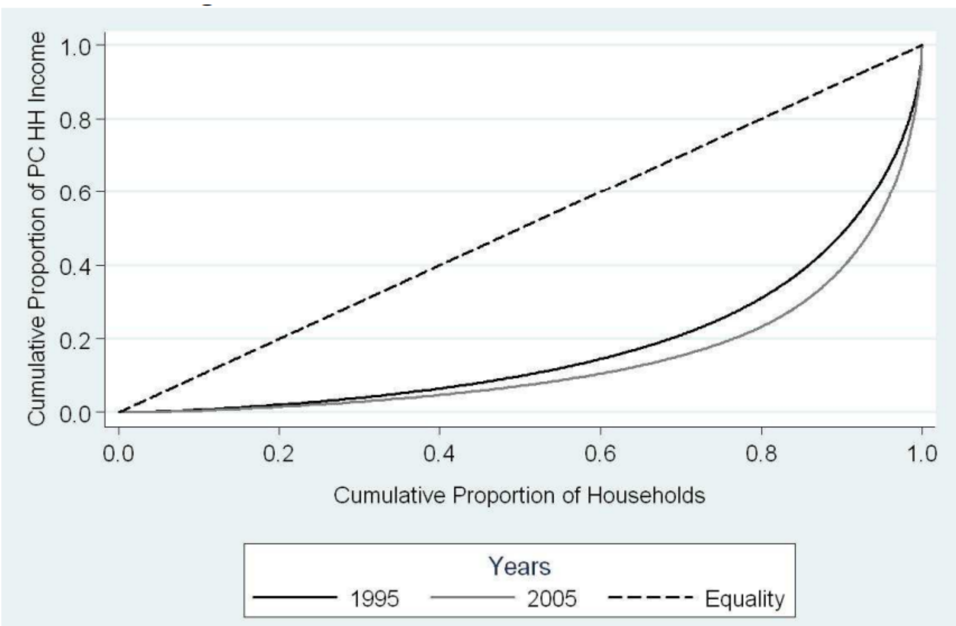


Figure 1.3: Lorenz curve of South Africa's Cumulative Population for 1995 and 2005 (Bhorat, van der Westhuizen and Jacobs, 2009:11)

Figure 1.3 shows the Lorenz curve of the population of South Africa (Bhorat, van der Westhuizen and Jacobs, 2009). Figure 1.3 shows the curve moving further away from the perfect line of equality between 1995 and 2005 (Bhorat, van der Westhuizen and Jacobs, 2009). The curve of 2005 indicates that a greater share of the population holds less income as compared to 1995 (Bhorat, van der Westhuizen and Jacobs, 2009).

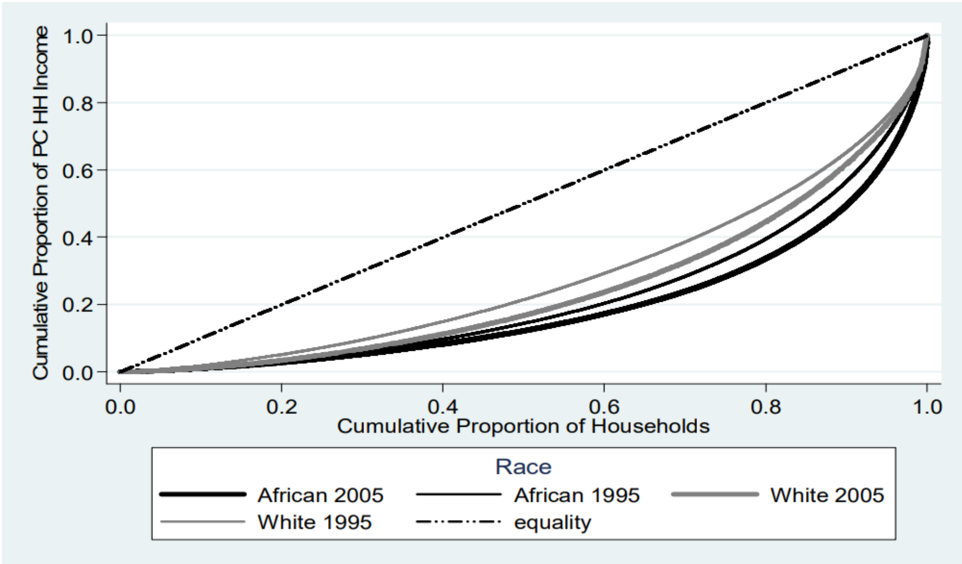


Figure 1.4: Lorenz curve of the White and African (Black) Population Groups of South Africa for 1995 and 2005 (Bhorat, van der Westhuizen and Jacobs, 2009:12)

Figure 1.4 shows the Lorenz curve of the white and African (black) population groups of South Africa (Bhorat, van der Westhuizen and Jacobs, 2009). Figure 1.4 shows the curve moving further away from the perfect line of equality for both population groups between 1995 and 2005 (Bhorat, van der Westhuizen and Jacobs, 2009). The curve of 2005 indicates that a greater share of the population holds less income as compared to 1995 for both population groups (Bhorat, van der Westhuizen and Jacobs, 2009). The black population experienced higher levels of inequality as compared to the white population(Bhorat, van der Westhuizen and Jacobs, 2009).

1.3.2 Theil Index for South Africa

The statistic mainly used to measure economic inequality and other economic phenomenon is known as the Theil index (Theil, 1967). the Theil index has also been used to measure racial segregation (Theil, 1967). This could indicate the inequality effects of racial discrimination concerning income due to apartheid within South Africa (Leibbrandt et al. 2001).

The Theil index has the advantage of separating the inequality that is between race groups and the inequality within race groups from the entire inequality of the country (Theil, 1967; Leibbrandt et al. 2001). The Theil-T statistic is defined as

$$T_T = T_B + \sum_{i=1}^{n} q_i T_i \tag{1.1}$$

with:

$$T_B = nq \sum_{i=1}^{n} T_i \tag{1.2}$$

where:

T_i measures the inequality within the ith group

q_i is the proportion of income accruing to the ith group

T_B measures the inequality between the different subgroups

T_B makes the assumption that that all the incomes within a group are the same (Leibbrandt et al. 2001).

Table 1.2: Theil Index by Race for South Africa 1995-2005 (Bhorat, van der Westhuizen and Jacobs, 2009:13)

By Race	**1995**	**2005**
Total Inequality (Theil-T)	0.87	1.14
Within	0.50	0.63
Between	0.37	0.51

Table 1.2 describes the Theil index income inequality of the different race groups within South Africa for the years 1995 and 2005 (Bhorat, van der Westhuizen and Jacobs, 2009). The year 1995 and 2005 describes the income inequality with regard to race one and 11 years after apartheid ended (Bhorat, van der Westhuizen and Jacobs, 2009). The Theil index components range from 0 to 1, where 0 means total equality and 1 describes total inequality (Theil, 1967; Leibbrandt et al., 2001). The higher the index, the higher the inequalities in the distribution of income are within the country's population and vice versa (Theil, 1967; Leibbrandt et al., 2001). As described in Table 1.2 the income inequality values have increased between 1995 and 2005 (Bhorat, van der Westhuizen and Jacobs, 2009). This would indicate that income inequalities regarding race groups have increased (Bhorat, van der Westhuizen and Jacobs, 2009).

1.4 Transportation Manifestations

The main public transportation modes are rail, bus, and MiniBus Taxi (MBT). The MBT originated in the 1950s and has a high correlation with the Groups Areas Act of 1950 (Behrens et al., 2016). The increased volume of people participating in public transport put an unexpected strain on the services they could provide to the public (Behrens et al., 2016). Many bus drivers lost their jobs after the 1955 municipal bus driver and boycott strikes (Mxolisis, 2006). These drivers started transporting people with their personal vehicles as the demand for public transportation services was still high and these drivers knew how the transportation network worked (Mxolisis, 2006). The vehicle size of transport services increased over the years to accommodate more passengers and resulted in the transport services that exist today (Mxolisis, 2006).

Over twenty years after apartheid laws were removed, the effects are still visible as a continuous spatial disparity between the location of work and where people live still has a significant financial strain on the non-white population up to today (Christopher, 2001). These factors contributes to the prevention of the non-white population from ascending financially and relocating closer to employment opportunities (Christopher, 2001).

With the locations of employment opportunities and facilities beginning and ending at similar times as peak travel periods throughout the day, heavy traveling congestion can occur during arriving at and departing from these locations (Christopher, 2001). The factors that vary within this situation are the amount of time and money people have time to spend to get from their homes to employment opportunities and facilities and travel back home again (Vuchic, 2005). Within recent times the rail transportation service has failed to facilitate the demand of the public due to its depreciating infrastructure which has led to decreased levels of accessibility and potential mobility of its users (Behrens et al., 2016). This lack of transportation alternatives leads to a sprawled city that does not have the infrastructure and services to facilitate the population that requires these facilities to reach their desired locations which are mainly the low-income population as they can not afford a personal vehicle and rely on public transportation services (Vuchic, 2005).

There are various limiting factors such as the cost, availability, accessibility, mobility, and operational times of the various transportation services (Christie & Collins, 1982; Orcutt, 2019). The MBT is regarded as being an informal public transportation service as it operates unscheduled and belongs to private business owners as opposed to the rail and bus transportation services which are scheduled and largely controlled by the government (Behrens et al., 2016). The MBT frequently operates during peak periods of the day and when there is a higher demand for their transportation services (Vuchic, 2005). When the demand for transit is low, during off-peak times, the MBT is unreliable as their service frequencies and provision are low (Vuchic, 2005). The rail and bus transportation services are more reliable than the MBT as they operate with a schedule concerning off-peak periods however, their frequencies in operations of services decrease which results in a decrease in accessibility and potential mobility of the transit for the transit users during these periods of the day (Behrens et al., 2016).

If the population groups that make use of public transit services would want to reach locations for employment opportunities and facilities during off-peak periods, they would most likely still have to travel during peak hours of the day as the transportation service frequencies and provision decreases during off-peak hours (Christie & Collins, 1982; Orcutt, 2019). Hence the importance of this research project's focus is to identify the various accessibility and mobility levels of various transport modes throughout the day within the city of Cape Town. The public transit users experience a decreased level of accessibility and potential mobility during off-peak hours as a result of the decrease in levels of access to these public transit services (Martens, 2017). Whereas someone with a personal vehicle that needs to be at the same location during off-peak hours could start traveling much closer to the time (Martens, 2017). Resulting in a reduced duration spent on traveling as compared to an individual who uses public transport (Martens, 2017). There are transport injustices in accessibility and mobility levels between a person with a vehicle and someone who depends on public transport (Martens, 2017). These injustices could be highlighted by analysing the accessibility and mobility levels throughout the day for the different races and income groups (Martens, 2017). Which this project aims to achieve.

Hence, the interest in metrics of injustices for transport as a result of the spatial, income, education, health care, and travel inequalities based on racial groups (Christie & Collins, 1982; Mabin, 1992; Orcutt, 2019). With the aid of geo-spatial time series data and geo-spatial data, it can become an opportunity to provide these metrics of transport injustices (Klopp et al., 2014, Martens, 2017).

1.5 Statement of the Problem

In an equal world, everyone should have equal access and availability to the various modes of public transport, but in reality this is not the case (Rawls, 1971).

Due to historical injustices in Cape Town, there is an inaccessibility and unavailability of the various public transport services at different locations throughout the day which leads to injustices of accessibility and mobility to opportunities (Christie & Collins, 1982; Martens, 2017; Orcutt, 2019).

Given the problems/disparities discussed above issues of justice of transport manifest themselves, both spatially, temporarily, and socially (for different groups). It is important that if we want to strive for a more just transport system and need indicators that speak to transport justice (DoT, 1996; Spatial Planning and Land Use Management Act, No. 16 of 2013, 2013).

A framework and database model for transport justice analysis in Cape Town is missing. Because of this, it is difficult to identify and quantify transportation disparities. To address transportation inequalities, there is a need for tools to provide such evidence. Hence, the main research problem in this thesis is how to develop a spatial-temporal geo-spatial database model for analysing transport justice, specifically looking at the various public transport services with respect to potential accessibility and mobility.

Using geo-spatial time series, geo-spatial, and non-spatial data will help in building the framework and database model for transport justice analysis in Cape Town by adopting Karel Marten's (Martens, 2017) Transport Justice theory.

1.6 Objective of the Study

The objective of this study is to develop and implement a concept, data model, and tools to facilitate spatial-temporal transport justice analysis within Cape Town. These transport justice analysis tools will then be tested within sub-regions of Cape Town to provide evidence that highlights the racial segregation and inequalities.

Transport justice analysis tools that were developed are a space-time cube (x, y, t, n) for the purpose of analysis with potential accessibility(x), potential mobility(y), and time of the day(t) being the three respective axes of this cube to represent the various modes of transport(n) at a given area. The second tool creates a temporal disaggregation of a two-dimensional scatter plot to represent the distribution of potential accessibility and mobility for all regions of Cape Town of the various modes of transport to represent a particular hour of the day. This will be done using GTFS

(General Transit Feed Specification) data, locational information, and census data within a GIS and coding environment.

Once these tools are created, the next step would be to implement and apply them to transport justice as another metric of transport injustice. This will be done by identifying a specific time of day and analysing the potential accessibility and mobility from origins to all possible destinations within a given time frame. These origins and destination pairs will be derived from the Transport Analyse Zones (TAZ) locations of Cape Town (Benenson et al., 2011). Each origin will be given a potential accessibility and mobility score based on methods of the cumulative sum of potential accessibility and mobility traveling to the various destinations for each transport mode. From here, it would be possible to plot the cumulative accessibility and mobility levels for all the regions of Cape Town and identify locations that fall below potential accessibility and mobility sufficiency thresholds. Based on these locations it could be determined if these locations are entitled to an increase in potential accessibility and mobility.

These transport justice analysis tools would allow us to study transport problems from a people and justice point of view.

1.7 Research Questions

This section will address the research questions that will be addressed in this research paper. The research questions for this investigation are listed below:

1. How can the potential accessibility and mobility of a region be quantified and visualised with respect to a public transport mode at various times of day?
2. How can GTFS data be used to model the potential accessibility and mobility of the various modes of public transport over time?
3. Do the current public transport system routes provide fair potential accessibility and mobility for the general public over the City of Cape Town?
4. Which time periods have the highest potential accessibility and mobility of the various modes of transport?
5. Which public modes of transport provide the highest potential accessibility and mobility?
6. Which locations have the least accessibility value for the various modes of transport?

7. What effect will potential mobility have on the accessibility of the various modes of transport?
8. Could this database model be used as a metric of transport justice for analysis?

1.8 Scope

This report only deals with 4 different modes of public transport, this being MyCiTi, MiniBus Taxis (MBT), Golden Arrow Bus Services (GABS) and MetroRail. The MyCiTi (2018), MBT(2018), GABS(2018) and MetroRail (2016) data dealt within this report will be over a time period within a year. Each respective year of the respective public transit is in brackets next to each public transit mode. Due to the time period of the available data, the temporal differences between years cannot be performed on the respective transport modes. Therefore, the effect on transport modes activities caused by changes in land use will not be monitored. The locations of these modes of transport will only be dealt with within the City of Cape Town.

1.9 Outline of Research Plan

This research plan was used to find solutions to the research questions identified in this report. The following workflow was chosen to create the transport justice database model and tools, this consists of the following steps:

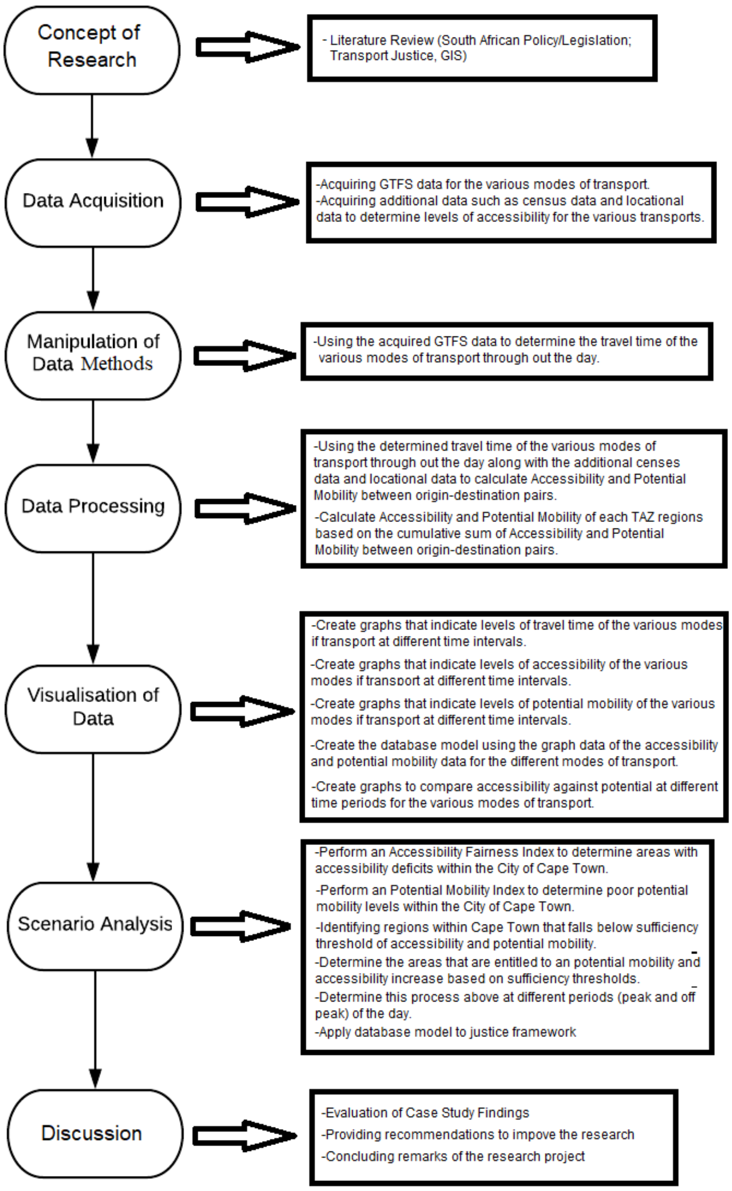


Figure 1.5: Work Flow Process of Research Plan

Figure 1.5 shows the workflow process of the research plan that achieved the outcomes for this research paper. The following headings of Concept of Research, Data Acquisition, Manipulation of Data, Data Processing, Visualisation of Data, Scenario Analysis, and Discussion will be further explained below within this section.

1.9.1 Concept of Research

The concept of research is to understand the problem of public transportation and the contributing factors that affect the potential accessibility and mobility of the population that makes use of public transit services. This was achieved by looking at previous work on public transport potential accessibility and mobility deficits, as well as additional literature that tries to implement an improvement of these deficits. In addition, identifies the negative connotations that these transport deficits have on the population that is experiencing these transport deficits and the reinforcements of these inequalities. This can be seen in chapter 3.

1.9.2 Data Acquisition

Based on the research review and the problems that were solved, various data sources were collected. Such data was aimed to quantify the potential accessibility and mobility levels of public transport. As mentioned before, the GTFS data for the various modes of public transport and other information such as census data was used to help solve these research problems. This can be seen in chapter 3.

1.9.3 Manipulation of Data Methods

The data acquired was then combined to give accurate values for potential accessibility and mobility of the various public transit services. The GTFS data was combined with a road network to form a network analyst layer to perform routing determinations. The main output of using the network analyst layer was to determine the travel time from origins to destinations. This was done in a GIS environment. This can be seen in chapter 4.

1.9.4 Data Processing

The travel time data was exported from the GIS environment and stored as txt files. The travel time data was combined with census opportunity data (i.e. working locations, education, and public facilities) so that the calculation of potential accessibility and mobility of the various public transit services could be determined. This was done in a Python coding environment. The cumulative potential accessibility and mobility levels for the various public transit services were determined for every

TAZ origin to all TAZ destinations for Cape Town. The potential accessibility and mobility levels were further refined by incorporating approximate operational times of the potential opportunities facilities, which affected the potential accessibility and mobility levels of public transit services throughout the day. This can be seen in chapter 4.

1.9.5 Visualisation of Data

Each TAZ location had its cumulative potential accessibility and mobility values determined for the various public transit modes throughout the day. The Space-Time cube tool was created and executed as it had all the necessary variables calculated to use the tool. Additional census data was attached to give more information about the population that resides in the TAZ of interest. Such additional information was about the population racial groups and income distribution of the TAZ of interest. This census data was displayed as bar graphs. The UDI (Urban Development Index) modal split of the public transport services was incorporated to produce a combined weighted public transportation score for potential accessibility and mobility of the TAZ of interest.

The Temporal Aggregation tool was then created and displayed a scatter plot of the potential accessibility and mobility of the combined weighted UDI modal split. This was done for the various public transit services of all the TAZ locations of Cape Town for a particular hour of the day. This can be seen in chapter 4.

1.9.6 Scenario Analysis

Additional graphs were used to display the number of TAZ and the share of the population that has potential accessibility and mobility deficits at various threshold values for Cape Town. The results obtained were then analysed and deductions were made of which regions were entitled to potential accessibility and mobility based on various criteria such as race, income, and the potential accessibility and mobility deficits within the various TAZ groups affected. This was done in the hopes of creating a new way to implement a framework towards transport justice. This can be seen in chapter 5.

1.9.7 Discussion

Finally, the results of this research were summarised. The implications of the results were discussed and recommendations on how to improve the research to obtain better results were presented. This can be seen in chapter 6.

1.10 Implications of Research

This project provides additional information on analysing the effect that the various modes of public transport have on potential accessibility and mobility which contributes to transport inequalities within the City of Cape Town. The land use patterns were highlighted based on the potential mobility and accessibility experienced by the TAZ regions within Cape Town. This information shows the fairness levels experienced by a user of the various modes of public transport and aids the government in creating a fairer system. In addition, this information would assist town planners in planning new routes to aid the increase of potential accessibility and mobility for the use of the various modes of public transport. This would lead to fairer potential accessibility and mobility experienced by a civilian regardless of their residential location. Ultimately, provides a measure for the aid of Transport Justice.

Chapter 2

Literature Review

This section describes the various factors and literature contributing to the importance, development, and implementation of the transport justice tools that were developed for this research project.

2.1 Land Use Transport Integration

Majority of developing countries and cities' economies that are developing focus on motorized and urban infrastructure development (UN Habitat, 2013). In the countries where this happens, accessibility is upheld to promote inclusively, equality and social sustainability in urban areas with regard to land use and transport resulting in income profits and efficiency (UN Habitat, 2013). When cities and residential locations are efficiently integrated with transportation infrastructure it can lead to sustainable mobility (UN Habitat, 2013). Taking into account the existing and potential developing land use, transport infrastructure, transport services, mobility, and accessibility is the objective of land-use and transport planning by integrating all these factors to promote social inclusion (UN Habitat, 2013). If existing locations have not made provision for potential expansion and begin to expand without the objective of land-use and transport planning, it may lead to unsustainable city growth and lower levels to access opportunities (Childers et al., 2014).

To have a sustainable city, there are certain factors that need to be kept in mind such as having mixed land use developments, medium densities in population, and promoting developments near locations that public transport services pass through so that it has high accessibility levels for the population (Banister, 2008). Integrating the land use development with transport polices will promote a decrease usage of a personal vehicle (Banister, 2008). Transport services change the shape and the land use form, while the land use affects the use and demand of the transport services, therefore the land use and transport are interdependent (Minter, 1997). While transport and land use are reliant on each other's success, often they are not taken into account in urban development leading to an unsustainable city with regards to mobility and accessibility (Minter, 1997).

People try to minimize their traveling distances to reach locations of opportunities, however, if traveling larger distances results in an increased level of accessibility they will compromise on the need to travel (Banister, 2007). Hence, sustainable mobility is necessary so that people can have optimal access to the land use through the available transport systems in cities (Banister, 2007).

2.2 Social Exclusion

Social exclusion deals with the inability and shortage of resources to participate in activities or normal relationships to which most of the population has access (Adato, Carter, May, et al., 2006; Lucas, 2012). There have been many studies that link social exclusion with transportation because transportation helps people access opportunities (Mackett, Achuthan & Titheridge, 2008). The role of transportation in overcoming social exclusion is crucial as the demand and supply of transportation need to be addressed because social exclusion is a complex situation (Church, Frost & Sullivan, 2000).

The demand aspect of transportation entails the accessibility to transportation systems and the effect it has on land use transportation planning. The supply aspect of transportation takes into account the level of accessibility to transportation systems and the quality of these systems. Keeping these factors of transportation in mind could show how social exclusion can be held fixed within a society (Currie & Delbosc, 2010; Lucas, 2012).

Infrastructure, economic, and social aspects of a society have been linked to social exclusion and transportation (Lucas, 2012). Social exclusion has been studied several times in most of the research indicates that multiple dimensions are linked to social exclusion (Lucas, 2012; Jianhong, Nesbitt, Daley, et al., 2016). The model provided by Lucas (2012) concisely depicts the various dimensions of social exclusion.

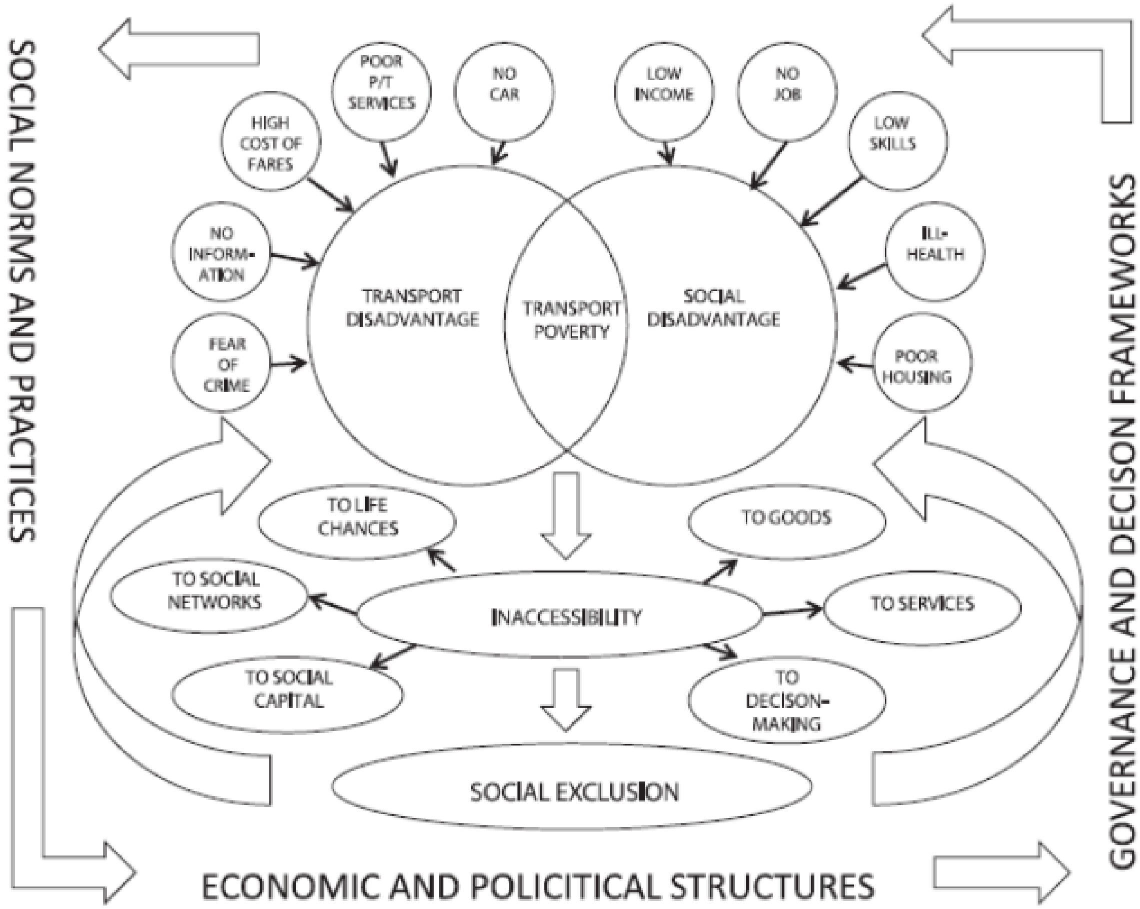


Figure 2.1: Linking social exclusion, social disadvantage and transport disadvantage (Lucas, 2012)

Figure 2.1 shows the model provided by Lucas (2012) that summarizes the various dimensions that are linked to social exclusion.

Dimensions that are related to transportation are mobility and accessibility as they are indicated to be the core reasons for poor levels of accessibility to opportunities and services which leads to social exclusion (Lucas, 2012). This mainly affects the low-income populations as their mobility is largely restricted and can not participate in what is considered normal opportunities or economic activities (Hernandez, Oviedo & Titheridge, 2016; Özkazanç, Nihan & Sönmez, 2017).

By looking at the diagram of Lucas (2012) it can be indicated how the role of transportation plays in social exclusion due to the unequal land use spatial planning of apartheid within South (Lucas, 2012; Schwanen, Lucas, Akyelken, et al., 2015). The apartheid spatial planning legacy still fundamentally continues and the inaccessibility to opportunities is largely felt by the low-income populations as they experience large traveling distances which result in higher traveling expenses, poor levels of accessibility to transportation, and poor living conditions which often make it unsafe for them to participate with the transportation networks provided (Lucas, 2011). In an attempt to provide low-cost housing for the poor population through policies, the South African government has reinforced the apartheid spatial planning as most of these housing locations are situated at the urban peripheries which have low accessibility and mobility levels resulting in these people being socially excluded (Turok, 2001).

One way these policies can be improved is to promote social inclusion when the policies are created to overcome the apartheid spatial planning patterns so that there is an increase in accessibility and mobility with respect to transportation (Stanley & Vella-brodrick, 2009).

2.3 Transportation Policies

This section aims to highlight the transport policies in South Africa to improve the past injustices that the country experienced as a particle result of apartheid with regard to accessibility and mobility.

2.3.1 White Paper on National Transport Policy (1996)

This is a transportation policy that promotes land transport planning for optimal traveling of people to improve on economic and social development of a country (DoT, 1996). Although this policy highlights various topics, some objectives apply to this research. Some of these objectives are the control of urban sprawl, improvement of accessibilities and mobility to opportunities for every population group, and the promotion of affordable and quality transportation systems (DoT, 1996). The policy tries to aim the land use transportation planning framework geared toward the users of public transport over those individuals that use a private vehicle (DoT, 1996). Overcoming social exclusion through the improvement of accessibility to transport systems and mobility are some of the primary objectives of this policy (DoT, 1996).

2.3.2 Spatial Planning and Land Use Management Act, No. 16 of 2013

The Spatial Planning and Land Use Management Act's (SPLUMA) main objectives are to make provision for fair distribution of spatial planning and land use management that works effectively for the country; offer land that is sustainable and makes efficient use of such land as well as to account of the spatial planning of the past to make sure that there is a fair distribution when there are further spatial development and land use management systems that are occurring (Spatial Planning and Land Use Management Act, No. 16 of 2013, 2013).

SPLUMA tries to promote principles like spatial efficiency, spatial sustainability, and spatial justice (Spatial Planning and Land Use Management Act, No. 16 of 2013, 2013).

The spatial efficiency aspect of SPLUMA aims at developing land that makes the most of the resources that are available to the land and the infrastructure that is already there (Spatial Planning and Land Use Management Act, No. 16 of 2013, 2013). This aspect of SPLUMA also tries to encourage social inclusion, spatial equity of land, and favorable land use patterns (Spatial Planning and Land Use Management Act, No. 16 of 2013, 2013).

Spatial sustainability tries to encourage suitable locations with regard to sustainability to be chosen for the development of land that also minimizes urban sprawl (Spatial Planning and Land Use Management Act, No. 16 of 2013, 2013).

The spatial justice aspect of this policy tries to enable disadvantaged communities to access land regardless of the value of the land (Spatial Planning and Land Use Management Act, No. 16 of 2013, 2013).

The principles that SPLUMA attempts to enforce are favorable for promoting mobility and accessibility through the role of transport (Spatial Planning and Land Use Management Act, No. 16 of 2013, 2013).

2.4 Transport Justice

Transport justice has a social aspect where accessibility to opportunities is concerned (Martens, 2017). The main aim of transport justice is to promote equal levels of accessibility to everyone, regardless of residential location and income groups (Martens, 2017). Transportation decision-makers should hold fair accessibility levels for everyone as their main criteria across all racial and income groups (Martens, 2017). Martens (2017) tries to highlight that transport falls within the region of being a social good and thus needs to be distributed fairly across the population however, the equality in transport is anticipated to fail due to the structure of the transportation networks that favour city centres and disregards the peripherals of the city (Martens, 2017)

Transportation injustice can lead to social exclusion (Lucas, 2012). Population groups that are transport disadvantaged, experience a lack of participation in social goods (Lucas, 2012). When a person experiences decreased access to transport services and access to opportunities they are implied to be transport disadvantaged (Kamruzzaman et al., 2016).

Accessibility takes into account a person's resources that are available to them, the land use, and mobility (Kamruzzaman et al., 2016). Accessibility varies over different instances of time as the factors that accessibility takes into account do not always remain static and become challenging to measure (Litman, 2011). Mobility is the measure of the transportation services, infrastructure, and efficiency in ease of movement to access opportunities (Kamruzzaman et al., 2016). Transport infrastructure design may limit the mobility and accessibility of transportation services and cause people to experience social exclusion as these transport services may have high travelling costs for accessing different opportunities, not enough transit stops that accommodate enough population groups and safety provisions for people that are waiting for transit services in dangerous locations (Church, Frost, & Sullivan, 2000).

The lack of infrastructure for non-motorized transportation leads to social exclusion in locations where people can not afford or lack access to participate in motorized transport services (Pendakur, 2005). This is visible in South Africa because transportation is geared to motorized transportation which poor people can not afford to participate in due to lack of a vehicle due to their low income (Pendakur, 2005). The population groups that are affected by this lack of access to participate in motorized transport services will lead to lower levels of accessibility and mobility to opportunities experienced towards them (Pendakur, 2005).

2.4.1 Transport in Africa

With regard to transportation services in African cities, both urban and rural areas experience disparities in mobility and accessibility (Pirie, 2009). Some population groups do have access to a personal vehicle, while others rely on transport services which involve them walking from their residential locations to locations where they have access to these transport services (Pirie, 2009). Often these walks are long and dangerous due to transport services being far away from their residential locations and lack of safety on the path they have to walk along, especially in rural areas (Jennings, 2015). Due to travelling distances being larger, it increases people's chances of experiencing unfavourable weather conditions and accidents while travelling (Jennings, 2015). The large travelling distances can have a negative effect on the population groups that experience this as travelling large distances can be exhausting with respect to both walking and using transport services (Pirie, 2009; Jennings, 2015).

According to WHO (2013), the least number of personal vehicles owned is held by Africa. WHO (2013) also highlights that Africa experiences an unusually high death rate figure in motorized vehicle accidents. Within Africa, 16% of deaths occur on the road and 38% of these road deaths are comprised of pedestrians, while Africa only owns 2% of cars in the world (WHO, 2013). Due to the lack of provisions for the safety and accommodation of non-motorised transportation, pedestrians often put their lives at risk when walking along the motorized transport infrastructure (WHO, 2013). In addition, the low provisions of public transport and non-motorised infrastructure lead to poor levels of mobility and accessibility to opportunities (Pirie, 2009). Often the design of transportation infrastructure is geared towards the desires of the investors whose goals are to promote a motorized transport infrastructure that promotes positive economic achievements for high-income earners as they own personal vehicles and disregards the transportation needs of those who experience transportation deficits (Lucas et al., 2016, Ciommo & Shiftan, 2017).

Studies from the United Nations indicate that 68% of the population of the world would have taken residence within urban areas by the year 2050 (UN, 2007). It is expected that by the year 2030, most megacities will be located in developing areas (UN, 2007).

Although urbanization has a positive correlation with economic growth, this may not always be the case as these factors can change without affecting the other (Chen, Zhang, Liu, et al., 2014; Jedwab & Vollrath, 2015). Locations that are experiencing increasing poverty and inequality levels such as developing countries are an example of how urbanization is not directly linked with economic growth (Barredo, Demicheli, Lavalle, et al., 2004). Rapid urbanisation has negative effects on cities, especially in developing countries such as countries within Africa, with respect to social, financial and environmental issues (Barredo, Demicheli, Lavalle, et al., 2004; Annez & Buckley, 2006; Glaeser, 2013). Urbanisation could likely have an impact towards factors such as accessibility to transportation systems, accessibility to opportunities and mobility (Wang, 2014; Rode, Floater, Thomopoulos, et al., 2017). Therefore it is important to understand how urbanisation affects the role of transport (Kammen & Sunter, 2016). Issues such as inequality, mobility and social exclusion could worsen if the amount of urbanisation exceeds the resources and infrastructure that the land can supply as displayed in Africa (Henderson, 2002; Cervero, 2013).

The population who can not afford to purchase a property close to these locations where urbanisation is occurring are expected to live in shack dwellings surrounding these areas (Cohen, 2006; UN, 2007). The cumulative population of shack dwellings is expected to be approximately 100 million people by 2020 (UN, 2007). In developing countries such as Africa, urbanisation has led to urban sprawl (Cervero, 2013; Jones, 2017). Such spatial development leads to inequality in service delivery, the promotion of social exclusion and puts a strain on the existing infrastructure of these locations (Cervero, 2013; OECD, 2018).

To overcome transport needs for the urban poor, it promoted that they own a personal vehicle as their mobility and accessibility levels are not satisfactory using public transport services (Banister, 2008). Owning a personal vehicle in Africa can be seen as an increase in social status as enables the person to not participate in public transport and implies that a person has a higher income level (Banister, 2008).

Due to the high supply of personal vehicles over the years, the cost has owning a vehicle has decreased enabling more people to purchase one in the hopes of escaping making use of public transportation services (Banister, 2008). This now leads to a new problem of congestion on the motorized transportation infrastructure and in cities, in addition to increased levels of air pollution from these vehicles (Banister, 2008). With the previously mentioned factors, now the high prices of fuel are negatively affecting their expenditures and do not seem as a financially feasible solution for escaping the use of public transportation, especially for the low-income population groups (Pirie, 2009, Litman, 2014).

Mobility and immobility are still to be seen present among high and low-income earners respectively. In South Africa, the segregation of racial population groups taking land use into account during apartheid is still experienced today in transportation (Jennings, 2015). The government reinforces this segregation structure by developing roads on the urban peripherals to locations of opportunities instead of providing an offer of housing closer to these high accessibility and mobility locations to opportunities (Jennings, 2015).

2.5 Transport in Cape Town

Many cities in South Africa still deal with the spatial planning of the past and the problems that come with living there in current times. This can seen in a city such as Cape Town where the land use and transport planning objectives of the past were to create a separation between residential locations that were populated by low-income groups and locations that had high levels of economic opportunities (Watson & Turok, 2001; Turok, 2011b).

Table 2.1 shows the population of Cape Town between 1940 and 2016, this table shows how the population has increased over time based on data collected by the government. The increase in population is contributed by many factors such as population growth, urban emigration, and migration from other countries (Turok & Borel-saladin, 2014). Rapid urbanisation has largely contributed to informal residential locations which have increased inequality levels of the low-income population

Table 2.1: Population of Cape Town 1940 - 2016 (Wilkinson, 2000; Western Cape Provincial Treasury, 2012; Western Cape Government, 2017)

Population by Year	**1940**	**1996**	**2011**	**2016**
Population(Per Person)	529,000	2,600,000	3,700,000	over 4,000,000

groups as they are forced towards the urban peripherals where there is minimal access to opportunities and transportation services (Turok & Borel-saladin, 2014). Due to gentrification, low-income population groups that have residential locations close to the city centers are forced to move toward the urban peripherals as well (Turok & Borel-saladin, 2014).

Table 2.2: Informal Housing in the City of Cape Town 1993 - 2014 (DoH, 2015)

Informal Housing by Year	**1993**	**2010**	**2014**
Informal Housing(Per Unit	50	230	379

Table 2.2 shows the number of informal housing in the City of Cape Town between 1993 and 2014. This table shows the increase of informal housing in the City of Cape Town over time. This increase shows that the low-income groups want increased levels of access to opportunities that living close to the city centers provides, however can not afford to own formal residency in these locations (Brown-Luthango, Makanga, & Smit, 2012).

Some of the low-income population groups have found informal housing to be a solution to increase levels of accessibility and mobility, while others that live towards urban peripherals have found owning a personal vehicle to be a solution to this problem (Brown-Luthango, Makanga, & Smit, 2012). Since the transport infrastructure in Cape Town is mainly comprised of motorised infrastructure, it would be understandable why the low-income population would want a personal vehicle to overcome the spatial separation between them and locations of opportunities (Bertolini & le Clercq, 2003; Reilly, Mara & Seto, 2009).

South Africa has started to make social inclusion a part of policies for all of the various population groups (DoT, 1996, Spatial Planning and Land Use Management Act, No. 16 of 2013, 2013). Having policies in place aids with the development of land use and the transport infrastructure that promotes to access opportunities and employment (DoT, 1996, Spatial Planning and Land Use Management Act, No. 16 of 2013, 2013).

Currently in Cape Town, the spatial patterns have low densities along transport routes which makes the operations of these transport systems very expensive to operate as there are not enough passengers on board to make a trip feasible (CoCT, 2012a). Cape Town has been identifying regions that face social exclusion and trying to integrate them into the public transport routes while holding low transportation costs for the public in mind (TDA, 2016). The affluent suburbs in Cape Town still have high levels of accessibility and mobility to opportunities as opposed to poor suburbs where low levels of accessibility and mobility occur. These disparities in accessibility and mobility between the low and high-income populations have yet to be fully addressed. Understanding the land use and transport infrastructure issues and the lack of integration of these two factors will aid in minimizing the inequalities citizens of Cape Town face.

2.5.1 Transportation Indicators

Measuring Accessibility

Accessibility is a measure of convenience and the ease with which opportunities for services and activities can be reached (Pirie, 1979; Thompson et al, 2012). The higher the number of surrounding opportunities the more accessibility increases for that region and vice versa (Martens, 2017). Distance also affects accessibility as the longer it takes a person to get to an opportunity, the less the accessibility becomes and vice versa (Martens, 2017).

Measuring accessibility with regard to transport justice is not a straightforward process as it is quantified from a social good aspect of the transportation sphere (Kwan, 1998). The population groups that have an accessibility deficit need to be identified so that they can be assisted accordingly with respect to transportation (Kwan, 1998). Various aspects of accessibility need to be considered when measuring accessibility, such as (van Wee et al., 2013):

- People aspect - take into account how the population's income, needs, education, interests, and racial groups influence their accessibility levels.
- Land-Use aspect - describes the supply and demand for opportunities and the spatial distribution of these opportunities.
- Temporal aspect - account for the time of the day and duration available for the population to engage in activities.

- Transportation aspect - focused on the transport system and the various services.

When deciding on measuring accessibility most times more than one of the above aspects of accessibility mentioned above are taken into account (van Wee et al., 2013). Some of the types of accessibility measures are:

1. Location-based accessibility measures - analysis of access to opportunities such as employment and public facilities between origins within a specified travel time interval such as gravity-based and cumulative accessibility measures (Makri & Carolin, 1969; Kwan, 1998).
2. Utility-based accessibility measures - focus on benefits people have due to the spatial distribution of activities. This accessibility measure is derived from the utility theory, i.e. the option that maximizes their utility will be the person's choice (Makri & Carolin, 1969; Kwan, 1998).
3. Person-based accessibility measures - focused on how easily a person can access activity locations (Kwan, 1998).
4. Infrastructure-based accessibility measures - only focus on the demand and supply of infrastructures and it is used in transportation planning (Kwan, 1998).

The various accessibility measures take into account aspects of accessibility and therefore produce different results for accessibility when measuring exactly the same data (Kwan, 1998; Bhat, et al., 2001).

The cumulative measures and gravity-based measures results can be closely correlated when using the same data, but differ in results from a space-time measure (Kwan, 1998; Bhat, et al., 2001). The cumulative measures and gravity-based measures have been criticized for assuming that the accessibility for the entire region is the same as the centroid of the region (Kwan, 1998; Bhat, et al., 2001). Another limiting factor of the cumulative measures and gravity-based measures is the identification of personal accessibility because of the aggregated nature of results (Kwan, 1998; Kamruzzaman et al., 2016).

Many studies on accessibility have been done on people's access to different services that involve different accessibility measures and indicators, with the main focus on transportation accessibility for the population and the relation between people's income and employment opportunities (Krizek et al., 2009; Golub & Martens, 2014; Guzman et al., 2017). The idea was to find the correlation between the income people have and the number of opportunities that could be reached based on their income allowed them to afford transportation (Krizek et al., 2009; Golub & Martens, 2014; Guzman et al., 2017).

This study aims to use Marten's (2017) approach to calculating accessibility, which is different from previous studies on this topic (see Chapter 3; Equation 1). This approach was chosen as it can accommodate multiple variables to calculate the measure of accessibility. Examples of these variables are a specific transport mode modal split and travel time using a specific transit mode. Martens (2017) uses accessibility thresholds to identify accessibility deficits at various levels. The methodology for calculating accessibility measures in this project is in line with the gravity measure, it is now combined with the potential mobility measure to produce a more reinforced result (Martens, 2017). The potential mobility measure shows the effect transportation systems have on the levels of accessibility for different population groups of different regions (Martens, 2017).

Martens (2017) introduced the Accessibility Fairness Index (AFI) which Martens developed. The AFI is used to rank the severity of accessibility deficit based on accessibility levels experienced and the number of populations affected (Martens, 2017). Marten's approach allows for the data to be less aggregated during the grouping stage based on the purposes of the task at hand as the accessibility measure largely aggregates the data (Martens, 2017).

Measuring Mobility

A lack of mobility, similar to an accessibility deficit, is a function of transportation disadvantage (Kamruzzaman et al., 2016). Mobility comprises many dimensions such as travel time, travel distance, infrastructure, public transport services, and owning a personal vehicle (Preston & Raje, 2007). Mobility is measured by quantifying travel speed, travel distance, traffic volume, and level of service data, which will be used to identify personal vehicle and transit speeds between origins and destinations (Litman, 2003). The means to travel are the goals of mobility measures (Litman, 2003). Measuring the level of service for transportation service between origin and destination pairs is insufficient when determining mobility and will not provide enough information to help improve access to locations (Litman, 2003). Mobility measures have taken into account the upgrade in infrastructure and the speed of the network to determine a better solution (Litman, 2003). Often the NMT and public transit services are disregarded when determining mobility measures as these measures favour the population that has a personal vehicle in its determination (Litman, 2003). Therefore, developments are often moved further away from the low-income population and closer to urban locations (Grengs, 2018). To help incorporate transport services into mobility measures, the accessibility that public transport provides to people and the level of service need to be taken into account (Grengs, 2018). The Potential Mobility Index (PMI) created by Martens

(2017) helps to show the contribution that transportation services have to mobility measures. Potential mobility is the ability and easement for a person to overcome distance through space (Sager, 2005; Martens, 2017). The PMI is derived by the aerial distance and travel time between origin and destination pairs (see Chapter 3; Equation 3) (Martens, 2017). Therefore, the higher the travel time between origin-destination pairs the lower the potential mobility experienced and vice versa (Martens, 2017). Regions that fall below a threshold of standard potential mobility should be eligible for a potential mobility increase in some manner (Martens, 2017). This is a better accessibility indicator than the level of service because it creates a correlation between the lowest possible distance between two points and travel time (Martens, 2017).

Quadrant System Framework of Martens (2017)

The potential mobility measure of Martens (2017) is used to complement the accessibility measure (Martens, 2017). Marten's (2017) goal is to define a framework that can assist with promoting a fair transportation system and highlight the role of transportation planning in addressing accessibility deficits. Marten's (2017) framework is intended to act as a key step in transportation planning based on principles of justice.

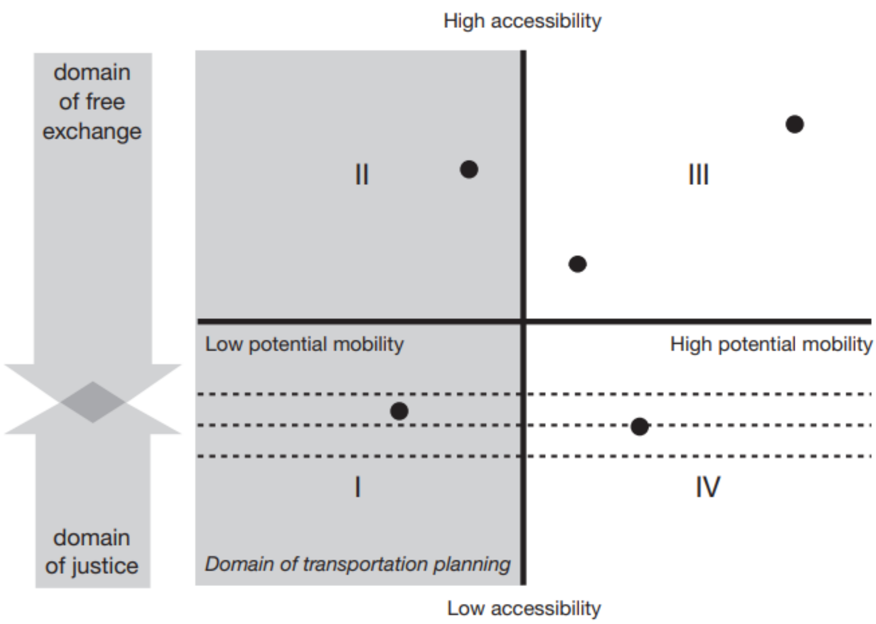


Figure 2.2: Quadrant System Framework (Martens, 2017)

Figure 2.2 shows accessibility and potential mobility are juxtaposed on a set of axes (Martens, 2017). Accessibility is represented on the vertical axis (which may be measured in different ways), and potential mobility is represented on the horizontal axis (measured in terms of PMI-scores) (Martens, 2017). The framework's coordinate system is defined by both axes which represent a continuum from a low to a high level in which various population groups can be positioned (Martens, 2017). The axes intersect in the middle, i.e. at the average level of accessibility and at the average level of potential mobility (Martens, 2017). This coordinate system enables the placement of population groups vis-à-vis both axes, based on the measurement of a group's accessibility and potential mobility (Martens, 2017). It also allows the introduction of a sufficiency threshold of accessibility, represented as dashed lines in Figure 2.2, which would most likely be positioned anywhere below the origin of the coordinate system by the real-life decision-makers involved in the process of setting the sufficiency standard (Martens, 2017). Firstly, the framework assists with establishing the role of transportation planning, which is confined to the left-hand side of the quadrant system (Martens, 2017). This section of the quadrant system is defined by the population groups with a below-average level of potential mobility, i.e. by population groups that experience relatively poor service delivery by transportation systems (Martens, 2017). Yet not all these population groups are entitled to improvements in the transportation system, the domain of justice is only reserved for the population groups with an insufficient level of accessibility (Martens, 2017). The sufficiency threshold is likely to be placed somewhere below the horizontal axis. This implies that the domain of justice also lies in the lower part of the coordinate system, whereas the domain of free exchange is located in the upper part of the diagram (Martens, 2017). This further reinforces the role of transportation planning based on principles of justice: that role is not simply defined to the section to the left of the vertical axis, but to the population groups that are positioned in the bottom-left part of the coordinate system (Martens, 2017). A poorly functioning transportation system plays a large factor in why these population groups experience a sub-standard level of accessibility (Martens, 2017). The plight of these population groups should first and foremost be addressed through transportation planning based on principles of justice (Martens, 2017).

2.6 Quantitative Approaches to Transport Justice

2.6.1 Databases

Transport data can be categorized into spatial and non-spatial data (Dondo & Rivett, 2004). The characteristics of non-spatial data are the data has no geographical positioning information and is largely obtained through the collection of survey data, an example being the number of people traveling from Mitchells Plain to Cape Town via a taxi (Dondo & Rivett, 2004). Some of the non-spatial transport data collected could be fares, service timetables, and the stops description of the transport service (Dondo & Rivett, 2004). The characteristics of spatial data are data that have geographical positioning information and can be displayed as a feature class on maps (Dondo & Rivett, 2004). The geographical position of the feature will be represented by x and y or latitude and longitude coordinates defined in a reference system (Dondo & Rivett, 2004). Some of the spatial transport data collected could be positions of transit stops, transit routes, and road networks (Dondo & Rivett, 2004). By integrating spatial and non-spatial data transport planners can obtain information that has meaning and can help with decision-making in the development and planning of transport services and infrastructure (Dondo & Rivett, 2004). In order to obtain the best results using spatial and non-spatial data, the data constantly needs to be updated as cities are constantly as well (Dondo & Rivett, 2004).

Most data sources can be very detailed and can be attributed to spatial data, an information system that deals with spatial and non-spatial data can integrate, update, store, perform analysis, and view this data. Since most of the data are voluminous and spatial in nature, a spatial information system facilitates the storage, integration, display, update, and analysis of the data (Dondo & Rivett, 2004).

Dondo & Rivett (2004) show that establishing a spatial information system for public transport data consists of three steps, which are creating a conceptual data model, a logical data model, and a prototype.

Conceptual Data Model

This was the first step that was done, it established the data that was used and how they were linked and related to each other (Dondo & Rivett, 2004). Dondo & Rivett (2004) uses the entity-relationship model, which is commonly used to represent the conceptual data model.

The entity-relationship model identifies the world as being comprised of entities and relationships (Dondo & Rivett, 2004). An entity is some identified element that is unique, an example could be a person's identity number (Chen, 1976). A group of similar entities is known as an entity type (Chen, 1976). The link between different entity types is known as a relationship (Chen, 1976).

There are three kinds of relationship degrees, namely the:

- one-to-one relationship
- one-to-many relationship

- many-to-many relationship.

The relationship degrees changes based on how each entity are participating with other entities (Chen, 1976). The data used to establish the conceptual model will be turn into entities and the relationships will be define to create the database (Dondo & Rivett, 2004).

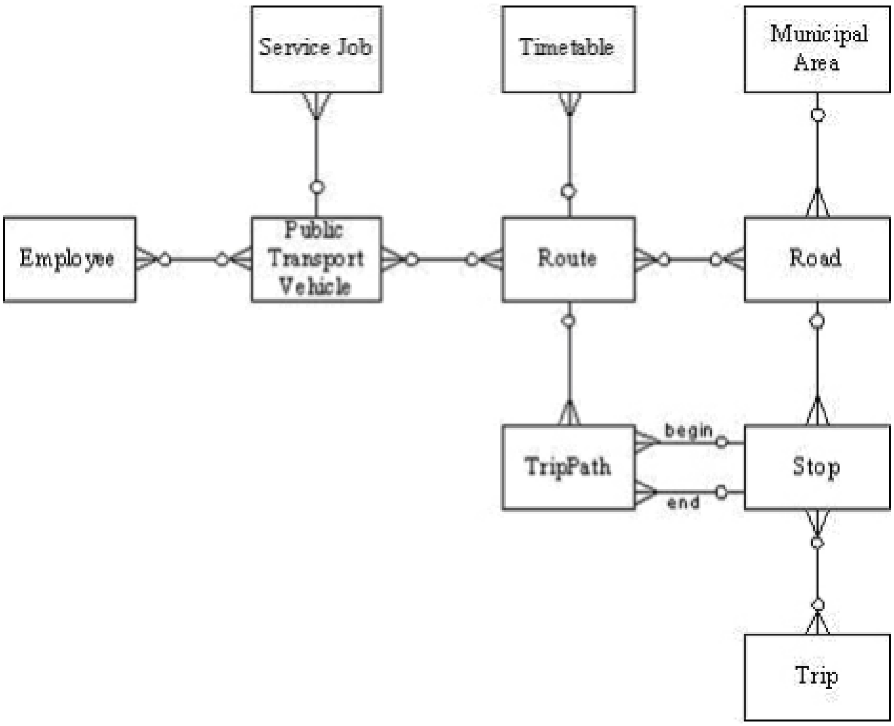


Figure 2.3: Identified Entities and Relationships display in the Conceptual Data Model (Dondo & Rivett, 2004)

Figure 2.3 show an example of a conceptual data model for transport data. The conceptual data model shows how spatial and non-spatial data interacts with one another through their relationship degrees (Dondo & Rivett, 2004).

Logical Data Model

The logical data model is a model based on the conceptual data model that can be applied in a chosen spatial information system program (Dondo & Rivett, 2004). Dondo & Rivett (2004) use the object-relational model to change from the conceptual data model to the logical data model showing it as an accumulation of tables that have assigned names in a database.

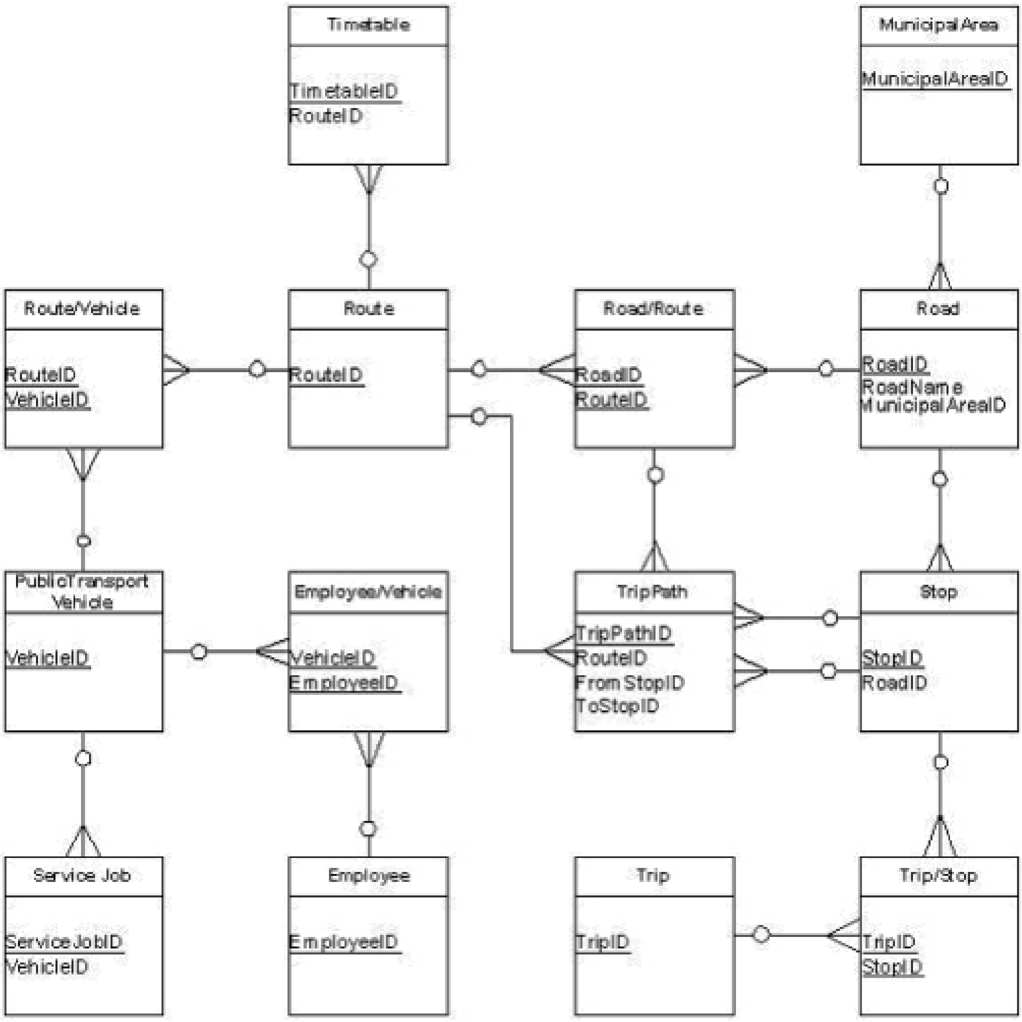


Figure 2.4: Logical data model for public transport information management (Dondo & Rivett, 2004)

Figure 2.4 represents an example of implementing a logical data model based on a conceptual data model for transport data. The object-relational model consists of named attribute types or columns and records or rows (Codd, 1970; Laurini and Thompson, 1996). Every row holds details relating to one entity and represents a unique identifier for the table (Codd, 1970; Laurini and Thompson, 1996). The intersection of each column and row in a table holds an attribute value (Codd, 1970; Laurini and Thompson, 1996).

Prototype

The final step was to create a prototype, which takes the logical data model design and implements it in the chosen software. The software chosen was ArcGIS and it can implement an object-relation model (Zeiler, 1999; Curtin, Noronha, Goodchild and Grise, 2001).

2.6.2 GTFS

GTFS (General Transit Feed Specification) data defines a common format for public transportation schedules and associated geographic information (GTFS, 2019). This data takes into account the variability in travel time at different times of the day for a public transit mode. Using this data with a network analyst layer enables users to identify how long a journey could be at any time of the day without having to use public transit to get this information. Using an OD Cost Matrix with this data allows for many origins to determine travel time to many destinations to be done simultaneously (Ashok, 2000). Depending on the completeness, correctness, and availability of this GTFS data, influence the choice selection of the region of interest as the correct determination of accessibility is dependent on these factors (Ashok, 2000; GTFS, 2019).

GTFS data is required for the determination of an impedance attribute for a time-based travel time of public transport. Depending on the completeness, correctness, and availability of this GTFS data, will influence the choice of selection of the region of interest as these factors mention that GTFS data needs to be good to achieve good results for potential accessibility and mobility determination. This also determines the software used to interpret and manipulate the GTFS data to produce the desired results of travel time duration of the various transport services throughout the day.

The GTFS data can be captured with real-time data or scheduled data. Real-time GTFS data is determined from a route taken by the transport nodes in real time. Scheduled GTFS data is determined from a route taken by the transport nodes from a planned route. The geographical location of the transport modes is captured by every transit stop on the various transport nodes. These transit stops have information about the routes of the public transit that passes these transit stops and the time that these public transport mode passes that transit stops during the day.

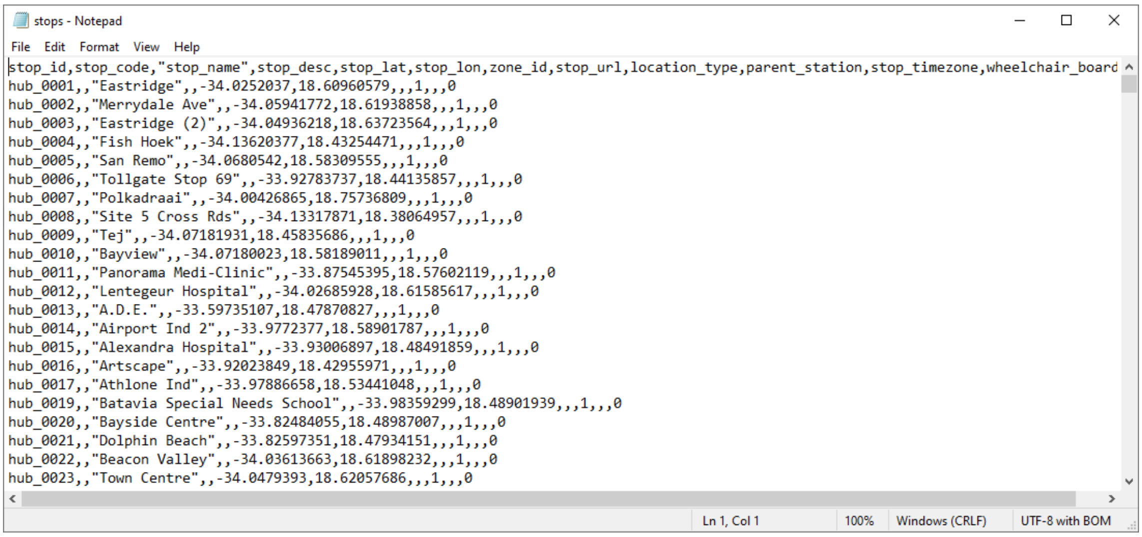


Figure 2.5: Sample Stops file of a GTFS data set

Scheduled GTFS data is easier to produce as it is based on estimated values from an approved schedule. Real-time GTFS data has the advantage over scheduled GTFS data as it produces a travel time duration that is more accurate to actual travel during the day as it is based on real-time data. However, real-time GTFS data is more difficult to produce as it requires the drivers of the transit modes to be trained to use software that captures the GTFS data.

2.6.3 Hägerstrand's Space-Time Geography Model

To understand travel behaviour in a more structured manner, accessibility measures used in many research papers have made use of Hägerstrand's space-time geography and its constraints (Miller, 1991; Kwan, 1998; Geurs & van Wee, 2004).

Person-specific space-time measures take into account the land-use system, individual space-time constraints, and spatial distribution of the transport system being used. Space-time measures are used to deduce different indications in the region of urban planning and highlight interpersonal discrepancies in the accessibility of travelers (Kwan, 1998). Most studies regarding mobility in urban environments within Sub-Saharan cities disregard Hägerstrand model on space-time geography. It would be beneficial if Sub-Saharan cities make use of Hägerstrand's approach as it has the ability to identify daily mobility among various social groups and takes into account time-space constraints more than just affordability.

Mobility constraints are analysed in a systematic manner using this approach, the low and high-income population group's potential mobility is compared and links potential mobility constraints with these population groups' expected standards of living. In the next part of Hägerstrand's approach, the identified constraints of the analysis were cross-examined with the policies that planners or decision-makers take into account when proceeding with the development of transportation. This assists with government transport policy development by using the results of this approach. This approach highlights regions and socio-economic groups that need accessibility improvements to economic and employment opportunities.

Hägerstrand's Three Time-Space Constraints

Time-space constraints created by Hägerstrand understand that movement to opportunities in a particular location comes from a person's need to participate in these opportunities (Hägerstrand, 1970). Hägerstrand created a quality of life measure by analysing a person's livelihood and daily activities which was derived from a person's daily mobility characteristics (Janusz, Kesteloot, Vermeiren, & Van Rompaey, 2019).

A person's livelihood opportunities, travel time, and travel distance, are largely limited because of three types of interacting constraints (Hägerstrand, 1970; Pred, 1977):

1. Capability Constraints
2. Coupling Constraints
3. Authority Constraints

Capability Constraints

Capability constraints deal with the available transportation systems a person can access and the person's physical capabilities to move in space. The factors this constraint determines are a person's speed and the time available to move around (Hägerstrand, 1970). Hägerstrand assumes people are traveling from home and need to return home again after they complete their task for traveling (Janusz, Kesteloot, Vermeiren, & Van Rompaey, 2019). This implies that people leave and return home within certain hours of the day so that they can rest at home after they access their desired opportunities (Hägerstrand, 1970; Janusz, Kesteloot, Vermeiren, & Van Rompaey, 2019). Speed and the time available within a day are used to create a prism in time-space that defines how many opportunities a person can reach from their origin. Any opportunities that fall outside of this prism imply that the person can not access these opportunities (Janusz, Kesteloot, Vermeiren, & Van Rompaey, 2019). Larger prisms will form due to an increase in speed and an increase in time available for moving around (Hägerstrand, 1970; Janusz, Kesteloot, Vermeiren, &

Van Rompaey, 2019). The prism is a three-dimensional representation of space, where speed and available time for moving are defined within the horizontal plane and the time of the day is represented in the vertical axis of the prism (Ellegård & Svedin 2012). The coupling and authority constraints will affect the space-time path of a person within the prism (Janusz, Kesteloot, Vermeiren, & Van Rompaey, 2019).

Coupling Constraints

This constraint relates a person to other people and the constraints that come with having this relationship with others (Hägerstrand, 1970; Janusz, Kesteloot, Vermeiren, Van Rompaey, 2019). An example of this is an employer and employee relationship (Janusz, Kesteloot, Vermeiren, Van Rompaey, 2019). Coupling constraints result in bundles within time-space, their individual paths in time-space show the need for co-presence when grouping their time-space time paths.

Authority Constraints

This constraint deals with authority over space that people are moving towards (Hägerstrand, 1970). Authority constraints are constraints that are in place to control people's behaviours in various locations within a particular time-space entity, leading to certain locations only being available to a limited amount of individuals (Hägerstrand, 1970; Janusz, Kesteloot, Vermeiren, & Van Rompaey, 2019). An example of this could be a pedestrian zone that does not allow vehicle drivers to drive in those locations. This time-space entity is known as a domain, access control within a domain could last for a certain periods of time or it can be permanent (Hägerstrand, 1970; Janusz, Kesteloot, Vermeiren, Van Rompaey, 2019). This possible access time constraint will form closed or open cylinders in time-space (Janusz, Kesteloot, Vermeiren, & Van Rompaey, 2019).

2.6.4 Urban Development Index (UDI) Modal Split

The UDI modal split measures the percentage contribution of the various forms of transport, including private, public transport, and NMT (City of Cape Town, 2019). This helped indicate the breakdown of the percentage of the population that makes use of the various transit services (City of Cape Town, 2019). It is recorded as a form of census data and was attached to the attribute table of the census centroid block of interest (City of Cape Town, 2019). The UDI modal split was used to form a weighted model for potential accessibility and mobility (City of Cape Town, 2019).

2.6.5 Operational Times of Opportunities

The operational times of opportunities helped to determine a refined and dynamic measure of accessibility (Järv et al., 2018). Without including the operational times of opportunities the potential accessibility measure will be a static value which will result in the same accessibility level throughout every hour of the day if no other restrictions on potential accessibility are in place (Järv et al., 2018). By incorporating operational times, the accessibility measure will change throughout the day instead of remaining constant. When an opportunity is operational then the accessibility for that opportunity will increase and vice versa. Using operational times of opportunities indicated a closer approximation of potential accessibility using the various public transit services (Järv et al., 2018).

Chapter 3

Method

This chapter describes the developed method used to create a spatial-temporal geo-spatial database model for transport justice using GTFS data and vector layers with census data. This research project's method workflow process is illustrated in Figure 3.1 below.

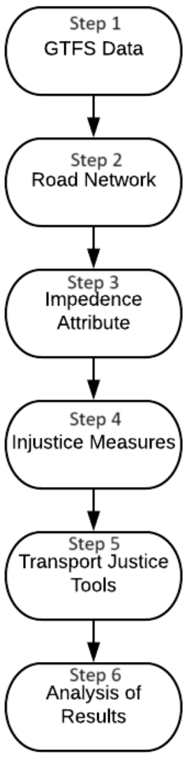


Figure 3.1: Work Flow Process of Method

The method consists of six steps. Firstly, acquire GTFS data of the various transport modes of interest for the project. Then an appropriate road network layer for the area where the GTFS data exists was chosen. This allowed the determination of travel time to be possible using the GTFS data. The GTFS data was joined with the road network to create a network analyst layer that is time-dependent based on the GTFS data of the respective public transport modes. The TAZ centroid points for Cape Town served as the origin and destination pairs of locations people traveled between while using the various public transit services. The TAZ centroids points were also used as the reference locations of the opportunity, population, and income data that has been derived from census data. Once the network layer is created, the impedance attribute of travel time is determined based on origin-destination pairs using the respective public transport modes. This travel time data was stored as text files to serve as a database of travel time of the various public transport modes between origin-destination pairs within Cape Town.

With the database of travel time created, potential accessibility and mobility were calculated for each TAZ centroid region. This database along with census data for opportunities, had all the necessary information required to determine the values for potential accessibility and mobility. The potential accessibility and mobility values were determined simultaneously with the transport justice tools created within a Python environment.

The values of potential accessibility and mobility were determined for an entire day and displayed in a Space-Time cube against each other. This Space-Time cube represented potential accessibility and mobility for every hour of the day. This was done for all of the respective public transport modes. Specific hours of potential accessibility and mobility analysis were chosen for a refined analysis. The Temporal Aggression transport justice tool was used to establish this refined analysis. The Temporal Aggression transport justice tool identified regions that fell below-specified threshold frequencies of potential accessibility and mobility values. This process was done for the various public transport services used in this research paper.

Finally, once these regions that fell below the specified threshold frequencies of potential accessibility and mobility values were determined, further analyses were done to establish the number of people residing within these groups/regions. The number of people and the various racial groups residing in these locations as well as their income levels determined if these groups were entitled to a potential accessibility increase for the respective public transport modes that were accessible to them.

3.1 Context

Impedance is defined as a measure of cost or resistance that a transit user experiences as they travel through a specific route (Niaraki & Kim, 2009; Kamruzzaman et al., 2011). Therefore, the higher the impedance measure, the greater the cost of resistance is to move and vice versa (Niaraki & Kim, 2009; Kamruzzaman et al., 2011). Some examples of transport impedance are travel distance, travel speed, or travel time duration (Niaraki & Kim, 2009).

A road network was combined with this GTFS data to determine the impedance attribute values of travel time duration at different times of the day. Combining these data sets created a geographical link between transit stops and a road network within the same frame of reference.

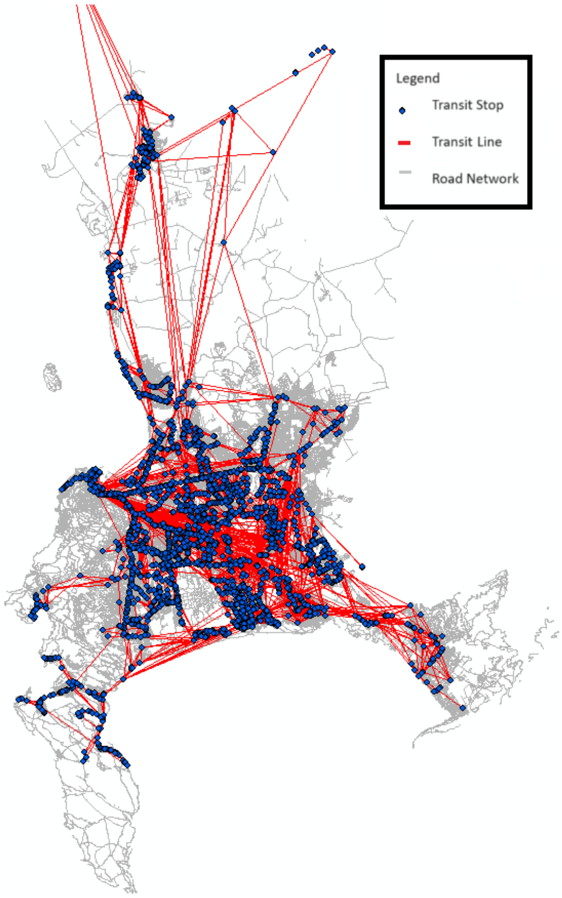


Figure 3.2: GTFS Transit Stops Connected To A Road Network

The region of interest for this research project was the City of Cape Town. The distribution of the road network and transit stops indicated the locations that were possible and impossible to reach within the region of interest while using the various public transport services (Farber, Morang and Widener, 2014). The possible locations that could be traveled from a specified origin were identified and allowed the travel time duration to be determined for the various public transportation modes throughout the day (Farber, Morang and Widener, 2014).

The main focus of this step was to create a method to travel from one TAZ location to another. The various impedance attributes such as travel time and traveling distance between the origin-destination pairs are then recorded in a database.

3.2 GTFS

The GTFS data captures the variability of the impedance attribute for travel time. Each public transport service that was analysed in this research project had its own set of GTFS data. Using this data combined with a network analyst layer, allowed the travel time to be determined at any period of the day without having to physically use a public transit service. The travel time was recorded at different hours of the day for each public transit mode.

Implementing the GTFS-enabled network analyst layer in an OD Cost Matrix allowed the travel time and travel distance of the origin-destination pairs to be calculated (Ashok, 2000). The OD Cost Matrix enables many origins-destination pairs to simultaneously be determined. This feature made it easier to handle large data sets and provided an automated process for the determination of travel time and travel distance.

The transit stops for each public transit service, capture the possible geographical locations that are capable of being traversed. The transit stops are represented as a point feature in a GIS environment. These transit stops have attribute information about the public transport service routes. The time that a specific transport service reaches the specific transit stop is one of these attributes. A public transport mode has many trips for the day. These trips use different routes across the region of interest where the transit stops are located.

3.3 Road Network

A road network layer was required to create the network analyst layer. This network analyst layer allowed the travel time duration to be calculated. The road network layer is a shape file that is represented as a line feature in a GIS environment.

These lines represented the real-life road network for the region of interest. One line represents the entire width of the road. The line is recorded from each end and at the centre of the width of the road. These lines are known as road centre lines. The road network layer was georeferenced so that any measurements made on this layer were true to scale.

The road network layer was combined with GTFS data of the various public transport modes to create the network analyst layer. Which allowed time-based travel time to be determined. Therefore, it can provide a simulation of travelling on public transit routes at different times of the day. This was required to determine the necessary travelling impedance attribute values to perform the transport justice analysis of the research project.

The road network layer is manually created by a GIS data service provider. The GIS specialist that creates the road network layer, may not digitise the road centre lines correctly. This leads to errors that accumulate in the road network layer. Examples of these errors are the length of the road digitised does not match the length of the actual road in real life or that a single road is digitised as multiple line segments. The line features of the road network layer need to intersect with each other to be traversable. If this intersection between lines does not occur in the road network layer, then the GIS software will implement an alternate route that joins these lines together. These errors in the road network layer will lead to incorrect values of travel time and travel distance. Therefore, when the road network layer matches the physical road network, it leads to a more precise calculation of the impedance attribute values for travel time duration and travel distance. The road network layer used in this project was acquired from OpenStreetMap.

3.4 Transit Analysis Zones (TAZ)

The transport analysis zones (TAZ) is a spatial distribution of a region, that allows the region of interest to be represented by different shapes and sizes of subsets within the region of interest (Farber, Morang and Widener, 2014). The centroids of the TAZ are similar to census data centroids, as they both are centroids point that has attribute information attached to their attribute table (Farber, Morang and Widener, 2014). The census attribute information was merged with the TAZ attribute table. This allowed potential accessibility to be determined while using the TAZ centroids (Farber, Morang and Widener, 2014).

3.4.1 Census Data

Census data is data collected by the governments of a country to provide information about the citizens and the land use (Census, 2011). This data is recorded in a database and is attached to a point feature class in a GIS environment. This point feature represents the centre of a census block. The region of interest is defined by multiple census blocks. These census blocks consist of various shapes and sizes. Census data includes categories such as income, race, age, gender and jobs of citizens as well as opportunities for the land (Census, 2011; Wang and Chen, 2015). These categories of census data capture the amount of people that fall within these categories for a specific census block (Census, 2011). Some subcategories examples are low and high income for the income census data and the different racial groups for the race census data (Census, 2011). The geo-spatial distribution of these subcategories for this census data indicated the inequalities of the different population groups in the region of interest (Leclerc et al., 2012).

Opportunity Data

Opportunities are the various facilities that are offered by locations (Dikec, 2009). These opportunities are made for public engagements (Dikec, 2009). Such opportunities are healthcare facilities, education facilities, shopping malls, and employment opportunities (Dikec, 2009). The opportunity data indicates the land use patterns of the region of interest. The distribution of opportunities illustrates how dense or sparse regions are in relation to opportunities (Dikec, 2009). This has a direct effect on the planning of transport infrastructure so that it accommodates these land use patterns. These transport infrastructures ultimately favours the population groups that reside closest to these opportunities and use public transport services to travel (Benenson et al., 2011). The geo-spatial distribution of opportunity data and the transport infrastructure highlighted the injustices of living far away from opportunities as opposed to living closer to these opportunities (Mattioli, 2014; Zhou & Long, 2016).

Population Data

Population data is derived from census data (Census, 2011). The population data illustrated how many people resided in a TAZ and the number of people within the different racial groups for that TAZ location (Census, 2011). The different racial groups in the census data were White, Black/African, Coloured, Indian or Asian and Other.

Income Data

Income data is derived from census data (Census, 2011). The income data determined how many people fell within the different income groups within a TAZ (Census, 2011). The four different income groups in the census data were Low, Low-Middle, Middle-High and High.

3.5 Network Analyst

A network analyst layer was used to create a georeferenced traversable road network. This process was done in ArcGIS. The network analyst layer was used to determine the travel time between origin-destination pairs, as well as to provide a geographical route for the total distance travelled. The travel time was calculated by using the travel speed and total distance travelled. Although the same route used at different times of the day will result in the same total distance, it may not result in the same travel time. This is due to the variations in travel speeds throughout the day. These travel speed variations are due to the increase or decrease in usage of a transit route (Haghani and Jung, 2005).

Using a network analyst layer that takes into account the variations in travel speeds throughout the day, will enable a transport model that imitates real-life travelling (Wong & Tong, 1998).

The GTFS data and the network analyst layer were both georeferenced within the same referenced system. This was done so that the two data sets were spatially correlated. The network analyst layer used in this project was created with GTFS data, which made it possible to perform travel time analysis at various times of the day. This process was done for each public transit service that was analysed in this project. Since the network analyst layer was enabled with GTFS data, it resulted in the travel time constantly changing at different times of the day.

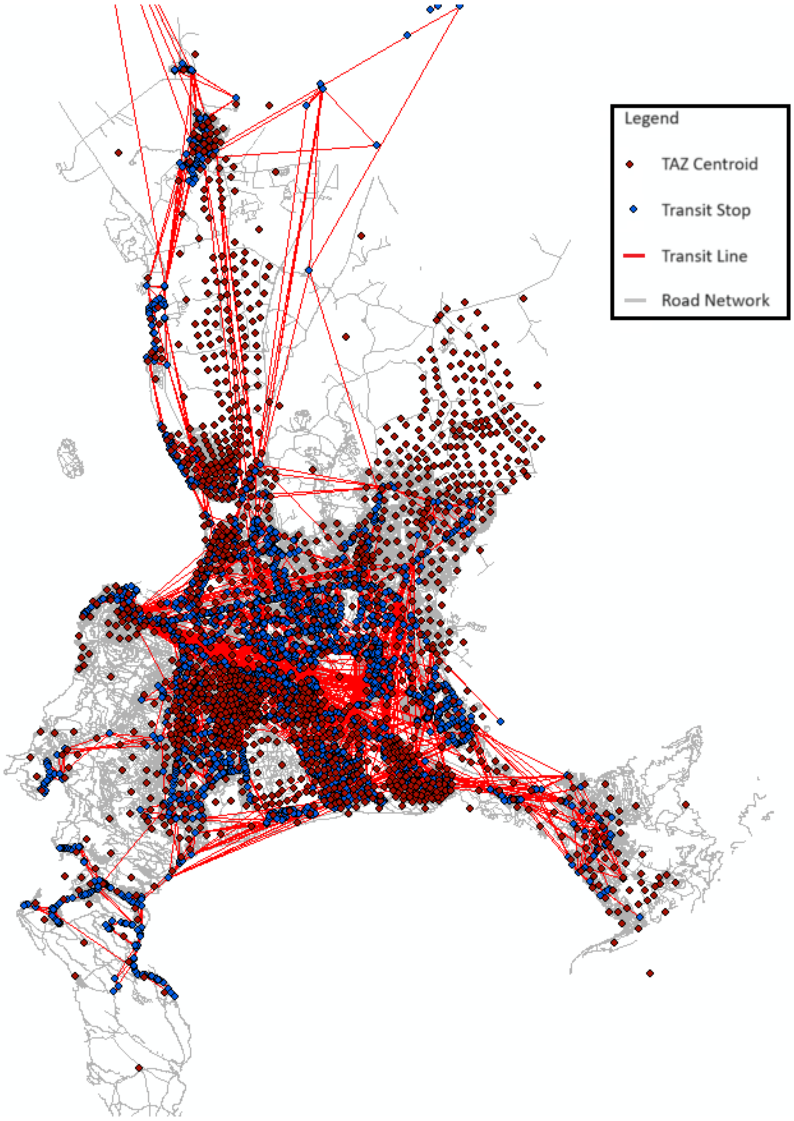


Figure 3.3: Network Analyst layer with GTFS data and a Road Network

The creation of the network analyst layer allowed the usage of an OD cost matrix. The OD cost matrix aided in determining the impedance attribute of travel time and travel distance of the chosen origins-destinations pairs. These TAZ origins and destination locations both fell within the region of interest.

3.5.1 Impedance Attribute

The OD cost matrix will provide additional impedance attributes other than travel time. Some of these additional impedance attributes are shape length and total length. The shape length will provide the Euclidean distance between origin-destination pairs in meters. The total length will provide the traveling distance traversed on the road network layer between origin-destination pairs in meters. The shape length will always be shorter than the total length, this is due to the Euclidean distance being the shortest distance between two points, which is the definition of the shape length. There can be instances when shape length and total length are equal to each other. This occurs when the route taken to calculate the attribute of shape length and total length are the same. These facts about the length attributes were used as a check to ensure that the correct attribute lengths were being used in any further calculations.

Travel Time

The impedance attribute value of the total travel time is the summation of walking time and transit travel time. Walking time is the total time taken for a person to walk from their origin to a starting transit stop and from their ending transit stop to their destination. The walking time is calculated by walk speed divided by the total walking distance traveled on the network analyst layer. Transit travel time is the total amount of time a person spends on a public transport service while traveling. Usually, the transit travel time is calculated by the summation of the speed of the road segments on the transit route traveled divided by it corresponding route length segments between the starting and the ending transit stop. However, a GTFS-enabled network analysis layer was used in this research project and it calculates transit travel time differently in comparison to the usual way. GTFS data has attribute information about the time of the day a trip of a public transit service arrives at a transit stop. A GTFS-enabled network analysis layer calculates travel time by the time difference between the starting and ending transit stops of a trip route. The time difference will not be static throughout because a GTFS-enabled network analysis layer was used to calculate travel time (Farber, Morang and Widener, 2014).

The walking and transit travel time between each TAZ origin-destination pairs was calculated with the OD Cost Matrix of the Network Analyst layer in ArcGIS 10.6 (Farber, Morang and Widener, 2014). This tool uses Dijkstra's algorithm to determine the shortest route in the Network Analyst layer (Farber, Morang and Widener, 2014). The ArcMap tool "Add GTFS to a Network Dataset" enabled the incorporation of GTFS data to a network analyst layer so that the calculation of a non-static travel time could be determined by different origin-destination pairs at various times of the day (Farber, Morang and Widener, 2014). The walk speed value was set as a constant value for the distance a person spent walking on the network analyst layer (Farber, Morang and Widener, 2014). The OD Cost Matrix calculated the total travel time by adding the walk and transit travel time together.

The travel time attribute value was used in both the injustice measures of potential accessibility and mobility. The potential accessibility measure used travel time to calculate the variations in potential accessibility of the same origin-destination pairs public transit trips throughout the day. The potential mobility measure used travel time to determine the variations in the aerial speed measure of between origin-destination pairs throughout the day.

Euclidean Distance

The Euclidean distance is the join distance between the origin and destination pairs. This distance was calculated in ArcMap by using the OD Cost Matrix (Farber, Morang and Widener, 2014). The Euclidean distance is labeled as shape length in the ArcMap OD Cost Matrix attribute table. The Euclidean distance was used as the aerial distance to determine the injustice measure of the Potential Mobility Index (see Chapter 3; Equation 3). The aerial distance is a variable of the the potential mobility equation and needed to be established in order to perform the potential mobility calculation (Martens, 2017).

3.6 Injustice Measures

The measures of potential accessibility and mobility were used to create the two different transport justice tools. The two transport justice tools were the space-time cube and temporal aggregation. The measure of potential accessibility is listed as Accessibility and potential mobility is listed as Potential Mobility Index (PMI). The measures of potential accessibility and mobility are further discussed below.

3.6.1 Accessibility

This measure indicated how easy or difficult it was for a person to travel between their origin-destination pairs in order to reach jobs/opportunities while using public transport.

Using the summation of the potential accessibility values of an origin to all possible destinations determined the cumulative potential accessibility value for that particular TAZ region. The Accessibility measure is a scalar value and is unitless (Martens, 2017).

$$A_{vt} = \sum_{m=1}^{M} \sum_{j=1}^{J} X_{jv} M_{mv} f(t_{vjmv}) \text{ Vi} \tag{3.1}$$

Assumed:

$\mathrm{M}_{mv} = M_m$

$\mathrm{f}(\mathrm{t}_{vjmv}) = (t_{vjmv})^{-1} = \frac{1}{t_{vjmv}} = \frac{1}{t_{vjm}}$

Every (v,t) is observed

$$A_{vt} = \sum_{m=1}^{M} \sum_{j=1}^{J} X_{jv} M_m \frac{1}{t_{vjm}} \text{ Vi} \tag{3.2}$$

Although the travel time is measured in minutes, potential accessibility is measured as a scalar. X is the number of opportunities, health care, education, and public facilities within the desired destination. M represents the modal split of the transport mode and m is the transport mode. t represents the travel time of using a public transit mode between TAZ origin-destination pairs, i will be the particular hour of the day. The determination of travel time between origin-destination pairs is discussed in 3.5.1. The accessibility measure will vary during the day due to factors such as; the operating times of the various jobs/opportunities, the change in travel time between origin-destination pairs throughout the day, and the modal split for that particular TAZ region.

The Discount Function is the inverse of travel time as larger travel time has a negative effect on the number of opportunities a person can reach, hence the travel time being inverted and acting as a discount function in the calculation of potential accessibility. The population groups that live further away experience lower potential accessibility than the population groups that live closer to opportunities. The discount function aids in highlighting this fact by exponentially reducing the accessibility measure while an increase in travel is occurring.

3.6.2 Potential Mobility Index

The Potential Mobility Index (PMI) measures the aerial speed between origin-destination pairs. This speed is measured in km/h and will produce a PMI value. A PMI score can be calculated for each origin i by taking the average of the PMI values for all relevant destinations for that origin TAZ (Martens, 2017). The PMI score is the value the PMI measure produces which is unitless (Martens, 2017).

Assumed:

if $t_{vjmt} > 0$:

$$PMI_{it} = \sum_{m}^{M} \sum_{j=1}^{J} d^{e}{}_{vj} \frac{1}{t_{vjmt}} \ Vi,t \tag{3.3}$$

Assumed:

if $\mathrm{t}_{vjmt} = 0$

$$PMI_{it} = \sum_{m}^{M} \sum_{j=1}^{J} d^{e}{}_{vj} + t_{vjmt} \text{ Vi,t} \tag{3.4}$$

$$\mathrm{PMI}_{vt} = \frac{\sum_{n=1}^{N} PMI_{vt}}{N} \text{ Vi,t} \tag{3.5}$$

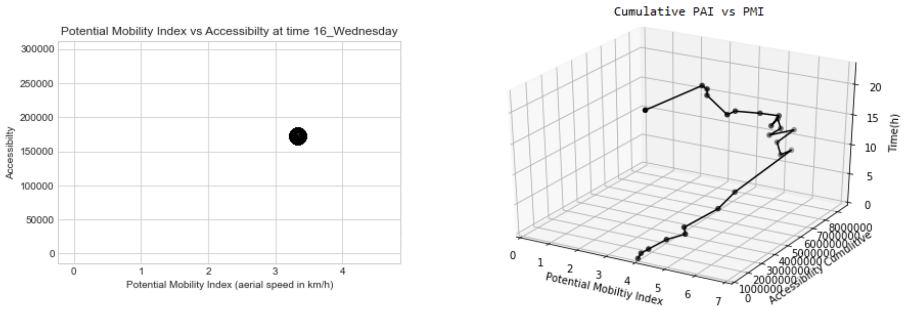
Using the OD cost matrix produced the necessary parameters for the determination of the PMI values for the various origin-destination pairs. The shape length provided the aerial distance (d) value and the travel time (t) was produced by the travel time of the OD cost matrix. However, the units produced by the OD cost matrix length and time meters and minutes respectively. These units of the OD Cost Matrix must be converted to kilometers and hours respectively in order to produce the correct PMI values of km/h dimensions. After all the PMI values are calculated then The PMI Score can be determined for a particular region. The PMI measure is unitless as it is an indicator of potential mobility for a particular region. The PMI measure will be negatively affected if the calculated PMI values are low or if they are equal to zero. Factors that will influence the result of a low PMI value are a very large travel time and with very short aerial distance between an origin-destination pair. This implies that a PMI value is directly proportional to aerial distance and inversely proportional to travel time. The aerial distance will always remain constant between an origin-destination pair however, travel time may vary throughout the day. This variation of travel time will result in different a PMI value for the same origin-destination pair, which will result in a different PMI score for the PMI measure. The transit stops and transit lines will define the service coverage that the public transit will be able to transport a person along. If a person wants to travel to a location outside the coverage of the transport network of the transport service they are using, then they would need to use an additional alternate form of transportation to complete the journey that includes this extended coverage. The coverage for the various public transport services used in this research project can be highlighted in Figure 4.9 in Chapter 4.

The travel time used to determine a PMI value for a specific public transport service uses the accumulation time of walking from an origin to a transit stop, waiting time to get on to public transit, travel time while on public transit, the time getting off a public transit to a public transit stop that is closest to the destination and the walking time from the public transit stop until the destination point. Waiting time to get on to public transit will decrease if public transit has a high availability at the public transit stop and vice versa. Travel time while on public transit decreases if the average travel speed of the transit increases and vice versa. Walk time will decrease if the transit stop is closer to the origin or destination location and vice versa. The PMI value will be low if all of these time accumulation factors are high, which will result in a lower PMI score in the PMI measure and vice versa.

3.7 Transport Justice Tools

3.7.1 Space-Time Cube

The space-time cube tool is the first of the two transport justice analyses tools that was created in this research project. The values of accessibility and PMI were plotted against each other at every hour of the day for the various transport modes. This cube focused on a single origin-destination pair of the various transport modes. The objective was to create a three-dimensional plot of the potential accessibility throughout the day via connected vector lines.



a. 2D Cut of Accessibility vs Potential Mobility at specified time of the day of an Origin location using Transit Services

b. Three Dimension cube of Accessibility vs Potential Mobility vs Time of an Origin location using Transit Services

Figure 3.4: 2D vs 3D plots of Accessibility vs PMI at different hours of the day

Figure 3.4 is an example result of what the transport justice tools will produce. Figure 3.4 is purely for demonstration purposes only and this example is not an actual result used within this project. Observing the trend of the line within the space-time cube, indicated the variations of potential accessibility and mobility throughout the day. The periods of the day where potential accessibility and mobility deviate the most could be identified by analysing the space-time cube. These deviations in potential accessibility and mobility were identified as an increase or decrease in the respective transport justice measures at particular periods of the day.

3.7.2 Temporal Aggregation

The temporal aggregation tool is the second of the two transport justice analyses tools that were created in this research project. This tool allowed for a specified hour of the day to be further analysed with respect to potential accessibility and mobility for TAZ regions. This tool could be used on various days of the week and transport modes. This tool allowed multiple TAZ regions to be represented on one graph with represent to their potential accessibility and mobility for a specific hour of the day.

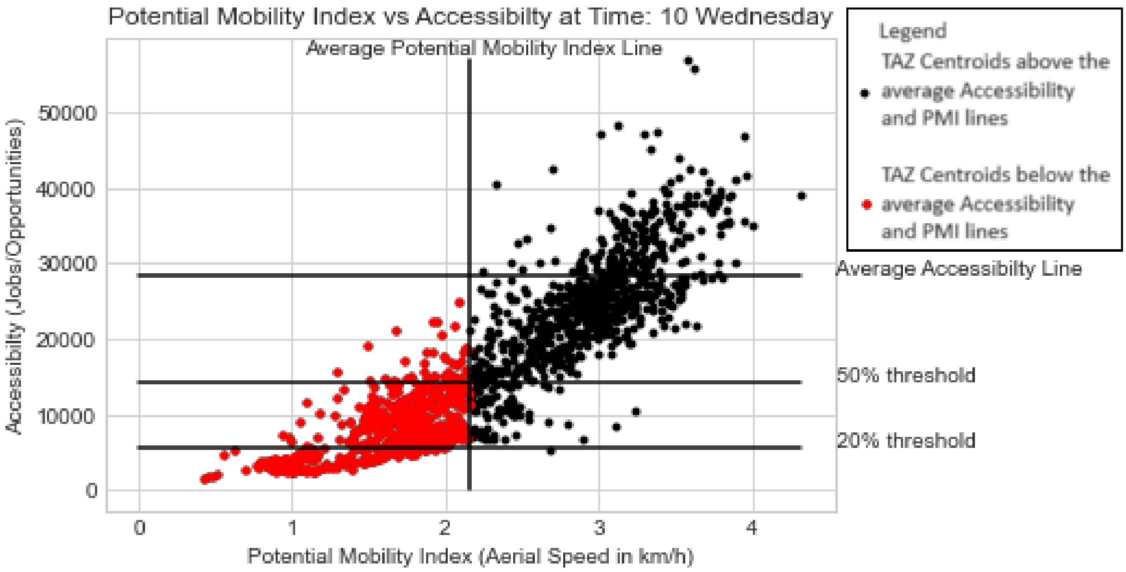


Figure 3.5: Temporal Aggregation Plot of Potential Mobility vs Accessibility

Figure 3.5 is a pre-result, i.e. example of what an actual result will appear within the context of the transport justice tool but not a result that is used within this project. Although a PMI value is determined using Aerial distance (km) divided by travel time (h) between an origin-destination pair, its result produces an indicator or score of potential mobility for a particular region. A PMI score can be calculated for each origin by taking the average of the PMI values for all relevant destinations for that origin (Martens, 2017).

In Figure 3.5 the red coloured points are the share of the TAZ points that fall below the average accessibility and potential mobility lines. Each point on the temporal aggregation plot represented a TAZ region's cumulative average potential mobility (x coordinate) and accessibility (y coordinate) for a particular hour of the day. The average value of accessibility and potential mobility of all the TAZ points on the temporal aggregation plot will become the origin point of the coordinate system while using the sufficiency thresholds for accessibility as described in Martens (2017). Based on the desired sufficiency thresholds for accessibility, it was possible to identify the TAZ points that fell below the desired sufficiency threshold of the particular transport mode. These areas were eligible for an accessibility increase based on the number of TAZ regions affected and the number of people affected within the particular regions that fell below the chosen sufficiency threshold.

3.8 Sufficiency Thresholds

Once the sufficiency thresholds have been established and all the TAZ regions that fell below the average accessibility and mobility lines have been identified, the regions that fell below the selected threshold frequencies were represented on a graph along with the number of people that were affected at that specific sufficiency percentage threshold (Martens, 2017).

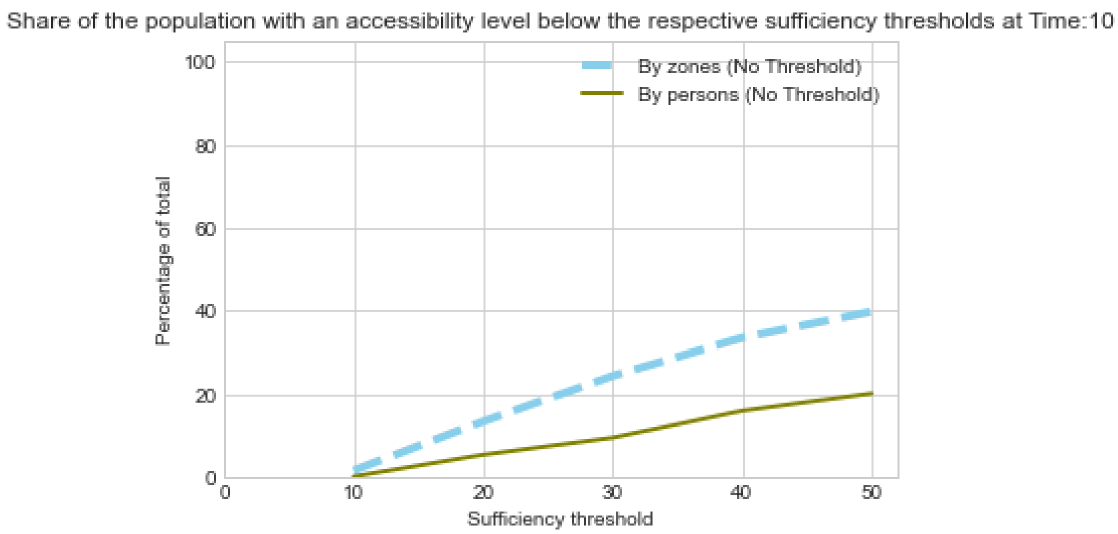


Figure 3.6: Share of the population with an accessibility level below the respective sufficiency thresholds

Figure 3.6 is a pre-result that highlights the share of the population with an accessibility level below the respective sufficiency thresholds. At each sufficiency threshold percentage, certain regions fell below a certain sufficiency threshold of accessibility based on their accessibility and potential mobility values for the given temporal dimension for a specific transport mode. The lower their accessibility and potential mobility values are the more the likelihood of these regions appearing at lower sufficiency threshold percentages. This indicates that these regions are the most in need of an accessibility and mobility increase.

Chapter 4

Implementation of Data and Methods

4.1 Introduction

This section describes the manner in which the method chosen was implemented, the selection of data and the software chosen to create the geo-spatial temporal database tools for transport justice. Regions with poor levels of accessibility were identified using the transport justice tools created for this research project. This was done using various:

1. TAZ locations and census attribute data
2. Transport services GTFS data
3. Time of day
4. Day of the week

4.2 Software Used

4.2.1 FeedValidator

The FeedValidator is a Python tool that verifies that the GTFS data meets the requirements defined in the GTFS-static reference (Watkins et al., 2011). The GTFS data used in this research project was tested using this tool. This tool assisted in understanding the correctness and completeness of the GTFS data for the various transit services (Watkins et al., 2011).

4.2.2 ArcMap 10.5.1

The majority of the spatial data analysis was done in the ArcMap program of ArcGIS. The georeferencing of the various spatial data was done in this program if data sets were not in the WGS84 coordinate system. The spatial coverage of the various spatial data sources was minimised to an extent over the region of interest by using the Extraction By Mask Tool of ArcMap to assist with minimising time processing time and aid in a better approximation value for potential accessibility and mobility. The creation of the Network Analyst Layer was done in ArcMap using the road network and GTFS data for the region of interest. The TAZ spatial reference layout for Cape Town was converted to centroids to serve as origins and destinations in the OD Cost Matrix for the Network Analyst Layer. Using the OD Cost Matrix tool provided the impedance attributes of travel time duration and Euclidean distances between the TAZ centroids of the various transit services throughout the day, which was used in the calculations of potential accessibility and mobility. The results obtained from the OD Cost Matrix were exported as txt files and serve as a database used by the transport justice tools for transport justice analysis.

4.2.3 Spyder(Anaconda3)

Spyder is a Python interpreter that allows the creation of Python programs to be created and executed. The creation of transport justice tool was created using this program. The TAZ centroids attribute table was combined with census data for every available TAZ centroid was done in Spyder and aided in the determination of potential accessibility. The opportunity data obtained from the census information was reinforced with operational times of these opportunities to provide a refined value for potential accessibility and mobility.

4.3 Data Source

4.3.1 GTFS

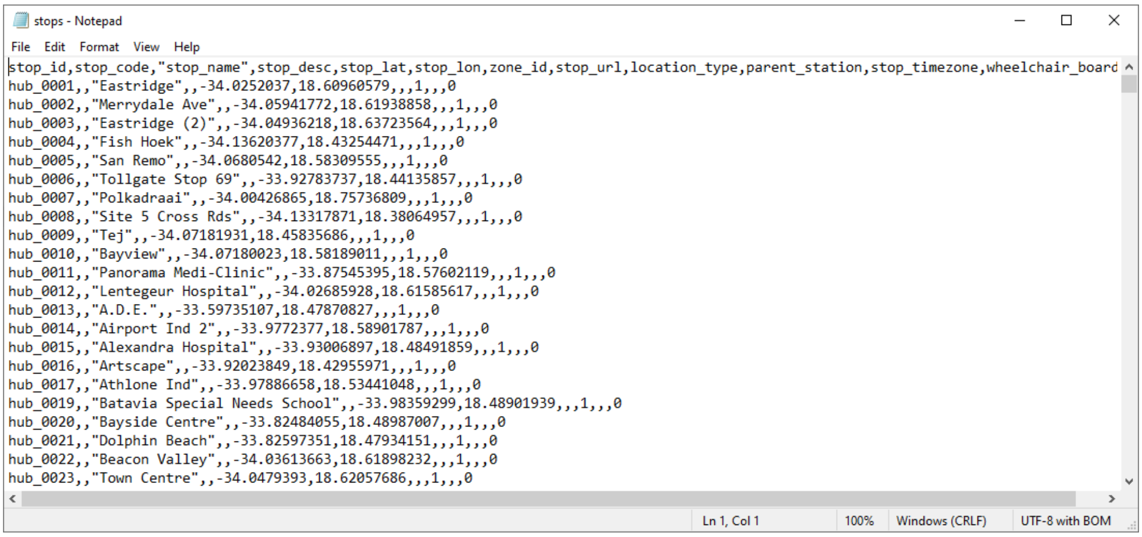
The GTFS data used for the development of the transport justice tool was static GTFS data based on the planned schedules of the various transport modes. These planned schedules are determined from interpolating times based on real-time data. These GTFS data sets have geographical locations for their transit stops embedded into their data source.

This guided the GTFS data on the actual geographical routes taken by the transit using the GTFS data for a query. The GTFS data is georeferanced and the geographical coverage for this data is the City of Cape Town. The various transport modes used were:

- CPT Taxi (2018) (Source: https://www.whereismytransport.com)
- Golden Arrow Bus Services (2018) (Source: http://www.gabs.co.za)
- Metrorail (2016) (Source: http://www.metrorail.co.za)
- MyCiTi (2016) (Source: http://www.myciti.org.za)

Due to the temporal differences between the GTFS data captured for the various transport modes, an explicit comparison between the GTFS data of the various transport modes could not be made. However, since all the GTFS datasets are georeferanced in WGS84 and have the same coverage, an initial comparison could be done. This can be later refined using GTFS which has the same coverage as well as the same temporal resolution across the various transport modes of interest by using the tools developed in this project.

The GTFS data are comprised of multiple .txt files. Each of these files was checked if they were stored in the UTF-8 encoding. If they are not stored in this format ArcGIS would not be able to read the GTFS datasets and the creation of the network analyst would not be possible.



```
stop_id,stop_code,"stop_name",stop_desc,stop_lat,stop_lon,zone_id,stop_url,location_type,parent_station,stop_timezone,wheelchair_board
hub_0001,,"Eastridge",,-34.0252037,18.60960579,,,1,,,0
hub_0002,,"Merrydale Ave",,-34.05941772,18.61938858,,,1,,,0
hub_0003,,"Eastridge (2)",,-34.04936218,18.63723564,,,1,,,0
hub_0004,,"Fish Hoek",,-34.13620377,18.43254471,,,1,,,0
hub_0005,,"San Remo",,-34.0680542,18.58309555,,,1,,,0
hub_0006,,"Tollgate Stop 69",,-33.92783737,18.44135857,,,1,,,0
hub_0007,,"Polkadraai",,-34.00426865,18.75736809,,,1,,,0
hub_0008,,"Site 5 Cross Rds",,-34.13317871,18.38064957,,,1,,,0
hub_0009,,"Tej",,-34.07181931,18.45835686,,,1,,,0
hub_0010,,"Bayview",,-34.07180023,18.58189011,,,1,,,0
hub_0011,,"Panorama Medi-Clinic",,-33.87545395,18.57602119,,,1,,,0
hub_0012,,"Lentegeur Hospital",,-34.02685928,18.61585617,,,1,,,0
hub_0013,,"A.D.E.",,-33.59735107,18.47870827,,,1,,,0
hub_0014,,"Airport Ind 2",,-33.9772377,18.58901787,,,1,,,0
hub_0015,,"Alexandra Hospital",,-33.93006897,18.48491859,,,1,,,0
hub_0016,,"Artscape",,-33.92023849,18.42955971,,,1,,,0
hub_0017,,"Athlone Ind",,-33.97886658,18.53441048,,,1,,,0
hub_0019,,"Batavia Special Needs School",,-33.98359299,18.48901939,,,1,,,0
hub_0020,,"Bayside Centre",,-33.82484055,18.48987007,,,1,,,0
hub_0021,,"Dolphin Beach",,-33.82597351,18.47934151,,,1,,,0
hub_0022,,"Beacon Valley",,-34.03613663,18.61898232,,,1,,,0
hub_0023,,"Town Centre",,-34.0479393,18.62057686,,,1,,,0
```

Figure 4.1: Stops .txt file of a GTFS data set

All the other .txt files of the GTFS data set will have this structure. They are linked by the headings of the comma-separated columns. If certain .txt files have the same headings as another .txt document, then those files are linked and share information when they are queried.

4.3.2 FeedValidator

The GTFS data needs to be tested to check the correctness and completeness of the data, which was done using the FeedValidator Python tool created by Google (Watkins et al., 2011).

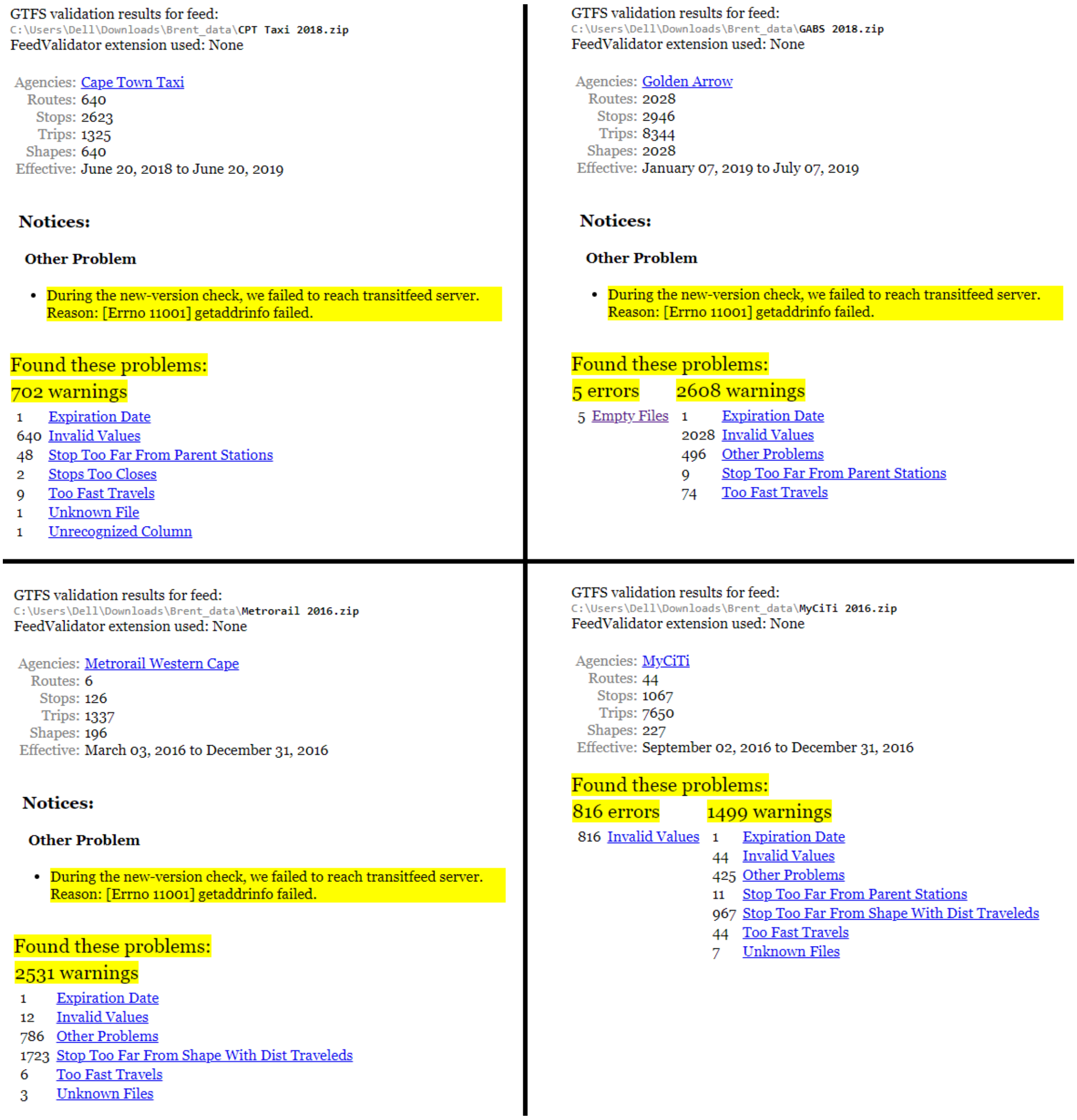


Figure 4.2: FeedValidator Results for the various transit services GTFS data

The figure above shows the FeedValidator results for the various transit services GTFS data. Although all the data sources have warnings, this does not imply the GTFS data is unusable (Watkins et al., 2011). From the results obtained it is implied that the GTFS data is complete for the purposes of network analysis, however may lack the correctness in results within certain regions of the GTFS data source. The Empty Files and Invalid Values errors do not affect the routing capabilities of using the GTFS data source within ArcGIS (Watkins et al., 2011).

4.3.3 Road Network

A road network was required to enable GTFS network analyst layer to be created. The road network needs to be georeferenced to the same coordinate system as the GTFS transit stops. The road network should have the same coverage as the GTFS data, this being the City of Cape Town. This data was acquired using OpenStreetMap.

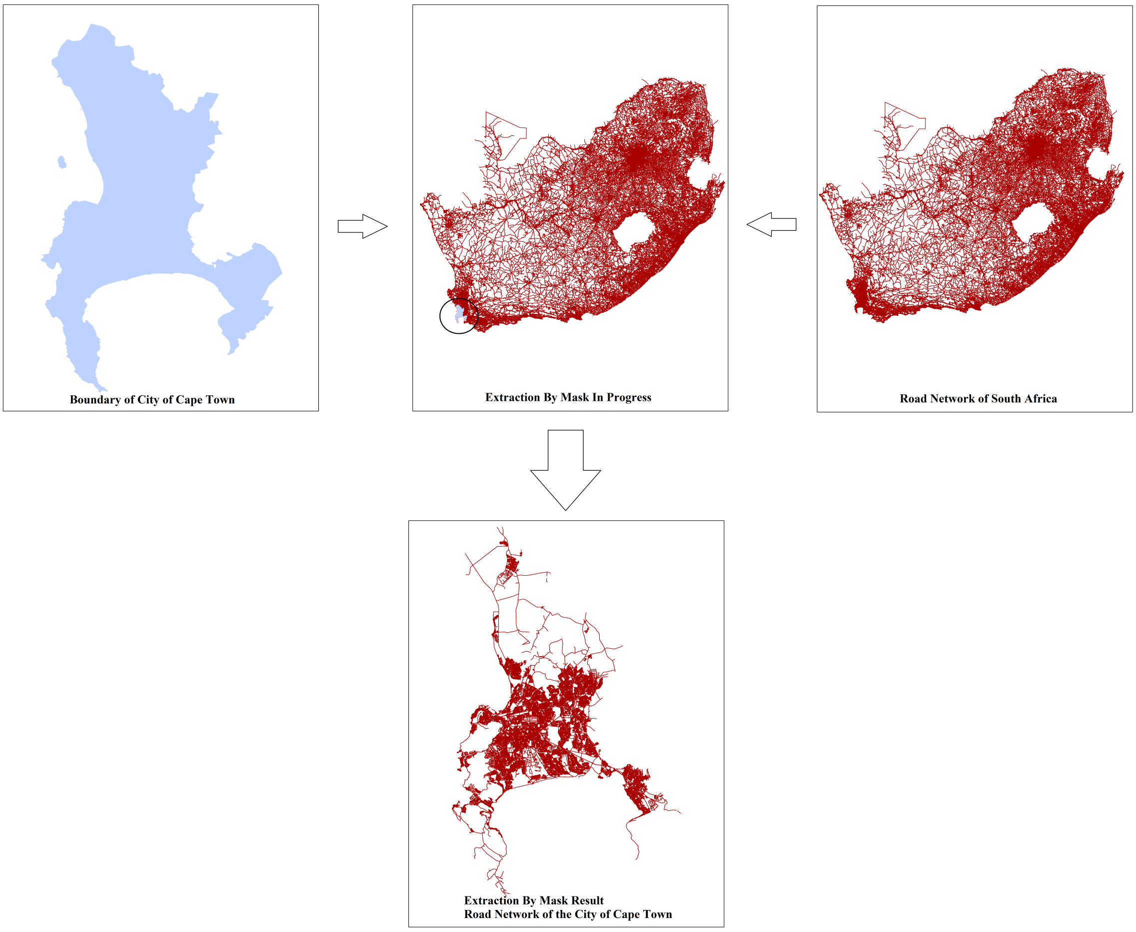


Figure 4.3: Extraction of The road network for the City of Cape Town using the Extraction By Mask tool (ESRI, 2018a)

The OpenStreetMap road network is large and only the road network for the City of Cape Town needed to be extracted. This can be executed by using the ArcGIS tool Extraction By Mask (ESRI, 2018a). This tool used a shape file that covers the region of interest, the City of Cape Town, and the road network that a road network subset is being extracted from, which is visualised in figure 4.3. The wards polygon shape file for the City of Cape Town was used as a coverage for the City of Cape Town to extract the road network from the OpenStreetMap road network data set.

4.3.4 Census Data

Opportunity Data

The opportunity/jobs data assisted in the determination of the accessibility levels of a particular area. Opportunity data is described in section 2.6.2. The higher the number of opportunities to a location and the closer the opportunities are to a location, the higher the accessibility levels for that location become (Martens, 2017). The opportunity data is recorded per TAZ region in 2013 which is based on the census data of 2011 (Census, 2011). There were 15 opportunities used in this research project defined as Office, Retail, Manufacturing, Service, Warehouse, Community, Construction, Transport, Parking, Recreation, Personel(School), Scholars(School), Personel(Tertiary), Student(Tertiary) and Agriculture. These 15 opportunities have variations in the distributions of their locations and frequencies within the different TAZ as seen in Figure 4.4.

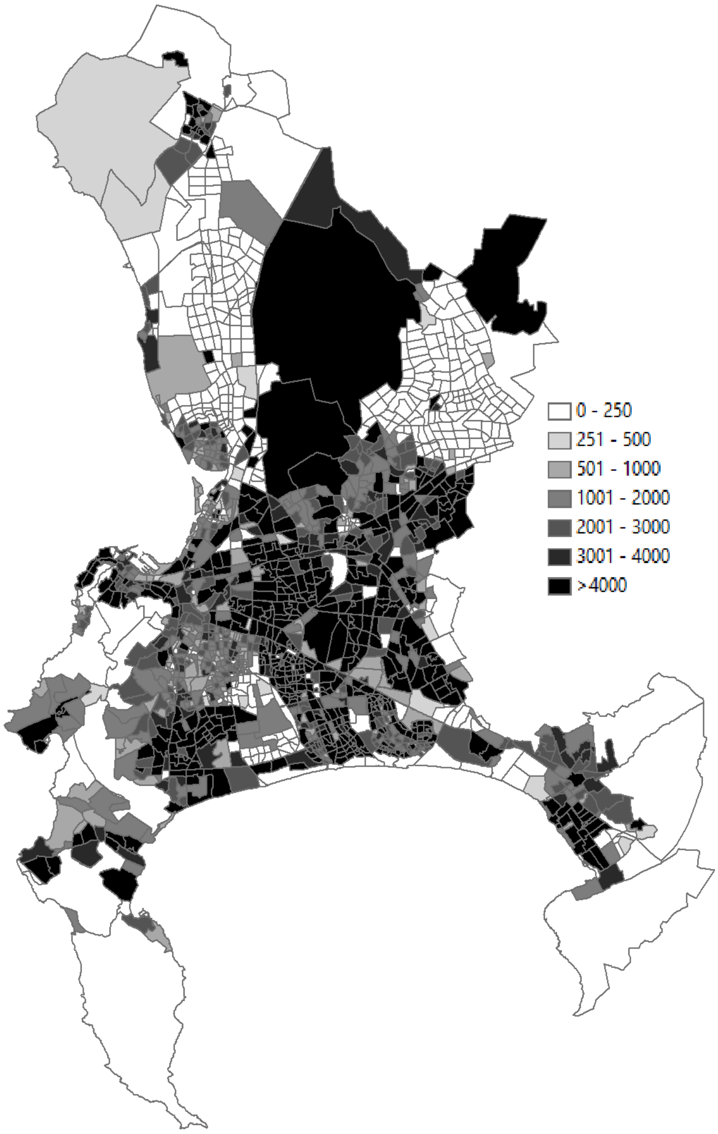


Figure 4.4: Distribution of Cumulative Opportunities for the City of Cape Town Per TAZ

Figure 4.4 shows the distribution of the cumulative opportunities per TAZ region for the City of Cape Town. The darker regions are the TAZ with higher opportunity levels and the lighter areas are regions with lower opportunity levels. This is under the assumption that these opportunities are all open and operating at the same time. These opportunities consist of Office, Retail, Manufacturing, Service, Warehouse, Community, Construction, Transport, Parking, Recreation, Personel (School), Scholars (School), Personel (Tertiary), Student (Tertiary), and Agricultural.

A matrix was created that contained operating times for these 15 opportunities. This allowed a temporal dimension to be added to the accessibility measure.

These operational times were approximated by analysing the open and closing times provided by Google Maps APIs for these types of opportunities.

Operating_Time	Office	Retail	Manufactur	Service	Warehouse	Community	Constructio	Transport	Parking	Recreation	Personel (S	Scholars (S	Personel (T	Student (Te	Agricultural
0:00:00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:00:00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2:00:00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3:00:00	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
4:00:00	0	0	0	0	0	0	0	1	1	1	0	0	0	0	1
5:00:00	0	0	0	0	0	0	0	1	1	1	0	0	0	0	1
6:00:00	0	0	1	0	1	1	0	1	1	1	0	0	0	0	1
7:00:00	0	0	1	0	1	1	1	1	1	1	1	0	0	0	1
8:00:00	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
9:00:00	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10:00:00	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11:00:00	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12:00:00	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13:00:00	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14:00:00	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15:00:00	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16:00:00	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
17:00:00	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1
18:00:00	0	1	0	1	1	1	1	1	1	0	0	0	0	0	1
19:00:00	0	1	0	1	1	0	0	1	1	0	0	0	0	0	0
20:00:00	0	1	0	1	1	0	0	1	1	0	0	0	0	0	0
21:00:00	0	1	0	1	1	0	0	1	1	0	0	0	0	0	0

Figure 4.5: Operating Hours of Opportunities for the City of Cape Town Per TAZ

The value of 0 represented that this particular opportunity was closed and not accessible at that time. The value of 1 represented that this particular opportunity was open and accessible at that time. The operating times of the various opportunities varied throughout the day and affected the result of accessibility. Having this matrix with 1 and 0 helped make the calculations of accessibility much easier as opportunities with the value of 0 were not included in the calculation of accessibility.

Income Data

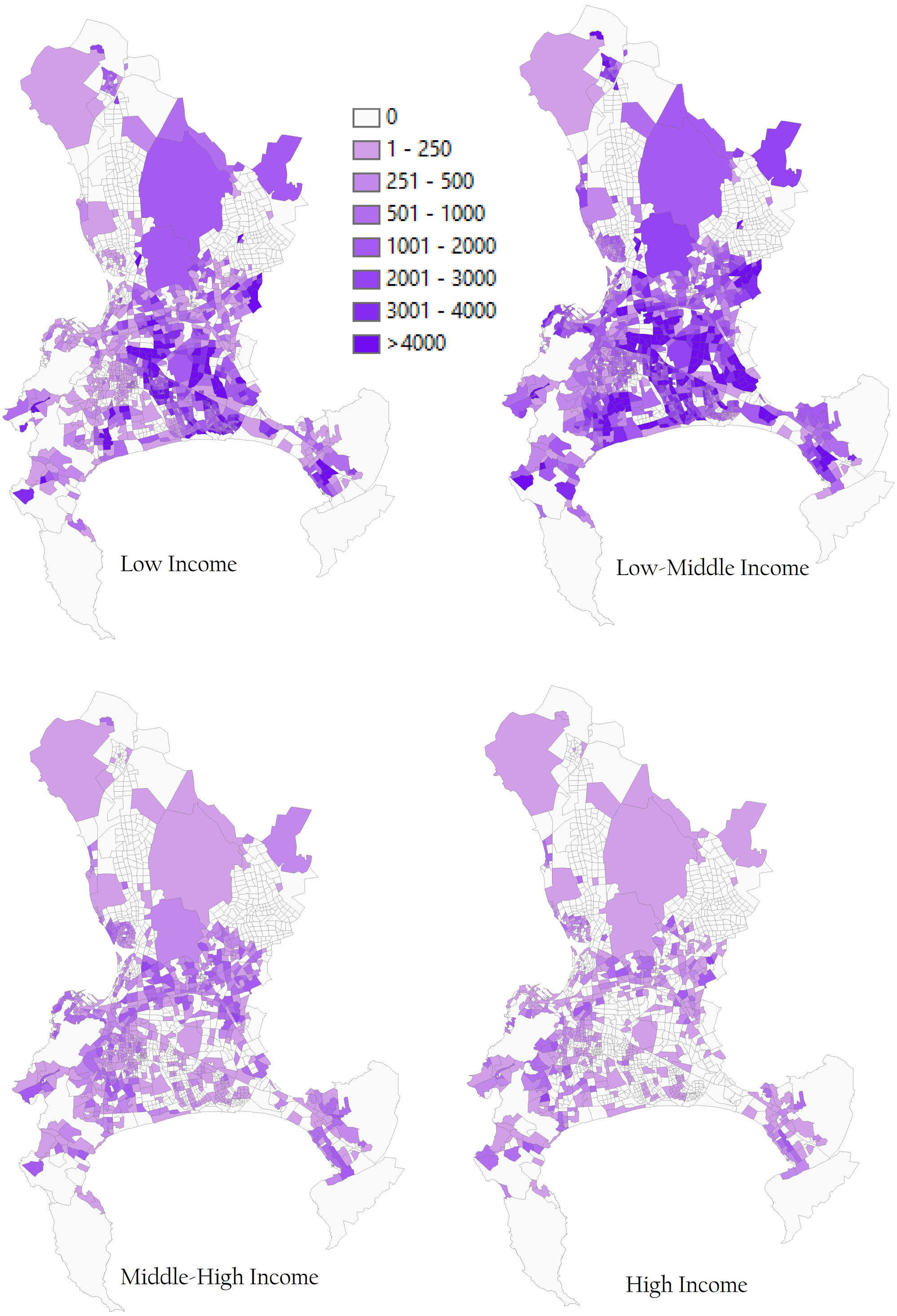


Figure 4.6: Income Distribution of the Population for the City of Cape Town

The population data helped to determine in the event of potential accessibility and mobility deficits in a particular region, if this region was entitled to an accessibility and potential mobility increase. This population and income data was based on the census data of 2011 (Census, 2011). If a potential accessibility and mobility deficit occurs, the region that experienced this deficit and the population numbers are too low to qualify for a potential accessibility and mobility increase.

Not only does this population data have the number of people that reside in a TAZ, but it also splits up in the income levels of the populations in the different regions. The population is split into four different income categories: Low, Low-Middle, Middle-High, and High Income. This income distribution among the population can be seen in Figure 4.6. This income distribution could potentially be used to aid in the decision on increasing a region's accessibility and potential mobility levels, i.e. poor communities are more entitled to an accessibility and potential mobility increase over richer communities. The income distribution implementation for this project can be seen in 4.7.3.

Demographic Data

The demographic data shows the distribution of the various races within the City of Cape Town as displayed in Figure 4.7. The darker regions regions more densely populated than the lighter regions for the given demographic group. This demographic data was based on the census data of 2011 (Census, 2011). Looking at the distribution of the different demographic groups could be used as another dimension for the determination if a particular is entitled to an accessibility and potential mobility increase or not. This was done by analysing the distribution of the demographic groups within a chosen TAZ region under analysis. It was in the favour of a TAZ that is mainly populated by demographic groups that were previously disadvantaged due to apartheid, to get an accessibility and potential mobility increase if they fall below the sufficiency threshold.

The distribution of the different demographic groups area was largely affected by the Group Areas Act of 1950 (Loveland, 1999). This act made locations of higher accessibility and potential mobility levels to be in the possession of the white population. As seen in Figure 4.6, the white population has kept this spatial distribution by their own decision to keep living in those locations. Whereas, the coloured and black population became more evenly dispersed across the city after the Group Areas Act of 1950 was removed post-apartheid. If the white population and high-income spatial distribution are compared, a strong correlation is seen between these factors. People with high incomes generally have access to a car and choose not to participate in the use of public transport (Martens, 2017).

Therefore, certain regions may appear to have poor potential accessibility and mobility levels using public transport, they may have extremely high accessibility and potential mobility levels using their personal vehicle.

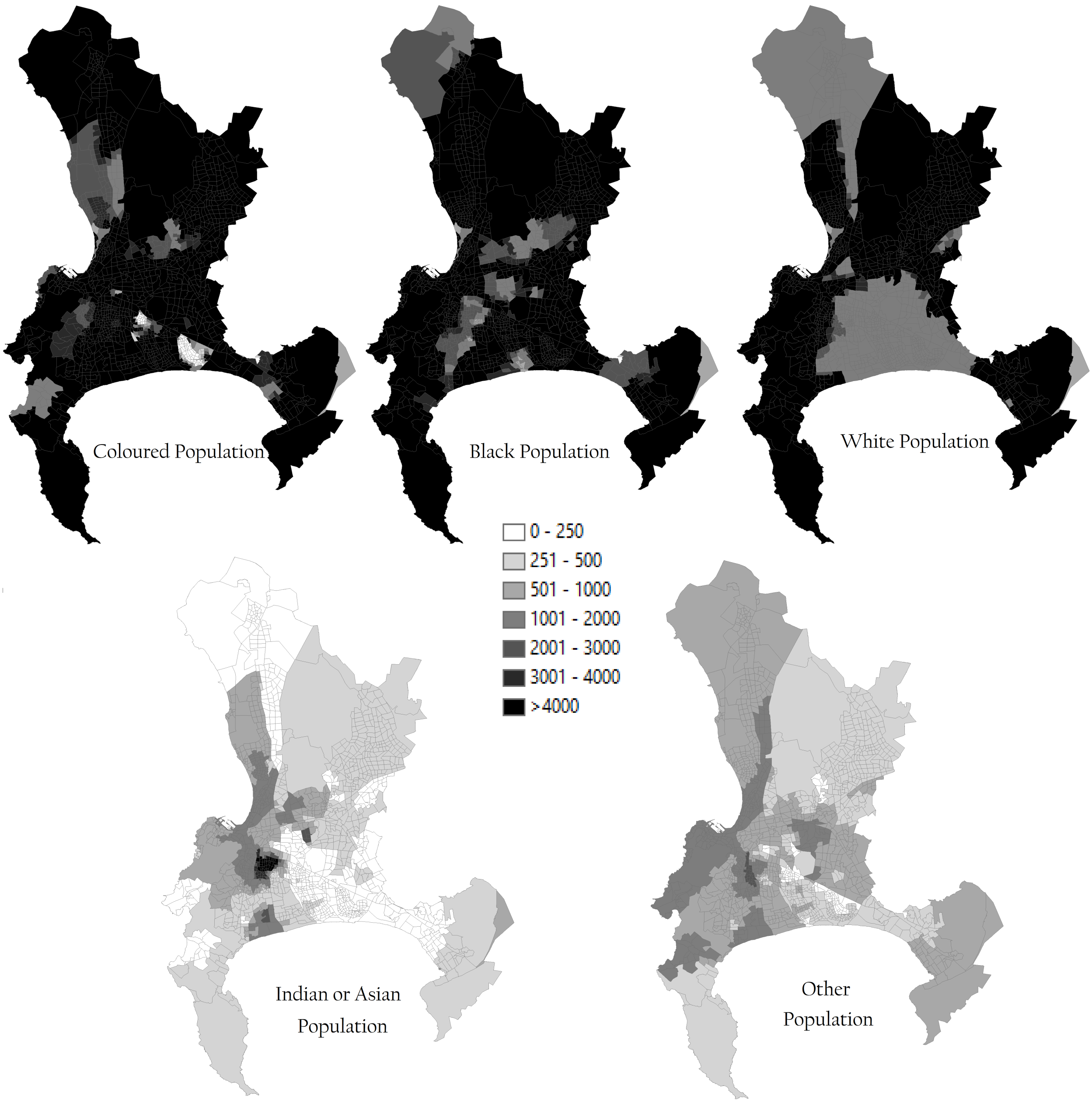


Figure 4.7: Demographic Distribution for the City of Cape Town

The above figure shows the spatial distribution of the different racial groups of Cape Town based on the cumulative population for the census 2011 data (Census, 2011)

4.4 Urban Development Index (UDI) Modal Split

The UDI modal split measures the percentage contribution of the various forms of transport, including private transport (City of Cape Town, 2019). However, for the purposes of this project, only the public transport UDI will be used.

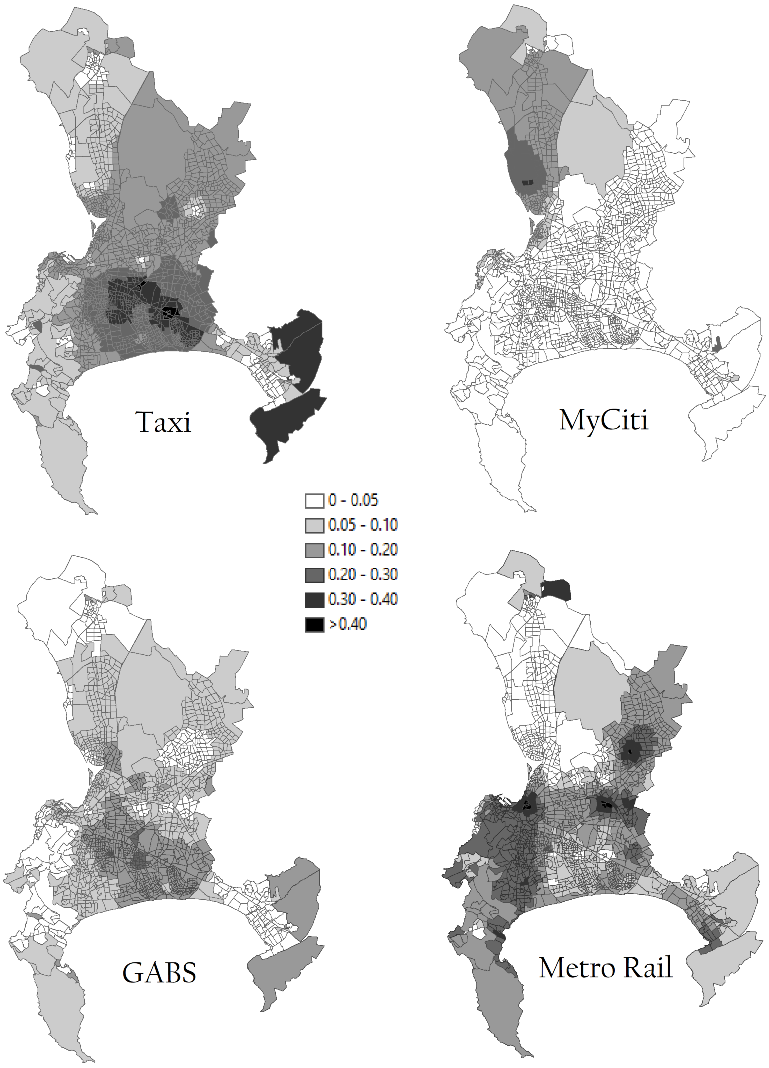


Figure 4.8: Urban Development Index (UDI) Modal Split of the Various Transit Modes for the City of Cape Town (City of Cape Town, 2019)

The UDI distribution of the various transport modes can be seen in Figure 4.8. The darker regions indicate the regions use that particular transport mode more than the lighter regions within the City of Cape Town. The UDI modal split will be used to determine a weighted average in potential accessibility and mobility in the Space-Time Cube transport justice tool for all four public transport systems used in this project. Implementation of the UDI modal split can be seen in chapter 5.1.

4.5 Network Analyst Layer

Traditionally only a road network is required to create a network analyst layer. However, since a time-dependent network analyst layer was created using GTFS data, additional steps were required. Using the ArcGIS tool "Add GTFS to a Network Dataset"(ESRI, 2018b) assisted with connecting the four public transport GTFS datasets to the network analyst layer.

4.5.1 Transit Stops and Lines

As mentioned before, the transit stops contain the geographical positions of the transit stops of the particular transport mode. These stops determined the coverage a person could travel to. Using the "Add GTFS to a Network Dataset" tool created a georeferanced point feature class for the transit stops and a line feature class that connects consecutive transit stops on various routes that a transit used (ESRI, 2018b).

As illustrated in Figure 4.9, the transit stops were connected using the line feature class. These connections are necessary for the network analyst layer to make use of the GTFS time schedules and make the travel time using the network analyst time dependent. The number of stops and the locations of these stops affected the potential accessibility and mobility level a user of public transit would experience. The more dispersed these transit stops are the more accessible locations become. This leads to an increase in potential accessibility and mobility, and vice versa.

With respect to the four public transport GTFS datasets, the GABS and Taxi transport modes have high density and evenly dispersed transit stops which indicates that these transport modes should have a higher overall level of potential accessibility and mobility. Whereas the MyCiti and Metro Rail transport modes have minimal transit stops as compared to the GABS and Taxi transit modes. The Metro Rail transit mode has a more even distribution compared to the MyCiti transit mode, however has much fewer transit stops. This made it more difficult for users of the transit to use this transit mode which decreased the levels of potential accessibility and mobility. The MyCiti transit stops have clustered transit stops around the CBD region. This makes these regions have a high potential accessibility and mobility level but experience extremely low levels of potential accessibility and mobility for the rest of the regions of the City of Cape Town.

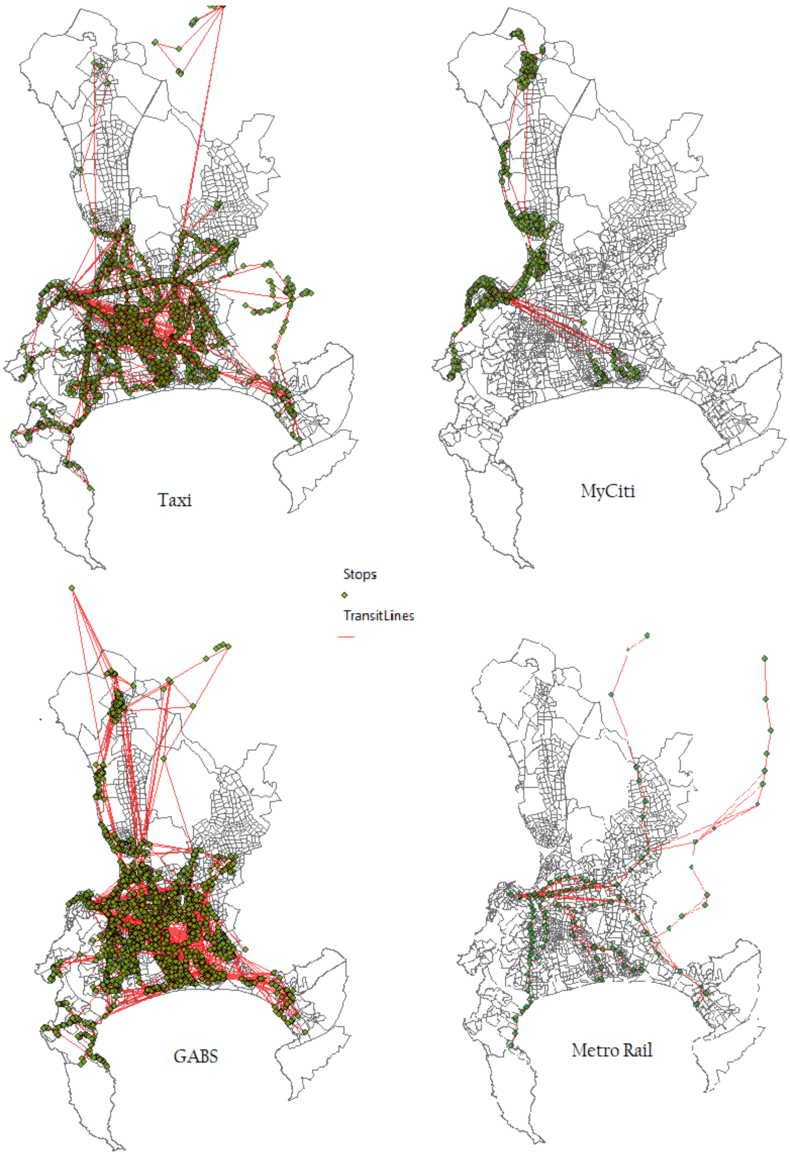


Figure 4.9: Transit Stops and Lines of the Various Transport Modes for the City of Cape Town

4.5.2 Transit Stops to Streets Connectors

An additional link needed to be created between the transit stop and the road network to complete the parameters required for the network analyst layer. Using the "Add GTFS to a Network Dataset" tool allowed the link between the transit stops and the road network to be created. This link was a line feature class. This connector is created using a perpendicular line between the transit stop and the road network.

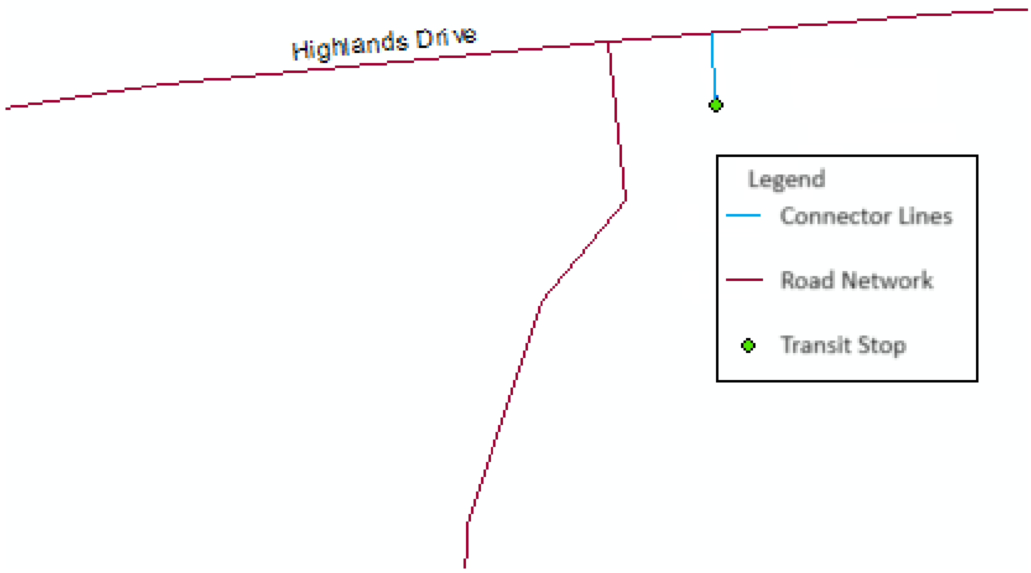


Figure 4.10: Transit Stop and Road Network with a stop to street connector

Not only does this link make the GTFS network analyst layer possible, but it also serves as a time interval for a user of the public transit system to get on and off a transit mode. This tries to make the simulation of travel time using the network analyst layer closer to reality. The time chosen to get on and off the public transit was set to be 0.25 minutes based on the research of (Farber, Morang and Widener, 2014), which used this value for their Network Analyst Layer to achieve the same impedance attributes values that this research paper is determining (ESRI, 2018b).

4.5.3 Network Analyst Layer

The network analyst layer could finally be created after the previous step are completed. All the data mention above needs to be stored within the same feature dataset. The network analyst was available by right clicking on the feature dataset and selecting "New", then "Network Dataset" (ESRI, 2018b).

Connectivity

It was crucial that the connectivity between the feature classes are as follows in figure 7.1 in Appendices to create the network analyst layer. If it does follow the connectivity displayed above, it produced incorrect travel time values and many locations will not be traversable.

Parameters

The "Use Specific Dates" parameter made it possible for the network analyst layer to use different days of the week as well as different hours of the day to analyse. The "Walk Speed" parameter determines how fast the user walks when the user is walking on the network analyst layer. This was recorded in meters per minute. The current speed has been set to 80 meters per minute (ESRI, 2018b). Only the Walk Speed parameter formed part of the evaluators.

Evaluators

Figure 7.3 in Appendices shows the selection of the type and value fields for the different sources to create the Travel Time With Transit evaluators. Using these values ensured that the sequence of using a transport mode was simulated in the correct order and true to a real-life scenario using this evaluator (ESRI, 2018b).

4.5.4 Network EIDs

Although the network analyst has been successfully connected, it was not time-dependent because the GTFS schedule was not linked to the network analyst layer. Using the "Add GTFS to a Network Dataset" tool allowed the network analyst layer to be linked to the GTFS schedules and make the time of day travel time possible. Use the "Get Network EIDs" feature on the "Add GTFS to a Network Dataset" tool (ESRI, 2018b).

4.5.5 TAZ Points

TAZ (Transport Analysis Zones) points served as the origin and destination locations from which the population groups travel to and fro using public transportation systems.

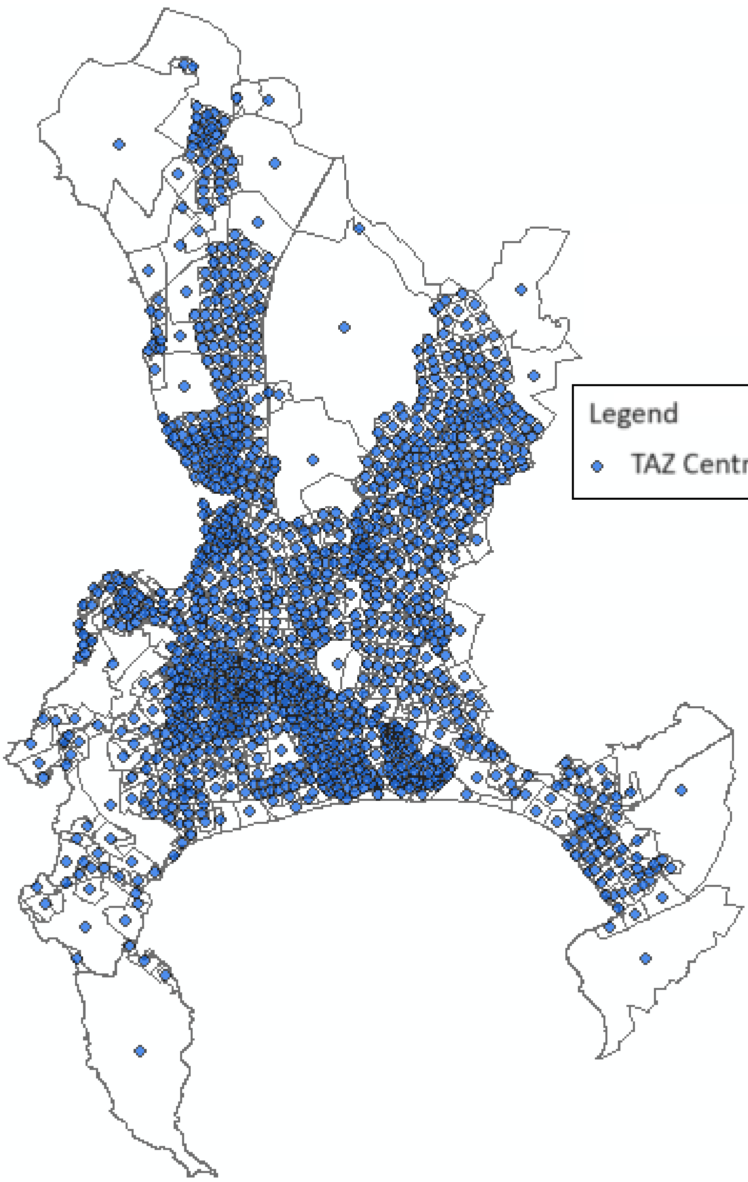


Figure 4.11: TAZ Points of the City of Cape Town

Figure 4.11 illustrates the distribution of the TAZ point which was used for the location points in the Network Analyst Layer. These points are the centroids of the TAZ polygon features. This was done because the Network analyst layer did not utilise polygon features in its calculations. This was overcome by using the Feature to Point tool in ArcGIS(ESRI, 2018c). This tool took the TAZ polygon feature and converted it to a point feature so that the TAZ locations could be used in the OD Cost Matrix.

When importing the origins and destination points into the OD cost matrix, the origins and destination points have the same point name as the points of TAZ.

The "Sort Field" sorted the origins and destinations of the OD cost layer according to the names of the TAZ points. This helped to know the exact order the OD cost matrix calculated the origin-destination pairs. Knowing the order helped with querying the data to get exactly what you are trying to find. The "Name" property ensured that the origins and destinations of the OD cost matrix layer had the same names as the TAZ points. The Field of "Name" needed to be the same as the selection for "Sort Field" as shown in Figure 7.4 in Appendices.

4.6 Impedance Attribute

4.6.1 OD Cost Matrix

The OD cost matrix allowed for the simultaneous origin-destination pair travel time calculation. This was calculated using all 1616 TAZ regions as their origins and destinations for every hour of the day. The days Tuesday, Wednesday, and Thursday were selected to perform analysis. These days are representative of a typical working day.

Once the OD cost matrix was created the Layer Properties tab of the OD cost matrix needed to be adjusted for use with a GTFS network analyst layer.

As displayed in Figure 7.5 in Appendices, only the "Streets Use this One" shapefile needs to be checked, everything else remains unchecked. If this was not done then some origins may not be able to link to the network analyst layer and would not be able to reach any destinations. The "Search Tolerance" needs to be a high number if the origin is very far from the network analyst layer. If it is not high enough the origin will not feature in the calculations of the OD cost matrix.

4.6.2 Transit Network Analysis Tools

Using the OD cost matrix to calculate the desired Impedance attribute could only be done at one single-hour iteration. Since the analysis needs to be done for an entire day it was not practical to run the analysis 24 times to complete the analysis for the day. The "Transit Network Analysis Tools" has a function called "Calculate Travel Time Statistics" which calculates the travel time at each hour of the day (ESRI, 2018d). This output produced a line feature shape file that contained the travel times between origins and destination in the attribute table. This attribute table information was the data that was required for the transport justice tools and needed to be exported as a .txt document. This was achieved using the "Table To Table" tool in ArcGIS.

4.7 Transport Justice Tools

4.7.1 UDI modal split

As described in Chapter 3, the Space-Time Cube calculated the accessibility (equation 3.2) and PMI (equation 3.3) of a given region for a 24-hour day for all public transit services. The TAZ regions that were chosen for analysis were Khayelitsha, Mitchells Plain, and Bishops Court for a Tuesday.

The UDI modal split will also be used to display a weighted Space-Time Cube.

Table 4.1: UDI Modal Split for Public Transit Contribution

TAZ	**Region**	**Metro(%)**	**MyCiti(%)**	**GABS(%)**	**Taxi(%)**
3413	Khayelitsha	10.0	2.6	13.2	30.9
3071	Mitchells Plain	7.5	0.5	18.3	26.5
1601	Bishops Court	24.9	0.6	4.2	13.1

Table 4.1 shows the UDI Modal Split for Public Transit Contribution for the areas of analysis using the transport justice tools. The remaining percentage of the UDI Modal Split was complied by NMT and private transport vehicles.

4.7.2 Location of Selected Case Studies

The locations of Khayelitsha, Mitchells Plain and Bishops Court where chosen to be case studies using the transport justice tools created. These locations were largely affected by the Groups Area Act of 1950. The racial group allocations of these locations during apartheid can still be seen today. The selection of these locations aims to highlight the various accessibility and potential mobility levels experienced by the different racial groups, as their residential locations have become correlated to their race partly due to apartheid.

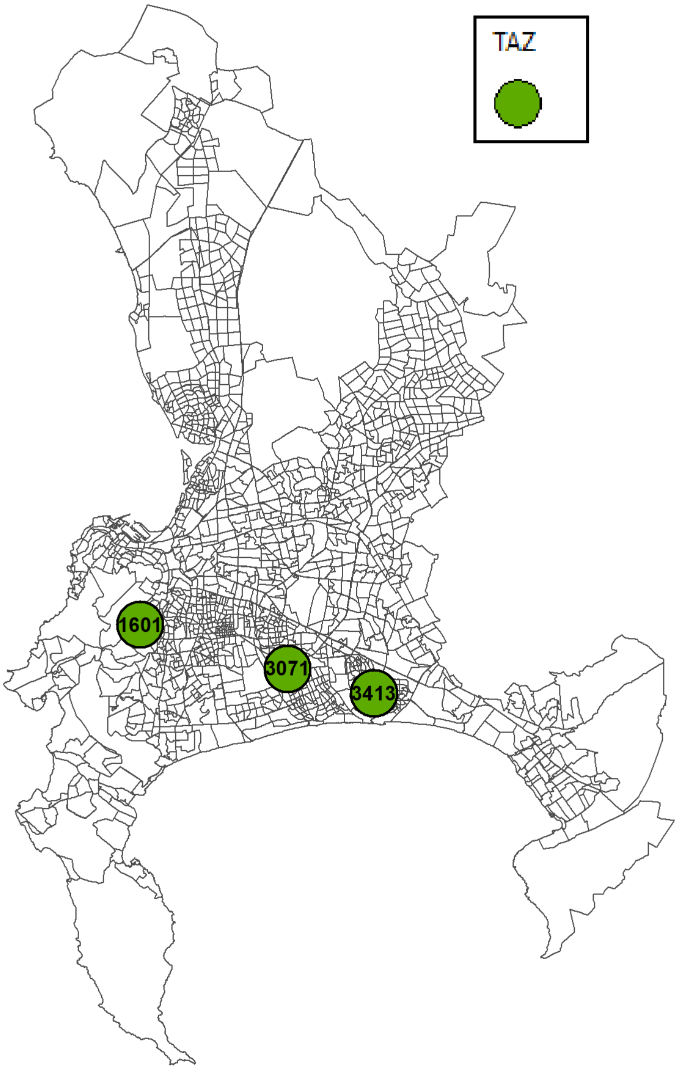


Figure 4.12: Selected TAZ Points Location for Analyses

The figure 4.12 shows the graphical locations of the selected TAZ for the analyses. These locations were regions part of the group areas act of 1950, where TAZ 1601 was declared a White only area, TAZ 3071 was declared a Coloured only area and TAZ 3413 was declared a Black-African only area.

4.7.3 Income Distribution in Case Study Areas

This section highlights the income distribution of the locations chosen in the case studies. The income distribution in these locations was categorized in 4 different classes. These classes are Low, Low-Middle, Middle-High, and High. The income distribution of these case studies aimed to indicate income inequalities between the case studies. Specifically, emphasizing the population groups and size that live within the different case studies locations. Using the parameters of income, population size and population groups will aid in determining if an area qualifies for a potential accessibility and potential mobility increase or not. A potential accessibility and potential mobility increase will be in the favour of case studies that have low income, high population size, and large number of previously disadvantaged population groups residing within the case studies locations. This statement mentioned above will only apply if the case studies location has low potential accessibility and low mobility levels.

Khayelitsha

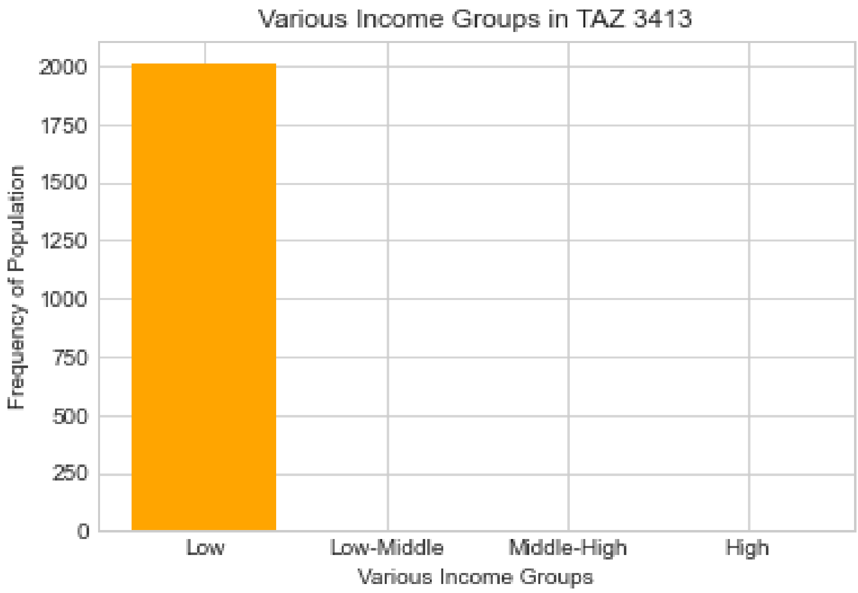


Figure 4.13: Income Distribution for 3413 Khayelitsha

Khayelitsha was one of the areas allocated to be occupied by the Black-African population only. During apartheid, the black population had the worst pay level due to the unjust laws of apartheid, which led to this population group still having a low-income level up to today as illustrated in Figure 4.13.

Mitchells Plain

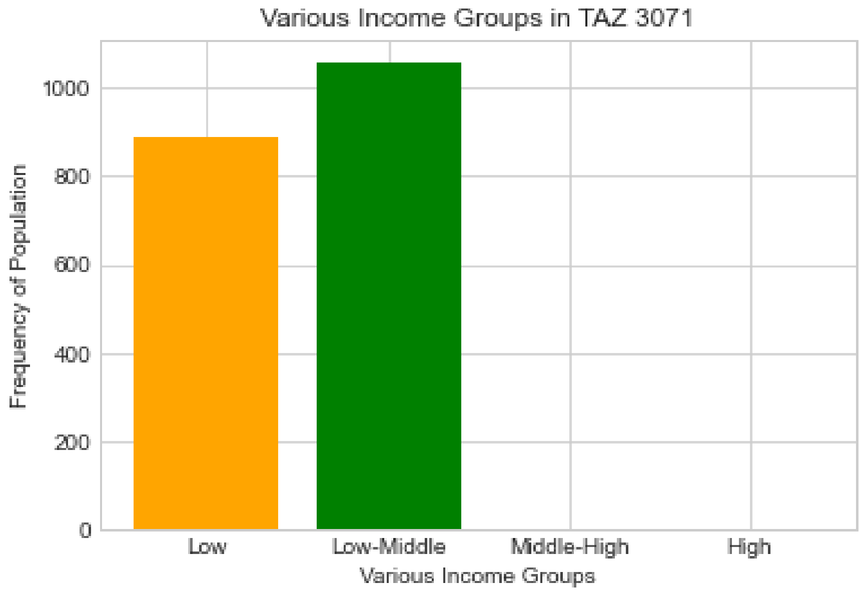


Figure 4.14: Income Distribution for 3071 Mitchells Plain

Mitchells Plain was one of the areas allocated to be occupied by the Coloured population only. During apartheid the Coloured population had a pay level less than the white population but higher than the Black population due to the unjust laws of apartheid, which lead to this population group to still have a low income and low-middle level up to today as illustrated in Figure 4.14.

Bishops Court

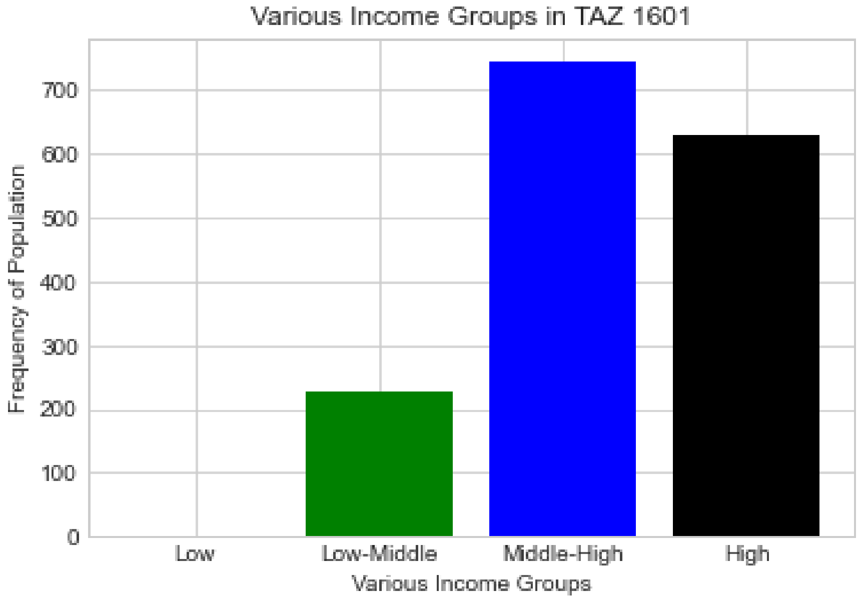


Figure 4.15: Income Distribution for 1601 Bishops Court

Bishops Court was one of the areas allocated to be occupied by the White population only. During apartheid, the White population had the best pay level due to the unjust laws of apartheid, which led to this population group still having a high-income level up to today as illustrated in Figure 4.15.

4.7.4 Demographic Groups in Case Study Areas

Although after the unjust apartheid laws that segregated races were removed, the population groups that were displaced have built communities and personal relationships within these displaced locations. The population groups who were displaced to these racially segregated locations still reside there today. Also these previously disadvantaged population groups, i.e. black and coloured populations, the majority still had lower income levels as compared to the white population groups. They still can not afford to move to a previously white-only area due to the high cost of residing in these locations. This demographic data will be used to aid in the decision-making process to identify if a TAZ location qualifies for potential accessibility and potential mobility increase. The demographic groups that were previously disadvantaged will be favoured more for potential accessibility and potential mobility increase if their TAZ falls within the parameters of qualifying for potential accessibility and potential mobility increase.

Khayelitsha

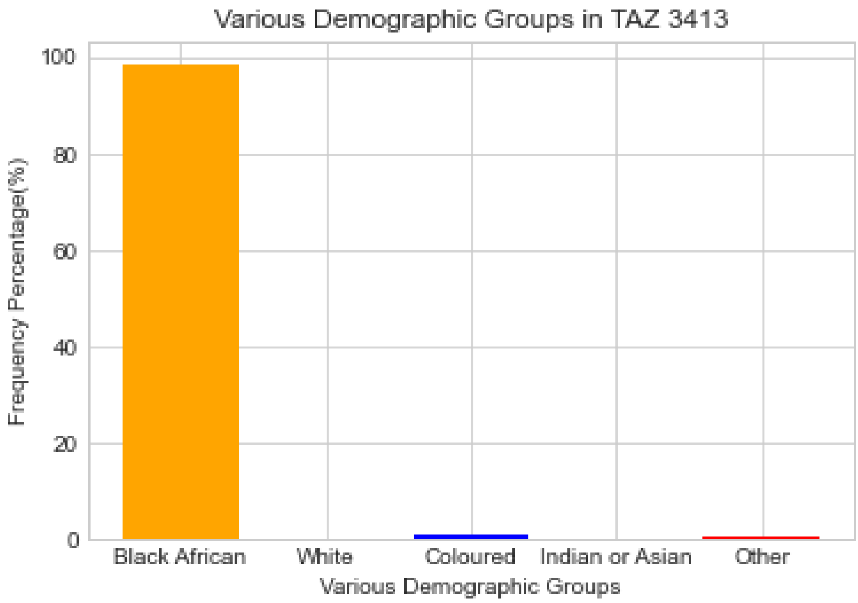


Figure 4.16: Demographic Distribution for TAZ 3413 Khayelitsha

The demographic groups which reside in Khayelitsha are still the majority of the Black African population. This region is still considered a poor area which makes taking up residence in Khayelitsha easy to live while still having a low income level as indicated in Figure 4.13.

Mitchells Plain

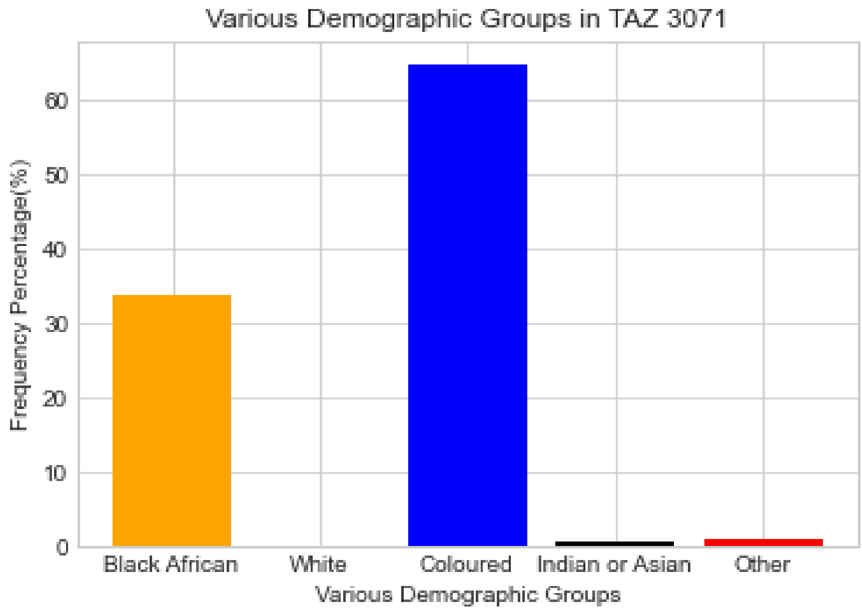


Figure 4.17: Demographic Distribution for TAZ 3071 Mitchells Plain

The demographic groups that reside in Mitchells Plain are still predominantly Coloured. This region is still considered a poorer region when compared to a previously white-only area, which makes taking up residence in Mitchells Plain easier while still having a lower income level.

Bishops Court

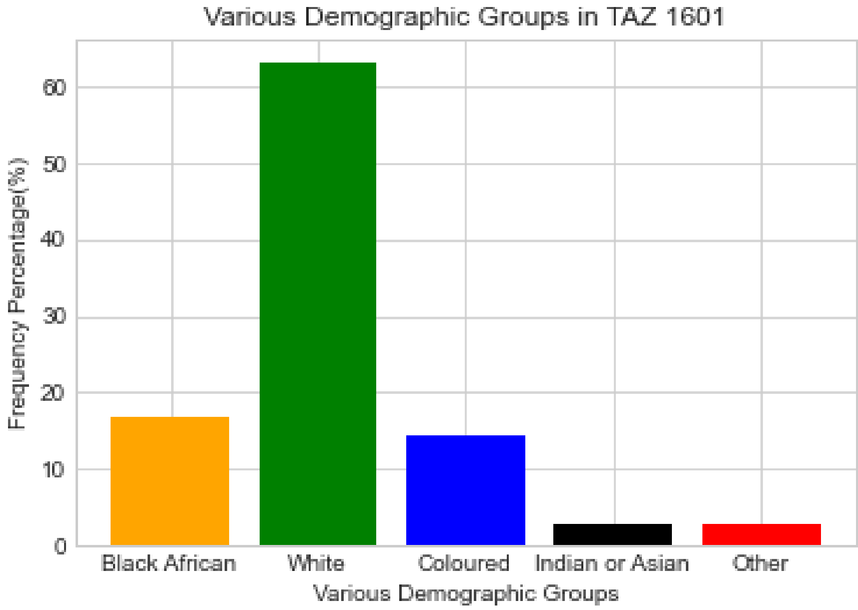


Figure 4.18: Demographic Distribution for TAZ 1601 Bishops Court

The demographic groups that resides in Bishops Court are still by a majority the White population. This region is still considered a richer region since it was previously a white-only area, which makes taking up residence in Bishops Court difficult to live there while still having a lower income level. Therefore, the non-white population groups that resided here post-apartheid had the financial capability to relocate to regions that were previously allocated for the white population groups only, due to the non-white population groups now being able to earn higher incomes.

Chapter 5

Results

5.1 Space-Time Cube

The goal of the Space-Time Cube was to illustrate the Accessibility and PMI levels that a TAZ region experiences throughout the day. Accessibility is a scalar value and PMI is a score value. Both Accessibility and PMI are unitless. The goal was to identify if a TAZ region has an Accessibility and PMI deficit by monitoring these measures throughout the day. The Space-Time Cube which was a spatial-temporal exhibition of potential mobility and accessibility, was tested on the three chosen TAZ case studies using various public transport systems. The various public transport systems analysed are GABS, Metro Rail, Taxi, and MyCiti. Once the Space-Time Cubes for each of these public transport systems are established, a final Space-Time Cube called Weighted Average, is then created based on weights from the UDI Modal Split in table 4.1 for each TAZ case study.

5.1.1 Khayelitsha

GABS

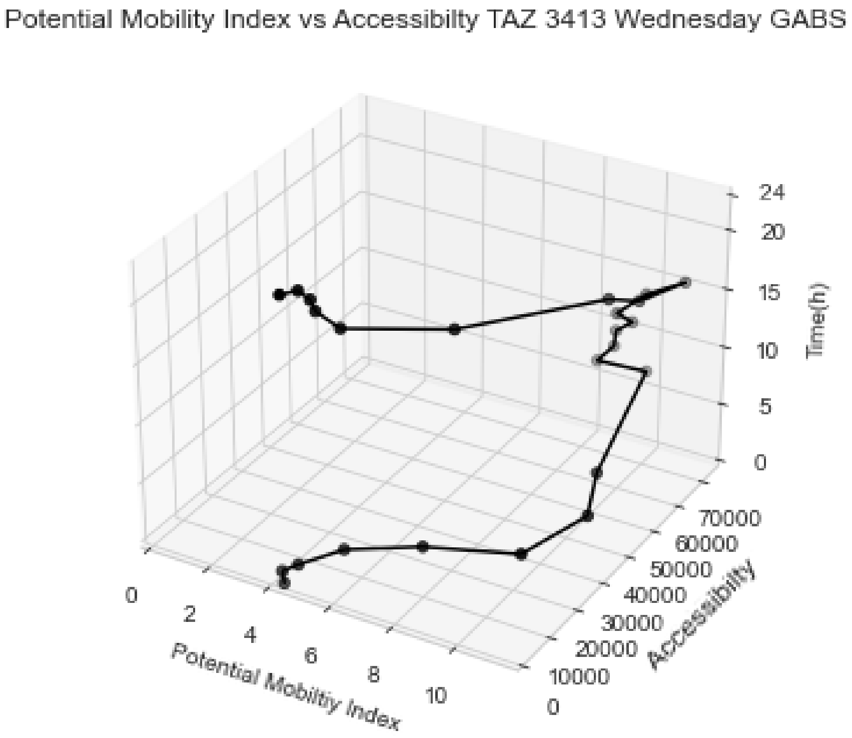


Figure 5.1: PMI VS Accessibility TAZ 3413 GABS Wednesday

The Accessibility levels increase rapidly from 0h00 to 5h00, then stabilise at 6h00 increase gradually until 15h00, then decreases rapidly until around 20h00, and gradually decreases until the day ends. The Accessibility levels experienced by the residents of Khayelitsha are good using the GABS public transit as good levels of Accessibility are reached throughout the day relative to the other transit modes.

The PMI levels increase rapidly from 0h00 to 4h00, then stabilize at 4h00 to 7h00 then decreases and stabilise until 15h00, then decreases rapidly until the day ends. The PMI levels experienced by the residents of Khayelitsha are good using the GABS public transit, as very good levels of PMI are reached throughout the day compared to the other transit modes.

Metro Rail

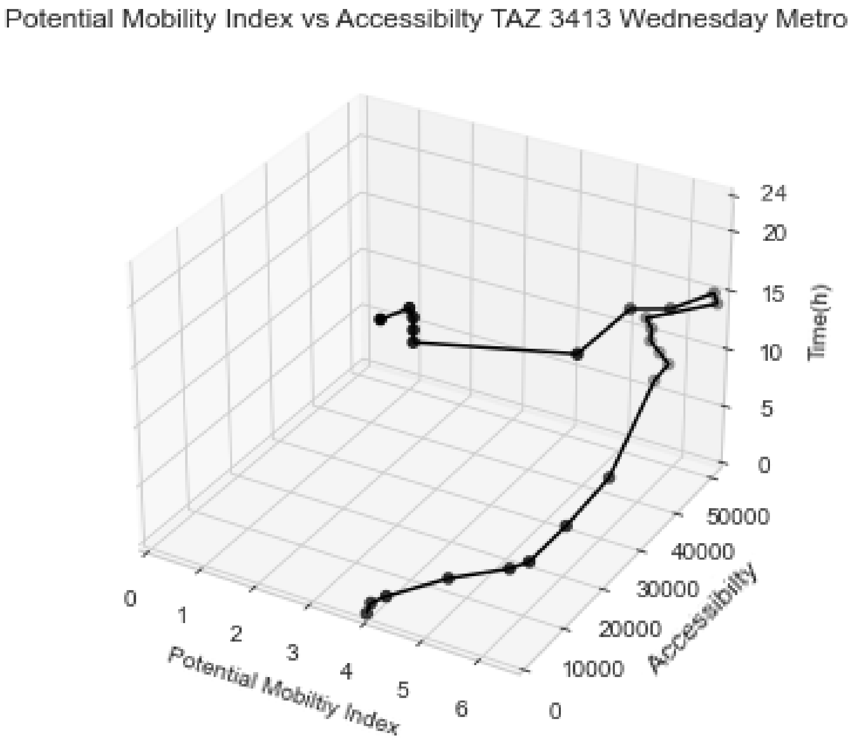


Figure 5.2: PMI VS Accessibility TAZ 3413 Metro Rail Wednesday

The Accessibility levels increase gradually from 0h00 to 5h00, then increases rapidly at 6h00 increase gradually until 15h00, then decreases rapidly until around 20h00, and gradually decreases until the day ends. The Accessibility levels experienced by the residents of Khayelitsha are good using the Metro Rail public transit as good levels of Accessibility are reached throughout the day relative to the other transit modes.

The PMI levels increase gradually from 0h00 to 4h00, then remain stable at 4h00 to 7h00 then increase gradually until 15h00, then decreases slowly until around 20h00, and rapidly decreases until the day ends. The PMI levels experienced by the residents of Khayelitsha are good using the Metro Rail public transit, as very good levels of PMI are reached throughout the day compared to the other transit modes.

Taxi

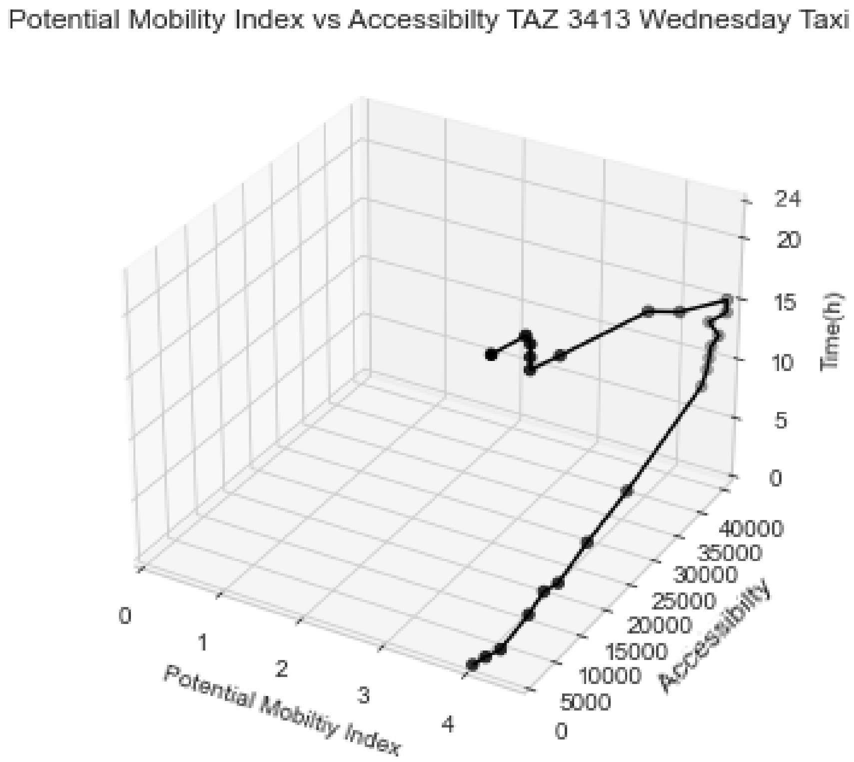


Figure 5.3: PMI VS Accessibility TAZ 3413 Taxi Wednesday

The Accessibility levels increase gradually from 0h00 to 5h00, then increases rapidly at 6h00 increase gradually until 15h00, then decreases rapidly until around 20h00, and gradually decreases until the day ends. The Accessibility levels experienced by the residents of Khayelitsha are good using the Taxi public transit as good levels of Accessibility are reached throughout the day.

The PMI levels remain very stable throughout the day, which means that the PMI level experienced will be achieved at any time of the day using this public transit for Khayelitsha. The PMI levels experienced by the residents of Khayelitsha are good using the Taxi public transit as good levels of PMI are reached throughout the day.

MyCiti

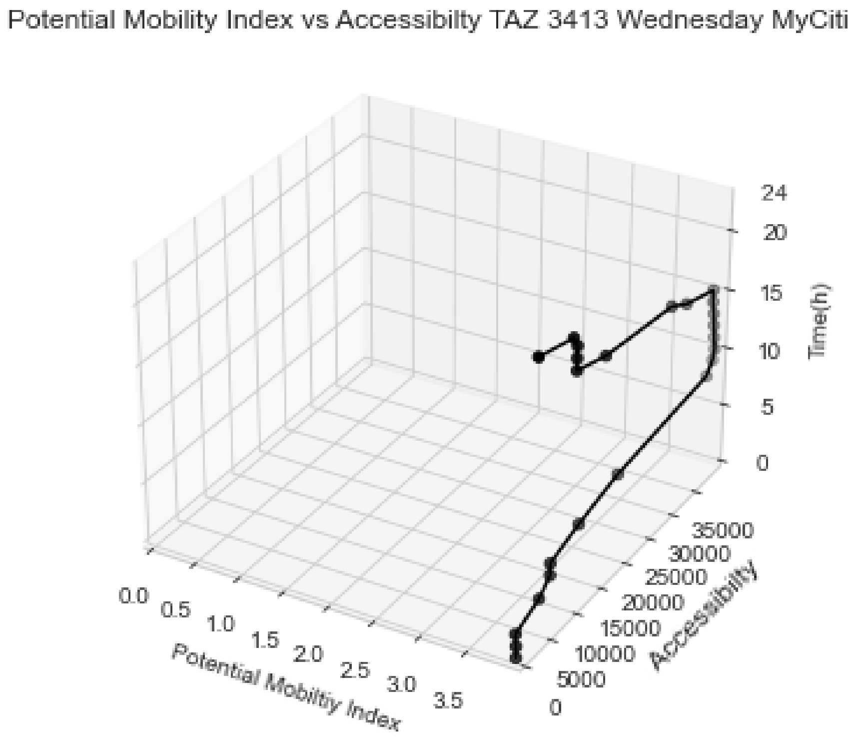


Figure 5.4: PMI VS Accessibility TAZ 3413 MyCiti Wednesday

The Accessibility levels increase gradually from 0h00 to 5h00, then increases rapidly at 6h00, then stabilize until 15h00, then decreases rapidly until the day ends. The Accessibility levels experienced by the residents of Khayelitsha are good using the MyCiti public transit as good levels of Accessibility are reached throughout the day.

The PMI levels remain very stable throughout the day, which means that the PMI level experienced will be achieved at any time of the day using this public transit for Khayelitsha. The PMI levels experienced by the residents of Khayelitsha are good using MyCiti public transit as good levels of PMI are reached throughout the day.

Weighted Average

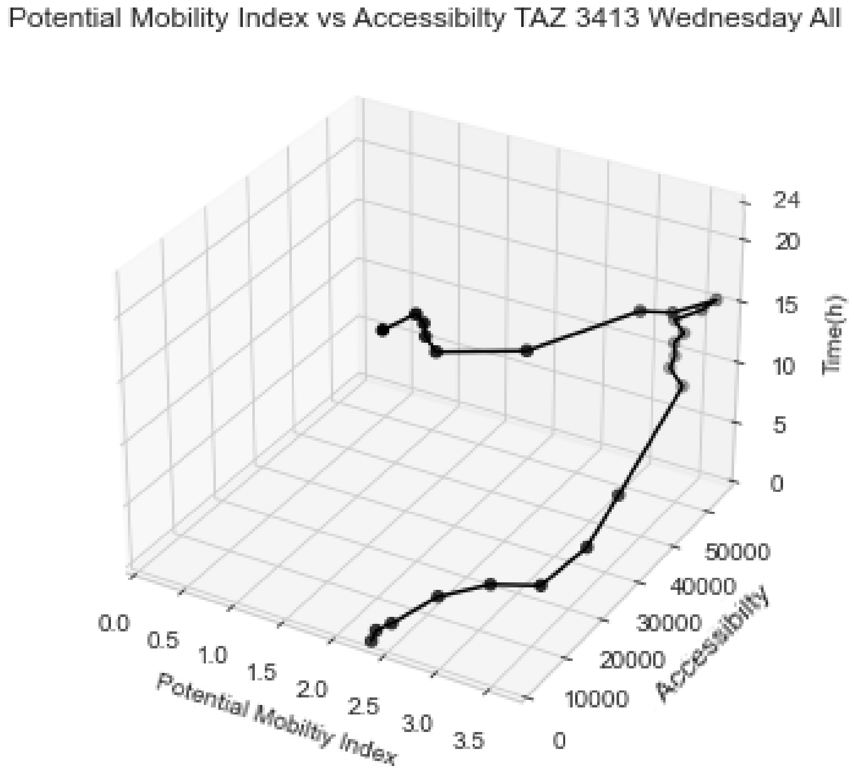


Figure 5.5: PMI VS Accessibility TAZ 3413 Weighted Average Wednesday

The Accessibility levels increase gradually from 0h00 to 5h00, then increases rapidly at 6h00, then stabilize until 15h00, then decreases rapidly until the day ends. The Accessibility levels experienced by the residents of Khayelitsha are good using the Weighted Average public transit as good levels of Accessibility are reached throughout the day.

The PMI levels increase gradually from 0h00 to 5h00, then stabilize at 5h00 to 15h00, then decreases rapidly until the day ends. The PMI levels experienced by the residents of Khayelitsha are good using the Weighted Average public transit as good levels of PMI are reached throughout the day.

5.1.2 Mitchells Plain

GABS

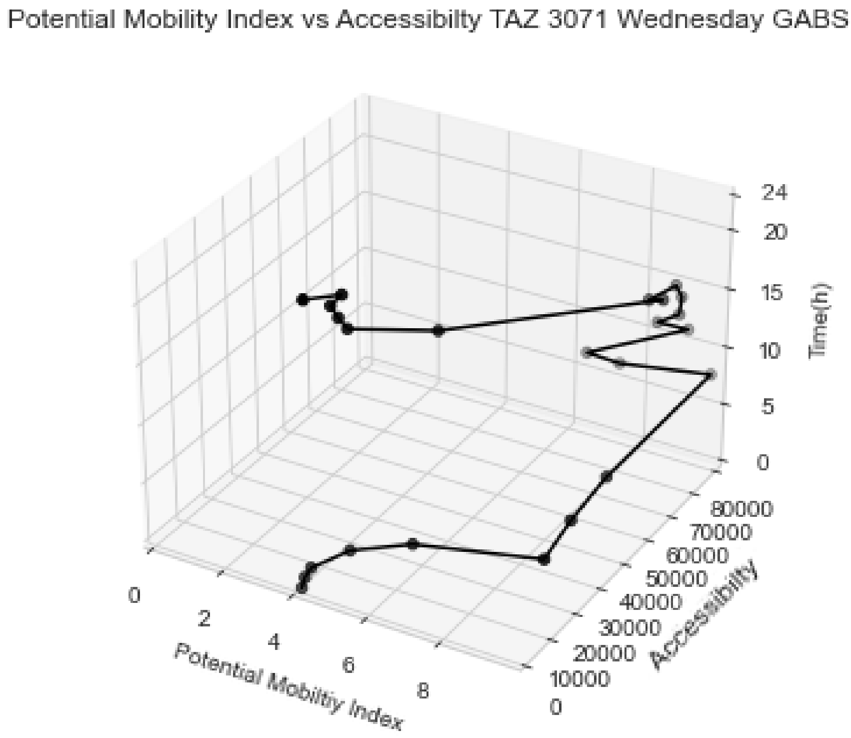


Figure 5.6: PMI VS Accessibility TAZ 3071 GABS Wednesday

The Accessibility levels increase gradually from 0h00 to 5h00, then increases rapidly at 6h00 increase gradually until 15h00, then decreases rapidly until around 20h00, and gradually decreases until the day ends. The Accessibility levels experienced by the residents of Mitchells Plain are good using the GABS public transit as good levels of Accessibility are reached throughout the day exceeding levels of Accessibility levels of over 80000[-].

The PMI levels increase gradually from 0h00 to 4h00, then increases rapidly at 4h00 to 7h00 stabilize until 15h00, then decreases rapidly until around 20h00, and gradually decreases until the day ends. The PMI levels experienced by the residents of Mitchells Plain are good using the GABS public transit as good levels of PMI are reached throughout the day.

Metro Rail

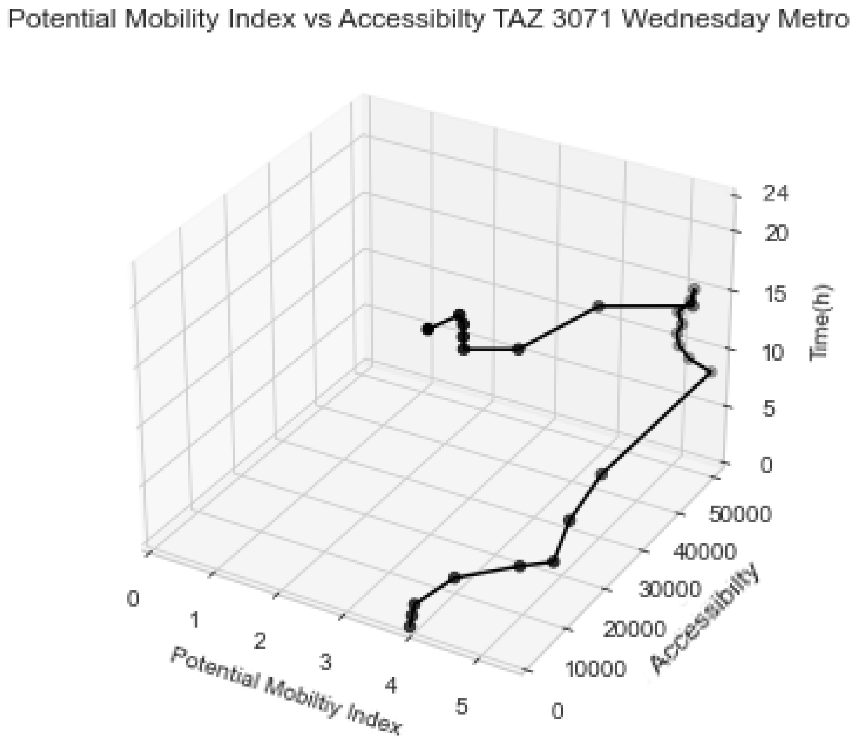


Figure 5.7: PMI VS Accessibility TAZ 3071 Metro Rail Wednesday

The Accessibility levels increase gradually from 0h00 to 5h00, then increases rapidly at 6h00 increase gradually until 15h00, then decreases rapidly until around 20h00 and gradually decreases until the day ends. The Accessibility levels experienced by the residents of Mitchells Plain are good using the Metro Rail public transit as good levels of Accessibility are reached throughout the day.

The PMI levels increase gradually from 0h00 to 4h00, then stabilize at 4h00 to 7h00 then decrease minimally until 15h00, then decreases rapidly until around 20h00, and gradually decreases until the day ends. The PMI levels experienced by the residents of Mitchells Plain are good using the Metro Rail public transit as good levels of PMI are reached throughout the day.

Taxi

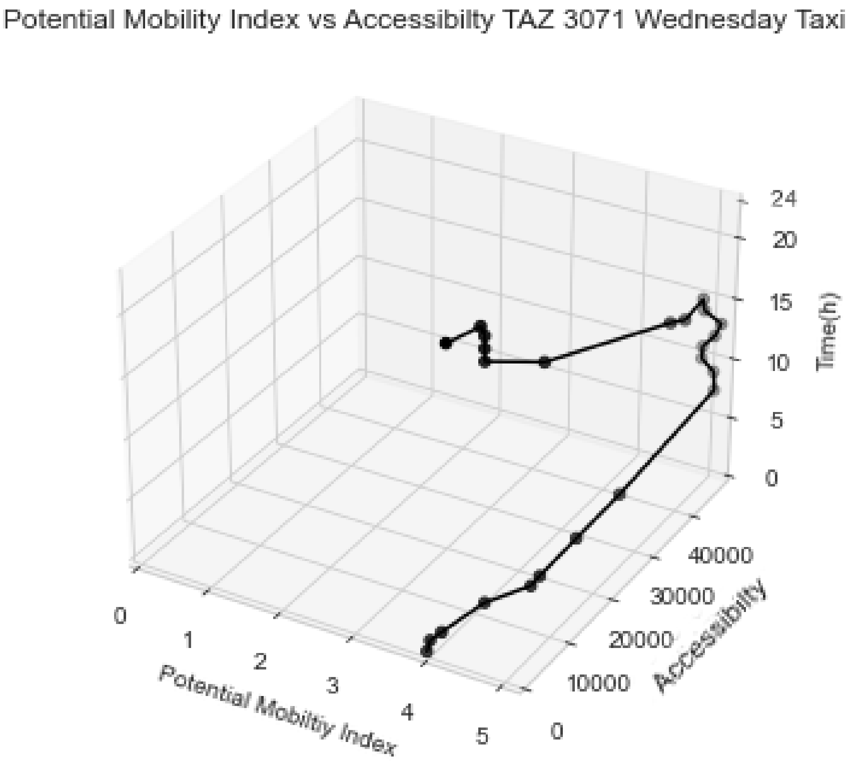


Figure 5.8: PMI VS Accessibility TAZ 3071 Taxi Wednesday

The Accessibility levels increase gradually from 0h00 to 5h00, then increases rapidly to 6h00, then stabilize until 15h00, then decreases rapidly until around 20h00, and gradually decreases until the day ends. The Accessibility levels experienced by the residents of Mitchells Plain are good using the Taxi public transit as good levels of Accessibility are reached throughout the day.

The PMI levels are very stable, staying around 4[-] throughout the day. The PMI levels experienced by the residents of Mitchells Plain are good using the Taxi public transit as good levels of PMI are reached throughout the day.

MyCiti

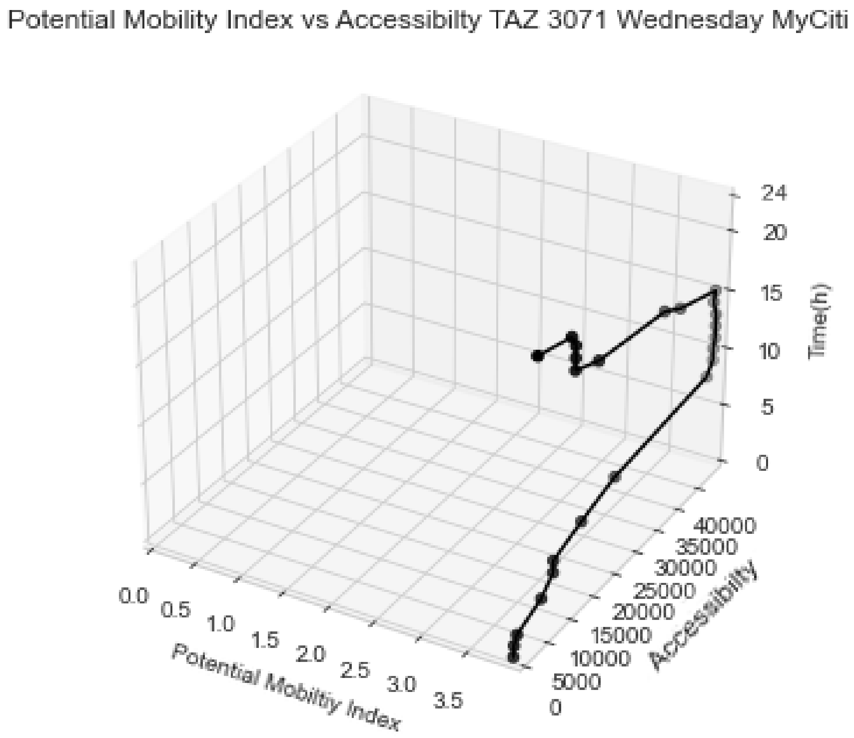


Figure 5.9: PMI VS Accessibility TAZ 3071 MyCiti Wednesday

The Accessibility levels increase gradually from 0h00 to 5h00, then increases rapidly to 6h00, then stabilize until 15h00, then decreases rapidly until around 20h00 and gradually decreases until the day ends. The Accessibility levels experienced by the residents of Mitchells Plain are good when using the MyCiti public transit as good levels of Accessibility are reached throughout the day.

The PMI levels are very stable, staying around 3.8[-] throughout the day. The PMI levels experienced by the residents of Mitchells Plain are good when using the MyCiti public transit as good levels of Accessibility are reached throughout the day.

Weighted Average

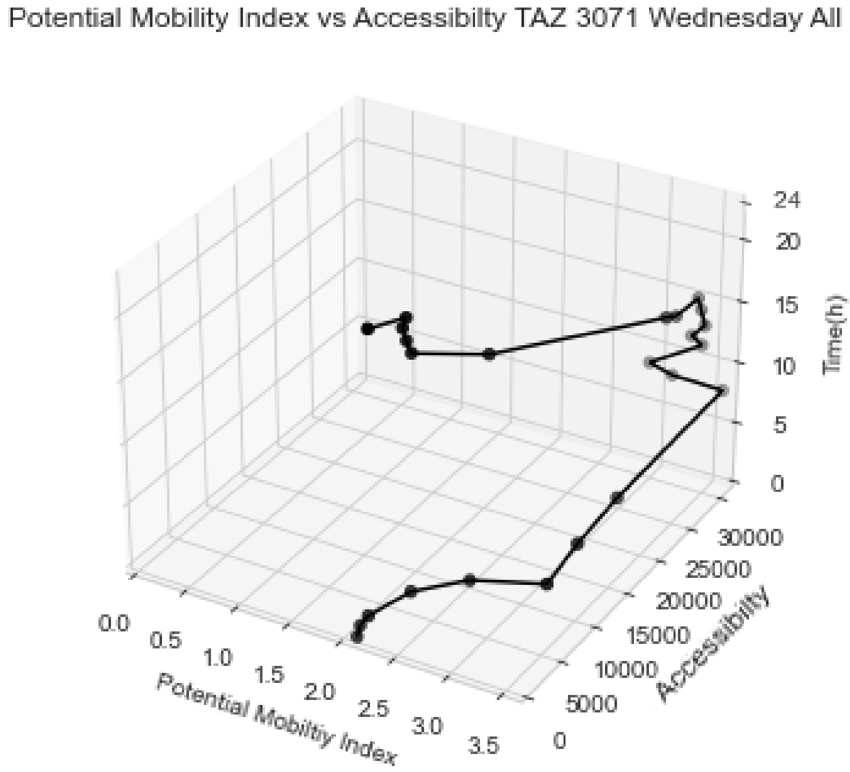


Figure 5.10: PMI VS Accessibility TAZ 3071 Weighted Average Wednesday

The Accessibility levels increase gradually from 0h00 to 5h00, then increases rapidly at 6h00 increase gradually until 15h00, then decreases rapidly until around 20h00, and gradually decreases until the day ends. The Accessibility levels experienced by the residents of Mitchells Plain are good using the Weighted Average public transit as good levels of Accessibility are reached throughout the day.

The PMI levels increase gradually from 0h00 to 4h00, then increase rapidly at 4h00 to 7h00 then stabilize until 15h00, then decreases rapidly until around 20h00, and gradually decreases until the day ends. The PMI levels experienced by the residents of Mitchells Plain are low using the Weighted Average public transit as values between 2[-] and 3.2[-] levels of PMI are reached throughout the day.

5.1.3 Bishops Court

GABS

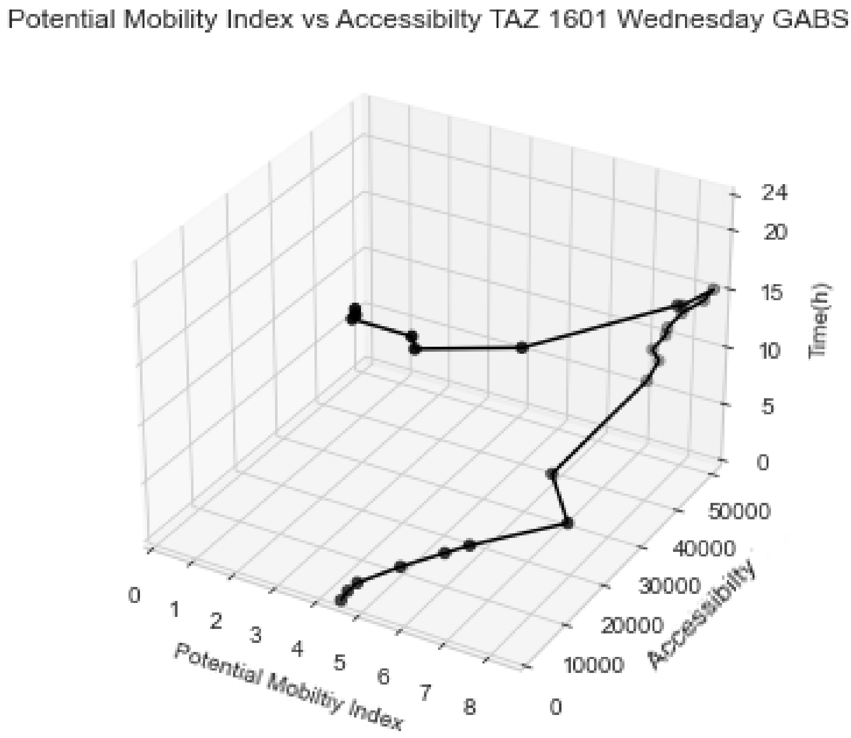


Figure 5.11: PMI VS Accessibility TAZ 1601 GABS Wednesday

The Accessibility levels increase slowly from 0h00 to 5h00, then increases rapidly at 5h00 to 15h00, then increase gradually until 15h00, then decreases rapidly until around 20h00 and gradually decreases until the day ends. The Accessibility levels experienced by the residents of Bishops Court are good using the GABS public transit as good levels of Accessibility are reached throughout the day.

The PMI levels increase gradually from 0h00 to 7h00, then stabilize until 15h00, then decreases rapidly until the day ends. The PMI levels experienced by the residents of Bishops Court are good using the PMI public transit as good levels of Accessibility are reached throughout the day.

Metro Rail

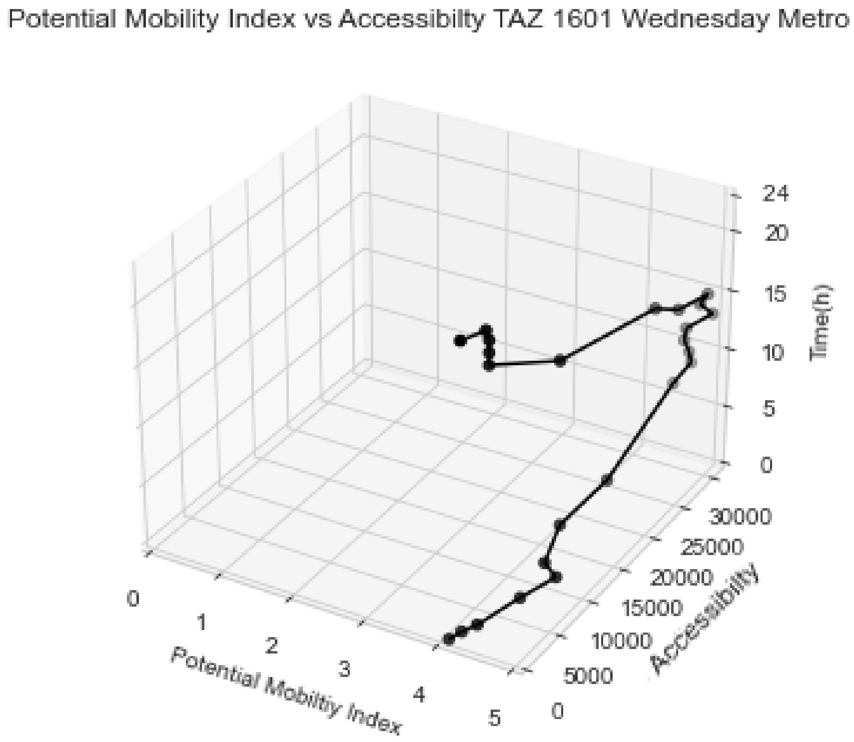


Figure 5.12: PMI VS Accessibility TAZ 1601 Metro Rail Wednesday

The Accessibility levels increase gradually from 0h00 to 5h00, then increases rapidly at 6h00 increase gradually until 15h00, then decreases rapidly until around 20h00 and gradually decreases until the day ends. The Accessibility levels experienced by the residence of Bishops Court are low using the Metro Rail public transit as low levels of Accessibility are reached throughout the day.

The PMI levels is stable, around 4[-] throughout the day. The PMI levels experienced by the residence of Bishops Court are good using the Metro Rail public transit as good levels of PMI are reached throughout the day.

Taxi

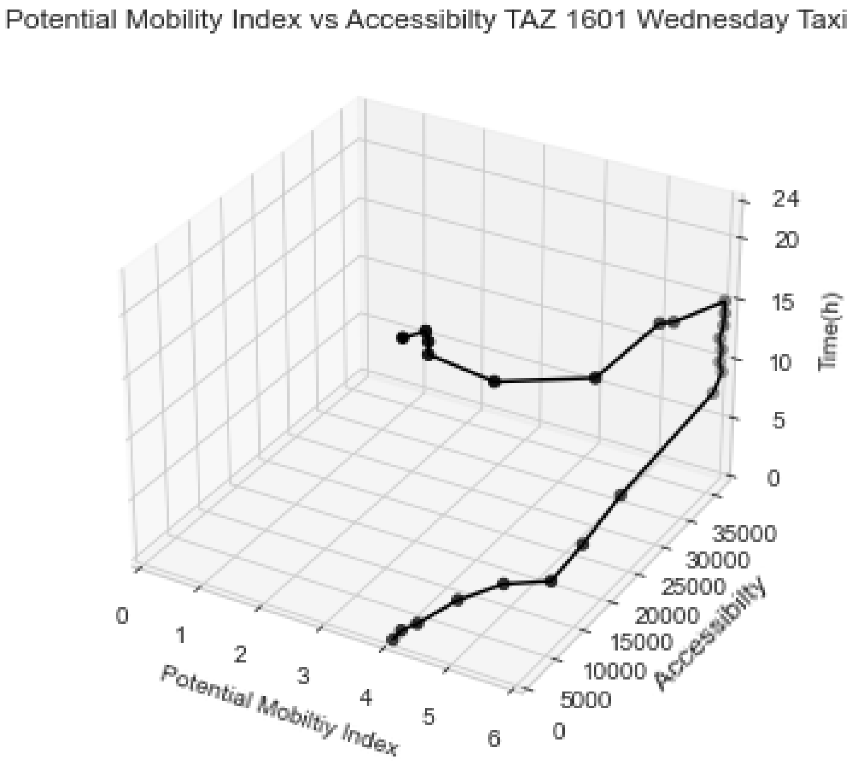


Figure 5.13: PMI VS Accessibility TAZ 1601 Taxi Wednesday

The Accessibility levels increase gradually from 0h00 to 5h00, then increases rapidly at 5h00 to 7h00, then stabilize until 15h00, then decreases rapidly until around 20h00 and gradually decreases until the day ends. The Accessibility levels experienced by the residents of Bishops Court are good using the Taxi public transit as good levels of Accessibility are reached throughout the day.

The PMI levels are stable, around 4[-] throughout the day. The PMI levels experienced by the residents of Bishops Court are good using the Taxi public transit as good levels of Accessibility are reached throughout the day.

MyCiti

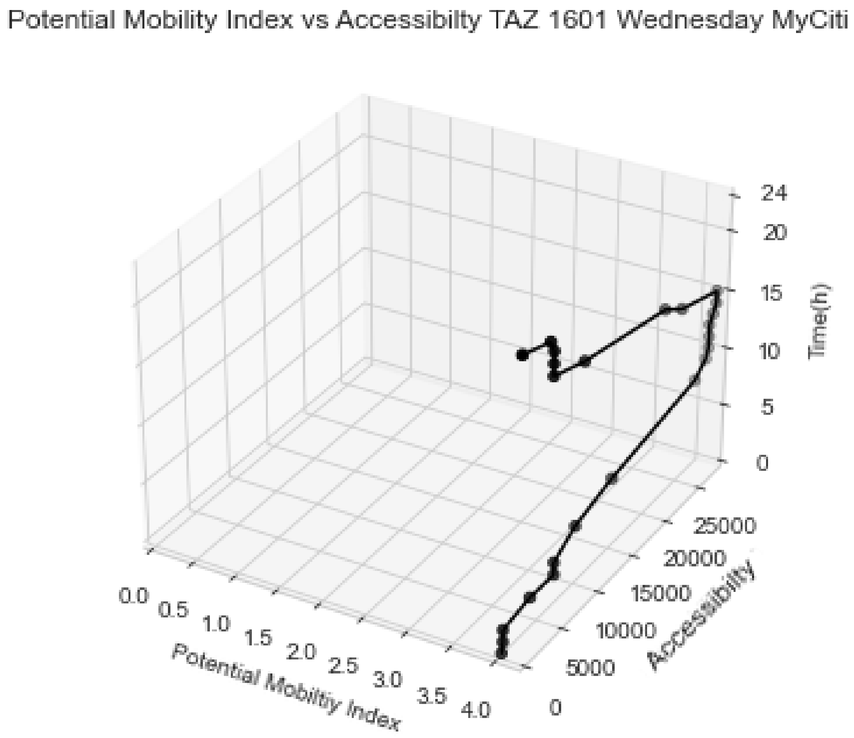


Figure 5.14: PMI VS Accessibility TAZ 1601 MyCiti Wednesday

The Accessibility levels increase gradually from 0h00 to 5h00, then stabilize at 6h00 to 15h00, then decreases rapidly until around 20h00 and gradually decreases until the day ends. The Accessibility levels experienced by the residents of Bishops Court are good using the MyCiti public transit as good levels of Accessibility are reached throughout the day.

The PMI levels are stable, around 4[-] throughout the day. The PMI levels experienced by the residents of Bishops Court are low using the MyCiti public transit as low levels of Accessibility are reached throughout the day.

Weighted Average

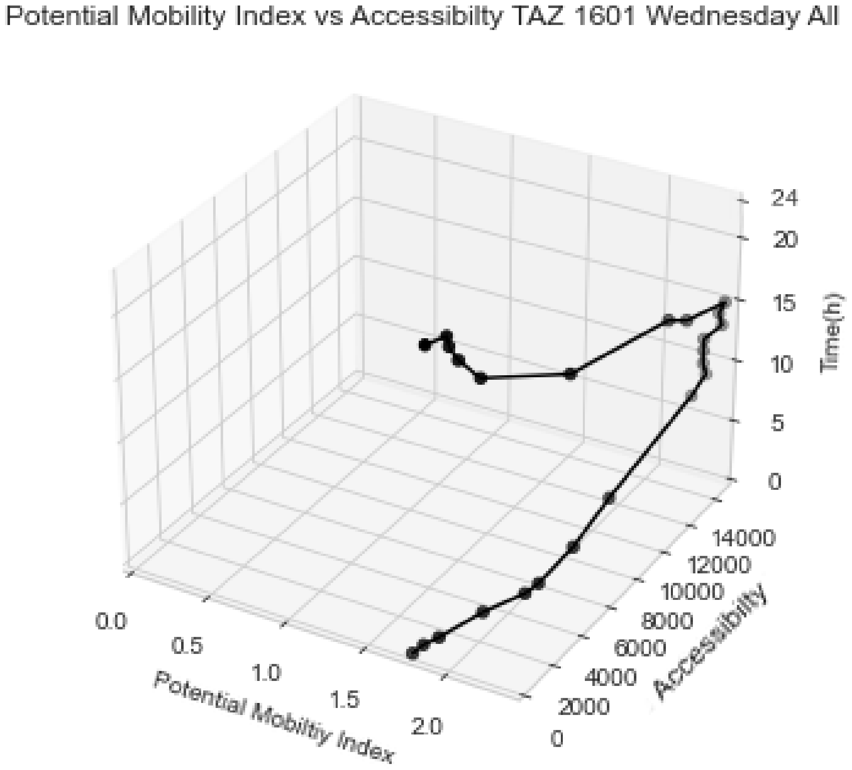


Figure 5.15: PMI VS Accessibility TAZ 1601 Weighted Average Wednesday

The Accessibility levels increase gradually from 0h00 to 5h00, then increases rapidly at 6h00 increase gradually until 15h00, then decreases rapidly until around 20h00, and gradually decreases until the day ends. The Accessibility levels experienced by the residents of Bishops Court are good using the Weighted Average public transit as good levels of Accessibility are reached throughout the day.

The PMI levels increase gradually from 0h00 to 4h00, then increases rapidly at 4h00 to 7h00 then increase gradually until 15h00, then decreases rapidly until around 20h00, and gradually decreases until the day ends. The PMI levels experienced by the residents of Bishops Court are poor using the Weighted Average public transit as poor levels of PMI of between 1.6[-] to 2.3[-] are reached throughout the day.

5.1.4 Comparison of Case Studies

This section will highlight the differences in accessibility and potential mobility between all case studies using various transport modes throughout the day. Finally, the weighted average differences in accessibility and potential mobility between all case studies using various transport modes throughout the day will be highlighted as well.

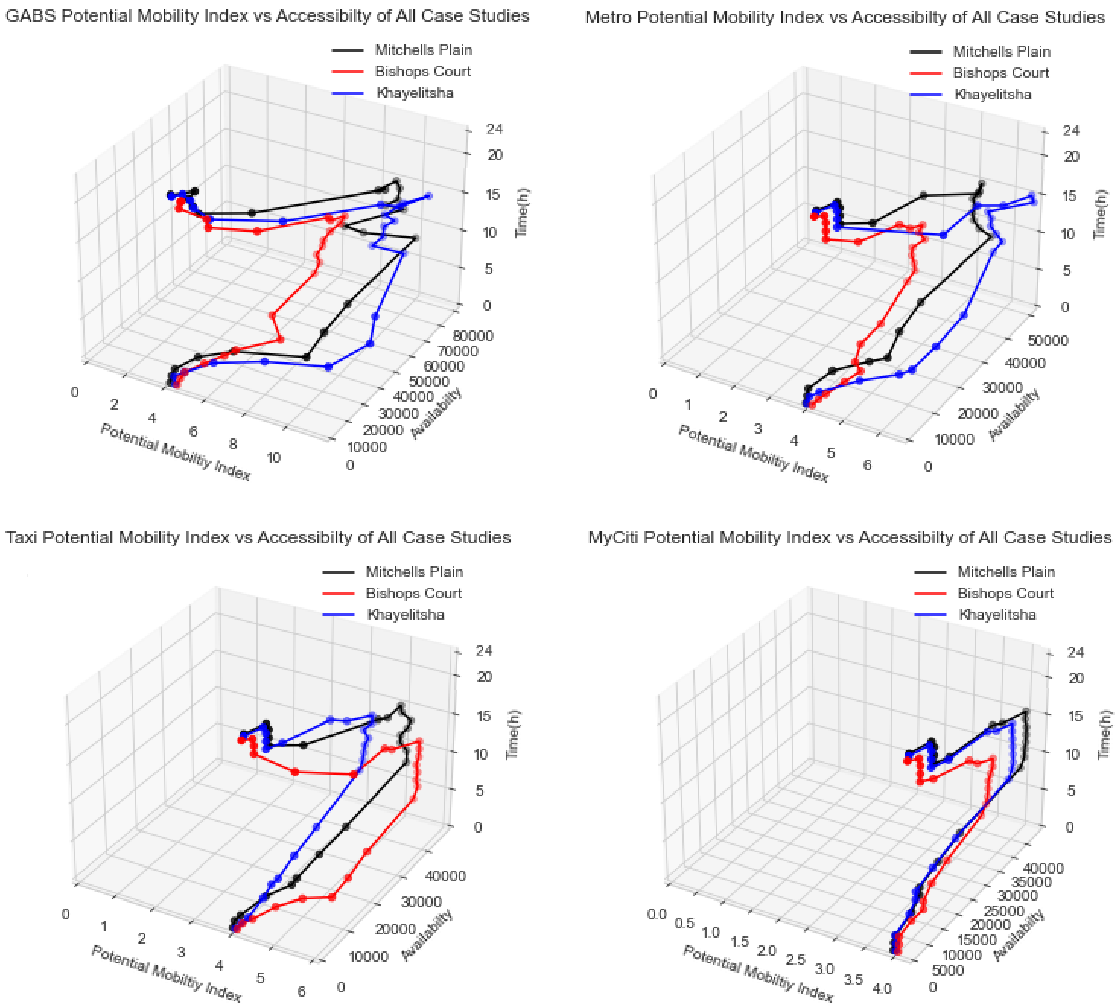


Figure 5.16: PMI VS Accessibility Comparison of Case Studies All Transport Modes Wednesday

GABS

Khayelitsha appears to have a better PMI than the other case studies. Bishops Court appears to have the worst PMI compared to the other case studies. Mitchells Plain appears to have better accessibility levels than the other case studies. Bishops Court appears to have the worst accessibility levels compared to the other case studies.

Metro Rail

Khayelitsha appears to have a better PMI than the other case studies. Bishops Court appears to have the worst PMI compared to the other case studies. Mitchells Plain appears to have better accessibility levels than the other case studies. Bishops Court appears to have the worst accessibility levels compared to the other case studies.

Taxi

Bishops Court appears to have a better PMI than the other case studies. Khayelitsha appears to have the worst PMI compared to the other case studies. Mitchells Plain appears to have better accessibility levels than the other case studies. Bishops Court appears to have the worst accessibility levels compared to the other case studies.

MyCiti

All case studies appears to have similar PMI levels. Mitchells Plain appears to have better accessibility levels than the other case studies. Bishops Court appears to have the worst accessibility levels compared to the other case studies.

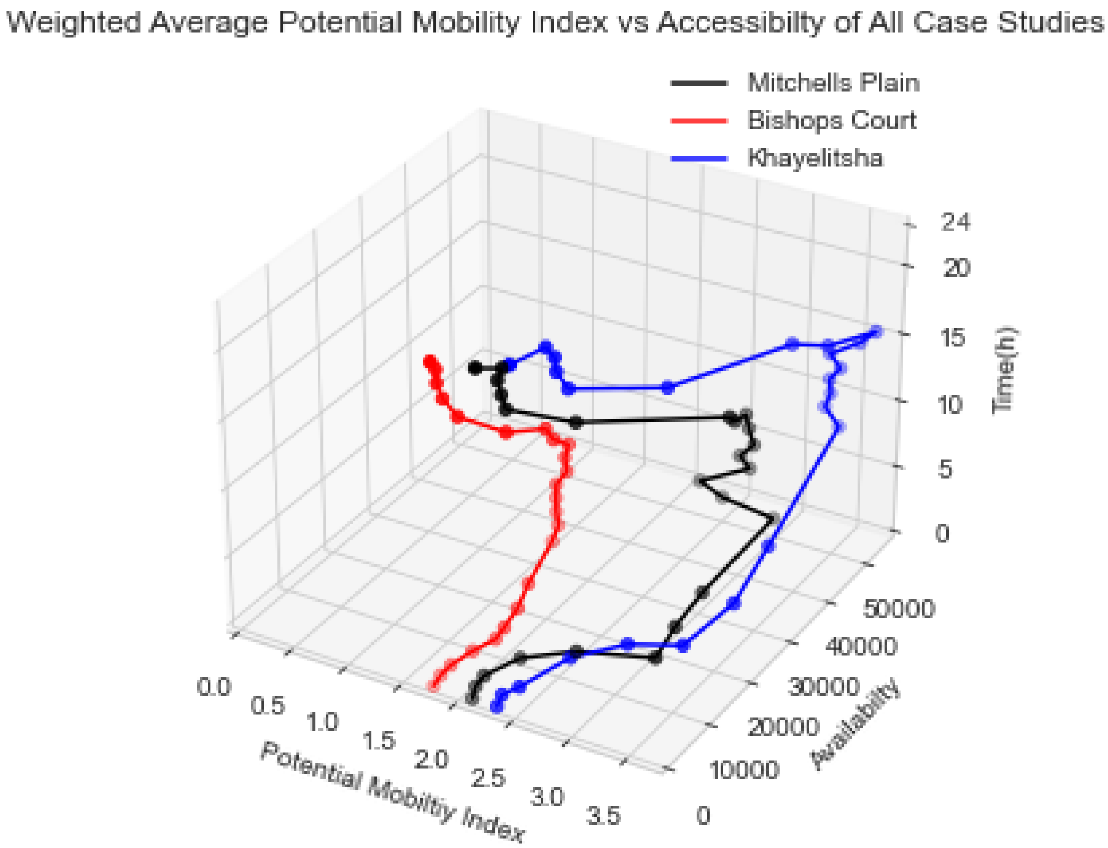


Figure 5.17: PMI VS Accessibility TAZ 1601 Weighted Average Wednesday

Weighted Average

Khayelitsha appears to have a better PMI than the other case studies. Bishops Court appears to have the worst PMI compared to the other case studies. Khayelitsha appears to have better accessibility levels than the other case studies. Bishops Court appears to have the worst accessibility levels compared to the other case studies

The results of Figure 5.17 for the different case studies are due to the UDI Modal Split of these locations. This can be seen in Table 4.1. Although certain locations can achieve better PMI and accessibility levels than others using the various public transport modes available to them, their actual PMI and accessibility levels will be bounded by which public transport mode they prefer to use the most.

5.2 Temporal Aggregation

The Temporal Aggregation tool was used to display the Accessibility and PMI of all 1616 TAZ centroids for the City of Cape Town at a desired hour instance. The average PMI score value and the average Accessibility value experienced by the Temporal Aggregation will become the origin of the coordinate system to identify the TAZ that falls below the Accessibility sufficiency threshold. The regions that have an Accessibility and PMI deficit are the regions that fall below the average Accessibility line and the average PMI line.

The chosen times for analysis were 06h00 (Morning Peak), 10h00 (Off Peak), and 16h00 (Evening Peak) on a Wednesday. These times were chosen to analyse the Accessibility and PMI levels experienced by the City of Cape Town when transport is most in demand and the latter. The UDI modal split weights of the various public transport modes of the TAZ were used to produce the Temporal Aggregation graph.

Morning Peak (06h00)

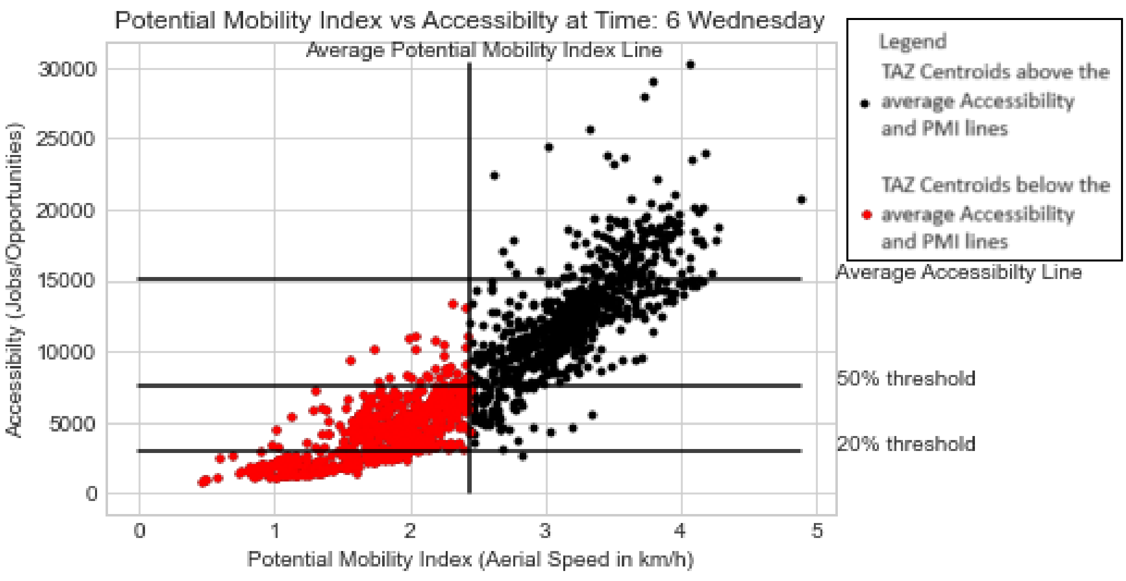


Figure 5.18: Cumulative TAZ PMI vs Accessibility 6h00 Wednesday

The black points represent the TAZ that falls below the average PMI line and the average Accessibility line. The red points are the TAZ that falls below the average PMI line and the average Accessibility line.

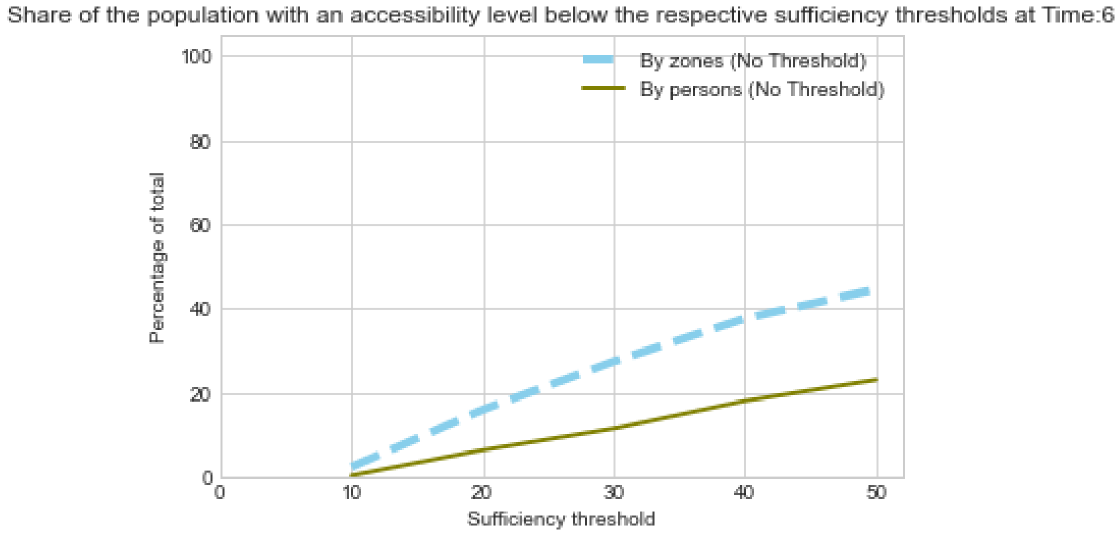


Figure 5.19: Population and TAZ up to 50% Threshold 6h00 Wednesday

Table 5.1: Population and TAZ below 50% Threshold 6h00 Wednesday figures

TAZ	**10%**	**20%**	**30%**	**40%**	**50%**
Affected-Zones-No. (No Threshold)	40	259	444	609	721
Affected-Population (No Threshold)(%)	0.44	6.48	11.58	18.13	23.12

The share of the affected population gradually decreases from 50 to 20 percent sufficiency then decreases much faster from 20 to 10 percent threshold. The share of the affected zones follows a similar pattern as the pattern of the affected population.

Off Peak (10h00)

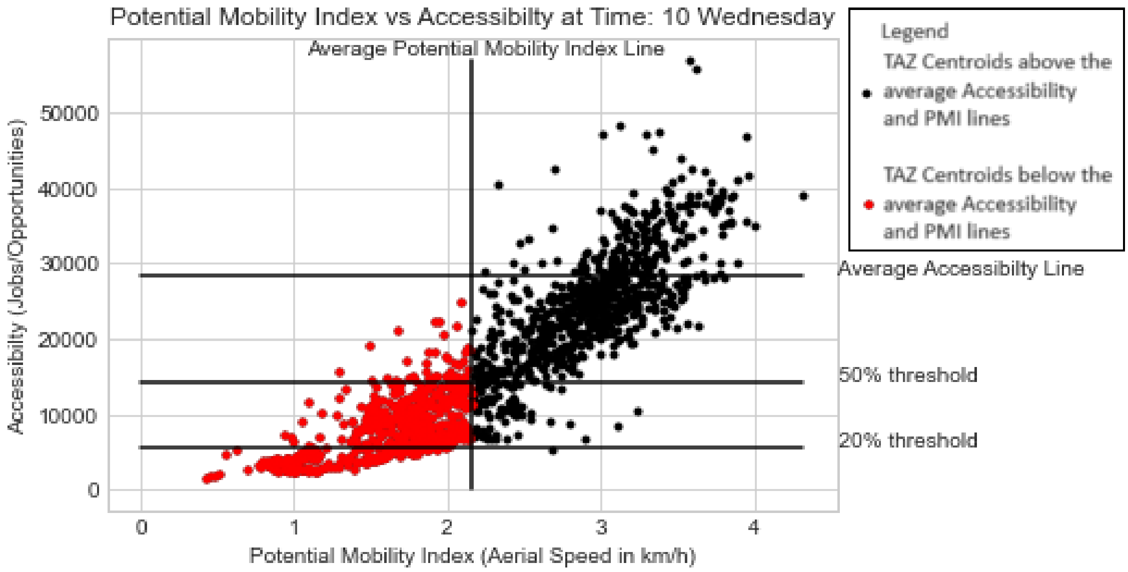


Figure 5.20: Cumulative TAZ PMI vs Accessibility 10h00 Wednesday

The black points represent the TAZ that falls below the average PMI line and the average Accessibility line. The red points are the TAZ that falls below the average PMI line and the average Accessibility line.

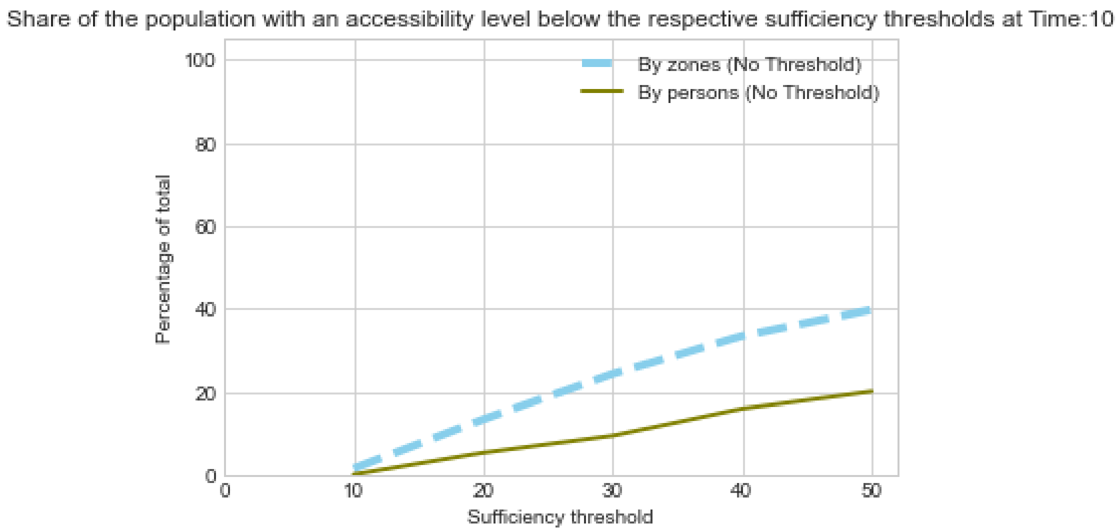


Figure 5.21: Population and TAZ up to 50% Threshold 10h00 Wednesday

Table 5.2: Population and TAZ below 50% Threshold 10h00 Wednesday figures

TAZ	10%	20%	30%	40%	50%
Affected-Zones-No. (No Threshold)	28	219	395	543	644
Affected-Population (No Threshold)(%)	0.31	5.53	9.61	16.10	20.23

The share of the affected population gradually decreases from 50 to 20 percent sufficiency and then decreases much faster from the 20 to 10 percent threshold. The share of the affected zones follows a similar pattern as the pattern of the affected population.

Evening Peak (16h00)

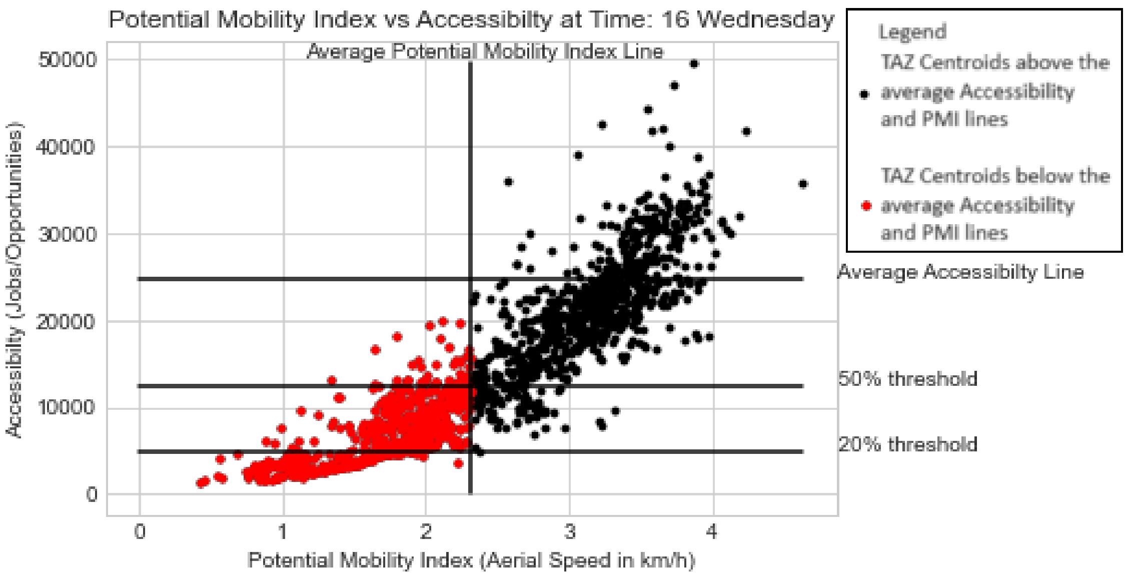


Figure 5.22: Cumulative TAZ PMI vs Accessibility 16h00 Wednesday

The black points represent the TAZ that falls below the average PMI line and the average Accessibility line. The red points are the TAZ that fall below the average PMI line and the average Accessibility line.

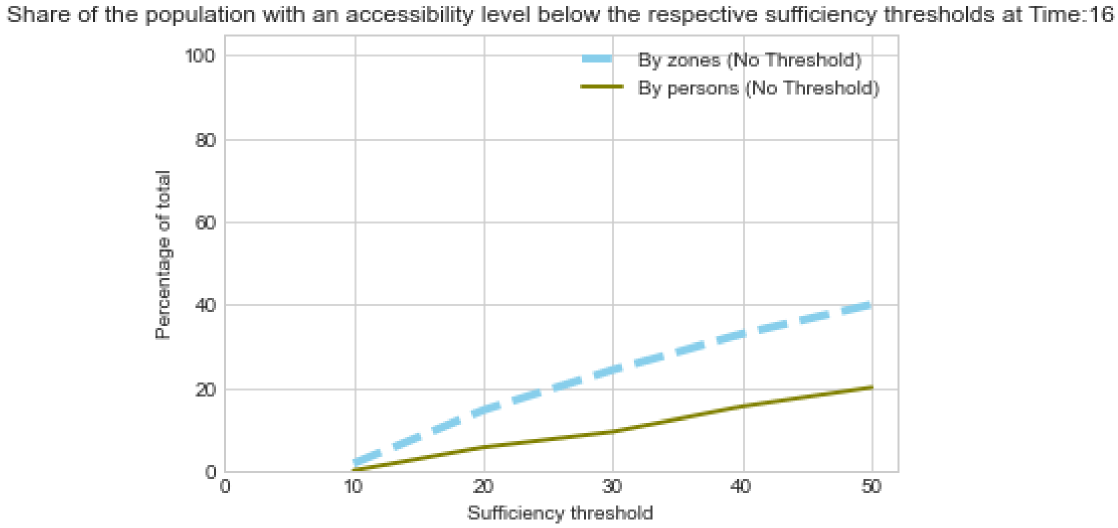


Figure 5.23: Population and TAZ up to 50% Threshold 16h00 Wednesday

Table 5.3: Population and TAZ below 50% Threshold 16h00 Wednesday figures

TAZ	**10%**	**20%**	**30%**	**40%**	**50%**
Affected-Zones-No. (No Threshold)	32	238	395	534	648
Affected-Population (No Threshold)(%)	0.28	5.87	9.60	15.63	20.22

The share of the affected population gradually decreases from 50 to 20 percent sufficiency then decreases much faster from 20 to 10 percent threshold. The share of the affected zones follows a similar pattern as the pattern of the affected population.

5.2.1 Spatial Location of Affected Groups

The TAZ's that fall below the 50 percent threshold can be displayed to identify the distribution of which regions experience this PMI and Accessibility deficit. These locations then can be compared to others with other factors such as demographic and income data to decide if these locations are in a demand of an Accessibility and PMI increase.

Morning Peak (06h00)

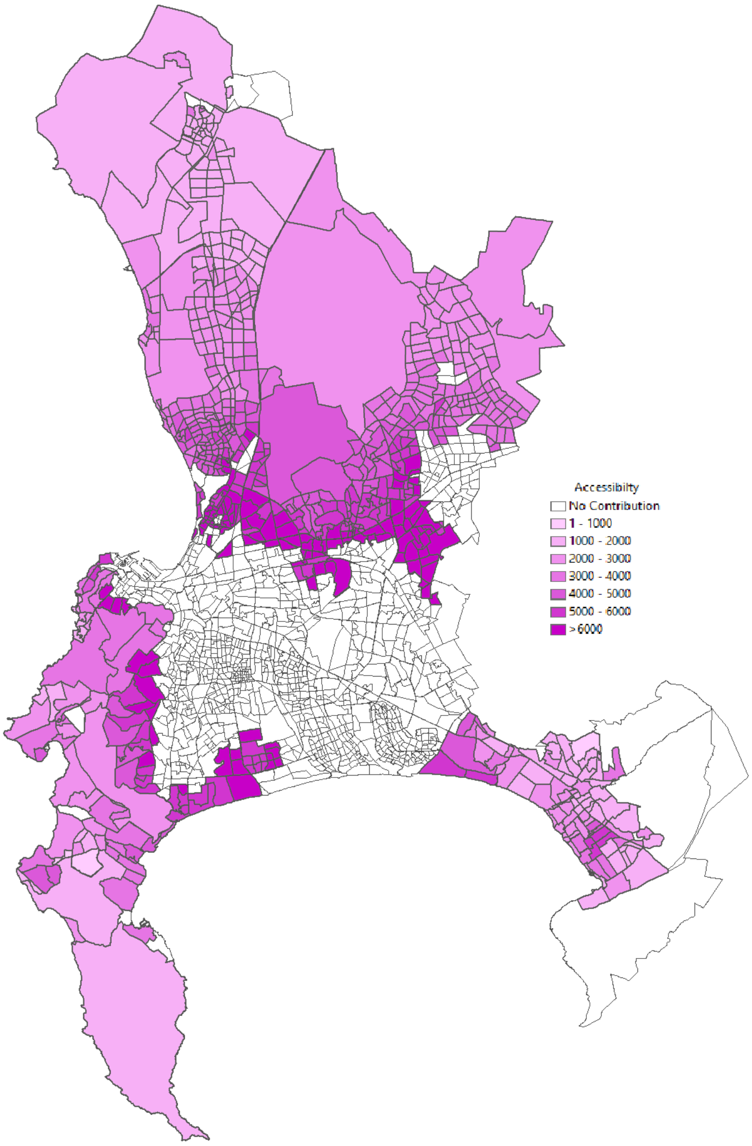


Figure 5.24: TAZ Distribution below 50% Threshold 6h00 Wednesday

Off Peak (10h00)

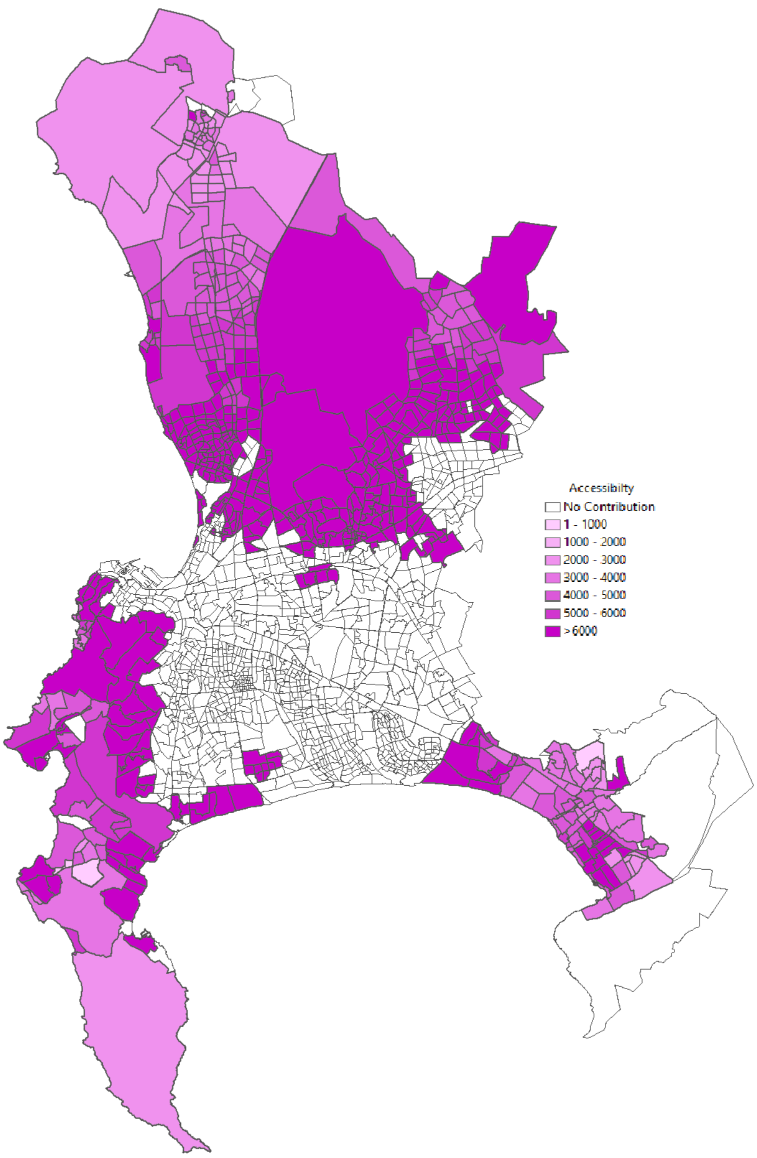


Figure 5.25: TAZ Distribution below 50% Threshold 10h00 Wednesday

Evening Peak (16h00)

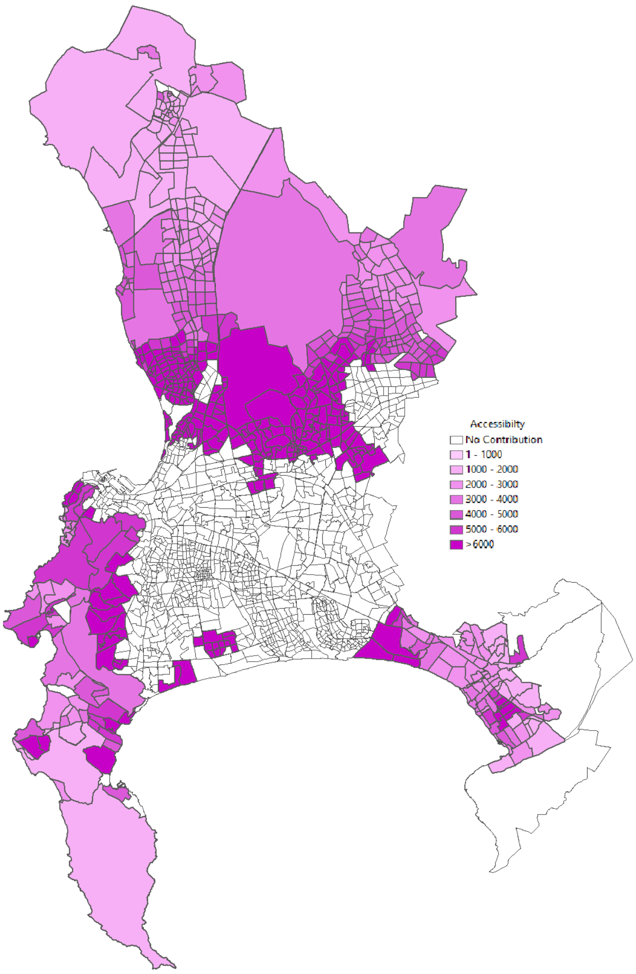


Figure 5.26: TAZ Distribution below 50% Threshold 16h00 Wednesday

5.2.2 Demographic and Income Threshold

The Temporal Aggregation tool was used to refine the TAZ with an Accessibility and PMI deficit by taking into account factors such as demographic and income data. This section showed how this was used on previously disadvantaged groups i.e. Coloured and Black-African demographic groups. A simple constraint of a minimum of 30 percent of the selected demographic group and a minimum of 500 people that fall in the category of Low income for the TAZ that falls below the sufficiency threshold.

Morning Peak (06h00)

Black African Population

Figure 5.27: TAZ Distribution below 50% Threshold Black-African Population 6h00 Wednesday

Table 5.4: TAZ Distribution below 50% Threshold Black-African Population 6h00 Wednesday Figures

TAZ	**10%**	**20%**	**30%**	**40%**	**50%**
Affected No.	0	9	14	17	19

Coloured Population

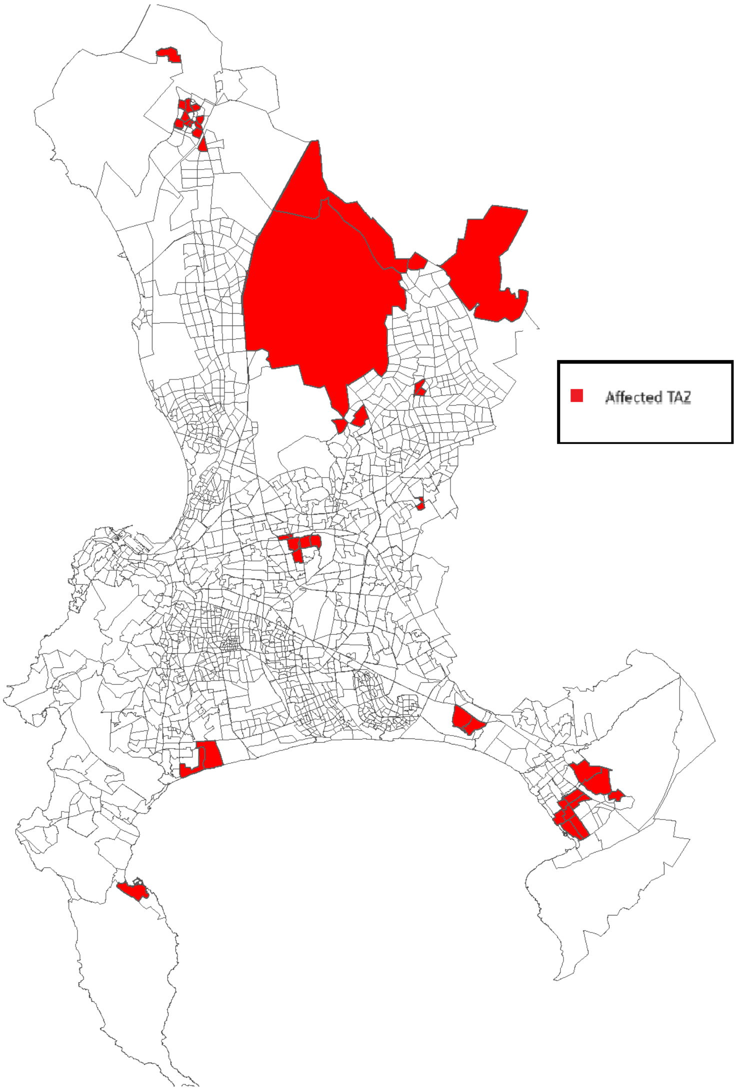


Figure 5.28: TAZ Distribution below 50% Threshold Coloured Population 6h00 Wednesday

Table 5.5: TAZ Distribution below 50% Threshold Coloured Population 6h00 Wednesday Figures

TAZ	**10%**	**20%**	**30%**	**40%**	**50%**
Affected No.	31	40	46	51	52

Off Peak (10h00)

Black African Population

Figure 5.29: TAZ Distribution below 50% Threshold Black-African Population 10h00 Wednesday

Table 5.6: TAZ Distribution below 50% Threshold Black-African Population 10h00 Wednesday Figures

TAZ	**10%**	**20%**	**30%**	**40%**	**50%**
Affected No.	0	6	12	17	18

Coloured Population

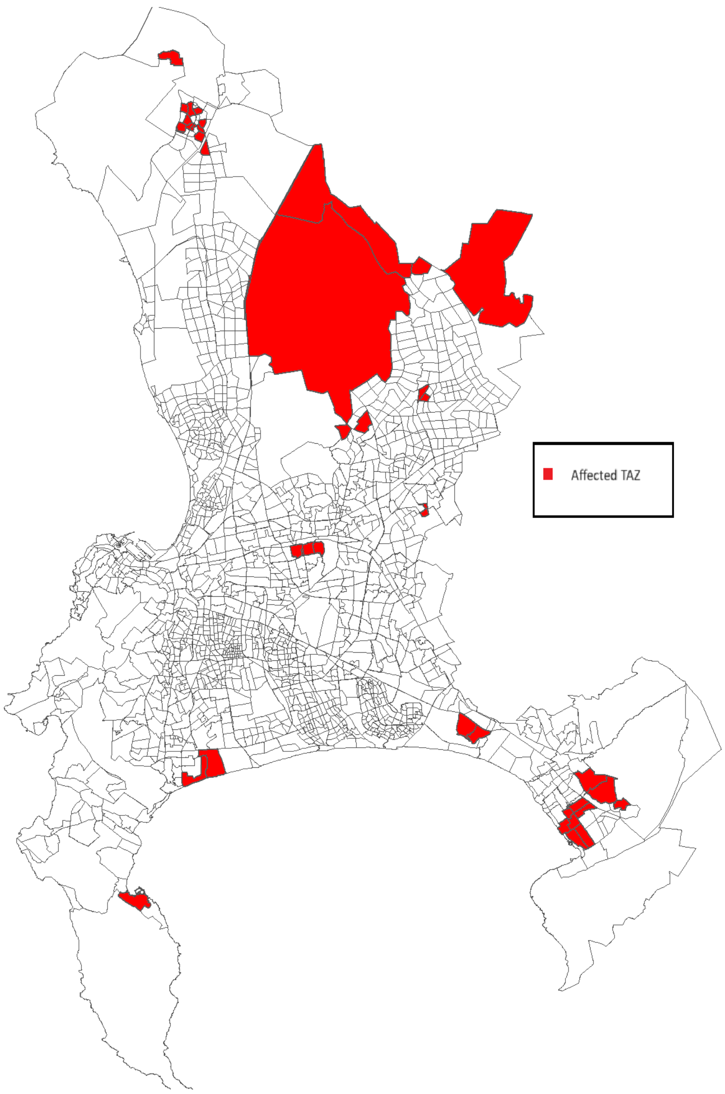


Figure 5.30: TAZ Distribution below 50% Threshold Coloured Population 10h00 Wednesday

Table 5.7: TAZ Distribution below 50% Threshold Coloured Population 10h00 Wednesday Figures

TAZ	**10%**	**20%**	**30%**	**40%**	**50%**
Affected No.	0	16	23	32	35

Evening Peak (16h00)

Black African Population

Figure 5.31: TAZ Distribution below 50% Threshold Black-African Population 16h00 Wednesday

Table 5.8: TAZ Distribution below 50% Threshold Black-African Population 16h00 Wednesday Figures

TAZ	**10%**	**20%**	**30%**	**40%**	**50%**
Affected No.	0	7	12	17	17

Coloured Population

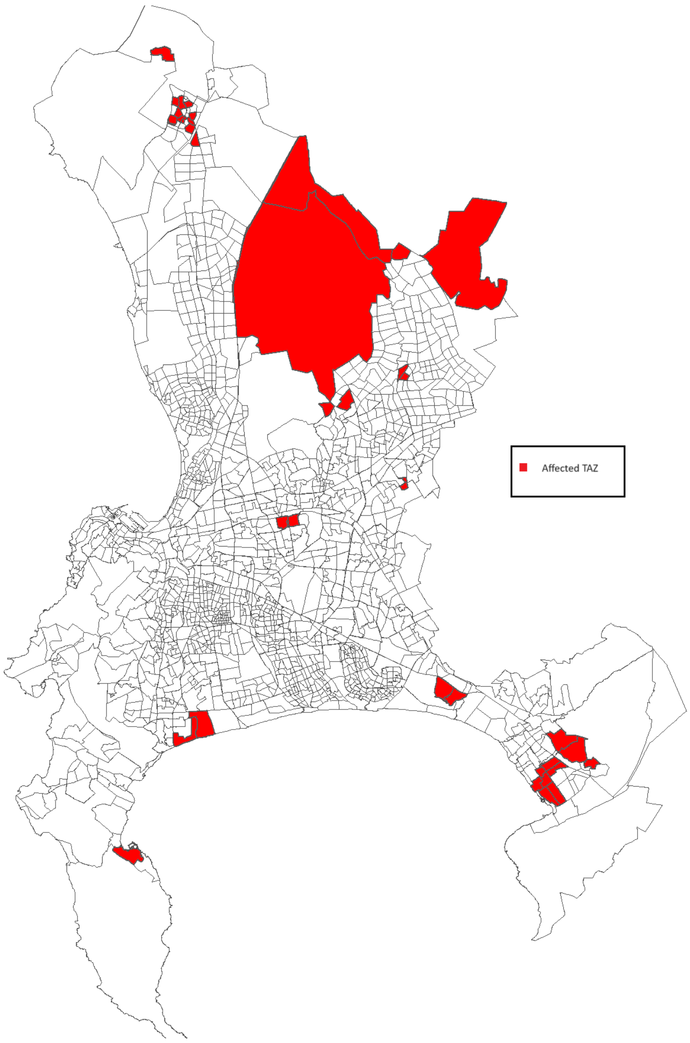


Figure 5.32: TAZ Distribution below 50% Threshold Coloured Population 16h00 Wednesday

Table 5.9: TAZ Distribution below 50% Threshold Coloured Population 16h00 Wednesday Figures

TAZ	**10%**	**20%**	**30%**	**40%**	**50%**
Affected No.	0	16	23	28	32

Chapter 6

Analysis

6.1 Introduction

This chapter discusses the results obtained using the method chosen for this project. The main focus was implementing the two transport justice tools, Space-Time Cube and Temporal Aggregation, that were created for this project and the results it produced. In addition, it highlights the possible factors that affected the PMI and Accessibility results obtained.

6.2 Data Used

With regard to the income data and demographic data, certain TAZ regions had null data related to these locations. This could actually mean that no people are living in these TAZ and there were no opportunities in those regions or the data was not captured for these regions. The null data was replaced with the value of 0 for these TAZ regions to participate in the potential accessibility and PMI calculations. This null data directly affected the values of potential accessibility experienced by the other TAZ that actually had real values captured. Since the null data was recorded as 0, it meant that travelling from any TAZ to a TAZ region with null data would result in a potential accessibility value of 0. However, this null data does not affect the PMI levels as PMI is purely calculated from the impedance attribute values of shape length and travel time between an origin-destination pair in the network analyst layer. The operational hours of the opportunities/jobs were incorporated in the potential accessibility measure as a constraint function. This constraint helped to display how accessibility levels change throughout the day based on the time of day. This was done to provide a closer potential accessibility level determined as it was experienced in a real-life scenario.

The income data indicated that a low amount of the population fell within the high-income group for the region of interest. The high-income population groups' geo-spatial distribution was highly correlated to regions with a higher accessibility level. The high-income population groups most likely resided in these locations as they were surrounded by more jobs/opportunities. Most of the population fell within the low and low-middle-income groups. These geo-spatial regions tend to coincide with the regions where the black-African and coloured population groups reside. This was a negative effect that was largely due to the Groups Area Act of 1950 and mainly geo-spatially disadvantaged the non-white demographic groups.

The UDI modal split geo-spatial distribution largely coincided geo-spatial with the locations of the respective public transit stops. The locations of the transit stops of the public transit services indicated that the specific public transit services were more accessible than others within the TAZ regions. This leads the residents of these TAZ regions to make more use of these public transit services than the ones that are less available to them. Therefore, if the public transit services increased the spread of the geo-spatial distribution of the transit stops and were located in more TAZ regions, the more likely the residents of the TAZ regions would make use of these public transport modes. This would ultimately result in higher accessibility and PMI levels for a TAZ.

6.3 Network Analyst Layer

The network analyst layer was created to produce the impedance attribute of travel time from an origin TAZ to any other TAZ of interest. This was done to determine accessibility and PMI levels an origin TAZ experiences at the hours of the day. With the input of the 1616 TAZ origins and the destinations being the same as the 1616 TAZ origins points. This produces a combination of 2 611 456 lines of data with each line showing an origin-destination pairs attributes.

These attributes included the travel time using the public transit, the total distance travelled by the public transit between origin-destination pairs and the Euclidean distance between origin-destination pairs. All of these attributes were derived from georeferenced data, which helps the measurements to be held to scale. The travel time attribute data would include a person walking to the closest public transit stop, then waiting until the public transit arrives, making use of a public transit mode, and finally getting off at the closest transit stop and walking to the destination.

So the closer the transit stops are to the origin-destination pairs, it will result in the shorter the time that the user needs to spend walking. This will increase the

PMI and accessibility levels experienced. In addition, the less time a person waits for a public transit mode to arrive the faster the user can arrive at their destination, which will also increase their accessibility and PMI values. The network analyst layer ran for 6-hour instances, which took approximately 7 hours to complete in ArcGIS. The accessibility and PMI values were calculated using the Space-Time Cube and Temporal Aggregation transport justice tools. This was done by making use of the parameters of travel time created by the OD cost matrix, the Euclidean distance between origin-destination pairs, the number of opportunities/jobs in the TAZ for the region of interest, the operational times of the opportunities/jobs in these TAZ and the UDI modal split.

6.4 Space-Time Cube

This process took approximately 18 hours to run to produce the result of accessibility and PMI for an entire 24-hour day of a particular TAZ origin. Although the days of Tuesday, Wednesday, and Thursday were processed only Wednesday was chosen to be further analysed. This was done because the Accessibility and PMI results were identical to the eighth decimal place at each hour of the day for all the transit modes of the days analysed. The possible reason why these results identical were was because the GTFS data were static/scheduled, since these days were weekdays, the the transit modes tend to operate at the same time every other weekday. This could explain the identical Accessibility and PMI results obtained for these days. Depending on how well the particular transit mode performs in the PMI and Accessibility levels reflected on the weighted average graph in the Space-Time Cube. A transit mode's Accessibility levels depended on the distance more job/opportunities are to the TAZ origin-destination pairs in question. The closer a TAZ origin is to another TAZ destination with a high opportunity value, the higher the potential accessibility values will be for the origin-destination pairs. A transit mode PMI values depend on the travel time which is affected by the distribution of the transit stop. People had to walk to the closest transit stop get off the closest transit to the destination and walk until the destination was reached. If transit stops are closer to the origin-destination pairs then it will decrease walking time which will increase the PMI value, which will increase the PMI score of the PMI measure. Therefore, the higher and more even the distribution of the transit stops is, the higher the PMI score will be. This increased the accessibility levels experience. Accessibility values varied as travel time changed. The shorter the travel time between origin-destination pairs, the higher the potential accessibility values will be, and vice versa. However, the population within a TAZ makes use of a public transit mode due to certain factors that make it more desirable for them. The UDI modal split displays the various percentages of a population that makes use of a particular public transport mode within a TAZ. This UDI modal split contribution for the TAZ combined showed the true value of Accessibility and PMI experienced by a TAZ using a public transit mode. So if a public transit mode had a high Accessibility and PMI level within a TAZ in addition the more the population of a TAZ makes use of a particular public transit mode, the higher the overall Accessibility and PMI values of the TAZ was and vice versa.

6.4.1 Khayelitsha

Khayelitsha has a good weighted average cube of Accessibility levels reaching above 50000[-] and PMI levels between 2.3[-] to 3.5[-]. This is mainly due to the stable graph of the Taxis performance, GABS high PMI and Accessibility values, and contribution from the UDI modal split. It has the highest weighted average cube of the TAZ analysed.

6.4.2 Mitchells Plain

Mitchells Plain has a good weighted average cube of Accessibility levels reaching above 30000[-] and PMI levels between 2[-] to 3.2[-]. This is mainly due to the stable graph of the Taxis performance, GABS high PMI and Accessibility values, and contribution from the UDI modal split. It has the second-highest weighted average cube of the TAZ analysed.

6.4.3 Bishops Court

Bishops Court has a good weighted average cube of Accessibility levels reaching above 14000[-] and PMI levels between 1.6[-] to 2.3[-]. This was mainly due to the stable graph of the Taxis performance, Metro Rail stable PMI, and low Accessibility values and contribution from the UDI modal split. It has the lowest weighted average cube of the TAZ analysed.

6.5 Temporal Aggregation

This process takes about 60 days to run one instant, on a normal computer, to produce the result of accessibility and PMI of an hour of day for all 1616 TAZ origins. It would be recommended to make use of a supercomputer to reduce the processing time to produce this result. The same factors that affected the Space-Time Cube affected the Temporal Aggregation as well. However, the Temporal Aggregation just showed the weighted average of PMI and Accessibility for all the 1616 TAZ at an hour instance.

6.5.1 Temporal Aggregation (PMI vs Accessibility Graphs)

The Temporal Aggregation graphs showed the scatter plot of all 1616 TAZ Accessibility and PMI. The scatter plot has a straight line distribution which implied that the higher the PMI the higher the Accessibility levels was and vice versa.

06h00

The level of Accessibility was low because of the time of the day and opportunities/jobs were not open at that time. The PMI levels might be higher because of the operational levels of the public transit modes at this time, which will have a positive effect on PMI and Accessibility as well.

10h00

The level of Accessibility was higher because of the time of the day and opportunities/jobs were open at that time. The PMI levels might be lower because of the operational levels of the public transit modes at this time, which will have a positive effect on PMI and Accessibility as well.

16h00

The level of Accessibility was high because of the time of the day and opportunities/jobs were open at that time. The PMI levels might be higher because of the operational levels of the public transit modes at this time, which will have a positive effect on PMI and Accessibility as well.

6.5.2 Spatial Location of Affected Groups

The middle TAZ region of the City of Cape Town is less affected by an Accessibility deficit because of the high distribution of opportunities in that region. As well as the high PMI scores experienced in these regions because it is central and produced the best PMI values. This is due to the structure of the PMI aerial distance to other TAZ points in the calculations. The Accessibility levels of the affected TAZ increases from 06h00 to 10h00 then decreases from 10h00 to 16h00. The PMI levels of the affected TAZ decrease from 06h00 to 10h00 then increases at 10h00 to 16h00. These graphs can be seen in Chapter 5.2.1.

6.5.3 Demographic and Income Threshold

The Demographic and Income threshold was very useful in identifying which TAZ is the most crucial need for an Accessibility and PMI increase from the affected TAZ with an Accessibility and PMI deficit.

For both of the previously disadvantaged groups selected for further analyses, the TAZ identified is mainly in the same locations throughout all the times analysed. The Coloured population TAZ affected groups are more than double the Black-African population TAZ affected groups. This indicated that these Coloured population groups could need an Accessibility and PMI increase with a high priority based on the number of groups that fall below the refined threshold.

6.6 Review of Method

This method's main goal was to calculate accessibility and PMI values using GTFS data. This chosen method was capable of performing this task. Although producing results in ArcGIS and Python has been proven to be computational memory intensive, however correct results were produced. To calculate the travel time of 1616 origins to 1616 destinations it took approximately over an hour to produce. Calculating the PMI and Accessibility levels experienced by a TAZ for a day took approximately 18 hours to produce. Calculating the PMI and Accessibility level experienced by all TAZ for an hour instance took approximately 60 days to produce. The processes could have gone faster with better computational processor power and more knowledge of how Python code works to sort large data sets.

Chapter 7

Conclusion

7.1 Summary of Results

Developing the two transport justice tools, Temporal Aggregation and Space-Time Cube, aided in the ability to identify transport injustices in accessibility and potential mobility of various locations and at different times of the day of these locations within Cape Town.

A brief summary of the main results is using the two Transport Justice Tools for this research project. The first tool, Space-Time cube python tool created for this project, will show that Bishops Court has the lowest PMI and Accessibility levels throughout the day, followed by Mitchells Plain and Khayelitsha with the highest PMI and Accessibility levels. This is largely due to the UDI modal split in use, where a weighted average PMI and Accessibility values were produced based on which public transit modes are used the most in a particular TAZ.

Using the second Transport Justice Tool, the Temporal Aggregation python tool, created for this project showed Accessibility levels are low at 6h00, then increases at 10h00, then decreases at 16h00 again. This is due to taking operational hours of the opportunities into account in the accessibility calculation as well as how frequently the public transit operates in the peak and off-peak periods of the day. The PMI levels are high at 6h00, then decreases at 10h00, then increases at 16h00 again. This is due to how frequently the public transit operates in the peak and off-peak periods of the day. PMI value is the total distance over the total travel time. PMI measure is the average value of PMI values of different origin-destination pairs for a particular region. If a person has to wait for a public transit to arrive, the waiting time is added to the total travel time which will result in a lower PMI experienced by the user of public transport. This could indicate that a person could reach destinations much faster due to the increase of availability of a particular public transit mode in the peak hours, however less accessibility due to the day these periods occur.

If transit stops are closer to the origin-destination pairs then it will decrease walking time which will increase the PMI value, which will increase the PMI score of the PMI measure.

7.2 Implications of Results

Completing this research project proved that it was possible to calculate the travel time of public transit using GTFS data. Doing this also enables accessibility and PMI to be determined for a TAZ. The travel time varied throughout the day in a manner of how it would be in real life.

Determining the PMI and accessibility values of a particular transit mode throughout the day is possible. This helped the owners of the transit mode to aid in identifying regions that need an accessibility and PMI increase and at what time of the day it is necessary for these increases to occur. This will help the owners of the public transit to plan more efficiently to save on the cost of providing this transport service. It will also encourage the public to make use of their public transit mode, which will increase their overall accessibility and PMI within their TAZ.

Incorporating income and demographic data as constraint parameters allows decision-makers to make a multivariate decision for the identification of the TAZ groups that qualify for a potential accessibility and PMI increase. Using these parameters would aid in the reasoning behind decisions made for transport justice.

7.3 Recommendations for further research

Recommendations mainly speak to methods with little real-life application of the analyses and results in terms of planning and decision-making. This needs to be addressed.

The transport justice tools created could allow for refined potential accessibility and potential mobility based on specific decision-makers requirements. The potential accessibility and potential mobility measures could be determined for specific types of opportunities, e.g. health care only or a combination of specified opportunities, instead of a cumulative opportunity measure for potential accessibility and potential mobility. The Temporal Aggregation Python tool could incorporate more parameters or allow for a specific combination of parameters to be used in deciding if a TAZ is eligible or not for a potential accessibility and potential mobility increase.

It will allow the decision-makers to produce different results for the same problem based on different criteria that should be featured in decision-making while dealing with potential accessibility and potential mobility issues.

The GTFS data set used was static data/scheduled data. Using real-time GTFS data sets will produce a travel time closer to the truth which will also produce an adjusted accessibility and PMI score. The PMI measure could be normalised to display values between 0 and 1. This will help to make direct PMI measure comparisons between public transport services and will help better express small PMI scores.

Having a complete and up-to-date data set of the number of opportunities/jobs for the TAZ of the City of Cape Town will help to produce a more accurate accessibility distribution across the TAZ groups. There were some TAZ locations that had zero opportunities recorded in the opportunities data set. This could have been a result of people not having recorded information about these TAZ locations or these TAZ locations actually do not have any opportunities within them. This had a negative effect on the potential accessibility measure for these TAZ with zero opportunities and resulted in a potential accessibility value of zero, as this value is determined by the product of the number of opportunities and the accessibility of the TAZ.

Including a GINI coefficient measure could highly enrich the results as it is an income inequality measure. Including this inequality measure in this project was not possible as the correct income data to perform this income measure was not available for the TAZ grid of the City of Cape.

Certain data sets had to be spatially interpolated to match TAZ grid from their original spatial grid. Data that has the exact same spatial grid with all the data set will produce the most accurate result for accessibility and PMI.

Having a strong foundation in Python coding language will greatly assist in sorting large data sets. Understanding how the code calculates the accessibility and PMI measures could help in finding a way to increase the processing time of the tools that were created for this project. Using a computer that has high-speed ram, an SSD and a CPU processor with very high speeds will help to significantly reduce the computational time of the process to determine accessibility and PMI scores.

References

Alesina, A., & Ferrara, E. (2005). "Ethnic Diversity and Economic Performance" (PDF). Journal of Economic Literature: 762–800. Retrieved December 15, 2016.

Annez, P. & Buckley, R. (2006). Urbanization and Growth: Setting the Context. In Urbanisation and Growth. M. Spence, C. Annez, Patricia, & M. Buckley, Robert, Eds. Washington: The World Bank. 1–46.

Ashok, K., & Ben-Akiva, M. E. (2000). Alternative approaches for real-time estimation and prediction of time-dependent origin destination flows. Transportation Science, 34, 21-36.

Banister, D. (2008). The sustainable mobility paradigm. Transport Policy, 73-80.

Bantu Education Act, No. 47 of 1953. 1953. Pretoria: Government Printer.

Behrens, R., McCormick, D. & Mfinanga, D. (2016). Abingdon, Oxon, UK. Routledge.

Benenson, I., K. Martens, Y. Rofé, et al. (2011). Public transport versus private car: GIS-based estimation of accessibility applied to the Tel Aviv metropolitan area. The Annals of Regional Science, 47(3): 499–515.

Bhorat, H., van der Westhuizen, C. & Jacobs, T. (2009). Income and non-income inequality in post-apartheid South Africa: What are the drivers and possible policy interventions? SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2184183

Black, W.R. (2003). Transportation: a geographical analysis. New York, The Guilford Press.

Census, (2011). "A Geo-referenced Census Frame of Dwellings for the 2011 Census of the Republic of South Africa". United Nations Statistics Division (Census Knowledge Base).

Cervero, R. (2013). Mobility and urban form. In UN-Habitat, Planning and design for sustainable urban mobility: Global report on human settlements 2013. United

Nations Human Settlements Programme. Abington: Earthscan. 75-106

Chen, M., Zhang, H., Liu, W. & Zhang, W. (2014). The global pattern of urbanization and economic growth: Evidence from the last three decades. PLoS ONE. 9(8).

Chen, P.P.S. (1976). The entity relationship model – towards a unified view of data. Association of Computing Machinery Transactions on Database Systems 1(1):9-36.

Christie, P. & Collins, C. (1982). Bantu Education: Apartheid Ideology or Labour Reproduction?. Comparative Education. 18(1):55-75.

Christopher, A.J. (2001). Urban Segregation in Post-Apartheid South Africa. Urban Studies. 38(3):449-466.

Church, A., Frost, M., & Sullivan, K. (2000). Transport and social exclusion in London. Transport Policy, 7, 195-205.

Ciommo, F. D., & Shiftan, Y. (2017). Transport equity analysis. Transport Reviews, 37(2), 139-151.

City of Cape Town Transport and Urban Development Unit, (2019). Development of a Urban Development Index (UDI).

Codd, E.F. (1970). A relational model of data for large shared data banks. Communications of the ACM 13(6).

Coovadia, H. et al. (2009) 'Health in South Africa 1: The health and health system of South Africa: historical roots of current public health challenges.', Lancet, 374(9692), pp. 817–834.

Curtin, K., Noronha, V., Goodchild, M.F. and Grise, S. (2001). ArcGIS transportation data model (draft). Environmental Systems Research Institute, Inc.

Delobelle, P. (2013) 'The health system in South Africa. Historical perspectives and current challenges.'

Dikec, M. (2009). Justice and the spatial imagination. In: P. Marcuse, J. Connolly, J. Novy et al. (eds) (2011), Searching for the Just City: debates in urban theory and practice. London, Routledge.

Dondo, C., & Rivett, U. (2004). Spatial information systems in managing public transport information. South African Journal of Information Management, 6(2).

Ellegård, K. & U. Svedin (2012). Torsten Hägerstrand's Time-geography as the Cradle of the Activity Approach in Transport Geography. Journal of Transport Geography 23, pp. 17–25.

ESRI, 2018a. Extract by Mask.
URL http://desktop.arcgis.com/en/arcmap/10.5/tools/spatial-analyst-toolbox/extract-by-mask.htm

ESRI, 2018b. Add GTFS to a Network Dataset.
URL https://www.arcgis.com/home/item.html?id=0fa52a75d9ba4abcad6b88bb6285fae1

ESRI, 2018c. Feature to Point.
URL https://desktop.arcgis.com/en/arcmap/10.3/tools/data-management-toolbox/feature-to-point.htm

ESRI, 2018d. Transit Network Analysis Tools.
URL https://www.arcgis.com/home/item.html?id=23cc1b720eb0418db4d0e61955b8fe43

ESRI, 2018e. Table To Table.
URL https://desktop.arcgis.com/en/arcmap/latest/tools/conversion-toolbox/table-to-table.htm

Farber, S., Morang, M. and Widener, M., (2014). Temporal variability in transit-based accessibility to supermarkets. Applied Geography, 53, pp.149-159.

Fearon, J. (2003). Journal Of Economic Growth, 8(2), 195-222. doi: 10.1023/a:1024419522867

Geurs, K.T. & B. van Wee (2004). Accessibility Evaluation of Land-use and Transport Strategies: Review and Research Directions. Journal of Transport Geography 12, pp. 127–140.

Giliomee, H & Mbenga, B.(2007). New history of South Africa. Cape Town: Tafelberg

GINI, C. (1912). Variabilità e mutabilità, contributo allo studio delle distribuzioni e relazione statistiche. Studi Economico-Giuridici dell'Univ. di Cagliari 3, 1–158.

Giuliano, G. (1995) "Land Use Impacts of Transportation Investments: Highway and Transit", in S. Hanson (ed) The Geography of Urban Transportation, New York: The Guilford Press, p. 307.

Glaeser, E.L. 2013. A world of cities: The causes and consequences of urbanisation in poorer countries. (19745). Cambridge.

Golub, A., & Martens, K. (2014). Using principles of justice to assess the modal equity of regional transportation plans. Journal of Transport Geography, 10-20.

Google, (2016). Google Company: Our history in depth.

Grengs, J. (2018). On the way but not there yet: Making accessibility the core of equity planning in transportation. In N. Krumholz, & K. W. Hexter, Advancing Equity Planning Now (pp. 128-148). New York: Cornell University Press.

GTFS Static Overview l Static Transit l Google Developers. (2019). Retrieved from https://developers.google.com/transit/gtfs/

Group Areas Act, No. 41 of 1950. 1950. Pretoria: Government Printer.

GTFS in ArcGIS. (2018). Retrieved from https://esri.github.io/public-transit-tools/

Guzman, L., Oviedo, D., & River, C. (2017). Assessing equity in transport accessibility to work and study: The Bogotá region. Journal of Transport Geography, 236-246.

Hägerstrand, T. (1970), What about People in Regional Science? Papers of the Regional Science Association 24, pp. 7–21.

Haghani, A., & Jung, S. (2005). A dynamic vehicle routing problem with time-dependent travel times. Computers & Operations Research, 32, 2959-2986.

Henderson, V. (2002). Urbanization in Developing Countries. The World Bank Research Observer. 17(1):89–112.

Henry Lever, H. (1982). Ethnicity in south African society, Humboldt Journal of Social Relations, Vol. 10, No. 1, Race & Ethnic Relations: Cross-Cultural Perspectives, pp. 239-253.—

Janusz, K., Kesteloot, C., Vermeiren, K., & Van Rompaey, A. (2019). Daily Mobility, livelihoods and transport policies in Kampala, Uganda: A Hägerstrandian analysis. Tijdschrift Voor Economische En Sociale Geografie, 110(4), 412-427. doi:10.1111/tesg.12349

Järv, O., Tenkanen, H., Salonen, M., Ahas, R. and Toivonen, T., (2018). Dynamic cities: Location-based accessibility modelling as a function of time. Applied Geography, 95, pp.101-110.

Jennings, G. (2015). Public Transport Interventions and transport justice in South Africa: A literature and policy review. Southern African Transport Conference. 34, pp. 764-774. Johannesburg: Southern African Transport Conference.

Kammen, D.M. & Sunter, D.A. (2016). City-integrated renewable energy for urban sustainability. Science (New York, N.Y.). 352(6288):922–8.

Kamruzzaman, M., Hine, J., Gunay, B. & Blair, N. (2011). Using GIS to visualise and evaluate student travel behaviour. Journal of Transport Geography. 19(1):13-32. DOI:http://dx.doi.org.ezproxy.uct.ac.za/10.1016/j.jtrangeo.2009.09.004.

Kamruzzaman, M., Yigitcanlar, T., Yang, J., & Mohamed, M. A. (2016). Measures of transport-related social exclusion: A critical review of the literature. Sustainability, 8(7), 6-11.

Kari Edison Watkins, Brian Ferris, Alan Borning, G. Scott Rutherford, and David Layton (2011). "Where Is My Bus? Impact of mobile real-time information on the perceived and actual wait time of transit riders," Transportation Research Part A: Policy and Practice, Vol. 45 pp. 839–848.

Klopp, J., Mutua, J., Orwa, D., Waiganjo, P., White, A., & Williams, S. (2014). Towards a Standard for Paratransit Data: Lessons from Developing GTFS Data for Nairobi's Matatu System (No. 14-5280).

Krizek, K., Lacono, M., El-Geneidy, A., Liao, F. C., & Johns, R. (2009). Application of accessibility measures for non-Auto travel Modes. Saint Paul: Minnesota Department of Transportation Research Services Section.

Kwan, M.-p. (1998). Accessibility: A Comparative Analysis Using a Point-based Framework. Geographical Analysis, 30(3).

Laurini, R. and Thompson, D. (1996). Fundamentals of spatial information systems. Academic Press Limited

Leclerc, P. D., McLay, L. A., & Mayorga, M. E. (2012). Modeling equity for allocating public resources. Community-based operations research. Springer 97–118.

Leibbrandt, M., Woolard, I., & Bhorat, H. (2001). Understanding Contemporary Household Inequality in South Africa. In: Fighting Poverty: Labour Markets and Inequality in South Africa. HSRC Press, Cape Town, pp. 21-40.

Litman, T. (2003). Measuring Transportation: Traffic, mobility, and accessibility. ITE Journal, 73(10), 28-32.

Litman, T. (2014). Evaluating transportation equity: Guidance for incorporating distributional impacts in transportation planning. Victoria Transport Policy Institute.

Loveland, Ian (1999). By Due Process of Law: Racial Discrimination & the Right to Vote in South Africa 1855-1960. Hart. p. 242-3.

Lucas, K. 2012. Transport and social exclusion: Where are we now? Transport Policy. 20:105–113.

Lucas, K., Mattioli, G., Verlinghieri, E., Guzman, A. (2016). Transport and its adverse social consequences. Institute for Transport Studies. Leeds: ICE Publishing.

Mabin, A. (1992). Comprehensive Segregation: The Origins of the Group Areas Act and Its Planning Apparatus. Journal of Southern African Studies. 18(2):405-429.

Martens, K. (2017). Transport Justice (1st ed.). New York: Routledge

Makri, M.-c., & Carolin, F. (1969). Accessibility measures for analyses of land use and travelling with geographical information systems. Lund Institute of Technology, Department of Technology and Society. Lund University.

Mattioli, G. (2014). Where sustainable transport and social exclusion meet: households without cars and car dependence in Great Britain. Journal of Environmental Policy & Planning, 16(3): 379–400.

Miller, H.J. (1991). Modelling Accessibility Using Space-time Prism Concepts within Geographic Information Systems. International Journal of Geographical Information Systems 5, pp. 287–301

Moore, T., Thorsnes, P. & Appleyard, B. (2007). The Transportation/Land Use Connection. Chicago: American Planning Association Planning Advisory Service

Mxolisis, M. (2006) A Critical Analysis of the Process of the Taxi Recapitalization Policy. Master Degree in Development and Management, North West University, South Africa.

Niaraki, A.S. & Kim, K. (2009). Ontology based personalized route planning system using a multi-criteria decision making approach. Expert Systems with Applications. 36(2 PART 1):2250-2259. DOI:10.1016/j.eswa.2007.12.053.

Nengwekhulu, R, H. (1986). Race, Class and Ethnicity in South Africa, African Journal of Political Economy, Vol. 1, No. 1, Southern Africa in crisis, pp. 29-39.—

OECD. (2018). Rethinking Urban Sprawl Moving Towards Sustainable cities. Paris.

Orcutt, J. (2019). Transportation Struggles in the Post-Apartheid City. Retrieved from http://www.plannersnetwork.org/1997/09/transportation-struggles-in-the-post-apartheid-city/

Patsiurko, N., Campbell, J., & Hall, J. (2011). Measuring cultural diversity: ethnic, linguistic and religious fractionalization in the OECD. Ethnic And Racial Studies, 1-23. doi: 10.1080/01419870.2011.579136

Parnell, S. 1993. Creating Racial Privilege: The Origins of South African Public Health and Town Planning Legislation. Journal of Southern African Studies. 19(3):471-488.

Pendakur, V. S. (2005). Non motorized transport in African cities lessons from experience in Kenya and Tanzania. Sub-Saharan Africa Transport Policy Program.

Pirie, G. H. (1979). Measuring accessibility: a review and proposal. Environment and Planning A, 11: 299–312.

Pirie, G. H. (2009). Virtuous mobility: moralising vs measuring geographical mobility in Africa. Africa Focus, 22(1), 21-35.

Posel, D. (2001). "What's in a name? Racial categorisations under apartheid and their afterlife" (PDF). Transformation: 50–74. ISSN 0258-7696. Archived from the original (PDF) on 8 November 2006.

Preston, J., & Rajé, F. (2007). Accessibility, mobility, and transport-related social exclusion. Journal of Transport Geography, 15(3), 151-160.

Prohibition of Mixed Marriages Act, Act No. 55 of 1949. 1949. Pretoria: Government Printer.

Ramutsindela, M, F. (1997). National identity in South Africa: the search for harmony, GeoJournal, Vol. 43, No. 1, pp. 99-110. —

Rawls, J. (1971). A Theory of Justice. Cambridge, MA, Harvard University Press.

Rode, P., Floater, G., Thomopoulos, N., Docherty, J., Schwinger, P., Mahendra, A. & Fang, W. 2017. Accessibility in Cities: Transport and Urban Form. (November):239–273.

Rondganger, L. (2006). "Being an African makes me who I am". IOL. Retrieved 9 March 2013.

Ross, R. (2008). A concise history of South Africa

Sager, T. (2005). Footloose and forecast-free: Hypermobility and the planning of society. European Journal of Spatial Development, 1-23.

Theil, H. (1967). Economics and Information Theory. Chicago: Rand McNally and Company.

Thompson, E., Arafat, A., O'Dell, W., Steiner, R. & Zwick, P. (2012). Helping Put Theory into Practice for Planning Sustainable Communities: A GIS Tool for Measuring Transit Accessibility. Gainesville, United States of America: Shimberg Center for Housing Studies.

UN. 2007. The state of the World's cities Report 2006/7. Nairobi, Kenya.

UN Habitat. (2013). Planning and design for sustainable urban mobility: Global report on human settlements. UN Habitat, Asia.

van Wee, B., Annema, J. A., & Banister, D. (2013). The transport system and transport Policy: An introduction. (J. A. Bert van Wee, Ed.) Cheltenham: Edward

Elgar Publishing.

Van Den Berghe, P. L. (1966). Racial segregation in South Africa: Degrees and kinds. Cahiers D'études Africaines, 6(23), 408-418. doi:10.3406/cea.1966.3075

Vuchic, V. (2005). Urban Transit: Operations, Planning and Economics. Hoboken, New Jersey: John Wiley, pp. 445 - 448; 469 - 470.

Wang, C.H., Chen, N.A., 2015. A GIS-based spatial statistical approach to modeling job accessibility by transportation mode: case study of Columbus, Ohio. J. Transp. Geogr. 45, 1–11.

Wang, Q. 2014. Effects of urbanisation on energy consumption in China. Energy Policy.65:332–339.

WHO. (2013). Road traffic injuries. Retrieved 2019, from World Health Organization: https://www.who.int/violence$_{i}njury_{p}revention/road_{s}afety_{s}tatus/2013/report/factsheet_{a}fro$

Wong, S., & Tong, C. (1998). Estimation of time-dependent origin=destination matrices for transit networks. Transportation Research Part B: Methodological, 32, 35-48.

Zeiler, M. (1999). Modelling our world. The ESRI guide to geodatabase design. Environmental Systems Research Institute Inc.

Zhou, J., Long, Y., (2016). Losers and Pareto optimality in optimising commuting patterns. Urban Stud. 53 (12), 2511-2529.

Appendices

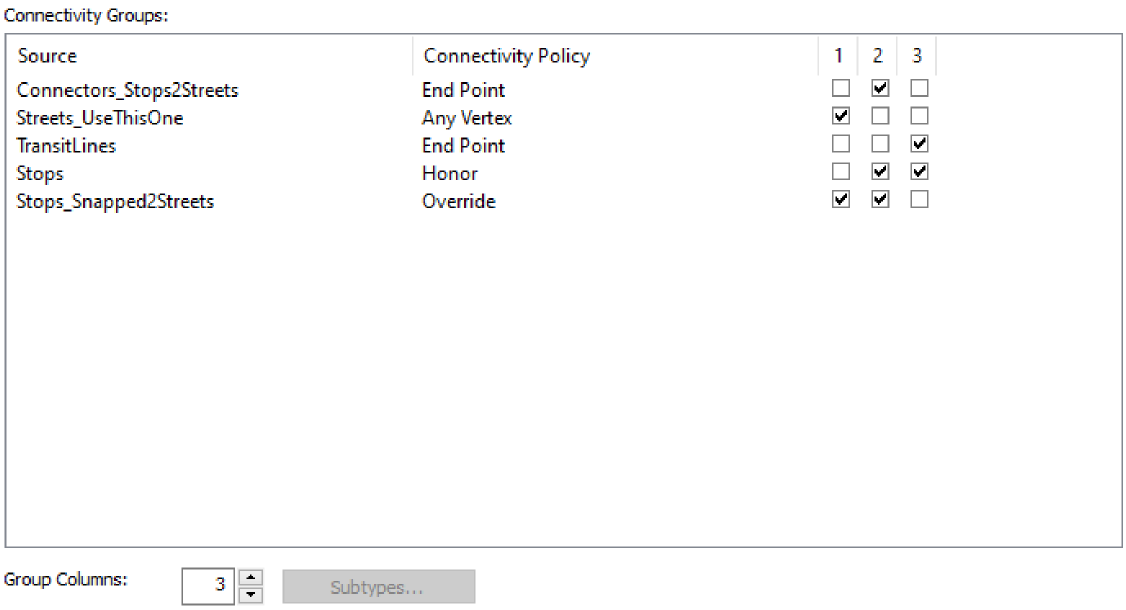


Figure 7.1: Connectivity of the Feature Classes for the Network Analyst Layer

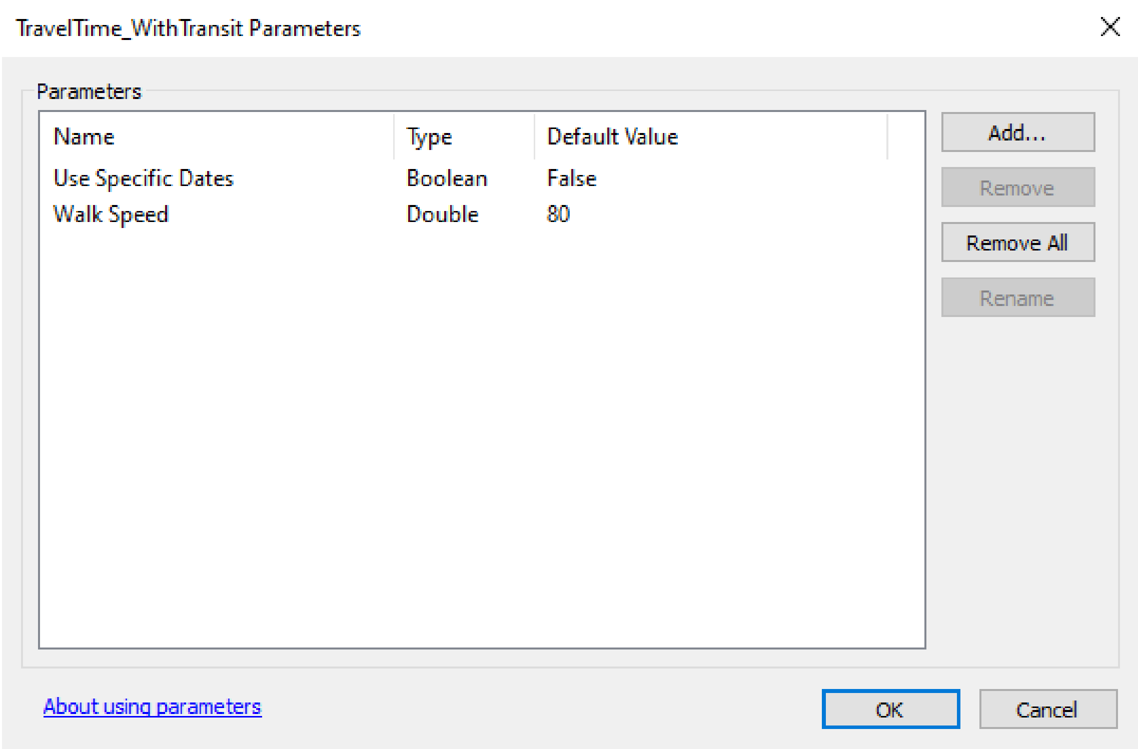


Figure 7.2: Parameters of for the Network Analyst Layer

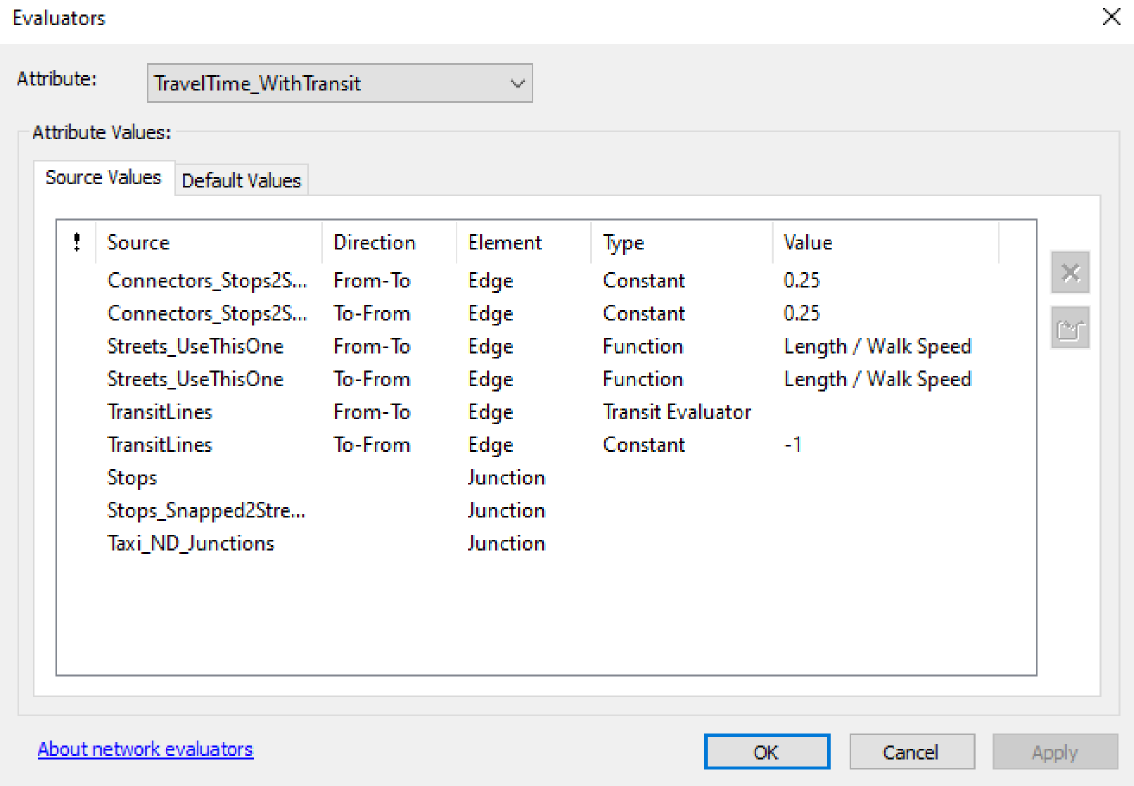


Figure 7.3: Evaluators of for the Network Analyst Layer

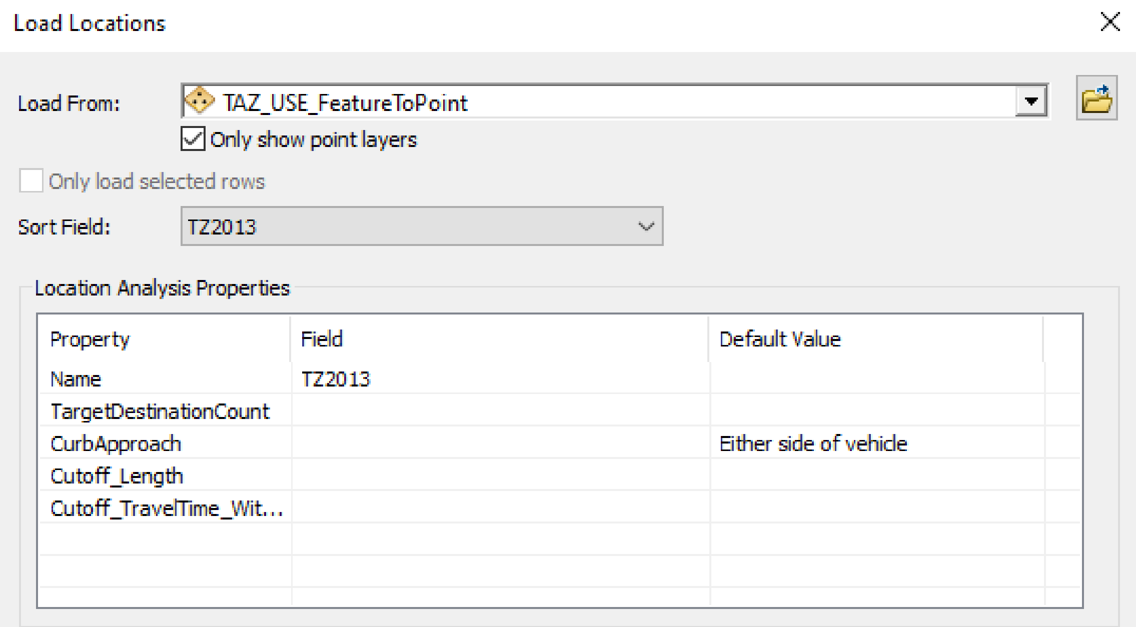


Figure 7.4: Load Locations of the OD Cost Matrix for the Network Analyst Layer

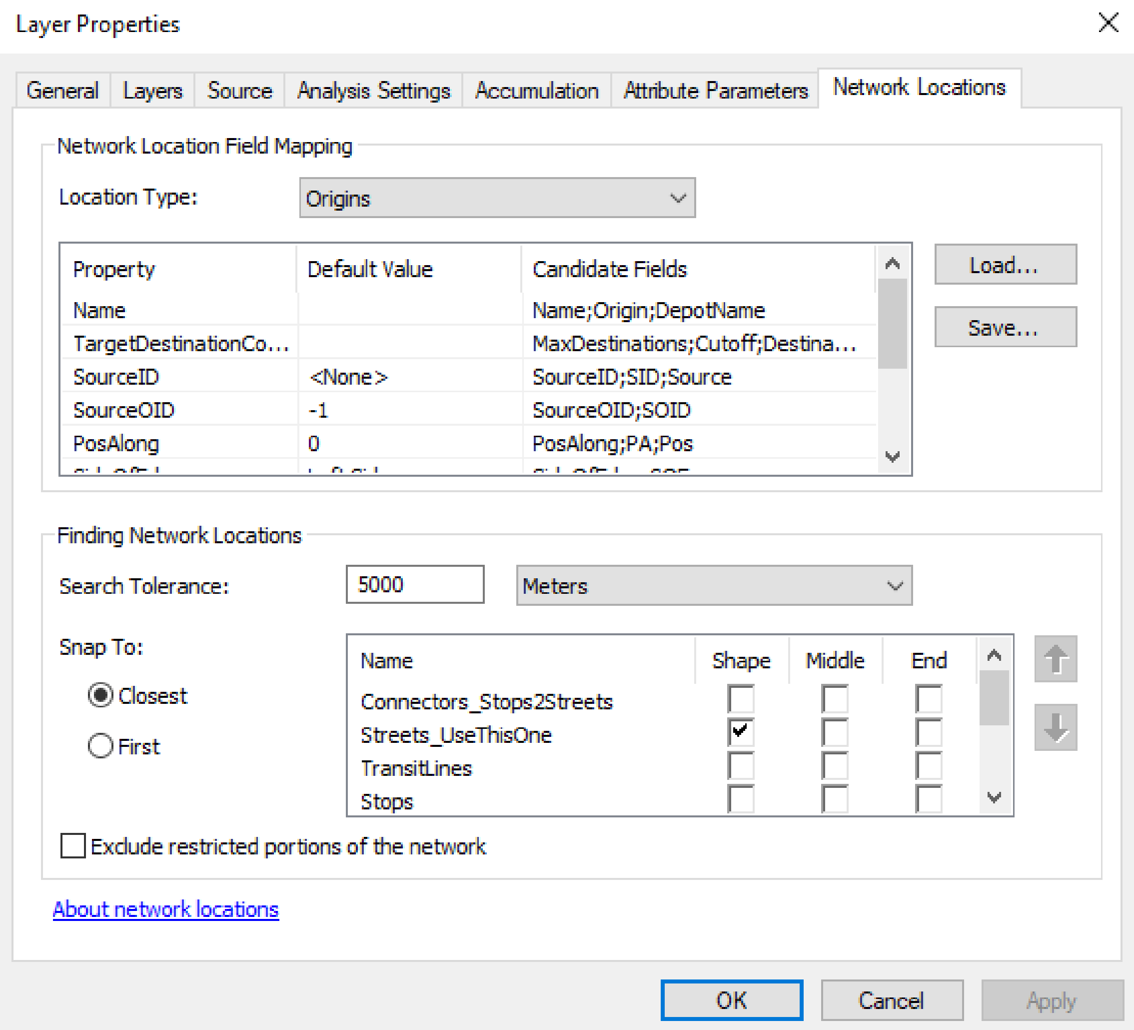


Figure 7.5: Layer Properties of the OD Cost Matrix for the Network Analyst Layer

```
// 3D Cube Plotter.py
# -*- coding: utf-8 -*-
"""
Created on Mon Jun 29 12:38:31 2020

@author: Brent Kotzee
"""
```

```
import pandas as pd
#import numpy as np
#from pandas import DataFrame
#from PMI import PMI
#from JUSTPMI import JustPMI
#from JUSTPAI import JustPai
#from BOTH import both
#from fucs import GetPMIcum,GetACCcum

#now plot parameter with time, accessibility and potential mobility as
    the axis

from mpl_toolkits import mplot3d
import matplotlib.pyplot as plt
plt.style.use('seaborn-whitegrid')

Hour = []
TAZ=[]
t=[]

#timeIN =
    [0:00:00,1:00:00,2:00:00,3:00:00,4:00:00,5:00:00,6:00:00,7:00:00,8:00:00,
9:00:00,10:00:00,11:00:00,12:00:00,13:00:00,14:00:00,15:00:00,16:00:00,17:00:00,
18:00:00,19:00:00,20:00:00,21:00:00,22:00:00,23:00:00]
timeIN=""
periodIN=""
orGININ=""


#timeIN = input("The hours you want to anal:\n")
#
#periodIN= input("The period you want to analyse in minutes:\n")

Transport = input("Please type in the
    Transport(BUS;Metro;MyCiti;Taxi):\n" )

Day = input("Please type in the day(Tuesday;Wednesday;Thursday):\n" )


#(h,m,s) = timeIN.split(':')
#Hour.append(int(h))
```

```
orGININ = input("Please type in the TAZ Origin:\n")
df =
    pd.read_csv("Cumulative_"+Transport+"_"+Day+"_"+str(orGININ)+'.txt',sep=";")


#MPTsub= df.loc[df["Travel_Time"] <= float(periodIN)]
accC = []
a = []
time=[]
for i in range(len(df)):
    o=df[ "PMI" ].values[i]
    p=df[ "Accessibilty" ].values[i]
    TAZ.append(df[ "TAZ" ].values[i])
    time.append(df[ "Time" ].values[i])
    a.append(o)
    accC.append(p)

print ("pmi\n")
fig = plt.figure()
ax = mplot3d.Axes3D(fig)
ax.set_xlim3d(0, max(a))
ax.set_ylim3d(0, max(accC))
ax.set_zlim3d(0, max(time))
#ax.view_init(30, 360)
ax.set_xlabel('Potential Mobiltiy Index ')
ax.set_ylabel('Accessibility')
ax.set_zlabel('Time(h)')

x_points = a
y_points = accC
z_points = time

ax.scatter3D(x_points, y_points, z_points, c="black", cmap='hsv');
ax.plot3D(x_points, y_points, z_points, c="black");

plt.show()


#chioce = input("Do you want to export this data (1 for Yes/ 0 for No)?
    :\n")
#if chioce == 1:
#    ProbArea.to_csv("PMI vs Accessibility
    "+Transport+"_"+Day+"_"+str(Hour)+".txt", header=True, index=False,
```

```
    sep=';', mode='a')
#else:
print("Done")
```

```
// Temporal Aggregation.py
# -*- coding: utf-8 -*-
"""
Created on Mon Jun 29 12:38:21 2020

@author: Brent Kotzee
"""
# import conda
import pandas as pd

# import cudf as pd
import cupy as np
#import numpy as np
from pandas import DataFrame
#from PMI import PMI
#from JUSTPMI import JustPMI
#from JUSTPAI import JustPai
#from BOTH import both
from fucs import GetPMIcumC,GetACCcum

#now plot parameter with time,availablity and accessibilty as the axis

#import matplotlib.pyplot as plt

Hour = []



t=[]

timeIN = "10:00:00"
Uier=[10]
#periodIN=""
#orGININ=[]

if ("," in timeIN):
        timeIN =timeIN.split(",")
#timeIN = input("The hours you want to anal:\n")
```

```
#if ("," in timeIN):
#    timeIN =timeIN.split(",")
#else:
#    timeIN =[timeIN]

#periodIN= input("The period you want to analyse in minutes:\n")
#periodIN= eval(periodIN)
#timeIN=eval(timeIN[0])


E=[]
Transport = input("Please type in the
    Transport(BUS;Metro;MyCiti;Taxi):\n" )

Day = input("Please type in the day(Tuesday;Wednesday;Thursday):\n" )

ACES=[]
APMI=[]
#for i in range(len(Uier)):
#    (h,m,s) = timeIN[i].split(':')
#    Hour.append(int(h))
df = pd.read_csv( Transport+"_"+Day+"_"+str(Uier[0])+'.txt',sep=";")
FOE=sorted( list( dict.fromkeys( df["Origin"].tolist() ) ) )

for l in range(len(FOE)) :
#    orGININ =[]
#    orGININ.append(str(l))
    a = []
    accC = []
#    if i ==0:
#
#
#        print( sorted( list( dict.fromkeys( df["Origin"].tolist() ) ) ) )
#        #TAZ=sorted( list( dict.fromkeys( df["Origin"].tolist() ) ) )
#        orGININ = input("Please type in the TAZ Origin:\n")
#        #Histo(orGININ)


    accC1=GetACCcum(str(FOE[l]), timeIN, df)
    a1=GetPMIcumC(str(FOE[l]), timeIN, df)

    accC = np.append(accC,accC1)
    a = np.append(a,a1)
```

```
    # accC=accC+accC1

    # a=a+a1
    # ACES.append(sum(accC))
    # APMI.append(sum(a))
    ACES = np.append(ACES,np.prod(accC1))
    APMI = np.append(APMI,np.prod(a1))
#E= sorted( list( dict.fromkeys( df["Origin"].tolist() ) ) )
tim=[]
for x in range(len(FOE)):
    # tim.append(Uier[0])
    tim = np.append(tim,Uier[0])

MPT = pd.DataFrame([FOE,APMI,ACES,tim]).transpose()
MPT.columns = ['TAZ','PMI','Accessibilty','Time']
#print (MPT)

for i in range(len(Uier)):
    MPT.to_csv("HourSlice_"+Transport+"_"+Day+"_"+str(Uier[i])+".txt",
        header=True, index=False, sep=';', mode='a')

#MPTsub= MPT[MPT["Travel_Time"] <= periodIN[0]]
#
#a= MPT[ "PMI" ].values[0]
#accC= MPT[ "Accessibilty" ].values[0]
#
#
#print ("PAI vs PMI\n")
#fig = plt.figure()
#
#fig.set_xlim(0, max(a))
#fig.set_ylim(0, max(accC))
#
##ax.view_init(30, 360)
#fig.set_xlabel('Potential Mobiltiy Index ')
#fig.set_ylabel('Accessibility')



#x_points = a
#y_points = accC
#fig.scatter(x_points, y_points, c="black", cmap='hsv');
```

```
#plt.xlabel("Leprechauns")
#plt.ylabel("Gold")
#plt.plot(x_points, y_points);
#
#plt.show()
#
#
#print(np.array([[round(num, 2) for num in a],[round(num,2) for num in
    accC]]).T)
#

print("DONE")
```

```
// Cumulative Plotter.py

# -*- coding: utf-8 -*-
"""
Created on Mon Jun 29 12:38:21 2020

@author: Brent Kotzee
"""

import pandas as pd
import numpy as np
#from pandas import DataFrame
#from PMI import PMI
#from JUSTPMI import JustPMI
#from JUSTPAI import JustPai
#from BOTH import both
#from fucs import GetPMIcum,GetACCcum

#now plot parameter with time,availablity and accessibilty as the axis


import matplotlib.pyplot as plt
plt.style.use('seaborn-whitegrid')

TAZ_POP = pd.read_csv('Population.csv', header ='infer')


Hour = []
```

```
#timeIN =
    [0:00:00,1:00:00,2:00:00,3:00:00,4:00:00,5:00:00,6:00:00,7:00:00,8:00:00,
9:00:00,10:00:00,11:00:00,12:00:00,13:00:00,14:00:00,15:00:00,16:00:00,17:00:00,
18:00:00,19:00:00,20:00:00,21:00:00,22:00:00,23:00:00]
timeIN=""
periodIN=""
orGININ=""


timeIN = input("The hours you want to anal:\n")

periodIN= input("The period you want to analyse in minutes:\n")

Transport = input("Please type in the
    Transport(BUS;Metro;MyCiti;Taxi):\n" )

Day = input("Please type in the day(Tuesday;Wednesday;Thursday):\n" )


(h,m,s) = timeIN.split(':')
Hour.append(int(h))

orGININ = input("Please type in the TAZ Origin:\n")
df =
    pd.read_csv("Cumulative_"+orGININ+"_"+Transport+"_"+Day+"_"+str(Hour[0])+'.txt',
sep=";")
#TAZ=sorted( list( dict.fromkeys( df["TAZ"].tolist() ) ) )

MPTsub= df.loc[df["Travel_Time"] <= float(periodIN)]


#TAZ_POP= TAZ_POP.loc[(df["TZ_2013"].isin(TAZ))]

accC = []
a = []
for i in range(len(MPTsub)):
    o=MPTsub[ "PMI" ].values[i]
    p=MPTsub[ "Accessibilty" ].values[i]
    a.append(o)
    accC.append(p)
```

```
#Threshold percentage
th5=MPTsub.loc[(df["PMI"]<= (0.5)*(max(a))) & (df["Accessibilty"]<=
    (0.5*0.5*100)*(max(accC))) ]
th4=MPTsub.loc[(df["PMI"]<= (0.5)*(max(a))) & (df["Accessibilty"]<=
    (0.4*0.5*100)*(max(accC))) ]
th3=MPTsub.loc[(df["PMI"]<= (0.5)*(max(a))) & (df["Accessibilty"]<=
    (0.3*0.5*100)*(max(accC))) ]
th2=MPTsub.loc[(df["PMI"]<= (0.5)*(max(a))) & (df["Accessibilty"]<=
    (0.2*0.5*100)*(max(accC))) ]
th1=MPTsub.loc[(df["PMI"]<= (0.5)*(max(a))) & (df["Accessibilty"]<=
    (0.1*0.5*100)*(max(accC))) ]
th0=MPTsub.loc[(df["PMI"]<= (0.5)*(max(a))) & (df["Accessibilty"]<=
    (0.5*100)*(max(accC))) ]

#Taz
p5=[]
p4=[]
p3=[]
p2=[]
p1=[]
p0=[]
for i in range(len(th5)):
    p5.append(th5[ "TAZ" ].values[i])
for i in range(len(th4)):
    p4.append(th4[ "TAZ" ].values[i])
for i in range(len(th3)):
    p3.append(th3[ "TAZ" ].values[i])
for i in range(len(th2)):
    p2.append(th2[ "TAZ" ].values[i])
for i in range(len(th1)):
    p1.append(th1[ "TAZ" ].values[i])
for i in range(len(th0)):
    p0.append(th0[ "TAZ" ].values[i])

#Pecenatage
pop5=[]
pop4=[]
pop3=[]
pop2=[]
pop1=[]
pop0=[]
for i in range(len(p5)):
    pop51=TAZ_POP.loc[TAZ_POP["TZ_2013"].isin( [str(p5[i])])]
```

```
    pop5.append(pop51["Population"].values[0])
for i in range(len(p4)):
    pop41=TAZ_POP.loc[TAZ_POP["TZ_2013"].isin( [str(p4[i])])]
    pop4.append(pop41["Population"].values[0])
for i in range(len(p3)):
    pop31=TAZ_POP.loc[TAZ_POP["TZ_2013"].isin( [str(p3[i])])]
    pop3.append(pop31["Population"].values[0])
for i in range(len(p2)):
    pop21=TAZ_POP.loc[TAZ_POP["TZ_2013"].isin( [str(p2[i])])]
    pop2.append(pop21["Population"].values[0])
for i in range(len(p1)):
    pop11=TAZ_POP.loc[TAZ_POP["TZ_2013"].isin( [str(p1[i])])]
    pop1.append(pop11["Population"].values[0])
for i in range(len(p0)):
    pop01=TAZ_POP.loc[TAZ_POP["TZ_2013"].isin( [str(p0[i])])]
    pop0.append(pop01["Population"].values[0])

#LIST of percentage population
pp5=((np.sum(pop5))*100)/np.sum(pop0)
pp4=((np.sum(pop4))*100)/np.sum(pop0)
pp3=((np.sum(pop3))*100)/np.sum(pop0)
pp2=((np.sum(pop2))*100)/np.sum(pop0)
pp1=((np.sum(pop1))*100)/np.sum(pop0)

#LIST of percentage population group
gp5=((np.sum((len(pop5))))*100)/np.sum(len(pop0))
gp4=((np.sum((len(pop4))))*100)/np.sum(len(pop0))
gp3=((np.sum((len(pop3))))*100)/np.sum(len(pop0))
gp2=((np.sum((len(pop2))))*100)/np.sum(len(pop0))
gp1=((np.sum((len(pop1))))*100)/np.sum(len(pop0))

pp=[pp1,pp2,pp3,pp4,pp5]
gp=[gp1,gp2,gp3,gp4,gp5]
pc=[10,20,30,40,50]

plt.xlabel("Potential Mobility Index (aerial speed in km/h)")
plt.ylabel("Accessibilty (jobs within "+periodIN+" minutes travel time)")
plt.title('Potential Mobility Index vs Accessibilty at time '+timeIN+" in
    TAZ "+orGININ)
plt.plot(a, accC, '.', color='black');

# draw vertical line
plt.plot([0.5*(max(a)) , 0.5*(max(a))], [0, max(accC)], 'k-',
```

```
    label='green')

# draw diagonal line from (70, 90) to (90, 200)
plt.plot([0 , max(a)], [0.5*(max(accC)),0.5*(max(accC)) ], 'k-',
    label='green')
plt.plot([0 , max(a)], [0.5*0.5*(max(accC)),0.5*0.5*(max(accC)) ], 'k-',
    label='green')
plt.plot([0 , max(a)], [0.2*0.5*(max(accC)),0.2*0.5*(max(accC)) ], 'k-',
    label='green')

#Lables
#Vertical
plt.annotate(xy=[0.5*0.5*max(a), max(accC)+500], s="Average Potential
    Mobility Index Line")
#Horizontial
plt.annotate(xy=[max(a), 0.5*max(accC)], s=" Average Accessibilty Line")
plt.annotate(xy=[max(a), 0.5*0.5*max(accC)], s=" 50% threshold")
plt.annotate(xy=[max(a), 0.2*0.5*max(accC)], s=" 20% threshold")
plt.show()

##########################
#plt.xlabel("Potential Mobility Index (aerial speed in km/h)")
#plt.ylabel("Accessibilty (jobs within "+periodIN+" minutes travel time)")
#plt.title('Potential Mobility Index vs Accessibilty at time '+timeIN+"
    in TAZ "+orGININ)
#plt.plot(a, accC, '.', color='black');
#
## draw vertical line
#plt.plot([0.5*(max(a)) , 0.5*(max(a))], [0, max(accC)], 'k-',
    label='green')
#
#
## draw diagonal line from (70, 90) to (90, 200)
#plt.plot([0 , max(a)], [0.5*(max(accC)),0.5*(max(accC)) ], 'k-',
    label='green')
#plt.plot([0 , max(a)], [0.5*0.5*(max(accC)),0.5*0.5*(max(accC)) ], 'k-',
    label='green')
#plt.plot([0 , max(a)], [0.2*0.5*(max(accC)),0.2*0.5*(max(accC)) ], 'k-',
    label='green')
#
##Lables
##Vertical
```

```
#plt.annotate(xy=[0.5*0.5*max(a), max(accC)+500], s="Average Potential
    Mobility Index Line")
##Horizontial
#plt.annotate(xy=[max(a), 0.5*max(accC)], s=" Average Accessibilty Line")
#plt.annotate(xy=[max(a), 0.5*0.5*max(accC)], s=" 50% threshold")
#plt.annotate(xy=[max(a), 0.2*0.5*max(accC)], s=" 20% threshold")
#plt.show()

# libraries
import matplotlib.pyplot as plt

import pandas as pd

# Data
dff=pd.DataFrame({'x': pc, 'y1': pp, 'y2': gp })

# multiple line plot
#plt.plot( 'x', 'y1', data=df, marker='o', markerfacecolor='blue',
    markersize=12, color='skyblue', linewidth=4)

plt.plot( 'x', 'y2', data=dff, marker='', color='skyblue', linewidth=4,
    linestyle='dashed', label="By zones")
plt.plot( 'x', 'y1', data=dff, marker='', color='olive',
    linewidth=2,label="By persons")
plt.xlim(0, 52)
plt.ylim(0, 105)
plt.xlabel("Sufficiency threshold")
plt.ylabel("Percentage of total")
plt.title('Share of the population with an accessibility level below the
    respective sufficiency thresholds, for accessibility to employment, a
    '+periodIN+' minutes travel time threshold, a cumulative opportunity
    measure, at '+timeIN )
plt.legend()
plt.show()

percent = input("Which threshold percentage value do you want to view
    (0-100)? :\n")

ProbArea=MPTsub.loc[(df["PMI"]<= (0.5)*(max(a))) & (df["Accessibilty"]<=
    ((int(percent))/100)*(max(accC))) ]
print(ProbArea)

chioce = input("Do you want to export this data (1 for Yes/ 0 for No)?
```

```
    :\n")
if chioce == 1:
    ProbArea.to_csv("PMI vs Accessibility
        "+Transport+"_"+Day+"_"+str(Hour)+".txt", header=True,
        index=False, sep=';', mode='a')
else:
    print("Done")
```

```
// Hello.java
import javax.swing.JApplet;
import java.awt.Graphics;

public class Hello extends JApplet {
    public void paintComponent(Graphics g) {
        g.drawString("Hello, world!", 65, 95);
    }
}
```

```
//fucs.py
# -*- coding: utf-8 -*-
"""
Created on Tue Jun 23 15:40:07 2020

@author: Brent Kotzee
"""
# import conda
# import cudf as pd
# import cupy as np
import pandas as pd
import numpy as np

import matplotlib.pyplot as plt
# import numba # We added these two lines for a 500x speedup

# @numba.jit
dataJob = pd.read_csv('Jobs.csv', header ='infer')
dataJobHour = pd.read_csv('Jobs_Hours.csv', header ='infer')
Income = pd.read_csv('Income.csv', header ='infer')
popPerct = pd.read_csv('TAZ_Population_Percentage.csv', header ='infer')

#get the data stored using the orgin-destinationtime key to gettrav time
```

```
    and length to calculate the accesibilty value
#the function is given orgin value, time[list],destiniation value,
    dictionary created in the main, user inputed name
def GetData(origin, time, destination, df, name):
        acc = []

        jobNum = dataJob.loc[(dataJob["TZ_2013"] == eval(destination)),
            name].values[0]

        for i in range(len(list(dict.fromkeys(time)))):

            #           data= df.loc[(df["Name"].isin( [str(origin) + " -
                "+ str(destination)] )) & (df["Time"].isin([str(time[i] ) ]
                ) )]
            data= df.loc[(df["Origin"].isin( [str(origin)])) &
                (df["Dest"].isin( [str(destination)] )) &
                (df["Time"].isin([str(time[i] ) ] ) )]
            jobHour =
                dataJobHour.loc[(dataJobHour["Operating_Time"].isin([str(time[i])]))]

            acc.append((jobHour[name].values[0])*jobNum *
                (data["Total_Trav"].values[0]))
        return acc

#def GetaccC(origin, time, destination, df, name):#can prolly remove name
#        accC = []
#        names =
    "Office,Retail,Manufacturing,Service,Warehouse,Community,Construction,Transport,Parkin
    (School),Scholars (School),Personel (Tertiary),Student
    (Tertiary),Agricultural"
#        names = names.split(",")
#        #jobHour =
    dataJobHour.loc[(dataJobHour["Operating_Time"].isin([str(time)]))]
#        for i in range(len(list(dict.fromkeys(time)))):
#            a = []
#            for j in names:
#                #data= df.loc[(df["Name"].isin( [str(origin) + " - "+
    str(destination)] )) & (df["Time"].isin([str(time[i] ) ] ) )]
#                data= df.loc[(df["Origin"].isin( [str(origin)])) &
    (df["Dest"].isin( [str(destination)] )) &
    (df["Time"].isin([str(time[i] ) ] ) )]
#                datasub = data[ "Total_Trav" ].values[0]
#                jobHour =
```

```
    dataJobHour.loc[(dataJobHour["Operating_Time"].isin([str(time[i])]))]
#               jobTotal = dataJob.loc[(dataJob["TZ_2013"].isin(
    [eval(destination)]))]
#               jobNum = (jobHour[j].values[0]) * ((jobTotal[j].sum()))
#               jobNum = jobNum + jobNum * datasub
#               a.append(jobNum)
#           accC.append(sum(a))
#
#       return accC
def GetaccC(origin, time, destination, df, name):#can prolly remove name
        accC = []
        names =
            "Office,Retail,Manufacturing,Service,Warehouse,Community,Construction,Transport
            (School),Scholars (School), Personel (Tertiary),Student
            (Tertiary),Agricultural"
        names = names.split(",")
        #jobHour =
            dataJobHour.loc[(dataJobHour["Operating_Time"].isin([str(time)]))]
        for i in range(len(list(dict.fromkeys(time)))):
            a = []
            for j in names:
                #data= df.loc[(df["Name"].isin( [str(origin) + " - "+
                    str(destination)] )) & (df["Time"].isin([str(time[i] )
                    ] ) )]
                data= df.loc[(df["Origin"].isin( [str(origin)])) &
                    (df["Dest"].isin( [str(destination)] )) &
                    (df["Time"].isin([str(time[i] ) ] ) )]
                datasub = data[ "Total_Trav" ].values[0]
                jobHour =
                    dataJobHour.loc[(dataJobHour["Operating_Time"].isin([str(time[i])]))]
                jobTotal = dataJob.loc[(dataJob["TZ_2013"].isin(
                    [eval(destination)]))]
                jobNum = (jobHour[j].values[0]) * ((jobTotal[j].sum()))
                jobNum = jobNum + jobNum * (1/datasub)
                a.append(jobNum)
            accC.append(sum(a))

        return accC

#def GetACCcum(origin, time, df):#can prolly remove name
##    for i in range(len(list(dict.fromkeys(time)))):
#           accC = []
#           TAZ=[]
```

```
#         names =
    "Office,Retail,Manufacturing,Service,Warehouse,Community,Construction,Transport,Parkin
    (School),Scholars (School),Personel (Tertiary),Student
    (Tertiary),Agricultural"
#         names = names.split(",")
#         for l in time:
#             s=[]
#             for z in names:
#                 jobHours =
    dataJobHour.loc[(dataJobHour["Operating_Time"].isin([str(time)]))]
#                 jobTotals = dataJob.loc[(dataJob["TZ_2013"].isin(
    [str(origin)]))]
#                 jobNums = (jobHours[z].values[0]) *
    ((jobTotals[z].sum()))
#                 s.append(jobNums)
#
#             o=[]
#             for k in sorted( list( dict.fromkeys( df["Dest"].tolist()
    ) ) ) :
#                     a = []
#                     for j in names:
#                         data= df.loc[(df["Origin"].isin(
    [str(origin)])) & (df["Dest"].isin( [str(k)] )) &
    (df["Time"].isin([str(time ) ] ) )]
#                         datasub = data[ "Total_Trav" ].values[0]
#                         tz= data[ "Dest" ].values[0]
#                         jobHour =
    dataJobHour.loc[(dataJobHour["Operating_Time"].isin([str(time)]))]
#                         jobTotal =
    dataJob.loc[(dataJob["TZ_2013"].isin( [str(k)]))]
#                         jobNum = (jobHour[j].values[0]) *
    ((jobTotal[j].sum()))
#                         jobNum = jobNum + jobNum *(1/datasub)
#                         a.append(jobNum)
#
#                     o.append(sum(a))
#                     TAZ.append(tz)
#
#             o.append(sum(s))
#             accC.append(sum(o))
#
#
#         return accC,TAZ
```

```
#

def GetACCcum(origin, time, df):#can prolly remove name

    accC = []
    b=[]


    for k in sorted( list( dict.fromkeys( df["Dest"].tolist() ) ) ) :
        o=[]
        if (origin==str(k)):
            jobHour =
                np.array((((dataJobHour.loc[(dataJobHour["Operating_Time"].isin([str(time)]
                (School)","Scholars (School)","Personel
                (Tertiary)","Student
                (Tertiary)","Agricultural"]]).values[0])
            jobTotal = np.array(((((dataJob.loc[(dataJob["TZ_2013"].isin(
                [str(int(origin))]))]))[["Office","Retail","Manufacturing","Service","Wareh
                (School)","Scholars (School)","Personel
                (Tertiary)","Student
                (Tertiary)","Agricultural"]])).values[0])
            jobNum = np.dot(jobHour,jobTotal)
            # o.append(np.prod(jobNum.astype(np.float64)))
            o = np.append(o,np.prod(jobNum.astype(np.float64)))
            # print("Sharp"+"_"+str(np.prod(jobNum)))


        else:

            data= df.loc[(df["Origin"].isin( [str(int(origin))])) &
                (df["Dest"].isin( [str(k)] )) &
                (df["Time"].isin([str(time)] ) )]
            jobHour =
                np.array((((dataJobHour.loc[(dataJobHour["Operating_Time"].isin([str(time)]
                (School)","Scholars (School)","Personel
                (Tertiary)","Student
                (Tertiary)","Agricultural"]]).values[0])
            jobTotal = np.array(((((dataJob.loc[(dataJob["TZ_2013"].isin(
                [str(k)]))]))[["Office","Retail","Manufacturing","Service","Warehouse","Com
                (School)","Scholars (School)","Personel
                (Tertiary)","Student
                (Tertiary)","Agricultural"]])).values[0])
            datasub = (1/(data[ "Total_Trav" ].values[0]))
```

```
            jobNum = np.dot(jobHour,jobTotal)
            # jobNum = jobHour.dot(jobTotal)
            # ja=(jobNum.astype(np.float64))*datasub
            # o.append(np.prod(ja))
            # print("Sharp"+"_"+str(k)+"_"+str(np.prod(ja)))
            o = np.append(o,np.prod((jobNum.astype(np.float64))*datasub))
            print(np.prod(o))
        b = np.append(b,np.prod(o))
    accC = np.append(accC,np.prod(b))
        # accC.append(sum(o))



    print(len(accC))
    return accC

def GetACCCcum(origin, time, df):#can prolly remove name
#    for i in range(len(list(dict.fromkeys(time)))):

            TAZ=[]
            accC = []
            names =
                "Office,Retail,Manufacturing,Service,Warehouse,Community,Construction,Trans
                (School),Scholars (School),Personel (Tertiary),Student
                (Tertiary),Agricultural"
            names = names.split(",")
            for l in np.sorted( list( dict.fromkeys( df["Origin"].tolist()
                ) ) ) :

                ACTAZ=[]
                tazz= [str(l)]
                for k in sorted( list( dict.fromkeys( df["Dest"].tolist()
                    ) ) ) :
                        a = []

#                   for z in names:
#                       jobHours =
    dataJobHour.loc[(dataJobHour["Operating_Time"].isin([str(time)]))]
#                       jobTotals = dataJob.loc[(dataJob["TZ_2013"].isin(
    [str(l)]))]
#                       jobNums = (jobHours[z].values[0]) *
    ((jobTotals[z].sum()))
#                       a.append(jobNums)
```

```
                        for j in names:
                            data= df.loc[(df["Origin"].isin( [str(l)])) &
                                (df["Dest"].isin( [str(k)] )) &
                                (df["Time"].isin([str(time ) ] ) )]
                            datasub = data[ "Total_Trav" ].values[0]
#                            tz= data[ "Dest" ].values[0]
                            jobHour =
                                dataJobHour.loc[(dataJobHour["Operating_Time"].isin([str(time
                            jobTotal = dataJob.loc[(dataJob["TZ_2013"].isin(
                                [str(k)]))]
                            jobNum = (jobHour[j].values[0]) *
                                ((jobTotal[j].sum()))
                            jobNum = jobNum + jobNum * (1/datasub)
                            a.append(jobNum)
                        ACTAZ.append(sum(a))



                accC.append(sum(ACTAZ))
                TAZ.append(tazz)
            return accC,TAZ

#def GetPMIcum(origin, time, df):#can prolly remove name
##    for i in range(len(list(dict.fromkeys(time)))):
#               a = []
#               t = []
#               for k in sorted( list( dict.fromkeys( df["Dest"].tolist()
    ) ) ) :
#                   data= df.loc[(df["Origin"].isin( [str(origin)])) &
    (df["Dest"].isin( [str(k)] )) & (df["Time"].isin([str(time) ] ) )]
#                   dataTrav = data[ "Total_Trav" ].values[0]
#                   dataLeng = data[ "Shape_Leng" ].values[0]
#                   t.append(dataTrav)
#                   if dataTrav or dataLeng == 0:
#                       PMIcum=0
#                   else:
#                       PMIcum =
    float(np.divide((dataLeng/1000),(dataTrav/60)))
#
#                   a.append(PMIcum)
#
#               return a,t
```

```
def GetPMIcumC(origin, time, df):#can prolly remove name
    a=[]

    b = []

    for k in sorted( list( dict.fromkeys( df["Dest"].tolist() ) ) ) :
        if (origin==str(k)):
            PMIcum=0.0
        else:
            PMIcum = float(np.divide((((df.loc[(df["Origin"].isin(
                [str(origin)])) & (df["Dest"].isin( [str(k)] )) &
                (df["Time"].isin([str(time) ] ) )])[ "Shape_Leng"
                ].values[0])/1000),( ((df.loc[(df["Origin"].isin(
                [str(origin)])) & (df["Dest"].isin( [str(k)] )) &
                (df["Time"].isin([str(time) ] ) )])[ "Total_Trav"
                ].values[0])/60)))
        b = np.append(b,PMIcum)
        # b.append(PMIcum)
        # print(str(PMIcum)+"_"+str(k))
    a = np.append(a,(np.divide(np.sum(b),1616.0 )))
    # a.append(float(np.divide((sum(b)),(1616) )))
    # print(str(PMIcum)+"_"+str(k)+"_"+str(a))
    print(a)
    print(len(b))
    return a

def GetCum(time, destination, df, name, origin):
    cum = []
    for i in range(len(time)):
#        data= df.loc[(df["Name"].isin( [str(origin) + " - "+
    str(destination)] )) & (df["Time"].isin([str(time[i] ) ] ) )]
        data= df.loc[(df["Origin"].isin( [str(origin)])) &
            (df["Dest"].isin( [str(destination)] )) &
            (df["Time"].isin([str(time[i] ) ] ) )]
        datasub = data[ "Total_Trav" ].sum()
        jobHour =
            dataJobHour.loc[(dataJobHour["Operating_Time"].isin([str(time[i])]))]
        sub = dataJob.loc[(dataJob["TZ_2013"] != eval(origin))]
        jobNum =(sub[str(name)].sum())*(jobHour[name].values[0])
        jobNum= jobNum + jobNum *(1/datasub)
        cum.append(jobNum)
    return cum
```

```
def GetAVA(origin, time, destination, df):
        AVA =[]
        for i in range(len(list(dict.fromkeys(time)))):
#            data= df.loc[(df["Name"].isin( [str(origin) + " - "+
    str(destination)] )) & (df["Time"].isin([str(time[i] ) ] ) )]
            data= df.loc[(df["Origin"].isin( [str(origin)])) &
                (df["Dest"].isin( [str(destination)] )) &
                (df["Time"].isin([str(time[i] ) ] ) )]
            AVA.append((data[ "Total_Trav" ].values[0]))
        return AVA

def GetPAI(origin, time, destination, df, name):
        pai =[]
        jobNum = dataJob.loc[(dataJob["TZ_2013"] == eval(destination)),
            name].values[0]
        for i in range(len(list(dict.fromkeys(time)))):
#            data= df.loc[(df["Name"].isin( [str(origin) + " - "+
    str(destination)] )) & (df["Time"].isin([str(time[i] ) ] ) )]
            data= df.loc[(df["Origin"].isin( [str(origin)])) &
                (df["Dest"].isin( [str(destination)] )) &
                (df["Time"].isin([str(time[i] ) ] ) )]
            jobHour =
                dataJobHour.loc[(dataJobHour["Operating_Time"].isin([str(time[i])]))]
            pai.append((jobHour[name].values[0])*jobNum * (1/(data[
                "Total_Trav" ].values[0])))
        return pai

def GETPMI(time, dist):

    return float(np.divide((dist/1000),(time/60)))

def GetJobs():
    for col in dataJob.columns:
        print(col)
def Histo(origin):
    Income = pd.read_csv('income.csv', header ='infer')
    popPerct = pd.read_csv('TAZ_Population_Percentage.csv', header
        ='infer')
    inrow = Income.loc[Income["TZ_2013"].isin([origin])]
    poprow = popPerct.loc[popPerct["TZ2013"].isin([origin])]
    names = ["Low",'Low-Middle',"Middle-High","High"]
```

```
    inlist =[inrow["1st"].values[0], inrow["2nd"].values[0],
        inrow["3rd"].values[0], inrow["4th"].values[0]]
    plt.bar(x=names, height =inlist,color=['orange',
        'green','blue','black'])
    plt.xlabel("Various Income Groups")
    plt.ylabel("Frequency of Population")
    plt.title('Various Income Groups in TAZ '+origin)
    plt.show()

    names = ["Black African",'White','Coloured',"Indian or Asian","Other"]
    inlist =[poprow["Black_African_%"].values[0],
        poprow["White_%"].values[0],
        poprow["Indian_or_Asian_%"].values[0],
        poprow["Other_%"].values[0],poprow["Coloured_%"].values[0]]
    plt.bar(x=names, height =inlist,color=['orange',
        'green','blue','black','red'])
    plt.xlabel("Various Demographic Groups")
    plt.ylabel("Frequency Percentage(%)")
    plt.title('Various Demographic Groups in TAZ '+origin)
    plt.show()
```

```
//BOTH.py
# -*- coding: utf-8 -*-
"""
Created on Wed Jun 24 15:25:01 2020

@author: Brent Kotzee
"""
import pandas as pd
import numpy as np
import re
from pandas.core.common import flatten
import datetime
import matplotlib.pyplot as plt
from mpl_toolkits import mplot3d
from fucs import GETPMI,GetData, GetJobs, GetCum, GetPAI

def both(df,orGININ,destinationIN,timeIN,i,name):
    stop = 1

    while stop < 3:
        pmi = []
```

```
Accs = []
OriginToAll = []
if stop == 1:
    print(df)
#--------------------------------------------------------------------------
#Get input parameters to display
if stop == 1:
    """
    print("Please type in the TAZ Origin:\n")
    print( sorted( list( dict.fromkeys( df["Origin"].tolist() ) )
        ) )
    orGININ = input()

    print("Please type in the TAZ Destination:\n")
    print( sorted( list( dict.fromkeys( df["Dest"].tolist() ) ) ) )
    destinationIN = input()

    print("Please type in the TimeOfDay(hh:mm:ss):\n")
    print( sorted( list( dict.fromkeys( df["Time"].tolist() ) ) ) )
    timeIN = list(re.split('[,,\s]',(input("Enter a multiple
        value: "))))
    """

"""
else:
    OrginQ = input("new orgin Type Y or N ?")
    destinationQ = input("new destination Type Y or N ?" )

    if (OrginQ == "Y") or (destinationQ == "Y") :
                OriginToAll= []
                print("Please type in the TAZ Origin:\n")
                print( sorted( list( dict.fromkeys(
                    df["Origin"].tolist() ) ) ) )
                orGININ = input()

                print("Please type in the TAZ Destination:\n")
                print( sorted( list( dict.fromkeys(
                    df["Dest"].tolist() ) ) ) )
                destinationIN = input()
                pmi = []#can prolly remove
                Accs = []#can proly remove
                timeIN = []
```

```
        print("Please type in the TimeOfDay(hh:mm:ss):\n")
        print( sorted( list( dict.fromkeys( df["Time"].tolist() ) ) ) )
        print("Do you want the entire day?")
        ans = input()
        if ans == "Y":
            #this can be replaced by the time list
            timeIN =
                ['0:00:00','1:00:00','2:00:00','3:00:00','4:00:00','5:00:00','6:00:00',
            '7:00:00','8:00:00','9:00:00','10:00:00','11:00:00','12:00:00','13:00:00','
            '18:00:00','19:00:00','20:00:00','21:00:00','22:00:00','23:00:00']
        else:
            timeIN = timeIN + list(re.split('[,,\s]',(input("Enter a
                multiple value: "))))
            timeIN = list(dict.fromkeys(timeIN))
    """
#extract the legth and time_trav to get and get the PMI value
#calculate the PMI for each time of the day available, via the use of
    dictionaries

    timeIN = [timeIN[i]]
    for j in range(len(timeIN)):
        data= df.loc[(df["Name"].isin( [str(orGININ) + " - "+
            str(destinationIN)] )) & (df["Time"].isin([str(timeIN[j] )
            ] ) )]
        pmi.append(GETPMI(data["Shape_Leng"].values[0],
            data["Total_Trav"].values[0]))
#calculates the PMI form the orgin to all other destincation for the
    users given input time or times
    """
    print("Please type in the TimeOfDay you want for cum anal
        (hh:mm:ss):\n")
    print(sorted( list( dict.fromkeys( df["Time"].tolist() ) ) ) )
    """
    timeI = timeIN
    for i in range(len(timeI)):
         sub = df.loc[df["Origin"].isin([str(orGININ)])
             &(df["Time"].isin([timeI[i]]))]
         t = sub['Total_Trav'].sum()
         d = sub['Shape_Leng'].sum()
         OriginToAll.append(GETPMI(t, d))

    Accs = Accs + GetData(orGININ, timeIN, destinationIN,df,name)
```

```
print(len(pmi))
#the graaphing section
Hour=[]
for i in range(len(timeIN)):
    (h,m,s) = timeIN[i].split(':')
    Hour.append(int(h))
"""
#---------------------------------------------------------------------------
#now plot parameter with time,availablity and accessibilty as the
    axis
from mpl_toolkits import mplot3d
#import numpy as np
import matplotlib.pyplot as plt

fig = plt.figure()
ax = mplot3d.Axes3D(fig)
ax.set_xlim3d(0, max(pmi))
ax.set_ylim3d(0, max(Accs))
ax.set_zlim3d(0, max(Hour))
#ax.view_init(30, 360)
ax.set_xlabel('Potential Mobiltiy Index ')
ax.set_ylabel('Availabilty')
ax.set_zlabel('Time')


z_points = Hour
x_points = pmi
y_points = Accs
ax.scatter3D(x_points, y_points, z_points, c="black", cmap='hsv');
ax.plot3D(x_points, y_points, z_points, c="black");

plt.show()
"""
pai = GetPAI(orGININ, timeIN, destinationIN, df, name)
"""
fig = plt.figure()
ax = mplot3d.Axes3D(fig)
ax.set_xlim3d(0, max(pai))
ax.set_ylim3d(0, max(Accs))
ax.set_zlim3d(0, max(Hour))
#ax.view_init(30, 360)
ax.set_xlabel('Potential Mobiltiy Index ')
ax.set_ylabel('Availabilty')
```

```
        ax.set_zlabel('Time')


        z_points = Hour
        x_points = pai
        y_points = Accs
        ax.scatter3D(x_points, y_points, z_points, c="black", cmap='hsv');
        ax.plot3D(x_points, y_points, z_points, c="black");

        plt.show()

        print("DONE")
        """
        stop = 3
        return Hour,pai,Accs,pmi
```

```
// Hello.java
import javax.swing.JApplet;
import java.awt.Graphics;

public class Hello extends JApplet {
    public void paintComponent(Graphics g) {
        g.drawString("Hello, world!", 65, 95);
    }
}
```

```
//Hourspliter.py
# -*- coding: utf-8 -*-
"""
Created on Tue Jun 30 00:20:51 2020

@author: Brent Kotzee
"""
import pandas as pd

#dataJob = pd.read_csv('Jobs.csv', header ='infer')
print("Please type in the Transport(BUS;Metro;MyCiti;Taxi):\n")
Transport = input( )
print("Please type in the day(Tuesday;Wednesday;Thursday):\n")
Day = input( )
#date =input("Pleaseinput the date for the day(dd/mm/yyyy like so
    1/4/1900) \n")
```

```
df = pd.read_csv( Transport+"_"+Day+'.txt',sep=",")

OandD= df["Name"].str.split(" - ", n = 1, expand = True)
df["Origin"]=OandD[0]
df["Dest"]=OandD[1]

time = df["TimeOfDay"].str.split(" ", n = 1, expand = True)
df["Time"] =time[1]
time = list( dict.fromkeys( df["Time"].tolist() ))

df = df[["Origin","Dest","Time","Total_Trav","Shape_Leng","Total_Leng"]]


Hour=[]
for i in range(len(time)):
            (h,m,s) = time[i].split(':')
            Hour.append(int(h))


#for i in range(len(Hour)):
#    Newdf = df.loc[ df["Time"].isin([str(time[i] ) ] ) ]
##    Newdf.drop(columns=['OID', 'OID_','OriGINID', 'Destinatio',
    'Destinat_1', 'TimeOfDay'])
##    del Newdf["OID"]
##    del Newdf[ "OID_"]
##    del Newdf['OriGINID']
##    del Newdf['Destinatio']
##    del Newdf['Destinat_1']
##    del Newdf["TimeOfDay"]
#    Newdf.to_csv(Transport+"_"+Day+"_"+str(Hour[i])+".txt", header=True,
    index=False, sep=';', mode='a')
#

for i in range(len(Hour)):
    (df.loc[ df["Time"].isin([str(time[i] ) ] )
       ]).to_csv(Transport+"_"+Day+"_"+str(Hour[i])+".txt", header=True,
       index=False, sep=';', mode='a')
```

```
// JUSTPAI.py
# -*- coding: utf-8 -*-
"""
Created on Wed Jun 24 14:37:09 2020
```

```
@author: Brent Kotzee
"""

import pandas as pd
import numpy as np
import re
from pandas.core.common import flatten
import datetime
import matplotlib.pyplot as plt
from mpl_toolkits import mplot3d
from fucs import GETPMI,GetData, GetJobs, GetCum, GetPAI,GetAVA

def JustPai(df,orGININ,destinationIN,timeIN,i,name):
    stop = 1

    while stop < 3:
        pmi = []
        Accs = []
        OriginToAll = []
        AV=[]
        if stop == 1:
            print(df)
        #----------------------------------------------------------------------------
        #Get input parameters to display
        if stop == 1:
            """
            print("Please type in the TAZ Origin:\n")
            print( sorted( list( dict.fromkeys( df["Origin"].tolist() ) )
                ) )
            orGININ = input()

            print("Please type in the TAZ Destination:\n")
            print( sorted( list( dict.fromkeys( df["Dest"].tolist() ) ) ) )
            destinationIN = input()

            print("Please type in the TimeOfDay(hh:mm:ss):\n")
            print( sorted( list( dict.fromkeys( df["Time"].tolist() ) ) ) )
            timeIN = list(re.split('[,,\s]',(input("Enter a multiple
                value: "))))
            """

        """
```

```
else:
    OrginQ = input("new orgin Type Y or N ?")
    destinationQ = input("new destination Type Y or N ?" )

    if (OrginQ == "Y") or (destinationQ == "Y") :
                OriginToAll= []
                print("Please type in the TAZ Origin:\n")
                print( sorted( list( dict.fromkeys(
                    df["Origin"].tolist() ) ) ) )
                orGININ = input()

                print("Please type in the TAZ Destination:\n")
                print( sorted( list( dict.fromkeys(
                    df["Dest"].tolist() ) ) ) )
                destinationIN = input()
                pmi = []#can prolly remove
                Accs = []#can proly remove
                timeIN = []

    print("Please type in the TimeOfDay(hh:mm:ss):\n")
    print( sorted( list( dict.fromkeys( df["Time"].tolist() ) ) ) )
    print("Do you want the entire day?")
    ans = input()
    if ans == "Y":
        #this can be replaced by the time list
        timeIN =
            ['0:00:00','1:00:00','2:00:00','3:00:00','4:00:00','5:00:00','6:00:00',
        '7:00:00','8:00:00','9:00:00','10:00:00','11:00:00','12:00:00','13:00:00','
        ,'18:00:00','19:00:00','20:00:00','21:00:00','22:00:00','23:00:00']
    else:
        timeIN = timeIN + list(re.split('[,,\s]',(input("Enter a
            multiple value: "))))
        timeIN = list(dict.fromkeys(timeIN))
"""
Accs = Accs + GetData(orGININ, timeIN[i], destinationIN, df,name)
AV = AV + GetAVA(orGININ, timeIN[i], destinationIN, df)
pai = GetPAI(orGININ, timeIN[i], destinationIN, df, name)
Hour=[]
(h,m,s) = timeIN[i].split(':')
Hour.append(int(h))

"""
fig = plt.figure()
```

```
        ax = mplot3d.Axes3D(fig)
        ax.set_xlim3d(0, max(pai))
        ax.set_ylim3d(0, max(Accs))
        ax.set_zlim3d(0, max(Hour))
        #ax.view_init(30, 360)
        ax.set_xlabel('Potential Mobiltiy Index ')
        ax.set_ylabel('Availabilty')
        ax.set_zlabel('Time')


        z_points = Hour
        x_points = pai
        y_points = Accs
        ax.scatter3D(x_points, y_points, z_points, c="black", cmap='hsv');
        ax.plot3D(x_points, y_points, z_points, c="black");

        plt.show()
        """
        print("DONE")

        stop = 3
        return pai,Accs,Hour,AV
```

```
// JUSTPMI.py
# -*- coding: utf-8 -*-
"""
Created on Wed Jun 24 12:21:55 2020

@author: Brent Kotzee
"""
import pandas as pd
import numpy as np
import re
from pandas.core.common import flatten
import datetime
import matplotlib.pyplot as plt
from mpl_toolkits import mplot3d
from fucs import GETPMI,GetData, GetJobs, GetCum, GetPAI

def JustPMI(df,orGININ,destinationIN,timeIN,i,name):
    stop = 1
```

```
    while stop < 3:
        pmi = []
        Accs = []
        OriginToAll = []
        if stop == 1:
            print(df)
        timeIN=[timeIN[i]]
        for j in range(len(timeIN)):
#             data= df.loc[(df["Name"].isin( [str(orGININ) + " - "+
    str(destinationIN)] )) & (df["Time"].isin([str(timeIN[j] ) ] ) )]
            data= df.loc[(df["Origin"].isin( [str(orGININ)])) &
                (df["Dest"].isin( [str(destinationIN)] )) &
                (df["Time"].isin([str(timeIN[j] ) ] ) )]
            pmi.append(GETPMI(data["Shape_Leng"].values[0],
                data["Total_Trav"].values[0]))
        Accs = Accs + GetData(orGININ, timeIN, destinationIN, df,name)
        Hour=[]
        for i in range(len(timeIN)):
            (h,m,s) = timeIN[i].split(':')
            Hour.append(int(h))
        stop = 3
        return Hour,pmi,Accs
```

```
// MainSelective - Cumulative
# -*- coding: utf-8 -*-
"""
Created on Mon Jun 29 12:38:21 2020

@author: Brent Kotzee
"""
import conda
# import cudf as pd
# import cupy as np
import pandas as pd
import numpy as np
from pandas import DataFrame
#from PMI import PMI
#from JUSTPMI import JustPMI
#from JUSTPAI import JustPai
#from BOTH import both
from fucs import GetPMIcumC,GetACCcum
```

```
#now plot parameter with time,availablity and accessibilty as the axis

#import matplotlib.pyplot as plt

Hour = []
accC = []

a = []
t=[]

timeIN =
   "0:00:00,1:00:00,2:00:00,3:00:00,4:00:00,5:00:00,6:00:00,7:00:00,8:00:00,
9:00:00,10:00:00,11:00:00,12:00:00,13:00:00,14:00:00,15:00:00,16:00:00,
17:00:00,18:00:00,19:00:00,20:00:00,21:00:00,22:00:00,23:00:00"
Uier=[0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18,
   19, 20, 21, 22, 23]
#periodIN=""
orGININ=""

if ("," in timeIN):
       timeIN =timeIN.split(",")
#timeIN = input("The hours you want to anal:\n")
#if ("," in timeIN):
#    timeIN =timeIN.split(",")
#else:
#    timeIN =[timeIN]

#periodIN= input("The period you want to analyse in minutes:\n")
#periodIN= eval(periodIN)
#timeIN=eval(timeIN[0])


E=[]
Transport = input("Please type in the
   Transport(BUS;Metro;MyCiti;Taxi):\n" )

Day = input("Please type in the day(Tuesday;Wednesday;Thursday):\n" )


for i in range(len(Uier)):
    (h,m,s) = timeIN[i].split(':')
    Hour.append(int(h))
    opt =1
```

```
    df = pd.read_csv( Transport+"_"+Day+"_"+str(Uier[i])+'.txt',sep=";")
    if i ==0:


        print( sorted( list( dict.fromkeys( df["Origin"].tolist() ) ) ) )
        #TAZ=sorted( list( dict.fromkeys( df["Origin"].tolist() ) ) )
        orGININ = input("Please type in the TAZ Origin:\n")
        #Histo(orGININ)

    if opt == 1:
        accC1=GetACCcum(orGININ, timeIN[i], df)
        a1=GetPMIcumC(orGININ, timeIN[i], df)


        accC = np.append(accC,accC1)
        a = np.append(a,a1)
    # accC=accC+accC1

    # a=a+a1

for x in range(len(timeIN)):
    E = np.append(E,orGININ)
    # E.append(orGININ)


MPT = pd.DataFrame([E,a,accC,Hour]).transpose()
MPT.columns = ['TAZ','PMI','Accessibilty','Time']
#print (MPT)

#for i in range(len(Hour)):
MPT.to_csv("Cumulative_"+Transport+"_"+Day+"_"+str(E[0])+".txt",
    header=True, index=False, sep=';', mode='a')

#MPTsub= MPT[MPT["Travel_Time"] <= periodIN[0]]
#
#a= MPT[ "PMI" ].values[0]
#accC= MPT[ "Accessibilty" ].values[0]
#
#
#print ("PAI vs PMI\n")
#fig = plt.figure()
```

```
#
#fig.set_xlim(0, max(a))
#fig.set_ylim(0, max(accC))
#
##ax.view_init(30, 360)
#fig.set_xlabel('Potential Mobiltiy Index ')
#fig.set_ylabel('Accessibility')



#x_points = a
#y_points = accC
#fig.scatter(x_points, y_points, c="black", cmap='hsv');
#plt.xlabel("Leprechauns")
#plt.ylabel("Gold")
#plt.plot(x_points, y_points);
#
#plt.show()
#
#
#print(np.array([[round(num, 2) for num in a],[round(num,2) for num in
    accC]]).T)
#

print("DONE")
```

```
// MainSelective.py
# -*- coding: utf-8 -*-
"""
Created on Mon Jun 29 12:38:21 2020

@author: Brent Kotzee
"""

import pandas as pd
import numpy as np
from PMI import PMI
from JUSTPMI import JustPMI
from JUSTPAI import JustPai
from BOTH import both
from fucs import GETPMI,GetData, GetJobs, GetCum,
    GetPAI,GetAVA,Histo,GetACCcum
```

```
#will be used with the loops
outter= 1
#going to be the constant data outside of the loop segment
#now plot parameter with time,availablity and accessibilty as the axis
from mpl_toolkits import mplot3d
#import numpy as np
import matplotlib.pyplot as plt
"""
times = input("The hours you want to anal:\n")
if ("," in times):
    times =times.split(",")
print("Please type in the Transport(BUS;Metro;MyCiti;Taxi):\n")
Transport = input( )
print("Please type in the day(Tuesday;Wednesday;Thursday):\n")
Day = input( )

df = pd.DataFrame()
for i in range(len(times)):
    dfinter = pd.ExcelFile(
        Transport+"_"+Day+"_"+str(times[i])[0]+'.xlsx').parse( 'Sheet1' )
    if i==0:
        df = dfinter
    else:
        df.append(dfinter)
#dataJob = pd.read_csv('Jobs.csv', header ='infer')
#date =input("Pleaseinput the date for the day(dd/mm/yyyy like so
    1/4/1900) \n")
"""
"""
OandD= df["Name"].str.split(" - ", n = 1, expand = True)
df["Origin"]=OandD[0]
df["Dest"]=OandD[1]

time = df["TimeOfDay"].str.split(" ", n = 1, expand = True)
df["Time"] =time[1]
"""
Hour = []
OriginToAll = []
asscum = []
Accs = []
AV =[]
pai = []
pmi = []
```

```
AccC=[]
timeIN =
   "0:00:00,1:00:00,2:00:00,3:00:00,4:00:00,5:00:00,6:00:00,7:00:00,8:00:00,
9:00:00,10:00:00,11:00:00,12:00:00,13:00:00,14:00:00,15:00:00,16:00:00,
17:00:00,18:00:00,19:00:00,20:00:00,21:00:00,22:00:00,23:00:00"

orGININ=""
destinationIN=""
while outter < 3:
#    timeIN = input("The hours you want to anal:\n")
    if ("," in timeIN):
        timeIN =timeIN.split(",")
    else:
        timeIN =[timeIN]
    print("Please type in the Transport(BUS;Metro;MyCiti;Taxi):\n")
    Transport = input( )
    print("Please type in the day(Tuesday;Wednesday;Thursday):\n")
    Day = input( )
    opt = eval(input("You want to do a full anal(1), PMI(2), PAI(3) or
        PMI and PAI(No cum)(4)""\n"))

    for i in range(len(timeIN)):

        (h,m,s) = timeIN[i].split(':')
        Hour.append(int(h))

        df = pd.read_csv(
            Transport+"_"+Day+"_"+str(Hour[i])+'.txt',sep=";")
        if i ==0:
            print("Please select the type of Opportunity in the TAZ
                Destination:\n")
            GetJobs()
            name = input()
            print("Please type in the TAZ Origin:\n")
            print( sorted( list( dict.fromkeys( df["Origin"].tolist() ) )
                ) )
            orGININ = input()
            Histo(orGININ)

            print("Please type in the TAZ Destination:\n")
            print( sorted( list( dict.fromkeys( df["Dest"].tolist() ) ) ) )
            destinationIN = input()
            Histo(destinationIN)
```

```
        while sorted( list( dict.fromkeys( df["Dest"].tolist() ) )
            ).count(eval(destinationIN)) == 0 :
            print("TAZ Destination was not in the list:\n")
            print("Please type in the TAZ Destination:\n")
            destinationIN = input()

    if opt ==2:
        print("Please select the type of Opportunity in the TAZ
            Destination:\n")
        GetJobs()
        name = input()
        Histo(orGININ)
    if opt == 1:
        Hour1,OriginToAll1,asscum1,Accs1, pai1,
            pmi1,AV1,AccC1=PMI(df,orGININ,destinationIN,timeIN,i,name)
    elif opt ==2:
        Hour1,pmi1,Accs1 =
            JustPMI(df,orGININ,destinationIN,timeIN,i,name)
    elif opt==3:
         pai1,Accs1,Hour1
             =JustPai(df,orGININ,destinationIN,timeIN,i,name)
    elif opt ==4:
        Hour1,pai1,Accs1,pmi1=
            both(df,orGININ,destinationIN,timeIN,i,name)


    OriginToAll = OriginToAll+OriginToAll1
    asscum = asscum + asscum1
    Accs = Accs + Accs1
    AccC = AccC + AccC1
    AV= AV + AV1
    pai = pai + pai1
    pmi = pmi + pmi1

if opt == 1:
    print ("PAI vs PMI\n")
    fig = plt.figure()
    ax = mplot3d.Axes3D(fig)
    ax.set_xlim3d(0, max(pmi))
    ax.set_ylim3d(0, max(Accs))
    ax.set_zlim3d(0, max(Hour))
    #ax.view_init(30, 360)
    ax.set_xlabel('Potential Mobiltiy Index ')
```

```
ax.set_ylabel('Accessibility')
ax.set_zlabel('Time(h)')


z_points = Hour
x_points = pmi
y_points = Accs
ax.scatter3D(x_points, y_points, z_points, c="black", cmap='hsv');
ax.plot3D(x_points, y_points, z_points, c="black");

plt.show()

print(np.array([[round(num, 2) for num in pmi],[round(num) for num
    in pai],timeIN]).T)


print ("PAI vs Availability\n")
fig = plt.figure()
ax = mplot3d.Axes3D(fig)
ax.set_xlim3d(0, max(pai))
ax.set_ylim3d(0, max(AV))
ax.set_zlim3d(0, max(Hour))
#ax.view_init(30, 360)
ax.set_xlabel('Potential Accessibility Index ')
ax.set_ylabel('Availabilty(min)')
ax.set_zlabel('Time(h)')


z_points = Hour
x_points = pai
y_points = AV
ax.scatter3D(x_points, y_points, z_points, c="black", cmap='hsv');
ax.plot3D(x_points, y_points, z_points, c="black");

plt.show()

print(np.array([pai,AV,timeIN]).T)


print ("Cumulative PAI vs PMI\n")
fig = plt.figure()
ax = mplot3d.Axes3D(fig)
ax.set_xlim3d(0, max(OriginToAll))
```

```
    ax.set_ylim3d(0, max(AccC))
    ax.set_zlim3d(0, max(Hour))
    #ax.view_init(30, 360)
    ax.set_xlabel('Potential Mobiltiy Index')
    ax.set_ylabel('Accessibility Cumulitive')
    ax.set_zlabel('Time(h)')


    z_points = Hour
    x_points = OriginToAll
    y_points = AccC
    ax.scatter3D(x_points, y_points, z_points, c="black", cmap='hsv');
    ax.plot3D(x_points, y_points, z_points, c="black");

    plt.show()

    print(np.array([pmi,AccC,timeIN]).T)

    print("DONE")
elif opt ==2:
    print ("pmi\n")
    fig = plt.figure()
    ax = mplot3d.Axes3D(fig)
    ax.set_xlim3d(0, max(pmi))
    ax.set_ylim3d(0, max(Accs))
    ax.set_zlim3d(0, max(Hour))
    #ax.view_init(30, 360)
    ax.set_xlabel('Potential Mobiltiy Index ')
    ax.set_ylabel('Availabilty')
    ax.set_zlabel('Time(h)')


    z_points = Hour
    x_points = pmi
    y_points = Accs
    ax.scatter3D(x_points, y_points, z_points, c="black", cmap='hsv');
    ax.plot3D(x_points, y_points, z_points, c="black");

    plt.show()
elif opt==3:
    print ("pai\n")
    fig = plt.figure()
    ax = mplot3d.Axes3D(fig)
```

```
    ax.set_xlim3d(0, max(pai))
    ax.set_ylim3d(0, max(Accs))
    ax.set_zlim3d(0, max(Hour))
    #ax.view_init(30, 360)
    ax.set_xlabel('Potential Mobiltiy Index ')
    ax.set_ylabel('Availabilty')
    ax.set_zlabel('Time(h)')


    z_points = Hour
    x_points = pai
    y_points = Accs
    ax.scatter3D(x_points, y_points, z_points, c="black", cmap='hsv');
    ax.plot3D(x_points, y_points, z_points, c="black");

    plt.show()
elif opt ==4:
    print ("pmi\n")
    fig = plt.figure()
    ax = mplot3d.Axes3D(fig)
    ax.set_xlim3d(0, max(pmi))
    ax.set_ylim3d(0, max(AccC))
    ax.set_zlim3d(0, max(Hour))
    #ax.view_init(30, 360)
    ax.set_xlabel('Potential Mobiltiy Index ')
    ax.set_ylabel('Accessibility')
    ax.set_zlabel('Time(h)')


    z_points = Hour
    x_points = pmi
    y_points = Accs
    ax.scatter3D(x_points, y_points, z_points, c="black", cmap='hsv');
    ax.plot3D(x_points, y_points, z_points, c="black");

    plt.show()
    print ("pai\n")
    fig = plt.figure()
    ax = mplot3d.Axes3D(fig)
    ax.set_xlim3d(0, max(pai))
    ax.set_ylim3d(0, max(Accs))
    ax.set_zlim3d(0, max(Hour))
    #ax.view_init(30, 360)
```

```
        ax.set_xlabel('Potential Mobiltiy Index ')
        ax.set_ylabel('Availabilty')
        ax.set_zlabel('Time(h)')


        z_points = Hour
        x_points = pai
        y_points = Accs
        ax.scatter3D(x_points, y_points, z_points, c="black", cmap='hsv');
        ax.plot3D(x_points, y_points, z_points, c="black");

        plt.show()
        Histo(destinationIN)

    outter = eval(input("You want to redo (1) or you want to stop(3),
        different ervice(2)""\n"))
```

```
// PMI.py
# -*- coding: utf-8 -*-
"""
Created on Tue Jun 23 17:40:46 2020

@author: Brent Kotzee
"""
import pandas as pd
import numpy as np
import re
from pandas.core.common import flatten
import datetime
import matplotlib.pyplot as plt
from mpl_toolkits import mplot3d
from fucs import GETPMI,GetData, GetJobs, GetCum,
    GetPAI,GetAVA,GetaccC,GetACCcum,GetPMIcum
#takes the data thats availiable to print and takes the orgin and
    destination and time
#thse values can be set to defualt
#name from jobs input
def PMI(df,orGININ,destinationIN,timeIN,i,name):
    stop = 1
1
    while stop < 3:
        pmi = []
```

```
    Accs = []
    AccC = []
    OriginToAll = []
    AV = []


    data= df.loc[(df["Origin"].isin( [str(orGININ)])) &
        (df["Dest"].isin( [str(destinationIN)] )) &
        (df["Time"].isin([str(timeIN[i] ) ] ) )]
    print(data)
    pmi.append(GETPMI(data["Total_Trav"].values[0],data["Shape_Leng"].values[0]
        ))
#calculates the PMI form the orgin to all other destincation for the
    users given input time or times
    """
    print("Please type in the TimeOfDay you want for cum anal
        (hh:mm:ss):\n")
    print(sorted( list( dict.fromkeys( df["Time"].tolist() ) ) ) )
    """
    timeI = [timeIN[i]]
    timeIN=[timeIN[i]]
    if ("," in timeI):
        timeI =timeI.split(",")
    for i in range(len(timeI)):
         sub = df.loc[df["Origin"].isin([str(orGININ)])
             &(df["Time"].isin([timeI[i]]))]
         data= df.loc[(df["Origin"].isin( [str(orGININ)])) &
             (df["Dest"].isin( [str(destinationIN)] )) &
             (df["Time"].isin([str(timeI[i] ) ] ) )]
         t = sub['Total_Trav'].sum()
         d = sub['Shape_Leng'].sum()
         OriginToAll.append(GETPMI(t, d))
         AV.append((data[ "Total_Trav" ].values[0]))

    Accs = Accs + GetData(orGININ, timeIN, destinationIN,df,name)
    AccC = AccC + GetaccC(orGININ, timeIN, destinationIN,df,name)

    #the graphing section
    Hour=[]
    for i in range(len(timeIN)):
        (h,m,s) = timeIN[i].split(':')
        Hour.append(int(h))
    #--------------------------------------------------------------------------
```

```
#now plot parameter with time,availablity and accessibilty as the
    axis
#from mpl_toolkits import mplot3d
#import numpy as np
# import matplotlib.pyplot as plt
"""
fig = plt.figure()
ax = mplot3d.Axes3D(fig)
ax.set_xlim3d(0, max(pmi))
ax.set_ylim3d(0, max(Accs))
ax.set_zlim3d(0, max(Hour))
#ax.view_init(30, 360)
ax.set_xlabel('Potential Mobiltiy Index ')
ax.set_ylabel('Availabilty')
ax.set_zlabel('Time')


z_points = Hour
x_points = pmi
y_points = Accs
ax.scatter3D(x_points, y_points, z_points, c="black", cmap='hsv');
ax.plot3D(x_points, y_points, z_points, c="black");

plt.show()
"""
pai = GetPAI(orGININ, timeIN, destinationIN, df, name)
"""
fig = plt.figure()
ax = mplot3d.Axes3D(fig)
ax.set_xlim3d(0, max(pai))
ax.set_ylim3d(0, max(Accs))
ax.set_zlim3d(0, max(Hour))
#ax.view_init(30, 360)
ax.set_xlabel('Potential Mobiltiy Index ')
ax.set_ylabel('Availabilty')
ax.set_zlabel('Time')


z_points = Hour
x_points = pai
y_points = Accs
ax.scatter3D(x_points, y_points, z_points, c="black", cmap='hsv');
ax.plot3D(x_points, y_points, z_points, c="black");
```

```
plt.show()
"""
asscum = GetCum(timeI, destinationIN, df, name, orGININ)
Hour=[]
for i in range(len(timeI)):
    (h,m,s) = timeI[i].split(':')
    Hour.append(int(h))

"""
#import numpy as np
fig = plt.figure()
ax = mplot3d.Axes3D(fig)
ax.set_xlim3d(0, max(OriginToAll))
ax.set_ylim3d(0, max(asscum))
ax.set_zlim3d(0, max(Hour))
#ax.view_init(30, 360)
ax.set_xlabel('Potential Mobiltiy Index (all directions)')
ax.set_ylabel('Availabilty cum')
ax.set_zlabel('Time')


z_points = Hour
x_points = OriginToAll
y_points = asscum
ax.scatter3D(x_points, y_points, z_points, c="black", cmap='hsv');
ax.plot3D(x_points, y_points, z_points, c="black");

plt.show()

print("DONE")
"""

print(Hour)
return Hour,OriginToAll,asscum,Accs,pai,pmi,AV,AccC
stop = 3

if stop == 1:
    print(df)
#---------------------------------------------------------------------------
#Get input parameters to display
#if stop == 1:
    """
```

```
    print("Please type in the TAZ Origin:\n")
    print( sorted( list( dict.fromkeys( df["Origin"].tolist() ) )
       ) )
    orGININ = input()

    print("Please type in the TAZ Destination:\n")
    print( sorted( list( dict.fromkeys( df["Dest"].tolist() ) ) ) )
    destinationIN = input()

    print("Please type in the TimeOfDay(hh:mm:ss):\n")
    print( sorted( list( dict.fromkeys( df["Time"].tolist() ) ) ) )
    timeIN = list(re.split('[,,\s]',(input("Enter a multiple
       value: "))))
    """

"""
else:
    OrginQ = input("new orgin Type Y or N ?")
    destinationQ = input("new destination Type Y or N ?" )

    if (OrginQ == "Y") or (destinationQ == "Y") :
            OriginToAll= []
            print("Please type in the TAZ Origin:\n")
            print( sorted( list( dict.fromkeys(
               df["Origin"].tolist() ) ) ) )
            orGININ = input()

            print("Please type in the TAZ Destination:\n")
            print( sorted( list( dict.fromkeys(
               df["Dest"].tolist() ) ) ) )
            destinationIN = input()
            pmi = []#can prolly remove
            Accs = []#can proly remove
            timeIN = []

    print("Please type in the TimeOfDay(hh:mm:ss):\n")
    print( sorted( list( dict.fromkeys( df["Time"].tolist() ) ) ) )
    print("Do you want the entire day?")
    ans = input()
    if ans == "Y":
        #this can be replaced by the time list
        timeIN =
           ['0:00:00','1:00:00','2:00:00','3:00:00','4:00:00','5:00:00','6:00:00',
```

```
            '7:00:00','8:00:00','9:00:00','10:00:00','11:00:00','12:00:00','13:00:00','
            '17:00:00','18:00:00','19:00:00','20:00:00','21:00:00','22:00:00','23:00:00
        else:
            timeIN = timeIN + list(re.split('[,,\s]',(input("Enter a
                multiple value: "))))
            timeIN = list(dict.fromkeys(timeIN))
        """
    #extract the legth and time_trav to get and get the PMI value
    #calculate the PMI for each time of the day available, via the use of
        dictionaries
```

```
//Population Plotter.py
import pandas as pd
#import numpy as np
#from pandas import DataFrame
#from PMI import PMI
#from JUSTPMI import JustPMI
#from JUSTPAI import JustPai
#from BOTH import both
#from fucs import GetPMIcum,GetACCcum

#now plot parameter with time,availablity and accessibilty as the axis


import matplotlib.pyplot as plt
plt.style.use('seaborn-whitegrid')

Hour = []
TAZ=[]
t=[]

#timeIN =
    [0:00:00,1:00:00,2:00:00,3:00:00,4:00:00,5:00:00,6:00:00,7:00:00,8:00:00,
9:00:00,10:00:00,11:00:00,12:00:00,13:00:00,14:00:00,15:00:00,16:00:00,
17:00:00,18:00:00,19:00:00,20:00:00,21:00:00,22:00:00,23:00:00]
timeIN=""
periodIN=""
orGININ=""


timeIN = input("The hours you want to anal:\n")
```

```
periodIN= input("The period you want to analyse in minutes:\n")

Transport = input("Please type in the
    Transport(BUS;Metro;MyCiti;Taxi):\n" )

Day = input("Please type in the day(Tuesday;Wednesday;Thursday):\n" )


(h,m,s) = timeIN.split(':')
Hour.append(int(h))

orGININ = input("Please type in the TAZ Origin:\n")
df =
    pd.read_csv("Cumulative_"+orGININ+"_"+Transport+"_"+Day+"_"+str(Hour[0])+'.txt',sep=";


MPTsub= df.loc[df["Travel_Time"] <= float(periodIN)]
accC = []
a = []
for i in range(len(MPTsub)):
    o=MPTsub[ "PMI" ].values[i]
    p=MPTsub[ "Accessibilty" ].values[i]
    a.append(o)
    accC.append(p)



plt.xlabel("Potential Mobility Index (aerial speed in km/h)")
plt.ylabel("Accessibilty (jobs within "+periodIN+" minutes travel time)")
plt.title('Potential Mobility Index vs Accessibilty at time '+timeIN+" in
    TAZ "+orGININ)
plt.plot(a, accC, '.', color='black');

# draw vertical line
plt.plot([0.5*(max(a)) , 0.5*(max(a))], [0, max(accC)], 'k-',
    label='green')


# draw diagonal line from (70, 90) to (90, 200)
plt.plot([0 , max(a)], [0.5*(max(accC)),0.5*(max(accC)) ], 'k-',
    label='green')
plt.plot([0 , max(a)], [0.5*0.5*(max(accC)),0.5*0.5*(max(accC)) ], 'k-',
```

```
    label='green')
plt.plot([0 , max(a)], [0.2*0.5*(max(accC)),0.2*0.5*(max(accC)) ], 'k-',
    label='green')

#Lables
#Vertical
plt.annotate(xy=[0.5*0.5*max(a), max(accC)+500], s="Average Potential
    Mobility Index Line")
#Horizontial
plt.annotate(xy=[max(a), 0.5*max(accC)], s=" Average Accessibilty Line")
plt.annotate(xy=[max(a), 0.5*0.5*max(accC)], s=" 50% threshold")
plt.annotate(xy=[max(a), 0.2*0.5*max(accC)], s=" 20% threshold")
percent = input("Which threshold percentage value do you want to view
    (0-100)? :\n")

ProbArea=MPTsub.loc[(df["PMI"]<= (0.5)*(max(a))) & (df["Accessibilty"]<=
    ((int(percent))/100)*(max(accC))) ]
print(ProbArea)

chioce = input("Do you want to export this data (1 for Yes/ 0 for No)?
    :\n")
if chioce == 1:
    ProbArea.to_csv("PMI vs Accessibility
        "+Transport+"_"+Day+"_"+str(Hour)+".txt", header=True,
        index=False, sep=';', mode='a')
else:
    print("Done")
```