# STUDY ON THE EFFECTIVENESS OF MODE CHOICE BETWEEN PRIVATE VEHICLES AND PUBLIC BUS: CASE STUDY ON THIRUVANANTHAPURAM CITY CIRCULAR BUSES

PROJECT REPORT

submitted by

**RABIA PARVEEN R**

**TKM21CETE14**

to

the APJ Abdul Kalam Technological University

in partial fulfillment of the requirements for the award of the Degree

of

Master of Technology

in

*Transportation Engineering*


**DEPARTMENT OF CIVIL ENGINEERING**

TKM College of Engineering, Kollam

MAY 2023

# DECLARATION

I, Rabia Parveen R hereby declare that, this project report entitled 'Study on The Effectiveness of Mode Choice Between Private Vehicles and Public Bus: Case Study on Thiruvananthapuram City Circular Buses' is bonafide work of mine carried out under the supervision of Prof. Karthik S, Assistant Professor, Department of Civil Engineering. I declare that, to the best of my knowledge, the work reported herein does not form part of any other project report or dissertation on the basis of which a degree or award was conferred on an earlier occasion to any other candidates. The content of this report is not being presented by any other student to this or any other university for the award of the degree.

Kollam RABIA PARVEEN R

10-05-2023

## DEPARTMENT OF CIVIL ENGINEERING

## T.K.M. COLLEGE OF ENGINEERING, KOLLAM

# CERTIFICATE

Certified that this report entitled **'STUDY ON THE EFFECTIVENESS OF MODE CHOICE BETWEEN PRIVATE VEHICLES AND PUBLIC BUS: CASE STUDY ON THIRUVANANTHAPURAM CITY CIRCULAR BUSES'** is the report of the project presented by **RABIA PARVEEN R, Reg. No. TKM21CETE14** during **2022-2023** in partial fulfilment of the requirements for the award of the Degree of Master of Technology in Transportation Engineering of the A P J Abdul Kalam Technological University.

| Guide | Project Coordinator | Head of the Department |
|---|---|---|
| **Prof. Karthik S** | **Dr. Adarsh S** | **Dr. Sajeeb R** |
| Assistant Professor | Professor | Professor |
| Dept. of. Civil Engineering | Dept. of. Civil Engineering | Dept. of. Civil Engineering |
| TKMCE Kollam | TKMCE Kollam | TKMCE Kollam |

# ACKNOWLEDGEMENT

I take this opportunity to express my deep sense of gratitude and sincere thanks to all who helped me to complete the project successfully.

First of all, I thank the almighty for giving me the strength and courage to do the project. I sincerely thank my respected principal **Dr T A Shahul Hameed** for his support towards the successful accomplishment of this course.

I am deeply indebted to my guide, **Prof**. **Karthik S**, Assistant Professor, Department of Civil Engineering, TKM College of Engineering for his excellent guidance, positive criticism and valuable comments.

I am grateful to **Prof. Anandu V G**, Assistant Professor, Department of Civil Engineering, TKM College of Engineering and **Prof. Karthika M G**, Assistant Professor, Department of Civil Engineering, TKM College of Engineering for serving as my evaluators for the project.

I am greatly thankful to my project coordinator, **Dr. Adarsh S**, Professor, Department of Civil Engineering, TKM College of Engineering for his constant supervision as well as for providing necessary information regarding the project.

I am greatly thankful to **Dr. Sajeeb R**, Professor and Head of the Department of Civil Engineering, TKM College of Engineering.

Finally, I thank my parents and friends who directly and indirectly contributed to the successful completion of my project.

**RABIA PARVEEN R**

# ABSTRACT

The private transportation system constitutes the major share of the world. As the population increases the demand for vehicles increases day by day. This makes the existing road inadequate, resulting in increased vehicle congestion and emission, ultimately leading to air pollution, especially in cities. In India, two-wheeler constitutes the major transportation mode among private vehicle. In order to reduce emissions and to provide sustainable transport movement, public transport systems are provided. Mode shift to public transport particularly to buses or walking inside or within the city could reduce the traffic congestion to a level. So, to increase the mode shift or make the public choose public transport, quality attributes related to public transport must be improved. The service attributes must be such that it should satisfy the needs of the customers who commute i.e., customer satisfaction is of utmost importance. Different types of public transport are available in Kerala, among that KSRTC launched City Circulars in Thiruvananthapuram city, which is limited within the city. The present study was focused on improvement required in the service attribute of the City Circular bus, and also to predict mode choice with individual characteristics. The study also focused on finding measures to promote these city circular buses so that further mode shift is encouraged which could help in reducing traffic congestion in the city. To classify different attributes, factor analysis was performed in the study. The study used the Ordinal logit model to obtain a relation between service attributes and user characteristics, along with that to identify the user's mode choice behaviour Multinomial logit model was also used. Findings revealed that service frequency, information provision, waiting time, and interchange are the important attributes to be improved. An increase in age and trip distance causes commuters to use the city circle bus. Occupation and gender also influence the user's mode choice.

***Keywords:*** City circular, Mode choice, Service attributes, Ticket price

# CONTENTS

| S.No. | List of Tables | Pg. No. |
|---|---|---|

| S.No. | List of Figures | Pg. No. |
|---|---|---|

# ABBREVIATIONS

| | |
|---|---|
| KSRTC | Kerala State Road and Transport Corporation |
| CC | City Circular |
| HOHO | Hop On Hop Off |
| RFID | Radio Frequency Identification |
| SPSS | Statistical Package for Social Science |
| MNL | Multinomial Logit Model |
| OLM | Ordinal Logit Model |
| EFA | Explanatory Factor Analysis |
| KMO | Kaiser-Meyer-Olkin |
| OR | Odds Ratio / Exp(B) |
| G | Gender |
| A | Age |
| O | Occupation |
| MI | Monthly Income |
| HH | Household size |
| VO | Vehicle Ownership |
| PM | Preferred Mode |
| TF | Trip Frequency |
| TP | Trip Purpose |
| TD | Trip Distance |
| BF | Bus Frequency |
| df | Degree of Freedom |

# CHAPTER 1

# INTRODUCTION

## 1.1 GENERAL

Private vehicles are considered the most preferred and used transportation system in the world. As the population increases, the demand for vehicles increases, as a result existing roads will become inadequate which results in increased vehicle congestion and vehicular emission, especially in cities. India ranks $6^{th}$ position in the countries with the greatest number of cars. Not only cars but also two-wheelers constitute the major transport mode. An increase in traffic can also increase traffic-related problems like congestion, accidents, etc. These are likely to be seen in the cities rather than in the countryside, as mentioned, one of the largest contributors to air pollution is traffic congestion. There are some ways to minimize these effects but prevention is quite difficult. Promoting the use of public transport is considered a sustainable way to minimize the congestion and pollution in cities where they are quite high. People choose a particular mode based on many factors. Some factors of private vehicles are advantageous when compared with public transport. So, to increase the mode choice of public transport, some factors must be improved and considered.

In Thiruvananthapuram, the capital city of Kerala, the Kerala State Road and Transport Corporation (KSRTC) launched City Circular Bus services. These runs along the city hoping to bring private vehicle users towards public transport. As these buses cover cities' important areas like schools, colleges, offices, etc. the policymakers in the city planned the service in such a way that there will be a mode shift towards these buses. By doing so a large amount of traffic congestion during peak times could be reduced to a minimum.

## 1.2 CITY CIRCULAR BUS – CC BUS

The city circular services were launched in Thiruvananthapuram city on November 2021 by the KSRTC. The city circulars shuttle and moves around the city. It is the first public transportation system in Kerala to have Hop-On Hop-Off (HOHO) model. The HOHO model is well-favoured among tourists around the world. But for city circulars, this is mainly meant for routine commuters.

One of the unique features of these buses is the unique colour provided, through which travel routes could be identified by the users. The seven distinct routes are represented by different colours mainly blue, green, yellow, red, magenta, brown, orange, and violet. The city circle

connects almost all the major routes inside or within the city. The seven distinct routes interlink and connects all the major points in the city, which connect government offices, hospitals, tourist spots, and other major areas. Another major feature of these city circulars is that it moves along the route where no previous transport route was present, it also moves along the existing transport route. The service is quite suitable for the general public who visit schools, shops, offices, hospitals, etc. within the city frequently or for daily commuters. Sometimes they also connect major tourist spots as well as international airport. Figure 1.1 shows the seven city circulars.

Figure 1.1 City circulars
(Source: www.citycircular.keralartc.com)

There are about 90 city circular buses covering 278 bus stops. The service frequency is 15 minutes during peak time and 30 min during non-peak time. The ticket price is affordable. The ticket cost is Rupees 10 for each trip you make. No matter what the trip distance is, the ticket cost only Rupees 10. Good day tickets and travel cards with Radio Frequency Identification (RFID) technology are also available. The interchange points present are major stops where two or more circular services intersect or meet. Passengers can switch to different routes along these points. Currently, there are 24 interchange points. The major interchange points are Eastfort, Palayam, and Peroorkada. Initially, diesel-powered city circulars were available but now e-buses (figure 1.2) as city circulars are also available. The electric buses are compact and produce no pollution. The travel is quite comfortable and silent.

The service moves in clockwise and anticlockwise direction along the same route, which is represented on the city circle buses. The circles starting from Eastfort are air rail, Red, Blue,

Brown, Orange, and green, while circles starting from Peroorkada are Magenta, Yellow, and violet.

Figure 1.2 Electric city circular buses

The route map of the circulars is provided in google maps along with timings. Figure 1.3 shows the different city circular routes across the city.

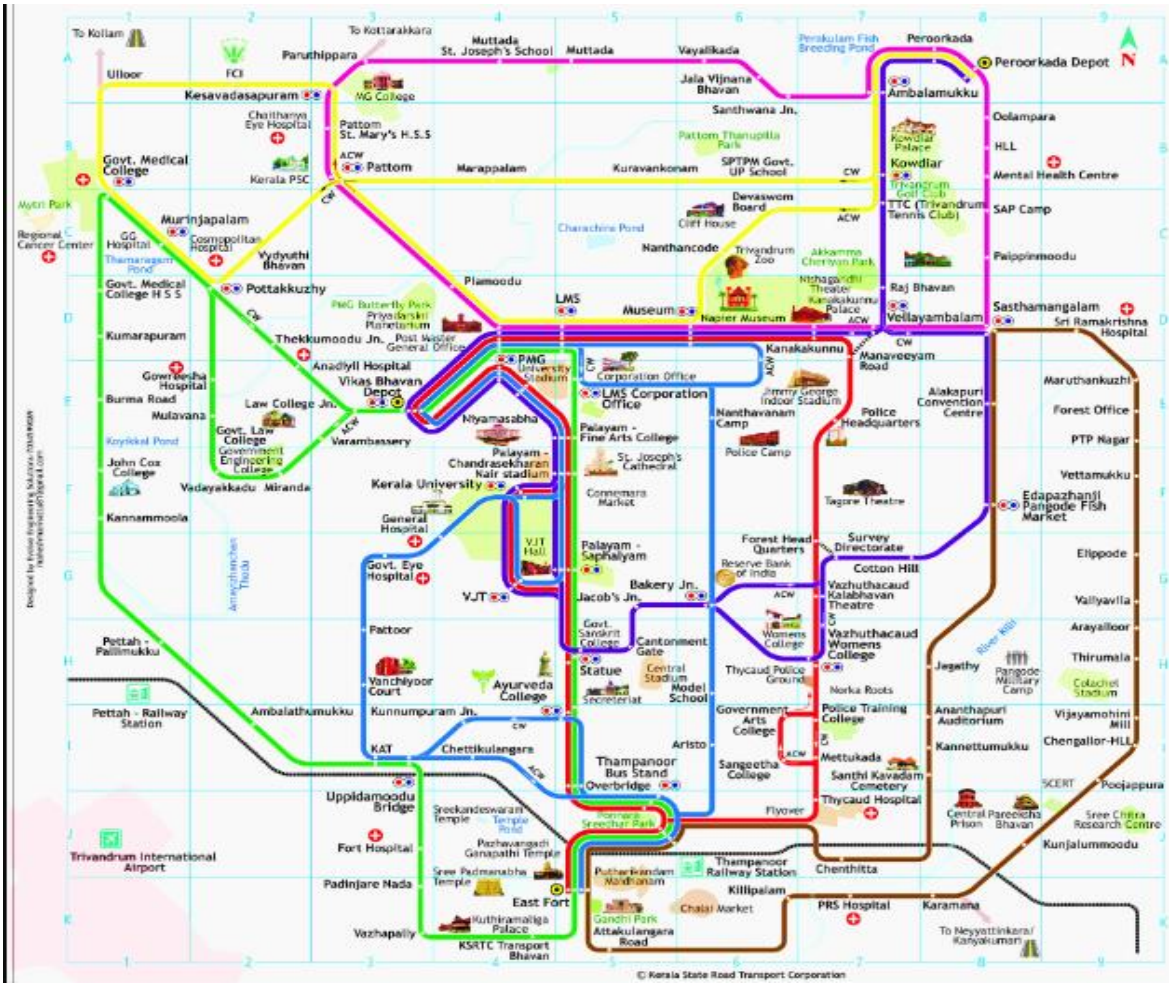



Figure 1.3 City Circular route map
(Source: www.team-bhp.com)

## 1.3 GAP IDENTIFIED

As the city circular services are new in Kerala, their effectiveness and utility need to be thoroughly understood so that these services could be further provided or introduced to other districts in Kerala. So, post-studies related to the overall performance of the CC bus need to be done.

- To identify how satisfied the service attributes are with respect to socio-demographic characteristics.
- To study the mode choice between private vehicles and CC buses.

## 1.4 OBJECTIVES OF THE STUDY

Since its launch in 2021, still many are not aware of the existence of these buses. The city gets congested as traffic increases especially the private vehicles. So, the presence of these buses in cities could reduce traffic congestion if a potential mode shift to these buses is utilised. Also, the existing users of these city circulars may often find information provision and transfers a factor for not using these buses. People not utilising the services provided by the government even if one or two services have demerits is a point to be noted. Not all the services are having 100% effectiveness, some have minor defects, so these defects should be identified and further improvements must be taken.

The objectives of the study are as follows:

- To conduct a pilot study to identify the perception attributes with regard to the CC bus.
- To identify how the satisfaction level of service attribute is influenced by the socio-demographic and trip characteristic of the private vehicle users.
- To develop a model using SPSS software to identify the impact of socio-demographic and trip characteristics on mode choice and also to predict the mode choice.
- To provide measures and identify areas of improvement.

## 1.5 SCOPE

The scope of the study is as follows:

- The study is limited to the perception of commuters in Thiruvananthapuram city i.e., inside the city.
- Includes analysis of user perception data towards city circulars in satisfaction level.
- Recommendations and improvements are based on survey and analysis.

## 1.6 ORGANIZATION OF REPORT

This report is organized into five chapters as follows:

- Chapter 1 Introduction: This section presents the overview of the research work which includes the objective, scope, gap identified etc.
- Chapter 2 Literature Review: This section presents the review of relevant literature related to the study concerning main areas like public transport service attributes of public transport and mode choice analysis.
- Chapter 3 Methodology: This chapter provides the methodology followed in the study, which includes pilot study, data collection, analysis techniques, etc.
- Chapter 4 Results and Discussions: This chapter mainly focuses on data analysis and its discussions.
- Chapter 5 Conclusions: This chapter presents the conclusion obtained from the study, especially from data analysis. This chapter also contains recommendations and future scope related to the study.
- References

# CHAPTER 2

# LITERATURE REVIEW

## 2.1 GENERAL

The following section consists of review of the literature studied for the project. It includes different service attributes of public transport that users prefer, the effect of socio-demographic and trip characteristics on choosing a mode, different analysis techniques adopted to identify mode choice and users' perception, etc.

## 2.2 SERVICE ATTRIBUTES.

There are different service attributes and its level for different modes of transport, each varying differently. Some of the attributes could be classified or grouped as availability, reliability, safety, and comfort. The importance of attributes is seen differently in different journals. Some journals focus on safety while others focus on reliability and comfort. Table 2.1 shows different service attributes. Service attributes are generally measured or checked for public mode or public transport rather than private mode. As the satisfaction level of each attribute increases, the chance of choosing public transport also increases. The publics choice of using public transport depends on the bus service attributes. Better the attribute. Better the chance of using public transport.

Table 2.1 Bus service attributes (Hu et al., 2015)

| **Bus service attributes** | | |
|---|---|---|
| Route layout | Transfer of waiting time | Driver behaviour |
| Easy to transfer and connect | Service frequency | Conditions at stops |
| Easy to access bus stop | Timetable information | Facility cleanliness |
| Ticket price | Safety and security while riding | Bus cleanliness |
| Ticket selling network | Stationarity of getting on/off of the bus | Odour and temperature on bus |
| Travel speed | Smoothness of riding | Free space on bus |
| Schedule delay | Safety facilities | |

Hu et al. **(**2015) in their study found that the reliability and comfort of the bus service have more impact on passenger's mode choice for bus when compared to availability and safety, while Schubert et al. **(**2020) in their study, found that bus cost was the important service attribute for choosing a bus over other private vehicles i.e., lower the ticket price and travel

cost, higher the chance of choosing public transport. For certain private vehicle users, frequency, speed, and intermodality were the important service attributes of public transport when compared to accessibility and individual space (Ona et al**.,** 2020**)**. Also, the frequency was considered as one of the important service attributes of public transport. It was also observed that other than service reliability and frequency, the important attribute that attract car users towards public transport are affective and connected individual perception, motivation, and context **(**Redman et al., 2013). The service attributes could be further classified as physical and perceived attributes. The perceived attributes must be observed in a direct manner when compared to physical attributes. Quality attributes that are generally prioritised include frequency, comfort, reliability, travel time, and network coverage, of which network coverage has an important effect when choosing public transport (Hansson et al., 2019).

Table 2.2 Definition of public transport service quality attributes (Redman et al., 2013)

| | **Attributes** | **Definition** |
|---|---|---|
| Physical | Reliability | How closely the actual service matches the routine timetable |
| | Frequency | How often the service operates during given period |
| | Speed | The time spent travel between specified points |
| | Accessibility | The degree to which public transport is reasonably available to as many people as possible |
| | Price | The monetary cost of travel |
| | Information provision | How much information is provided about route and interchanges |
| | Ease of transfer/interchange | How simple transport connections are including time spend waiting. |
| | Vehicle condition | The physical and mechanical condition of the vehicles. |
| Perceived | Comfort | How comfortable the journey is regarding the access to seat, noise level, driver handling, air conditioning etc. |
| | Safety | How safe from traffic accidents passengers feels during journey as well as personal safety. |
| | Convenience | How simple is the services is in use and how well it adds to one's ease of movement |
| | Aesthetic | Appeal of vehicle, station and waiting area to user's sense. |

## 2.3 MODE CHOICE BASED ON SOCIO-DEMOGRAPHIC AND TRIP CHARACTERISTICS

Mode choice or choosing a particular mode for a trip depends on various factors. Among them include socio-demographic factors and trip characteristics.

The socio-demographic factors like gender, monthly income, occupation, etc. also played a crucial role in choosing a particular mode. Ashalatha et al. **(**2013) after conducting a study on the mode choice behaviour of commuters of Thiruvananthapuram city, found that as age increases choosing or preferring car increases while the preference for two-wheelers decreases in comparison to public transport like bus. Also, as time and cost per distance increase, people tend to choose private vehicles over public vehicles. From gender-wise comparison of choosing a mode males preferred cars and bikes over bus while females preferred buses over cars and bikes. They also found that two-wheeler users preferred to travel by bus when compared to car users. Based on a qualitative study conducted in Porto, Cabral and Beirão (2007) found out that individual characteristics and lifestyle along with the type of journey and service performance influences mode choice. Other than that, the situational variables also affect the choice of a particular mode.

Paulley et al. (2006) found out that income and car ownership plays an important role in mode choice and often act as a background factor when compared to service attributes of public transport. Other than the normal socio-demographic factors, mode choice also depends upon the location of the users **(**Sun and Wandelt., 2021). Also, mode choice depends on the population density and location of transit vehicles (Sun and Wandelt., 2021). From the study conducted by Murugan and Marisamynathan (2022), on mode shift towards electric two-wheelers, it was observed that trip characteristics like existing travel mode, distance travelled in a day, fuel expenditure, and characteristics of electric two-wheelers play an important role in mode shift when compared to socio-demographic factors which have a minimal effect, especially gender, age and occupation. Other important factors that affect the user's travel habits and influence the mode choice were local land use and infrastructure (Redman et al., 2013).

## 2.4 ANALYSIS TECHNIQUES AND MODELS

Earlier the studies related to mode choice and mode shift were analysed qualitatively. For that purpose, software like NVivo 2.0 was used **(**Cabral and Beirão 2007**).** This type of analysis

was done by interviewing the respondents for about 30 minutes and mainly focused on mode choice for regular journeys and the influencing factors in choosing that mode.
Regression models were generally used for the mode choice analysis. Ordinal logit models, binary logit models, and Multinomial Logit (MNL) models were some common models used for analysis. MNL model was normally used for mode choice analysis as they were simple, easy to estimate & interpretation, and add or remove choices. Ashalatha et al. (2013) used MNL regression model and analysed the relationship between age, income, gender, time per distance, and vehicle ownership to mode preference between bus & car and bus & two-wheeler. MNL modelling could also be used to identify the impact of perception factors on the mode choice preference **(**Hu et al., 2015**).**

Logit models could be used for both revealed and stated preference data for mode choice analysis (Schubert et al., 2020). Ordered logit models could be used in situations where the dependent variable was ordinal. These models could be useful in explaining users' perceptions concerning service quality attributes (Ona et al., 2020). Using multinomial logit model analysis Schubert et al. (2020) found that the students would be interested in switching from their motor vehicles to other options especially buses i.e., public transport. Bus cost was the most important factor for mode shift. Yuan et al. (2020) used binary logit models to investigate the relationship between socio-demographic factors, trip factors, and other motivational factors to the mode shift. MNL was found quite useful in determining the mode choice of the users and are used in almost all the study conducted. Some software where MNL could be performed are STATA, SPSS, etc. The MNL was considered one of the prominent models.

For travel mode choice analysis, MNL was considered an important model. It is also a type of logistic model. It follows certain assumptions only then the analysis is to be made. They are generally used when we are expected to have more than 2 outcomes i.e. the model is used to predict the probabilities of the different possible outcomes of a categorically distributed dependent variable, with respect to a given set of independent variables. One of the assumptions is that there must be no correlation between independent variables. The analysis is interpreted based on the regression coefficient and its estimates, the odds ratio. The structural equation model (SEM) was also seen as an important analysis method to identify the impact of socio-demographic and trip characteristics on service attributes. The SEM is accompanied by EFA & CFA (explanatory factor analysis and confirmatory factor analysis).

Hu et al. (2015) used SEM to identify how mode choice preference was affected by the impact of bus service perceptions. He used EFA & CFA to extract important main factors from the bus service attributes for further study. By using EFA, service attributes of the transportation mode could be factored or classified into different factors i.e., EFA helps to extract factors among different attributes and groups them. While CFA gives structural stability among the grouped factors. From figure 2.1, we could see that 20 service attributes are grouped under 4 factors after performing EFA by Hu et al. (2015).

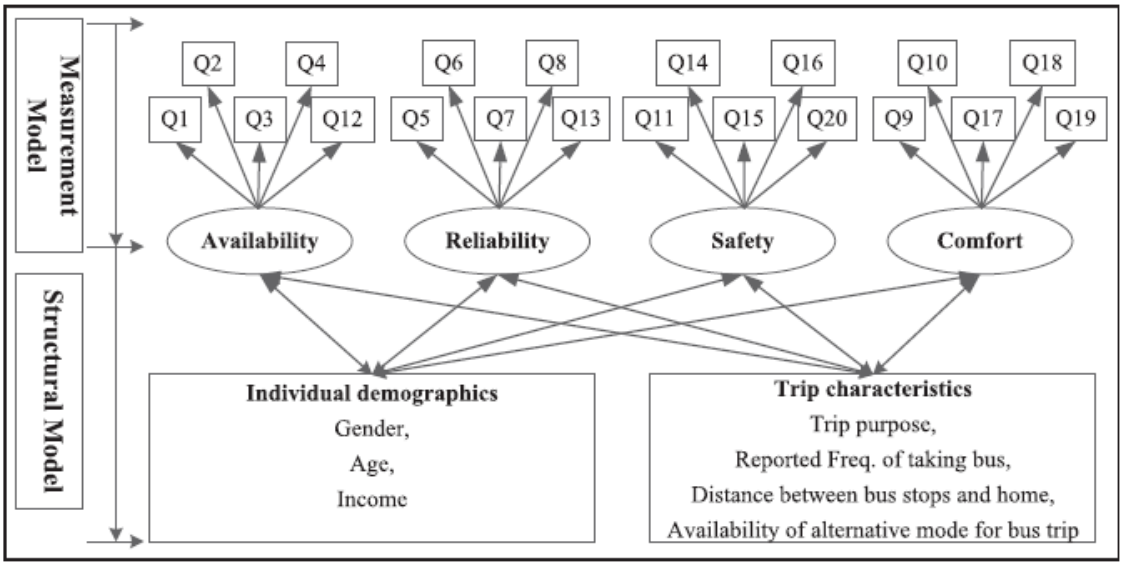



Figure 2.1 Structural Equation Model (Hu et al., 2015)

## 2.5 SUMMARY

From the above literature review, it was observed that to increase the mode shift or to promote mode choice to public transport, especially towards buses, the service attributes must satisfy commuters' requirements. The frequency, reliability, and low transfers were the main attribute. These physical attributes have the utmost importance regarding mode preference/ choice. Other physical attributes like speed, cost per trip, and time per trip are also factors responsible for mode choice. The perceived attribute has less importance when compared to physical attributes, out of which comfort and safety were important. Also, for mode choice socio-demographic and trip characteristic plays an important role when compared to service attribute. MNL is considered one of the best regression models to analyse and predict the mode choice

# CHAPTER 3

# METHODOLOGY

## 3.1 GENERAL

The methodology adopted in the study was staged in five as shown in figure 3.1.

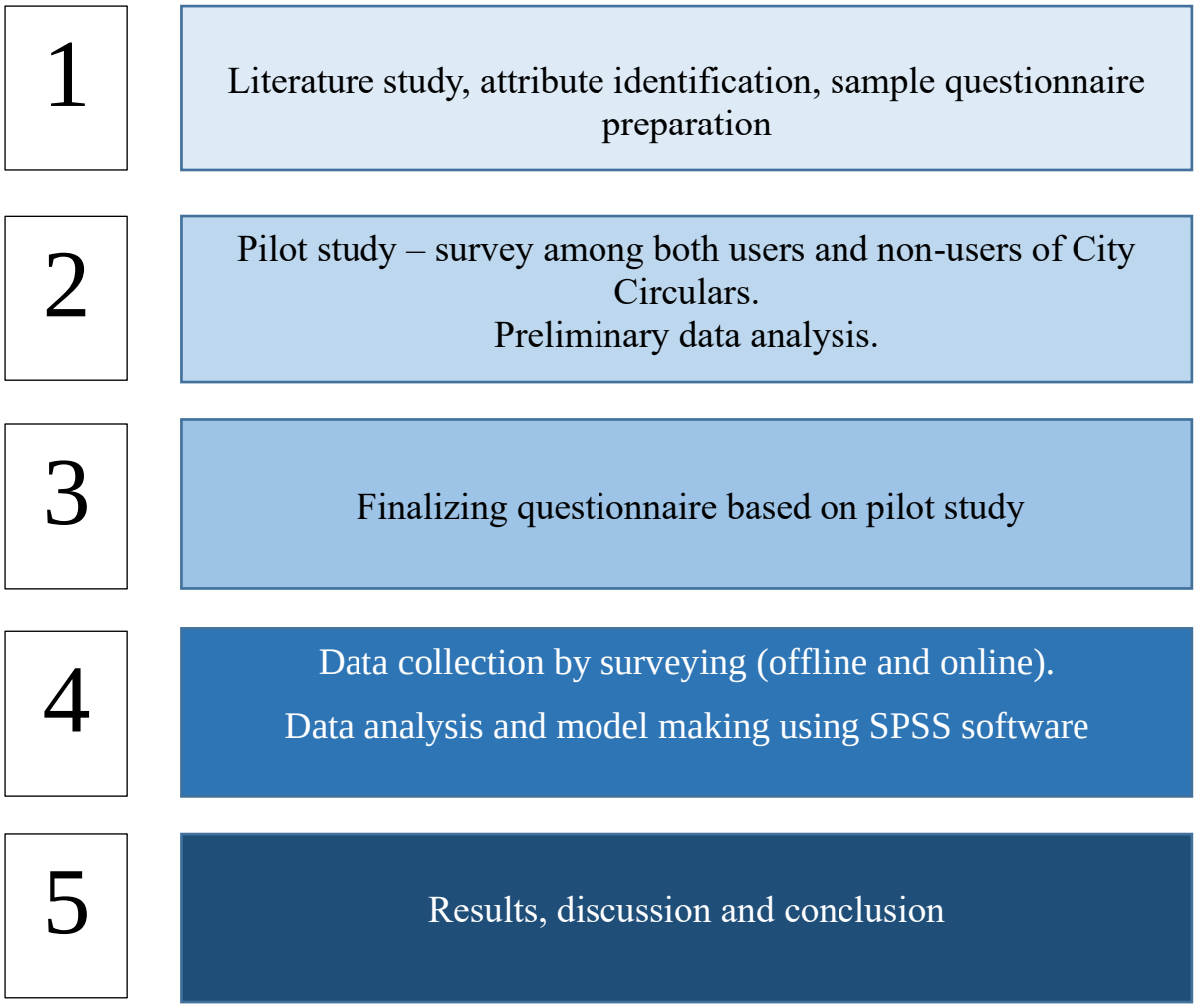



Figure 3.1 Flowchart of methodology

## 3.2 PILOT STUDY

A pilot study has been conducted in various bus stops in Thiruvananthapuram city like East fort bus stop, Thampanoor bus stop on November 2022. The stops are selected based on the route of city circular. About 40 responses were obtained from the pilot study. The study was done by asking questions and using google forms to the residents in the city. The questionnaire consisted of socio-demographic factors, trip characteristics, and service attributes of the city circular (in Likert scale). From the 40 responses obtained, 23 used city circulars while 17 didn't use the city circular. The reason for not using could mainly be seen as they didn't know about these buses, no services along their routes and confuse in the information regarding the buses.

Among the 23 users, majority of the users were highly satisfied about the ticket price and were least satisfied with transfers/ interchanges. Based on pilot study, the questionnaire was further improved and mode choice were added. And also, two other service attributes were included based on the comments of the respondents and two other existing attributes were combined.

## 3.3 STUDY AREA

As these city circular buses were launched only on Thiruvananthapuram city, the study area chosen was in this city itself. Figure 3.2 shows the map of Thiruvananthapuram city. Being the capital of Kerala, Thiruvananthapuram is one of the most populated cities in Kerala. Based on the existing population growth rate, presently the population of the city in the year 2022 is 9,94,000, while the population in the city and outer area overall is 11,01,000 (CENSUS 2011).

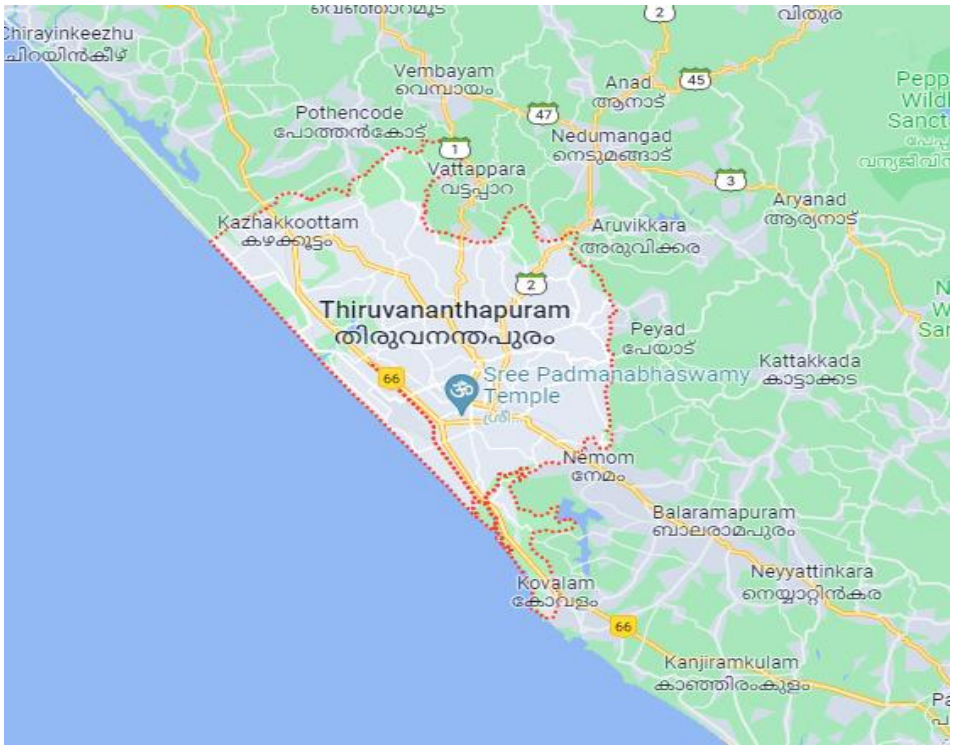



Figure 3.2 Map of Thiruvananthapuram City
(Source: www.googlemaps.com)

The city also consists of many migrants along with the natives. The city consists of many government as well as non-government offices, educational and research institution, industrial units, etc. Presently different modes are used by the people of the city, in which private vehicle constitutes the most, which leads to high traffic congestion during peak hour. It was observed that auto was considered as one of the most important transport modes within the city. The data collection was done from 27$^{th}$ January 2023 to 20$^{th}$ March 2023 covering the major bus stops, college, offices, parks, etc. within the city.

## 3.4 VARIABLES USED

The important variables used are socio-demographic characteristics, trip characteristics and service attributes. The variables are selected based on transportation planning.

### 3.4.1 Socio-demographic Factors

These included normal details like gender, age, occupation, monthly income, vehicle ownership, house hold size. These are the basic factors collected and here it is used as independent variable for the analysis.

### 3.4.2 Trip Characteristics

These include details that a commuter makes. It includes preferred mode of travel, trip purpose, trip distance travelled in a day, frequency of trip, and frequency of using buses. This variable could also be used as an independent variable for the analysis.

### 3.4.3 Service Attributes

There are currently 12 service attributes of the city circle has been given. The attributes selected were based on the city circle. The attributes include;

1. Accessibility (Ac) – public transport is available to as many people in the city.
2. Information system (I) – how much information is provided with related to bus routes, timing, etc.
3. Ticket price (T) – it consists of cost of ticket
4. Travel speed (S) – time spent traveling in CC bus
5. Service frequency (SF) – how often the CC operates given its frequency time.
6. Ease of transfers/ interchanges (In) – it is related to transfer connections and its time spent waiting.
7. Reliability (R) – how the given service of CC bus matches the timetable and whether its available always.
8. Waiting time (W) – time spent waiting for the CC bus
9. Comfort (C) – comfortable the journey is relating to seat, noise level, smell etc.
10. Safety (Sa) – how safe the passenger feels while riding the CC bus including personal safety.
11. Vehicle condition and aesthetic (VA) – it relates to the appeal of vehicles, stations, waiting areas, etc.
12. Attitude of the staff (AS) – behaviour of the staff towards passengers.

The attributes were in Likert scale as they were used to identify the satisfaction by the CC bus users. The scale being from 1 – 5, where 1 being dissatisfied, 2 – not satisfied, 3 – neutral 4 – satisfied and 5 being very satisfied (refer Appendix A).

## 3.5 SAMPLE SIZE AND DATA COLLECTION FOR THE STUDY

Sample size generally depends on the population. As these services are within the city, population taken to calculate sample size was the projected city population (N), which was found as 9,94,000 for the year 2022. Assuming population to be distributed normally. The equations for sample size calculation are given below (3.1, 3.2, 3.3, 3.4). The sample size was calculated from Cochran's formula (Cochran 1977).

$$n_0 = \frac{z^2 pq}{e^2} \quad (3.1)$$

$$q = 1 - p \quad (3.2)$$

$$n = \frac{\frac{z^2 pq}{e^2}}{1 + \frac{(\frac{z^2 pq}{e^2})}{N}} \quad (3.3)$$

Or,

$$n = \frac{n_0}{1 + \frac{(n_0)}{N}} \quad (3.4)$$

Where,

$n_0$ – infinite population sample size

n – finite population size

Z – 1.96; statistical parameter with respect to confidence level taken from z table

Confidence interval was taken as 95%

e – marginal error 5%

p – true proportion for population 0.5

By substituting the values in equation 3.3, the sample size obtained was 385. This sample size was taken as minimum size and for the present study, 542 samples were collected from major bus stops like Eastfort (figure 3.3), Palayam etc., government and non-government offices, colleges, parks, house surveys, etc. both directly interviewing and using online platform.

Figure 3.3 Eastfort bus stop

Among the collected samples, 42 respondents didn't used city circulars, so for further analysis 500 data were used. To improve the reliability of the study data collected were undergone data cleaning.

The reason for not using the service was also asked among the non-users as part of the survey. It was seen that lack of service was the major reason for not knowing about the city circular. Some non-users liked to use their own vehicles and some respondents didn't use city circle due to lack of information provided. That meant lack of awareness about the bus for commoners. Figure 3.3 shows the reason for not using city circulars among the 42 survey responses.

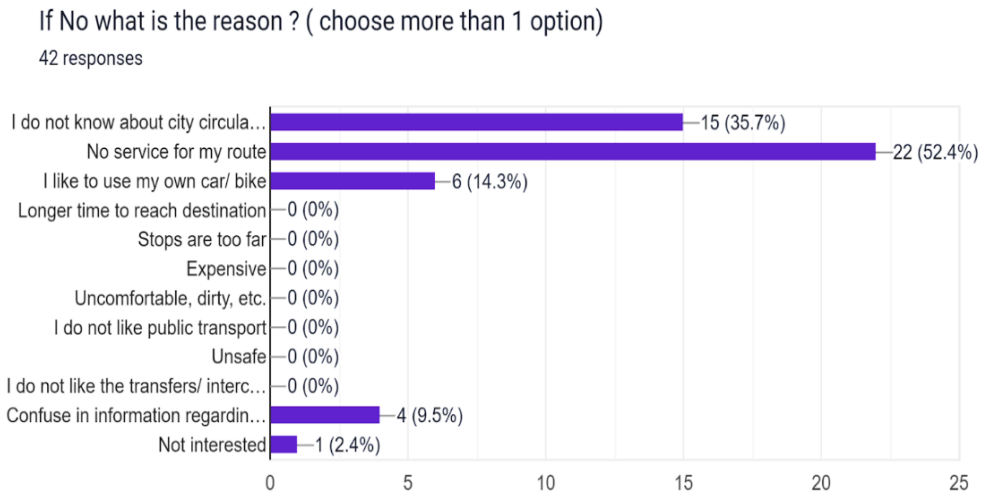



Figure 3.4 Reason for not using city circulars

## 3.6 EXPLORATORY FACTOR ANALYSIS (EFA)

They are generally a part of factor analysis and are based on some correlation among different variables. This is generally done to classify or group the attributes having the same importance together i.e., dividing the variables into groups that are slightly correlated to each other.

Between the different groups formed, there must not be any correlation. This could be done by factor analysis in the SPSS under dimension reduction. The reliability of the data is checked by analysing the Kaiser-Meyer-Olkin (KMO) result within EFA. Generally, KMO is a validation tool used generally to check sampling adequacy. The value of KMO lies between 0 – 1, of which values greater than 0.7 or nearer to 1 have a good relation. The KMO is followed by Bartlett's test of sphericity, which tests the chi–square.

KMO and Bartlett's test is followed by communalities. This is generally done using Principal component analysis. Extraction values of communalities need to be greater than 0.4. Using the same method, a component matrix is formed after explaining total variance in the form of eigen value. The grouping of factors is done by analysing the Eigen value. Eigen value represents the amount of variance contained by a component and is used to reduce dimension. Factors having Eigen value greater than 1 are then extracted. Also, the grouping is done based on correlations. There exists some correlation among some of the attributes.

Components having eigen value more than 1 are considered as factors to which other attributes having a small correlation fall under. Using the VARIMAX rotation technique, a rotated component matrix is formed which shows factors and the attributes under it. VARIMAX rotation technique and principal component analysis are considered as factoring methods. There are many types of methods for doing factor analysis other than VARIMAX and principal component. Another method is selected based on the condition of analysis (refer Appendix B).

## 3.7 ORDINAL LOGIT MODEL (OLM)

This model is generally done to identify the relation between the dependent variables of ordinal nature with different independent variables. These models are also called ordered logistic regression and proportional odds models. They are similar to binary logistic regression and can be called an extension to it. Here we assume the distribution is of logistic nature. The type of dependent variable includes scale natured data like satisfaction level, ratings, etc. Once the data set is made and imported into the software for analysis after running the model a warning table mentioning cells with zero frequencies was shown first. This is followed by a case processing summary showing the statistics of the attributes and data.

Model fitting information and goodness of fit show how well the data fit a distribution from a population. The goodness of fit follows chi-square as it is one of the best methods. Pseudo-$R^2$ when compared to $R^2$ has lesser importance and generally checks Mc Fadden value. Also, the pseudo-$R^2$ depends on the log likelihood value. When a model is having lower value of

likelihood, the log of the likelihood will be having a larger value. Hence, a small ratio of log likelihoods shows that the full model is more fit than the intercept model. The parameter estimates of the model show the probability of the variable under a given category of the dependent variable. The negative and positive sign associated with the estimate is interpreted as a linear regression. The positive sign falls under the higher category of dependent variable while the negative falls under the lower category. The test of parallel lines is also checked to find whether it's significant or not.

## 3.8 MULTINOMIAL LOGIT MODEL (MNL)

The MNL is considered one of the prominent models. For travel mode choice analysis, MNL is considered an important model. It is also a type of logistic model (can also be called an extension of the binary logistic model when the dependent variable has more than 2 categories). It follows certain assumptions only then the analysis is made. They are generally used when we are expected to have more than 2 outcomes i.e., the model is used to predict the probabilities of the different possible outcomes of a categorically distributed dependent variable, concerning a given set of independent variables. One of the assumptions is that there must be no correlation between independent variables. The analysis is interpreted based on the regression coefficient and its estimates, the odds ratio. While doing this analysis, one category of the dependent variable is set as a reference category to which another mode/ category is compared**.** Case processing summary is followed by model fitting and goodness of fit that shows how well the data fit a distribution from a population. Unlike OLM the Mc Fadden pseudo-$R^2$ value lies above 0.2 explaining the variance. MNL also contains a likelihood ratio test showing the significance of the independent variable. A classification table showing observed and predicted values is shown at last (refer Appendix D).

# CHAPTER 4

# RESULTS AND DISCUSSIONS

## 4.1 DATA ANALYSIS

### 4.1.1 Reliability check and descriptive statistics

The reliability of the perception factors was checked to find if they were reliable or suitable for the study. For this, reliability analysis was done under scale option in SPSS. From the total data collected, 500 data were used after checking their reliability. Also, along with that data validation was also done. This was done to increase the efficiency while making a model. The data collected were obtained as valid for the study. Table 4.1 shows the case processing summary of service attributes.

Table 4.1 Case processing summary

| | | N | % |
|---|---|---|---|
| **Cases** | Valid | 500 | 100.0 |
| | Excluded | 0 | 0 |
| | Total | 500 | 100.0 |

Cronbach's Alpha was used to check the reliability. It checks the internal consistency reliability of the unobserved variables. The value ranges from 0 – 1. The minimum value of Cronbach's alpha that is acceptable is 0.6 if it's above 0.7 it is good and acceptable, above 0.9 is excellent. Table 4.2 shows the Cronbach's alpha value of the 12 service attributes obtained after analysis; the value was greater than 0.7 so it was found good for further analysis. This was done to the service attributes as they were found to have a chance of correlation rather than on socio-demographic and trip characteristics which do not have a slight correlation with each other, making them suitable for regression analysis. The Cronbach's Alpha could be increased if we increase the variables used in the study. These were also based on average inter-item correlations. The values should be between 0.15 – 0.5. greater the value, the greater the correlation.

Table 4.2 Reliability test result of service attributes

| Cronbach's Alpha | N of Items |
|---|---|
| 0.772 | 12 |

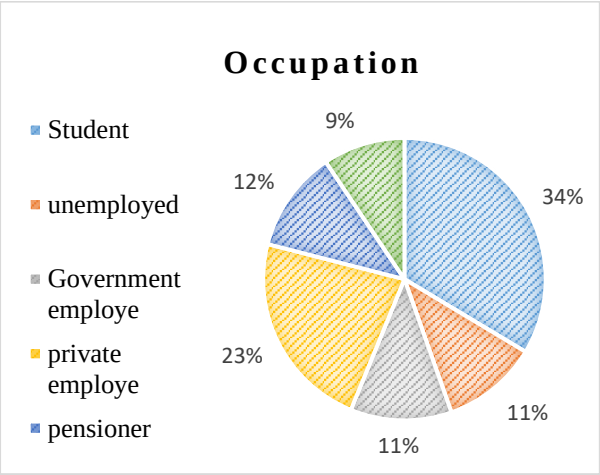
The descriptive statistic with the coding pattern of the collected data is shown in Table 4.3. The table includes socio-demographic and trip characteristics along with their count and

percentage. The total valid sample size was 500. The male-to-female ratio was 0.953. While 71.6% of the sample were frequent trip makers across the city of which 41.6% are work-related trips, which are followed by students.

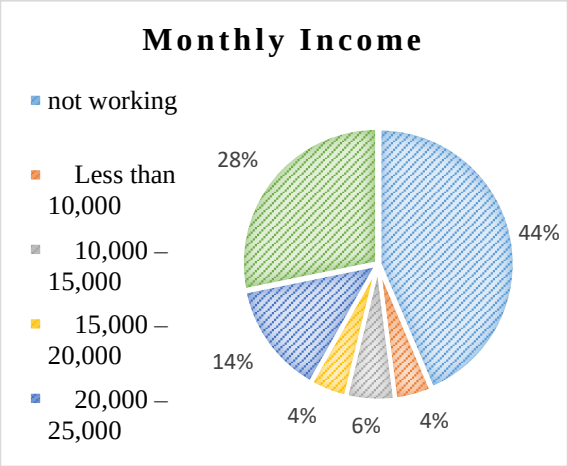
Table 4.3 Socio-demographic and trip characteristic and coding pattern adopted
(N- number of observations, P- percentage, C – code)

| **Socio-demographic characteristics** | **N = 500** | **P %** | **C** | **Trip characteristic** | **N = 500** | **P %** | **C** |
|---|---|---|---|---|---|---|---|
| Gender (G) | | | | Preferred mode (PM) | | | |
| Male | 244 | 48.8 | 1 | Car | 120 | 24 | 1 |
| Female | 256 | 51.2 | 2 | Bike/scooter | 136 | 27.2 | 2 |
| Age (A) | | | | Public bus | 179 | 35.8 | 3 |
| Less than 18 | 29 | 5.8 | 1 | Private bus | 24 | 4.8 | 4 |
| 18 – 40 | 293 | 58.6 | 2 | Auto | 41 | 8 | 5 |
| More than 40 | 178 | 35.6 | 3 | Other | 1 | 0.2 | 6 |
| Occupation (O) | | | | Trip frequency (TF) | | | |
| Student | 168 | 33.6 | 1 | Occasional | 142 | 28.4 | 1 |
| Unemployed | 55 | 11 | 2 | Frequent | 358 | 71.6 | 2 |
| Government employ. | 58 | 11.6 | 3 | Trip purpose (TP) | | | |
| Private employ | 114 | 22.8 | 4 | School/college | 163 | 32.6 | 1 |
| Pensioner | 58 | 11.6 | 5 | Work | 208 | 41.6 | 2 |
| Self-employed | 47 | 9.4 | 6 | Shop | 97 | 19.4 | 3 |
| Monthly income (MI) | | | | Home | 11 | 2.2 | 4 |
| Not working | 218 | 43.6 | 1 | other | 21 | 4.2 | 5 |
| Less than 10,000 | 22 | 4.4 | 2 | Trip distance (TD) | | | |
| 10,000 – 15,000 | 29 | 5.8 | 3 | Less than 5 km | 121 | 24.2 | 1 |
| 15,000 – 20,000 | 22 | 4.4 | 4 | 5 – 10km | 254 | 50.8 | 2 |
| 20,000 – 25,000 | 68 | 13.6 | 5 | 10 – 15 km | 50 | 10 | 3 |
| Above 25,000 | 141 | 28.2 | 6 | More than 15 km | 75 | 15 | 4 |
| Vehicle ownership (VO) | | | | Frequency of using bus (BF) | | | |
| Owns a car | 61 | 12.2 | 1 | Occasional | 253 | 50.6 | 1 |
| Owns a bike/scooter | 176 | 35.2 | 2 | Frequent | 245 | 49 | 2 |
| Owns car & bike/scooter | 155 | 31 | 3 | Nil | 2 | 0.4 | 3 |
| No vehicle | 103 | 20.6 | 4 | | | | |
| Other | 5 | 1 | 5 | | | | |

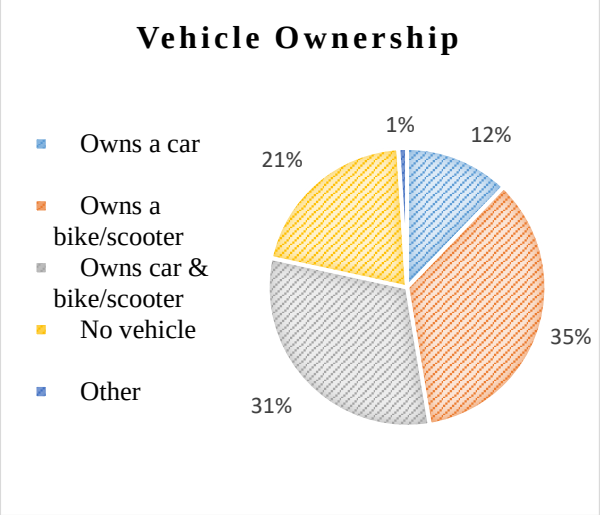
Figure 4.1 shows the frequency of socio-demographic and trip characteristics of certain factors of 500 respondents.



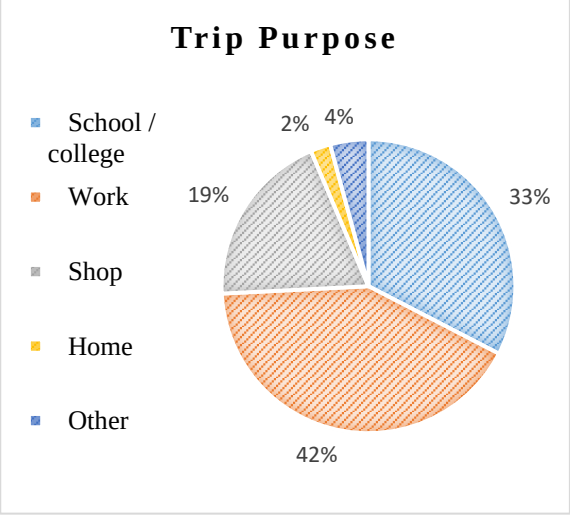

a. Frequency of occupation



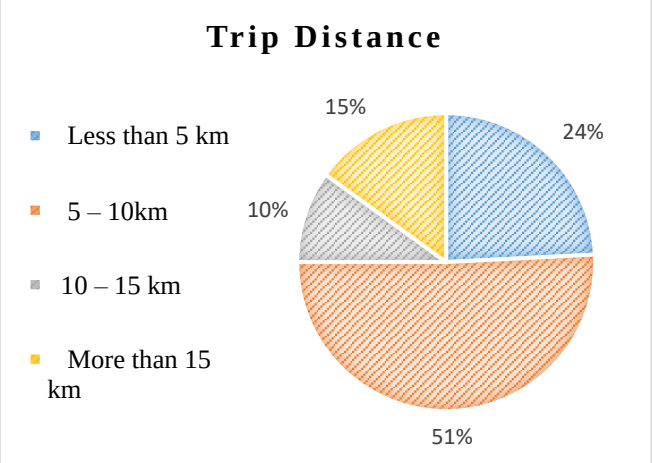

b. Frequency of monthly income




c. Frequency of vehicle ownership




d. Frequency of trip purpose




e. Frequency of trip distance

Figure 4.1 Frequencies of socio-demographic and trip characteristic

The gender and income-wise comparison of mode preference (Figures 4.2 and 4.3) shows that public transport was preferred among the users in the city. It was also seen that females prefer public buses when compared to males who prefer 2-wheelers. Car was also preferred by both almost same manner, which could also be seen in private bus preference. For auto, females tend to use them more than males. In the case of income-wise comparison, the public vehicle is preferred by non-working groups (students, unemployed, housewife, etc.). Public transport and 2-wheeler are preferred by the middle-income group, while cars were generally preferred among the higher-income group. Also, among the higher income especially among the pensioners, preference to auto could be seen. The general trend that could be seen is that public transport was being favoured within the city by both genders.

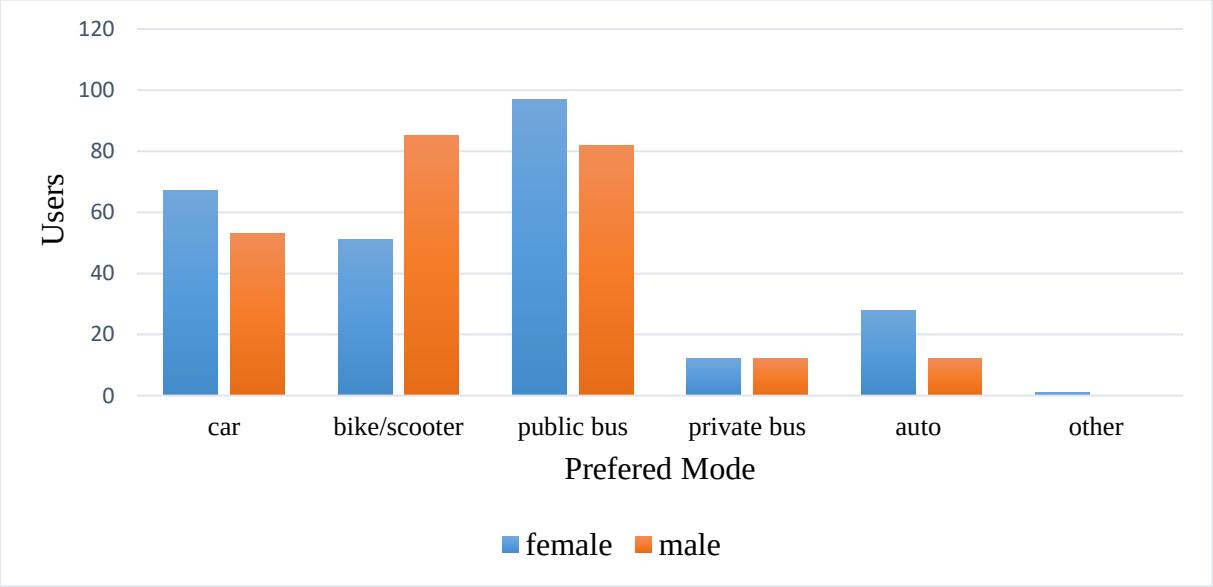



Figure 4.2 Mode preference based on gender

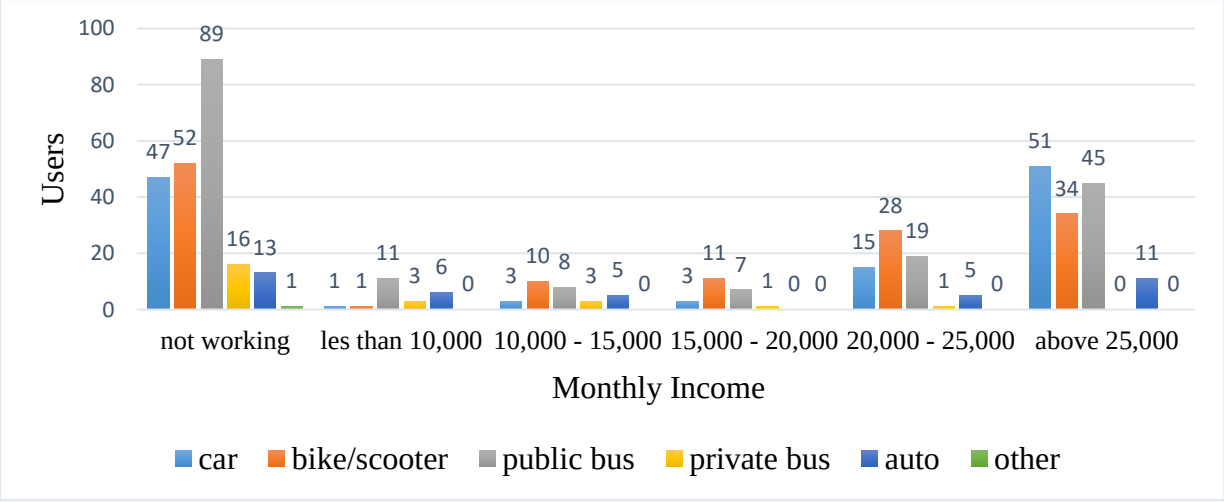



Figure 4.3 Income wise mode preference

From figure 4.4 we could see that public transport was preferred in all cases along with cars and 2 – wheelers, especially in the case where users didn't own any vehicle. From figure 4.5 we can observe that frequent trip makers preferred the public bus when compared to other and for occasional trip makers almost all the mode was preferred equally. The private bus was the least preferred among occasional trip makers, while the auto is least preferred among frequent trip makers.

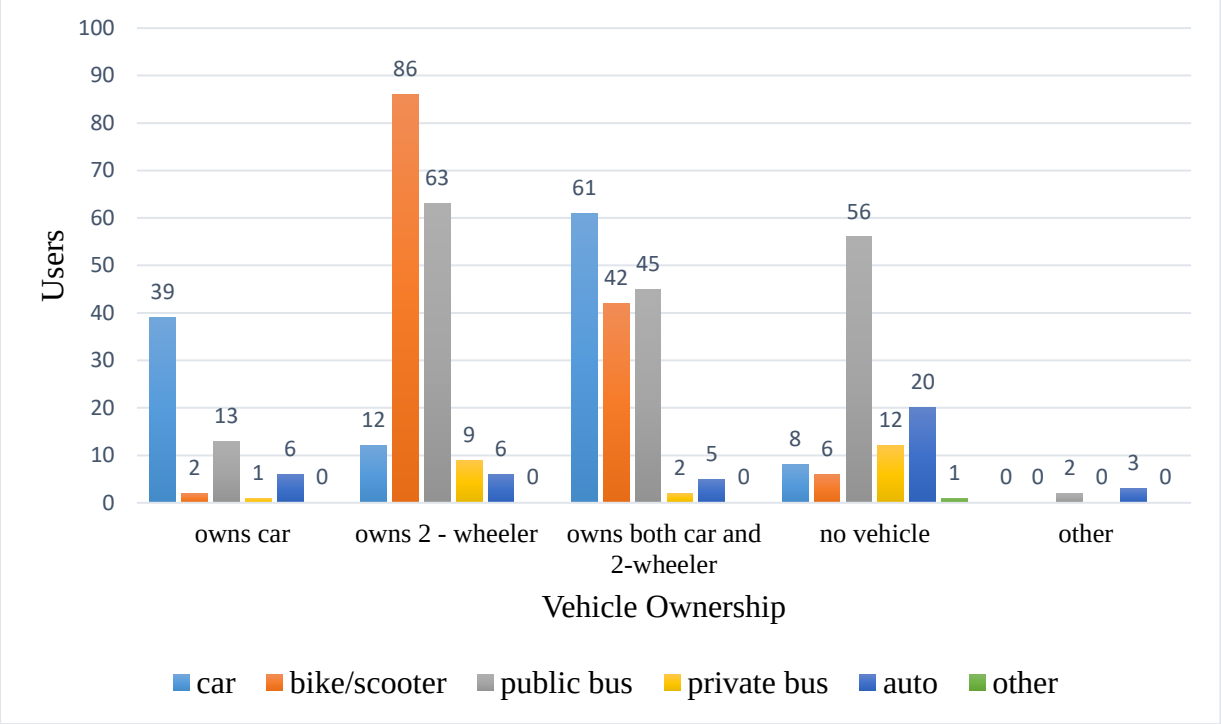



Figure 4.4 Mode preference based on vehicle ownership

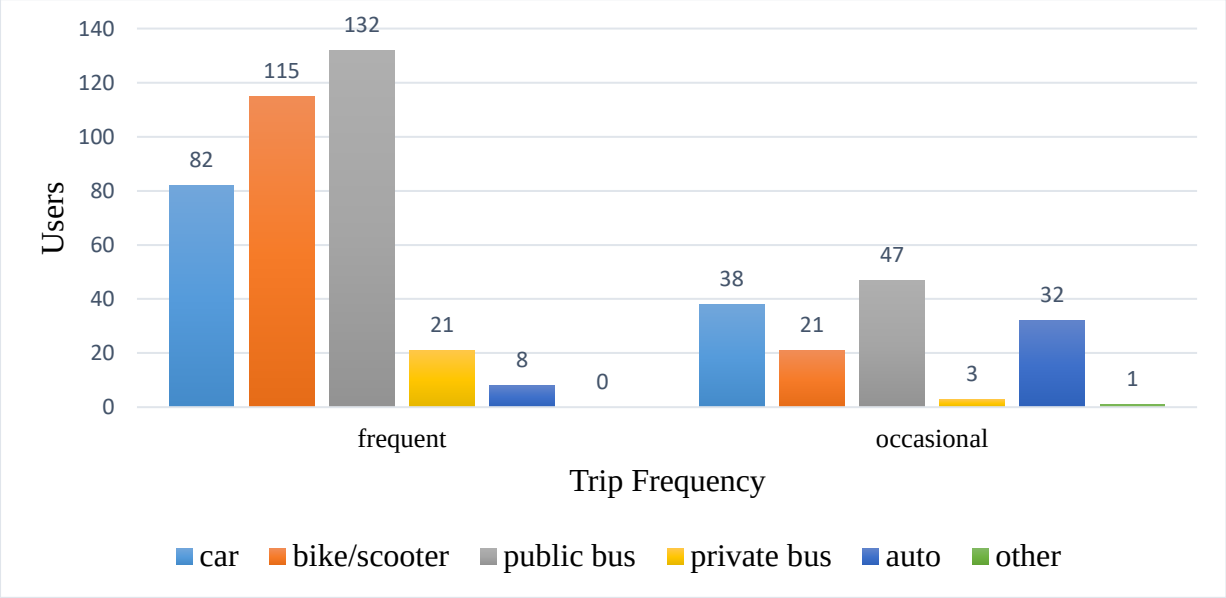



Figure 4.5 Mode preference based on trip frequency

From table 4.4, it was seen that the CC bus users were highly satisfied with the ticket price while they were least satisfied with the waiting time. Attributes like comfort, safety, etc. called perceived attributes were also seen as satisfied when compared to physical attributes. The attribute they find the least satisfactory other than waiting time was the current information system of the CC bus. This was particularly seen as the only factor that was least satisfied by all the age groups especially higher age groups. After the information system, the next attribute to have the least satisfaction was the ease of transfer/ interchanges. This was because of the lack of services on certain routes. Using 2 buses to reach a particular destination is often considered troublesome, especially for older age. Even though the CC bus has a scheduled service frequency i.e., 15 min during peak hour and 30 min during off-peak, the users were not satisfied with the frequency as they don't like to spend more time waiting. But some were satisfied with the waiting time.

Table 4.4 Descriptive statistics of the service attributes.
(Sd – Standard Deviation)

| Service Attributes | Mean | Sd | Service Attributes | Mean | Sd |
|---|---|---|---|---|---|
| (A) Accessibility | 4.14 | 0.699 | (W) Waiting time | 3.68 | 0.826 |
| (T) Ticket price | 4.73 | 0.554 | (C) Comfort | 4.27 | 0.738 |
| (S) Travel speed | 4.17 | 0.635 | (Sa) Safety | 4.40 | 0.684 |
| (SF) Service frequency | 3.96 | 0.763 | (VA) Vehicle condition and aesthetic | 4.34 | 0.702 |
| (In) Ease of transfer | 3.87 | 0.796 | (AS) Staff behaviour | 4.39 | 0.657 |
| (R) Reliability | 3.92 | 0.723 | (I) Information provision | 3.79 | 0.878 |

### 4.1.2 Explanatory Factor Analysis – EFA

These were part of factor analysis and were based on some correlation among different variables. This was done to classify or group the service attributes having the same importance together i.e., dividing the variables into groups that were slightly correlated to each other. Between the different groups formed, there must not be any correlation. This was done by factor analysis in the SPSS under dimension reduction, from which further analysis was done by the VARIMAX rotation technique and principal component analysis (factoring method). Out of the 12 attributes under the study, after performing EFA, 10 attributes were obtained to

have some correlation and are grouped into 2 factors. This was done by analysing the Eigen value. Factors having Eigen value greater than 1 were extracted which explains the 44.5% of the total variance. Ticket price and accessibility were found as the variable having no correlation among other variables and the remaining attributes were grouped. Table 4.5 shows the KMO test results. The KMO value showed the sampling adequacy. As the value was above 0.7, the data collected were adequate for further analysis if required. Bartlett's test compares the relation between the correlation matrix and the identity matrix. If the p-value was less than 0.05, then the data was suitable for analysis. The reliability of the 10 grouped attributes was also checked (Table 4.6). The reliability was checked by analysing the Cronbach's alpha value. It measures the internal consistency of the data; the value of alpha lies between 0 – 1 and must be greater than 0.7. Here, the value of alpha was 0.735, which was acceptable.

Table 4.5 KMO and Bartlett's test result

| Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) | | 0.780 |
|---|---|---|
| Bartlett's Test of Sphericity | Approx. Chi-Square | 844.324 |
| | df | 45 |
| | Sig. | .000 |

Table 4.6 Reliability test of grouped service attributes

| **Cronbach's Alpha** | **N** |
|---|---|
| 0.735 | 10 |

After analysis, two factors were found to have eigen value greater than 1 among the 10 attributes, explaining 45% of total variance. So other attributes were be grouped under the following based on the rotation matrix.

- Factor 1 consists of speed, comfort, safety, vehicle condition & aesthetics and attitude of the staff.
- Factor 2 consists of information provision, service frequency, ease of interchange, reliability and waiting time.

We can group them under these 2 factors having almost same influence and slight correlation

### 4.1.3 Ordinal Logit Modelling

Identify the relation between the dependent variable of ordinal nature with different independent variables.

These models are also called ordered logistic regression and proportional odds models. They are similar to binary logistic regression and can be called an extension of it. Here we assume the distribution is of logistic nature. The parameter estimates of the model show the probability of the variable under the given category of the dependent variable. The negative and positive sign associated with the estimate is interpreted as a linear regression. The positive sign falls under the higher category of dependent variable while the negative falls under the lower category. Based on the EFA, for preparing OLM each of the 10 service attributes was considered a dependent variable while the socio-demographic and trip characteristics were considered independent variables. The removed attribute was the satisfied attribute which was likely to have the same influence. Each variable was used to obtain its respective model to identify the relation between them. Not all the variables have a significant influence on service attribute satisfaction. The significant values were shaded in the table.

Also, for logistic regression, in the place of $R^2$, pseudo-$R^2$ was checked. They look like $R^2$ as their range is same (ranging from 0 to 1), but some pseudo $R^2$ never achieve 0 or 1. Higher values indicate better model fit, but they cannot be interpreted as one would interpret a normal regression $R^2$ and different pseudo-$R^2$ can arrive at very different values. For all the OLM obtained, the McFadden pseudo-$R^2$ obtained was lower.

Pseudo-$R^2$ cannot be interpreted independently across the datasets. They are valid and useful in evaluating multiple models which predicts the same outcome on the given dataset. It will only be meaningful only when compared to other pseudo-$R^2$ of the same type, on the same data, predicting the same outcome. So, the higher pseudo-$R^2$ indicates which model better predicts the outcome.

1. Information system Vs independent variables

The model fitting information was found significant ($p<0.05$) and the goodness of fit which showed that the chi-square value was also significant ($p>0.05$). The McFadden value of pseudo-$R^2$ showed 5.5% improvement in the prediction outcome.

From table 4.7, it was observed that age was the factor that was found as a significant variable influencing information system. There was a negative relation, which meant that with increasing age there was a lower satisfaction in the information system and hence it falls under the low satisfaction category. This could mean that older aged were not satisfied with the information system. This was also seen during the survey. A similar trend could be seen for

other variables especially trip purpose and vehicle ownership, even though it was found as non-significant.

All together this attribute should be improved to have satisfied responses from users. This attribute was also seen as the one to which the respondents responded negatively the most.

Table 4.7 OLM of Information system and independent variables

| | **Variables** | **Estimate** | **OR** |
|---|---|---|---|
| Threshold | [Information = 1] | -6.755 | 1.899 |
| | [Information = 2] | -5.435 | 1.881 |
| | [Information = 3] | -3.642 | 1.874 |
| | [Information = 4] | -1.064 | 1.865 |
| Location | Age | -0.500 | 0.230 |
| | Salary | 0.075 | 0.098 |
| | distance | -0.137 | 0.099 |
| | [Gender=1] | 0.101 | 0.182 |
| | [Gender=2] | 0[a] | |
| | [Occupation=1] | -0.369 | 0.735 |
| | [Occupation=2] | -0.040 | 0.614 |
| | [Occupation=3] | 0.184 | 0.465 |
| | [Occupation=4] | 0.401 | 0.399 |
| | [Occupation=5] | -0.764 | 0.571 |
| | [Occupation=6] | 0[a] | |
| | [purpose=1] | 0.686 | 0.784 |
| | [purpose=2] | -0.380 | 0.646 |
| | [purpose=3] | -0.312 | 0.481 |
| | [purpose=4] | -0.991 | 0.724 |
| | [purpose=5] | 0[a] | |
| | [Vehicles=1] | -0.759 | 0.940 |
| | [Vehicles=2] | -1.367 | 0.915 |
| | [Vehicles=3] | -1.434 | 0.918 |
| | [Vehicles=4] | -1.316 | 0.911 |
| | [Vehicles=5] | 0[a] | |

2. Speed Vs independent variables

The model fitting information was found significant ($p<0.05$) and the goodness of fit which shows the chi-square value was also significant ($p>0.05$). The McFadden value of pseudo-$R^2$ showed 5.6% improvement in the prediction outcome.

From table 4.8, it was observed that many variables influence the satisfaction of the travel speed. Age was also found to have a negative relation i.e., as age increases by a unit the satisfaction level for travel speed falls under the lower category. A similar trend could be seen for occupation, in which the occupation=3 (government employ) has a significant influence. Coming to vehicle ownership, they showed a positive relation and were found significant i.e., all fall under the high satisfaction category. For gender when compared to females, males were found satisfied to speed even if it was found as not significant.

Table 4.8 OLM of Speed and independent variables

| | **Variables** | **Estimates** | **OR** |
|---|---|---|---|
| Threshold | [Speed = 1] | -20.988 | 0.000 |
| | [Speed = 2] | -20.277 | 0.000 |
| | [Speed = 3] | -16.673 | 0.000 |
| | [Speed = 4] | -13.542 | 0.000 |
| Location | Age | -0.410 | 0.663 |
| | [Gender=1] | 0.056 | 1.058 |
| | [Gender=2] | 0[a] | |
| | [Occupation=1] | -0.491 | 0.612 |
| | [Occupation=2] | -0.404 | 0.668 |
| | [Occupation=3] | -0.980 | 0.375 |
| | [Occupation=4] | -0.661 | 0.516 |
| | [Occupation=5] | 0.102 | 1.107 |
| | [Occupation=6] | 0[a] | |
| | [Vehicles=1] | 2.437 | 11.434 |
| | [Vehicles=2] | 2.363 | 10.623 |
| | [Vehicles=3] | 2.904 | 18.242 |
| | [Vehicles=4] | 1.969 | 7.160 |
| | [Vehicles=5] | 0[a] | |

3. Service frequency Vs independent variables

The model fitting information was found significant ($p<0.05$) and the goodness of fit which shows the chi-square value was also significant ($p>0.05$). The McFadden value of pseudo-$R^2$ showed a 5.1% improvement in the prediction outcome.

Compared to information and speed, service frequency was influenced by many variables (table 4.9). Out of the variables, the trip distance was having a positive relationship which means that the higher the trip distance it falls under the higher satisfaction. A similar trend was seen in vehicle ownership and occupation type. For frequent trip users were also having positive relation and fall under higher satisfaction category. While trip mode was having negative

relation i.e., service frequency falls under lower satisfaction when compared to their private vehicle.

Table 4.9 OLM of Service frequency and independent variables

| | Variables | Estimate | OR |
|---|---|---|---|
| Threshold | [Service frequency = 1] | -19.690 | 0.000 |
| | [Service frequency = 2] | -18.362 | 0.000 |
| | [Service frequency = 3] | -16.348 | 0.000 |
| | [Service frequency = 4] | -13.504 | 0.000 |
| Location | distance | 0.266 | 1.304 |
| | [Vehicles=1] | 2.538 | 12.653 |
| | [Vehicles=2] | 2.574 | 13.116 |
| | [Vehicles=3] | 2.340 | 10.377 |
| | [Vehicles=4] | 1.623 | 5.067 |
| | [Vehicles=5] | 0[a] | |
| | [Mode=1] | -1.092 | 0.335 |
| | [Mode=2] | -1.161 | 0.313 |
| | [Mode=3] | -0.853 | 0.426 |
| | [Mode=4] | -1.235 | 0.290 |
| | [Mode=5] | 0[a] | |
| | [trip=1] | 1.031 | |
| | [trip=2] | 0[a] | |

4. Waiting time Vs independent variables

The model fitting information was found significant ($p<0.05$) and the goodness of fit which showed that the chi-square value was also significant ($p>0.05$). The McFadden value of pseudo-$R^2$ showed 3.1% improvement in the prediction outcome.

From table 4.10, it was seen that satisfaction level for waiting time was mainly influenced by the occupation type. Occupations 1 and 4 i.e., students and private employees had significance and showed a positive relation i.e., they fall under the high satisfaction group. Compared to occupation 6, students have a higher chance of falling under higher satisfaction. Even though age and monthly income showed a positive relation, there was no significant influence on the waiting system. But for mode preference, they have negative relation. From the survey, waiting

time was found as the least satisfying attribute, and it was not found satisfied for older-aged respondents.

Table 4.10 OLM of Waiting time and independent variables

| | Variables | Estimate | OR |
|---|---|---|---|
| Threshold | [Waiting = 1] | -3.325 | 0.0359 |
| | [Waiting = 2] | -1.208 | 0.298 |
| | [Waiting = 3] | 0.810 | 2.248 |
| | [Waiting = 4] | 3.265 | 26.17 |
| Location | [Occupation=1] | 1.289 | 3.627 |
| | [Occupation=2] | 0.243 | 1.274 |
| | [Occupation=3] | 0.351 | 1.42 |
| | [Occupation=4] | 0.760 | 2.137 |
| | [Occupation=5] | -0.087 | 0.916 |
| | [Occupation=6] | 0[a] | |
| | Age | 0.197 | 1.217 |
| | Salary | 0.053 | 1.054 |

5. Safety Vs independent variables

The model fitting information was found significant ($p<0.05$) and the goodness of fit showed that the chi-square value was also significant ($p>0.05$). The McFadden value of pseudo-$R^2$ showed a 5.9% improvement in the prediction outcome.

It was observed from table 4.11 that except trip purpose, all other variables were found insignificant which meant that all have almost the same influence with respect to safety. As safety was generally or commonly influenced by almost all the variables. Here trip purpose influences the safety parameter. Trip purpose 3 i.e., shopping is having significant influence and that too positively. But for age, there was a negative relation i.e., the older age group fall under the lower satisfaction group. Similarly, monthly income and gender have a negative relation and fall under the low satisfaction group.

Table 4.11 OLM of Safety and independent variables

| | Variables | Estimate | OR |
|---|---|---|---|
| Threshold | [Safety = 1] | -7.725 | 0.000 |
| | [Safety = 2] | -6.782 | 0.001 |
| | [Safety = 3] | -4.551 | 0.010 |
| | [Safety = 4] | -1.902 | 0.149 |

| | | | |
|---|---|---|---|
| Location | Age | -0.095 | 0.909 |
| | [Gender=1] | -0.218 | 0.804 |
| | [Gender=2] | 0[a] | |
| | [purpose=1] | 0.026 | 1.025 |
| | [purpose=2] | 0.298 | 1.347 |
| | [purpose=3] | 0.959 | 2.608 |
| | [purpose=4] | -1.084 | 0.338 |
| | [purpose=5] | 0[a] | |

6. Vehicle aesthetic Vs independent variables

The model fitting information was found significant ($p<0.05$) and the goodness of fit which showed that the chi-square value was also significant ($p>0.05$). The McFadden value of pseudo-$R^2$ showed a 4% improvement in the prediction outcome.

The satisfaction for vehicle aesthetics falls under the lowest category for occupation 2 (unemployed), which had a negative estimate. Age was also seen as having a negative estimate even though it was found as insignificant, which means lower age groups fall under the highest satisfaction while older age groups under lower. Almost all the variables were having negative estimates. Also, the above attribute was found quite okay for all the respondents.

Table 4.12 OLM of Vehicle Aesthetic and independent variables

| | **Variables** | **Estimate** | **OR** |
|---|---|---|---|
| Threshold | [Aesthetic = 1] | -8.546 | 0.00 |
| | [Aesthetic = 2] | -7.150 | 0.00 |
| | [Aesthetic = 3] | -4.381 | 0.012 |
| | [Aesthetic = 4] | -2.061 | 0.127 |
| Location | Age | -0.357 | 0.700 |
| | [Occupation=1] | -0.695 | 0.499 |
| | [Occupation=2] | -1.131 | 0.322 |
| | [Occupation=3] | 0.084 | 1.087 |
| | [Occupation=4] | 0.010 | 1.009 |
| | [Occupation=5] | 0.363 | 1.437 |
| | [Occupation=6] | 0[a] | |

7. Comfort Vs independent variables

The model fitting information was found significant ($p<0.05$) and the goodness of fit which showed that the chi-square value was also significant ($p>0.05$). The McFadden value of pseudo-$R^2$ showed a 5.8% improvement in the prediction outcome.

From table 4.13, it was observed that except for occupation, age, trip purpose, and mode preference, all other variables were having a negative estimate. For vehicle ownership i.e., for car owners, 2-wheeler owners, etc. they have a negative estimate which meant these variables positioned among lower satisfaction. It also meant that users owning private vehicles (car, bike\scooter, auto, etc.) were satisfied with their private modes of comfort rather than public transport.

Table 4.13 OLM of Comfort and independent variables

| | **Variables** | **Estimate** | **OR** |
|---|---|---|---|
| Threshold | [Comfort = 1] | -21.686 | 0.000 |
| | [Comfort = 2] | -20.813 | 0.000 |
| | [Comfort = 3] | -18.433 | 0.000 |
| | [Comfort = 4] | -15.969 | 0.000 |
| Location | Age | 0.251 | 1.2854 |
| | [Vehicles=1] | -1.730 | 0.177 |
| | [Vehicles=2] | -1.916 | 0.147 |
| | [Vehicles=3] | -1.823 | 0.1615 |
| | [Vehicles=4] | -2.647 | 0.0708 |
| | [Vehicles=5] | 0[a] | |

8. Interchange Vs independent variables

The model fitting information was found significant ($p<0.05$) and the goodness of fit which showed the chi-square value was also significant ($p>0.05$). The McFadden value of pseudo-$R^2$ showed a 4.4% improvement in the prediction outcome. From table 4.14, we could observe that the estimates between the interchange and independent variables. We could see that similar to service frequency, trip distance has an influence and that too negative estimate. As the trip distance increases by a unit, satisfaction falls under the lower category. This could mean that people tend to dislike having multiple interchanges for a particular trip.

Almost all the variables were having negative estimates even though they were insignificant. But compared to age and monthly income, they were having positive relation and belonged to the higher satisfaction category. As mentioned, trip-related service attributes like service frequency and interchanges were found as the least satisfactory along with waiting time while the survey. The interchange was also an important attribute that requires high improvement and also should be reduced to a minimum so that a passenger could travel with ease without spending more time. Interchanges are also associated with the waiting time.

Table 4.14 OLM of Interchange and independent variables

| | **Variables** | **Estimate** | **OR** |
|---|---|---|---|
| Threshold | [Interchange = 1] | -8.141 | 0.00 |
| | [Interchange = 2] | -5.615 | 0.003 |
| | [Interchange = 3] | -3.356 | 0.034 |
| | [Interchange = 4] | -0.917 | 0.399 |
| Location | Age | 0.097 | 1.101 |
| | Salary | 0.148 | 1.159 |
| | distance | -0.541 | 0.582 |
| | [Bus=1] | -2.136 | 0.118 |
| | [Bus=2] | -1.972 | 0.139 |
| | [Bus=3] | 0[a] | |

9. Reliability Vs independent variables

For this model fitting information ($p>0.05$) and goodness of fit was not found significant ($p>0.05$), so further analysis made will not be effective for analysis. Reliability was also one of the factors having low satisfaction. Almost all the parameter estimates were found as negative showing they follow under the low satisfaction group.

10. Attitude of staff Vs independent variables

For this, model fitting information ($p>0.05$) and goodness of fit ($p<0.05$) were not found significant, so further analysis made will be not effective. Attitude of the staff was also one of the factors falling in higher satisfaction. Socio demographic and trip characteristic have no influence on attitude of the staff. This attribute doesn't have any particular influence when compared to other attributes as all the respondents found it quite satisfied and no complaints were made by the respondents. Almost all the estimates have positive relation.

### 4.1.4 Multinomial logit modelling - MNL

was considered one of the prominent models. For travel mode choice analysis, MNL was considered an important model. It was also a type of logistic model. It follows certain assumptions only then the analysis was to be made. They are generally used when we are expected to have more than 2 outcomes i.e. the model is used to predict the probabilities of the different possible outcomes of a categorically distributed dependent variable, for a given set of independent variables. One of the assumptions is that there must be no correlation between independent variables. The analysis is interpreted based on the regression coefficient and its estimates, the odds ratio.

While doing this analysis, one category of the dependent variable was set as a reference category to which another mode was compared. The city circle users were asked about the mode they choose/prefer to travel a short distance (3km and 7 km) including the CC bus. Since CC circles around the city, many used them for short city trips having distances of 3km, 5km, etc., MNL was done to identify the influence of various factors influencing the switching behaviour of commuters from private vehicles to public transport and vice versa. Here the mode of transport was taken as the dependent variable and socio-demographic and trip characteristics were taken as the independent variable. Mode of transport used are- CC, Car, Auto, 2 wheelers, private bus/public bus. These are the common mode used by the citizens within the city. CC was taken as the reference category to which another mode was compared and analysed.

- For 3 km.

The goodness of fit, model fitting, and likelihood ratio tests were within the significance level, which could be used to conclude that the variable used in the model has a significant impact on predicting the mode choice behaviour (table 4.15 & 4.16). The pseudo-$R^2$ value of, Cox and Snell, Nagelkerke, McFadden was 0.742, 0.791, 0.488, which was found satisfied for the model and indicates the respective variance of the response. The value ranges from 0 – 1, of which values should be above 0.2 and when it's nearer to 1, it showed an excellent model.

Table 4.15 Model fitting information

| **Model** | **Model Fitting Criteria** | | | |
|---|---|---|---|---|
| Intercept Only | 1351.724 | | | |
| Final | 674.259 | 677.465 | 112 | 0.000 |

Table 4.16 Goodness of fit

| | **Chi-Square** | **df** | **Sig.** |
|---|---|---|---|
| Pearson | 1146.853 | 1664 | 1.000 |
| Deviance | 644.571 | 1664 | 1.000 |

Even though almost all the independent variables were obtained as significant in the likelihood ratio test, they were not found significant in the parameter estimates ($p>0.05$). This pattern was seen in this MNL model parameter estimate. But that doesn't mean that they were not influencing instead it means all the other variables have equal influence in predicting the model. The city circle was taken as the reference category of mode choice (this is because to find the mode choice between city circle with other modes), while the last category of each independent variable was taken as reference (default option). Significant ones among the parameters were listed below according to the mode.

Table 4.17 – 4.20. shows the parameter estimates of the significant variables. B value estimated was the estimated regression coefficient while Exp(B) was the odds ratio. The odds ratio is a quantity that gives the relation between two events (refer Appendix D). The parameters having a negative B value decrease the likelihood of the reference variable taken in the study. The likelihood ratio test showed household size, trip frequency, and bus frequency doesn't have any significant influence.

From table 4.17, it was observed that as age increases by a unit, the odds of choosing a car compared to a CC bus decrease by 58.1% i.e., the chance of choosing CC was high for the higher aged. A similar trend was observed in the trip distance as when trip distance increases a unit, the odds of choosing a car compared to a CC bus decreases by 31.1%. As the trip distance increases people tend to use the CC bus when compared to cars. But for monthly income, there was a positive trend i.e., if the income increases by a unit, there showed an increase in odds of 0.516 times of choosing a car. This showed that as income increases people tend to need more comfort so they prefer private vehicles. Similarly, when students were compared with self-employed in the case of occupation, it showed a positive trend i.e., compared to self-employed, students tend to prefer car over CC bus. This meant that self-employed preferred using CC buses compared to cars.

Table 4.17 Parameter estimate of choice between car and CC bus

| Response category | Predictor variables | B | Exp(B) |
|---|---|---|---|
| Car | Intercept | -60.229 | 0.000 |
| | A | -0.869 | 0.419 |
| | MI | 0.416 | 0.516 |
| | TD | -0.373 | 0.689 |
| | [O=1] | 3.489 | 32.739 |

In the case of 2 – wheelers and CC bus (table 4.18), except age, other variables were having positive relation and that means chance of choosing 2 – wheeler is high when compared to age. As age increases by a unit, the chance of choosing 2 – wheeler decreases by 65.8%. this could simply mean older aged people tends to prefer or choose CC bus over 2 – wheelers. For gender, when compared to females, males prefer 2 – wheelers 2.97 times. For students and government employ the chance of choosing 2 – wheelers over CC bus were 29.76 and 7.31 times that of self-employed.

Table 4.18 parameter estimate of choice between two-wheeler and CC bus

| Response category | Predictor variables | B | Exp(B) |
|---|---|---|---|
| 2 – wheelers | Intercept | -33.452 | 0.000 |
| | A | -1.073 | 0.342 |
| | [G=1] | 1.088 | 2.970 |
| | [O=1] | 3.394 | 29.796 |
| | [O=3] | 1.998 | 7.371 |
| | [TP=3] | 1.783 | 5.948 |

For private buses (table 4.19), almost all the significant and non-significant variables have negative relation, which meant the chance of choosing CC bus increases. But for the variable trip distance, there was a positive relation which means the chance of choosing private bus increases by 52.3% compared to a CC bus. This could be due to a lack of service routes for the city circle. Respondents who used private buses previously started using CC once it was launched.

Table 4.19 Parameter estimate of choice between private bus and CC bus

| Response category | Predictor variables | B | Exp(B) |
|---|---|---|---|
| Private bus/ Public bus | Intercept | - 47.649 | 0.000 |
| | A | -0.932 | 0.394 |
| | MI | -0.560 | 0.571 |
| | TD | 0.421 | 1.523 |

For auto, it was observed that (table 4.20) most of the significant values have negative coefficients. This meant that many preferred CC over auto. But in the case of pensioners[O=5], there was a positive relation i.e., when compared to self-employed, pensioners have a higher probability of choosing an auto with respect to the CC bus. Similarly, for monthly income, there was a positive relation i.e., income increases by a unit, the chance of choosing auto increases by 63.3%. we could directly relate that pensioners of high income tend to prefer auto on many occasions when compared to CC. They find auto comfortable over the CC bus. For preferred mode [PM], they were having negative relation which showed that compared to other preferred modes, other users preferred the CC bus over other modes for short distances, even though it was only about 0.2 – 4%.

Table 4.20 Parameter estimate of choice between Auto rickshaw and CC bus

| Response category | Predictor variables | B | Exp(B) |
|---|---|---|---|
| Auto rickshaw | Intercept | -4.586 | 0.010 |
| | MI | 0.490 | 1.633 |
| | TD | -0.757 | 0.469 |
| | [O=5] | 3.103 | 22.259 |
| | [PM=1] | -3.291 | 0.037 |
| | [PM=2] | -6.309 | 0.002 |
| | [PM=3] | -5.049 | 0.006 |

Table 4.21 given below shows the prediction accuracies of mode choice. The diagonal cells showed correct predictions while the other cell showed incorrect predictions. It was observed that out of 204 commuters who used city circular, the model predicted the number of city circle

users as 155 and wrongly predicted that 11 city circle users used cars, 31 used 2 – wheelers, 1 used private bus, and 6 used autos. The accuracy of prediction for city circular was 76%.

For 134 two–wheelers, the model wrongly predicted 25 two–wheeler users used city circle and 9 used cars and none used auto and private buses while correctly predicted 100 users used two–wheelers itself. The accuracy of prediction for two – wheelers was about 74.6%. Similarly, the other modes could be predicted. The overall accuracy of the model is 72%, which is okay. Higher the percentage higher the prediction.

Table 4.21 Prediction accuracy for 3km and actual count

| **Classification** | | | | | | |
|---|---|---|---|---|---|---|
| | Predicted | | | | | |
| Observed | City Circle | Car | Bike/ Scooter | Public/Private Bus | Auto | Percent Correct |
| City Circle | 155 | 11 | 31 | 1 | 6 | 76.0% |
| Car | 12 | 73 | 9 | 0 | 1 | 76.8% |
| Bike/Scooter | 25 | 9 | 100 | 0 | 0 | 74.6% |
| Public/Private Bus | 16 | 0 | 3 | 2 | 0 | 9.5% |
| Auto | 9 | 6 | 1 | 0 | 30 | 65.2% |
| Overall Percentage | 43.4% | 19.8% | 28.8% | 0.6% | 7.4% | 72.0% |

| Actual count | |
|---|---|
| city circle | 204 |
| car | 95 |
| bike/scooter | 134 |
| public/ private bus | 21 |
| auto | 46 |

- For 7 km

Similar to the mode choice analysis for 3 km, mode choice prediction analysis for 7km was also made. From tables 4.22 and 4.23 it was observed that the goodness of fit and model fitting and likelihood ratio tests were within the significance level, which could be used to conclude

that the variable used in the model has a significant contribution towards predicting the mode choice behaviour. But for parameter estimates, it was obtained that all the variables showed a non-significant (p>0.05) which means all the factors have equal importance in contributing to mode choice prediction. The pseudo R2 value of Cox and Snell, Nagelkerke, McFadden was0.468, 0.594, 0.407 respectively and was above 0.2, and it shows about 40.7% - 59.4% variance. The larger the pseudo $R^2$, the larger the variance explained by the model.

Table 4.22 Model fitting information

| **Model** | **Model Fitting Criteria** | | | |
|---|---|---|---|---|
| Intercept Only | 753.066 | | | |
| Final | 437.662 | 315.404 | 112 | 0.000 |

Table 4.23 Goodness of fit

| | **Chi-Square** | **df** | **Sig.** |
|---|---|---|---|
| Pearson | 685.869 | 1664 | 1.000 |
| Deviance | 418.829 | 1664 | 1.000 |

As Age and trip distance increase by a unit, the chance of choosing a car decreases. Also, the chance of choosing a car also decreases based on occupation. Gender, income, etc. have positive relation as it increases by a unit, the chance of choosing a car also increases. Age was having a negative relation. i.e., Age increases the chance of choosing 2-wheeler decreases and prefers CC. Also, for unemployed and pensioners, the chance of choosing a 2-wheeler decreases when compared to a CC bus. It was observed that income and gender have negative relation. As income increases, choosing private buses decreases. Also compared to females, males have less chance of choosing a private bus. A similar trend was seen among students, self-employed and unemployed. Age, income gender, and preferred mode had a negative relation. i.e., as these increase by a unit, the chance of choosing auto decreases and prefers CC bus. Also unemployed preferred CC over auto.

Table 4.26 given below showed the prediction accuracies of mode choice. The diagonal cells show correct prediction while the other cell shows incorrect predictions. We can see that out of 389 commuters who used city circular, the model predicted the number of city circle users as 376 and wrongly predicted that 8 city circle users used cars, 3 used 2 – wheelers, 0 used private buses, and 2 used autos. The accuracy of prediction for city circulars is 96.7%. for auto users correctly predicted 5 used auto and incorrectly predicted 2 used city circular among 7

auto users. The accuracy of prediction of auto users was found as 71.4% Similarly, the other modes could be predicted. The overall accuracy of the model is 82.2%.

Table 4.24 Prediction accuracy for 7km and the actual count

| **Classification** | | | | | | |
|---|---|---|---|---|---|---|
| Observed | Predicted | | | | | |
| | city circle | car | bike/ scooter | public/ private bus | auto | Percent Correct |
| city circle | 376 | 8 | 3 | 0 | 2 | 96.7% |
| car | 25 | 19 | 0 | 0 | 0 | 43.2% |
| bike/ scooter | 41 | 2 | 8 | 0 | 0 | 15.7% |
| public/ private bus | 6 | 0 | 0 | 3 | 0 | 33.3% |
| auto | 2 | 0 | 0 | 0 | 5 | 71.4% |
| Overall Percentage | 90.0% | 5.8% | 2.2% | 0.6% | 1.4% | 82.2% |

| Actual count | |
|---|---|
| city circle | 389 |
| car | 44 |
| bike/scooter | 51 |
| public/private bus | 9 |
| auto | 7 |

## 4.2 DISCUSSION

From the 542 data collected (table 4.3), the CC buses were used by 500 while the remaining did not used. Among the used users, 244 were male and 256 were female. Most were between 18 – 40 years old which contributed about 293 of the respondents. Higher-income and unemployed/not working were seen more among the respondents. The common household size seen among the respondents was with family rather than single. As the study area was the capital, many government employees as well as pensioners were seen. Among the vehicle ownership, bike owners were the most followed by cars. Some did not own any vehicle while some owned autos and used that as transport modes for daily necessities.

From the common preferred mode of travel, it was seen that the public bus constitutes the most preferred mode of transport constituting 35.8%, which was followed by 2 – wheelers and cars. Besides those, auto rickshaws were also used in the city. Pensioners were the ones who were seen preferring autos. 71.6% of the respondents were frequent trip makers, and the major trip purpose was work-related followed by students. Also, 49% of the respondents were frequent bus users (both private and public).

Based on income and gender-wise mode preference, it was seen that public transport was one of the most preferred modes. Females preferred public buses and cars the most while males preferred public buses and 2 – wheelers the most. Higher the income, higher the preference for cars while lower the oncome public bus and 2 – wheelers were preferred.

For the service attributes of the CC bus, ticket price and perceived attributes were the most satisfied when compared to other physical attributes. As these attributes were related to the CC bus, the waiting time and the information related to the bus were found to be least satisfactory and the users may find it confusing. Even though the KSRTC set the service frequency, the users find it not satisfied. Following this the users find it difficult when they were forced to use two or more buses to reach their destination i.e., transfers were one attribute that almost all the users find not satisfied be it CC bus or other public modes. The perceived attributes particularly the attitude of the staff was found satisfactory which was a great improvement since the staff of KSRTC especially the conductors previously had conflicts among the citizens.

After doing the factor analysis to classify the attributes, it was seen that ticket price and accessibility were found as the attributes having less influence on any model to be made. It could be based on satisfaction as these were highly satisfied. After classification, the 10 attributes could be classified under 2 factors. These 2 factors could overall be named Overall Reliability and Overall Safety. Attributes that fall under overall safety are speed, comfort, safety, vehicle condition & aesthetics, and attitude of the staff, and attribute that fall under overall reliability are information provision, service frequency, ease of interchange, reliability, and waiting time.

By doing ordinal logit models to identify the impact of socio-demographic on attributes, it was observed that the important bus attributes like service frequency and interchange was affected mainly by trip distance. The service frequency was found as the most influential attribute and requires improvement. Other than that, attributes like speed and information systems were affected by age. Vehicle ownership and occupation are the two factors that influenced the

attributes. The perceived attributes were found not that much getting influenced by the socio-demographic and trip characteristics, but some influence could also be seen.

For the multinomial logit model to predict mode choice for two different kilometres, the prediction accuracy obtained was 72 % and 82.2 % respectively for 3km and 7km. The MNL for 3km was found significant for its parameter estimates, while for 7km it was not, even though the log-likelihood and goodness of it were found significant. The pseudo-$R^2$ was also found within the range. From the model for 3 km, the CC bus was compared to other mode choices based on socio-demographic factors, of which age, gender, trip distance, and income were found to be having important influences. When modes were compared, there was a negative relation between the parameters for auto and private bus, which meant that auto and private bus users preferred the CC bus over the other mode. But in the case of 2 – wheelers, there was a positive relation and the users tend to prefer 2 – wheeler over the CC bus. Age and trip distance was found as the factor having negative relation for cars. The parameter estimates for MNL for 7km were found not significant, but that doesn't mean it has no influence but it shows that all parameters contribute equally.

# CHAPTER 5

# CONCLUSIONS

## 5.1 FINDINGS

The following conclusions were made from the models and analysis;

- The ticket price was the highly satisfactory service attribute and waiting time was the least satisfactory attribute for the city circular bus.
- After factor analysis, ticket price and accessibility were found as the attribute that does not fit other or that doesn't have any correlation with other attributes.
- Speed, comfort, safety, vehicle condition & aesthetic, and attitude of the staff could be grouped into a factor.
- Information provision, service frequency, ease of interchange, reliability, and waiting time could be grouped under a common factor.
- From ordinal logit models, service frequency, interchange, speed, and waiting time were considered important attributes concerning for the socio-demographic and trip characteristics.
- From mode choice prediction, it was seen that CC buses were preferred when the attribute trip distance increases. Also, as age increases the chance of choosing CC bus increases. Occupation and vehicle ownership also influence the mode choice.
- Monthly income plays an important role in choosing a CC bus when compared to car, private bus, and auto.
- For shorter trips, age plays an important role in choosing CC bus when compared to car, 2-wheeler, and private bus.
- Gender was not seen as an important factor in determining mode choice except for 2-wheeler, where males prefer 2-wheeler over CC buses for short trips.
- For mode choice along 7km, the parameter estimates were found not significant for all the modes which indicate all have equal influence in prediction.
- Generally, private bus users and auto users were seen among the users who now prefer the CC bus.
- Students and the unemployed prefer the CC bus, this could be due to low travel costs.
- When compared to socio-demographic and trip characteristics, the perception factor had less importance on mode choice.

Other than software analysis, by directly interviewing the respondents, various comments and remarks were obtained related to city circulars. The majority of the respondents have a positive

remark about the ticket price and a negative remark about the information system. Other than that, in the sense of economy, the launch of city circulars had a huge impact on auto-rickshaws and private bus owners. Many who used auto and private buses previously have started commuting via city circulars, this, in turn, affected the livelihood of auto rickshaw employees and private bus staff. If the service is further introduced into other routes, again it would affect the livelihood of those employees. So, a careful approach must be taken between KSRTC and the private workers so that a mutual system could be made. Other than that, it was also observed that after any traffic congestion, the frequency of the bus changed from 15 minutes to every 5-minute creating a lack of service for the next adjacent time.

Recommendations;

- It should be noted that there must be an improvement in the signage as well as the information system provided so that all age groups could be able to understand easily. This could be done by installing digital systems, guides, etc.
- Other than that, proper details regarding bus routes and interchanges must be provided within the bus stops in the form of a banner or board.
- Services should be increased in more routes so that a greater number of people could get access to it.
- A digital system should be provided inside the bus to show the present location (language both native and English) so that outsiders could able to identify the location.

By providing suitable improvements these circles could be launched in other important cities. Proper management and coordination result in the perfect movement of the buses which increases the public's satisfaction. As a result, more users will shift or choose city circles for short trips.

## 5.2 FUTURE SCOPE

- The sample size could be increased and future research can provide an extension to this work by adding more variables.
- Analysis can be done using machine learning with artificial intelligence.

# REFERENCES


Ashalatha, R., Manju, V. S., & Zacharia, A. B. (2013). Mode choice behaviour of commuters in Thiruvananthapuram city. *Journal of Transportation Engineering*, *139*(5), 494-502.

Beirão, G., & Cabral J. S. (2007). Understanding attitudes towards public transport and private car: A qualitative study. *Transport policy*, *14*(6), 478-489.

Cass, N., & Faulconbridge J. (2016). Commuting practices: New insights into modal shift from theories of social practice. *Transport policy*, *45*, 1-14.

Clark, B., Chatterjee, K., & Melia, S. (2016). Changes to commute mode: The role of life events, spatial context and environmental attitude. *Transportation Research Part A: Policy and Practice*, *89*, 89-105.

Cochran, W. G. (1977). *Sampling techniques*. John Wiley & Sons.

Davison, L. J., & Knowles, R. D. (2006). Bus quality partnerships, modal shift and traffic decongestion. *Journal of Transport Geography*, *14*(3), 177-194

De Oña, J., Estévez, E., & de Oña, R. (2020). Perception of public transport quality of service among regular private vehicle users in Madrid, Spain. *Transportation research record*, *2674*(2), 213-224.

Hansson, J., Pettersson, F., Svensson, H., & Wretstrand, A. (2019). Preferences in regional public transport: a literature review. *European Transport Research Review*, *11*(1), 38.

Hu, X., Zhao, L., & Wang, W. (2015). Impact of perceptions of bus service performance on mode choice preference. *Advances in mechanical engineering*, *7*(3), 1-11

Ma, X., Yuan, Y., Van Oort, N., & Hoogendoorn, S. (2020). Bike-sharing systems' impact on modal shift: A case study in Delft, the Netherlands. *Journal of Cleaner Production*, *259*, 1-13

Murugan, M., & Marisamynathan, S. (2022). Mode shift behaviour and user willingness to adopt the electric two-wheeler: A study based on Indian road user preferences. *International Journal of Transportation Science and Technology*.

Paulley, N., Balcombe, R., Mackett, R., Titheridge, H., Preston, J., Wardman, M., ... & White, P. (2006). The demand for public transport: The effects of fares, quality of service, income and car ownership. *Transport policy*, *13*(4), 295-306

Pike, S., & Lubell, M. (2016). Geography and social networks in transportation mode choice. *Journal of transport geography*, *57*, 184-193.

Redman, L., Friman, M., Gärling, T., & Hartig, T. (2013). Quality attributes of public transport that attract car users: A research review. *Transport policy*, *25*, 119-127.

Schubert, T. F., Henning, E., & Lopes, S. B. (2020). Analysis of the Possibility of Transport Mode Switch: A Case Study for Joinville Students. *Sustainability*, *12*(13), 5232.

Sun, X., & Wandelt, S. (2021). Transportation mode choice behaviour with recommender systems: A case study on Beijing. *Transportation research interdisciplinary perspectives*, *11*, 100408

Thrane, C. (2015). Examining tourists' long-distance transportation mode choices using a Multinomial Logit regression model. *Tourism Management Perspectives*, *15*, 115-121.

Urbanek, A. (2021). Potential of modal shift from private cars to public transport: A survey on the commuters' attitudes and willingness to switch–A case study of Silesia Province, Poland. *Research in Transportation Economics*, *85*, 101008.

# APPENDIX A

This appendix includes the questionnaire designed for developing the perception of the people in Thiruvananthapuram city towards city Circular Buses.

Questionnaire for identifying public's attitude towards City Circular Buses.

I'm a student of T.K.M College of Engineering, currently doing M. Tech in Transportation Engineering. The following questionnaire is prepared on the basis of M. Tech project related to City Circular Buses. I will be thankful to you if you spend a part of your precious time for filling this questionnaire. Your personal responses will be kept confidential and will only be used for academic purpose.

PART 1

This section consists of socio-demographic and travel characteristics.

**A. Socio-demographic Factors**

1) Gender
    i) Male
    ii) Female
    iii) Other
2) Age
    i) Less than 18 yrs.
    ii) 18-40 yrs.
    iii) Above 40 yrs.
3) Occupation
    i) Student
    ii) Unemployed
    iii) Government Employee
    iv) Private employee
    v) Pensioner
    vi) Self employed
4) Monthly Income
    i) Not Working
    ii) Less than 10,000
    iii) 10,000 – 15,000
    iv) 15,000 – 20,000
    v) 20,000 – 25,000
    vi) Above 25,000

5) House hold size (family members)
    i) 1
    ii) 2
    iii) 3
    iv) 4
    v) More than 4
6) Vehicle ownership
    i) Owns a car
    ii) Owns a Bike/ Scooter
    iii) Owns both car and bike/scooter
    iv) NO vehicle
    v) Other ----------
7) Number of working days in a week?

    _____

**B. Trip Characteristics**

1) Preferred mode of travel
    i) car
    ii) Bike/ Scooter
    iii) Public Bus
    iv) Private bus
    v) Auto rickshaw
    vi) Other ----------
2) Trip frequency by any mode
    i) Frequent
    ii) Occasional
3) Trip purpose
    i) School/College
    ii) Work
    iii) Home
    iv) Shop/ shopping
    v) Other ----------
4) Trip distance in a day
    i) Less than 5 km
    ii) 5 – 10 km

iii) 11 – 15 km

iv) More than 15 km

5) Frequency of using Bus

i) Frequent (more than 5 times a week)

ii) Occasional (less than 5 times a week)

PART 2.

City Circulars

Do you know about the city circular buses? Have you heard about them?

City Circular Service is an innovative and new type of transportation network/ services introduced by KSRTC in Thiruvananthapuram. They connect all the major routes inside the city, interlinking all major points like hospitals, schools, government offices, tourist spots, etc.

For KSRTC City Circular service, there are 7 routes. These are operated in circular routes at fixed intervals. Bus services would be available, in both clockwise and anti-clockwise directions. So, a passenger can board the bus from a stop, and then can alight at another stop or in between any stops. This would get repeated until the passenger reaches the desired destination. One of the best features is that the passenger will not be forced to wait at a stop for more than 15 to 30 minutes, depending on the time of travel.


**Figure** 1 Red City Circular

The unique feature of this transport is that all the city circular services will have a unique colour. A passenger can identify the destination of the bus by the colour of the circle. There are currently 8 colours which represents 7 different routes.

- The ticket price is always 10 Rs till a stop.
- Electric city circulars are also available for users.
- Blue -Green- Yellow- Red- Magenta- Brown- Violet- Orange

1) Have you travelled by city circular bus?

- Yes
- No

1.a. If NO, what is the reason? (Choose more than 1 option)

  - I do not know about city circular buses
  - No service for my route
  - I like to use my own car/ bike
  - Longer time to reach destination
  - Stops are too far
  - Expensive
  - Uncomfortable, dirty, etc.
  - I do not like public transport
  - Unsafe
  - I do not like the transfers/ interchanges
  - Confuse in information regarding city circular bus
  - Other reason ----------

(Go to part 3)

1.b. If yes

The following section consist of questions related to city circular bus. Service attributes are provided, the options for each are provided in satisfaction level ranging from 1 - 5

(5- very satisfied, 4 – satisfied, 3 – neutral, 2 – not satisfied, 1 – dissatisfied).

i. Accessible – ease of entry and exit from stops and bus
   - 5
   - 4
   - 3
   - 2
   - 1

ii. Information system – bus timing, signage, route information etc.
   - 5
   - 4
   - 3
   - 2
   - 1

iii. Ticket price – cost of the ticket for the journey
   - 5
   - 4
   - 3
   - 2

- 1

iv. Travel speed

- 5
- 4
- 3
- 2
- 1

v. Service frequency – frequency or number of services or how often the service operates

- 5
- 4
- 3
- 2
- 1

vi. Ease of transfer / interchange – ease in transferring to other mode of transport to reach destination

- 5
- 4
- 3
- 2
- 1

vii. Reliability – available when needed

- 5
- 4
- 3
- 2
- 1

viii. Waiting time – time spent at the stops

- 5
- 4
- 3
- 2
- 1

ix. Comfort – travel comfort, noise effects etc.

- 5
- 4
- 3
- 2
- 1

x. Safety and security

- 5
- 4
- 3

- 2
- 1

xi. Vehicle condition and aesthetic – inside and outside the bus.

- 5
- 4
- 3
- 2
- 1

xii. Attitude of the staff – behaviour of the staff towards commuters

- 5
- 4
- 3
- 2
- 1

PART 3

Mode choice set.

This section consists of question which compares city circular to our choice of mode.

1. Frequently used mode of transport to travel short distance of about 3 km and 7 km within the city
   - Bike / Scooter
   - Car
   - Auto
   - Private bus / Public bus

a. choose the mode you use

| Mode | Distance km | Travel Time min | Travel Price Rs | Waiting Time min | |
|---|---|---|---|---|---|
| Bike / scooter | 3 | 7 | 9 | - |  |
| City Circular | | 10 | 10 | 15 |  |
| Car | | 10 | 30 | - |  |
| Private Bus | | 10 | 13 | 10 |  |
| Auto Rickshaw | | 8 | 45 | 5 |  |

- Bike/ Scooter
- City circular
- Car
- Private bus
- Auto

b. Choose the mode you use

| Mode | Distance km | Travel Time min | Travel Price Rs | Waiting Time min | |
|---|---|---|---|---|---|
| Bike/ scooter | 7 | 15 | 21 | - |  |
| City Circular | | 25 | 10 | 15 |  |
| Car | | 20 | 70 | - |  |
| Private Bus | | 20 | 18 | 10 |  |
| Auto Rickshaw | | 15 | 100 | 5 |  |

- Bike/ Scooter
- City circular
- Car
- Private bus
- Auto

# APPENDIX B

## Explanatory Factor Analysis

(Source: www.statisticssolutions.com)

It is a statistical technique that is used to reduce data to a smaller set of summary variables and to explore the underlying theoretical structure of the phenomena. It is used to identify the structure of the relationship between the variable and the respondent. Exploratory factor analysis can be performed by using the following two methods:

- R-type factor analysis: When factors are calculated from the correlation matrix, then it is called R-type factor analysis.
- Q-type factor analysis: When factors are calculated from the individual respondent, then it said to be Q-type factor analysis.

Driving factor:

There are two methods for driving factor, these two methods are as follows:

1. Principle component factor analysis method: This method is used when we need to drive the minimum number of factors and explain the maximum portion of variance in the original variable.
2. Common factor analysis: This method is used when the researchers do not know the nature of the factor to be extracted and the common error variance.

Selection of factors to be extracted: Theory is the first criteria to determine the number of factors to be extracted. From theory, we know that the number of factors extracted does make sense. Most researchers use the Eigenvalue criteria for the number of factors to be extracted. Value of the percentage and variance explained method is also used for exploratory factor analysis. We can use the scree test criteria for the selection of factors. In this method, Eigenvalue is plotted on a graph and factors are selected.

Orthogonal rotation: In this method, axis is maintained at 90 degrees, thus the factors are uncorrelated to each other. In orthogonal rotation, the following three methods are available based on the rotation:

A. QUARTIMAX: Rows are simplified so that the variable should be loaded on a single factor.

B. VARIMAX: Used to simplify the column of the factor matrix so that the factor extracts are clearly associated and there should be some separation among the variables.

C. EQUIMAX: The combination of the above two methods. This method simplifies row and column at a single time.

Criteria for Practical and Statistical Significance of Factor Loadings: Factor loading can be classified based on their magnitude:

Greater than:

+ .30 — minimum consideration level

+ .40 — more important

+ .50 — practically significant

Power and significance level: The researcher can determine the statistical power and significance level. For instance, in order to achieve a factor loading of .55 with a power of .80, a sample of 100 is needed.

Assumptions:

1. Variables used should be metric. Dummy variables can also be considered, but only in special cases.
2. Sample size: Sample size should be more than 200. In some cases, sample size may be considered for 5 observations per variable.
3. Homogeneous sample: A sample should be homogenous. Violation of this assumption increases the sample size as the number of variables increases. Reliability analysis is conducted to check the homogeneity between variables.
4. In exploratory factor analysis, multivariate normality is not required.
5. Correlation: At least 0.30 correlations are required between the research variables.
6. There should be no outliers in the data.

# APPENDIX C

# Ordinal Logistic Model

(Source: en.wikipedia.org)

In statistics, the ordered logit model (also ordered logistic regression or proportional odds model) is an ordinal regression model—that is, a regression model for ordinal dependent variables—first considered by Peter McCullagh. For example, if one question on a survey is to be answered by a choice among "poor", "fair", "good", "very good" and "excellent", and the purpose of the analysis is to see how well that response can be predicted by the responses to other questions, some of which may be quantitative, then ordered logistic regression may be used.

Ordinal logistic regression is an extension of logistic regression, where the logit (i.e., the log odds) of a binary response is linearly related to the independent variables. If instead the response variable has k levels, then there are k-1 logits. A major assumption of ordinal logistic regression is the assumption of proportional odds: the effect of an independent variable is constant for each increase in the level of the response. Hence the output of an ordinal logistic regression will contain an intercept for each level of the response except one, and a single slope for each explanatory variable.

There are several ways in which an ordinal regression model can be parameterized and different statistical software packages use different parameterizations. Thus, great care should be taken when interpreting the output from ordinal regression models. The model only applies to data that meet the proportional odds assumption, the meaning of which can be exemplified as follows. Suppose there are five outcomes: "poor", "fair", "good", "very good", and "excellent". We assume that the probabilities of these outcomes are given by $p_1(x)$, $p_2(x)$, $p_3(x)$, $p_4(x)$, $p_5(x)$, all of which are functions of some independent variable(s) x. Then, for a fixed value of x, the logarithms of the odds (not the logarithms of the probabilities) of answering in certain ways are given in figure 1:

$$\text{poor}: \quad \log \frac{p_1(x)}{p_2(x)+p_3(x)+p_4(x)+p_5(x)},$$

$$\text{poor or fair}: \quad \log \frac{p_1(x)+p_2(x)}{p_3(x)+p_4(x)+p_5(x)},$$

$$\text{poor, fair, or good}: \quad \log \frac{p_1(x)+p_2(x)+p_3(x)}{p_4(x)+p_5(x)},$$

$$\text{poor, fair, good, or very good}: \quad \log \frac{p_1(x)+p_2(x)+p_3(x)+p_4(x)}{p_5(x)}$$

Figure 1

The proportional odds assumption states that the numbers added to each of these logarithms to get the next are the same regardless of *x*. In other words, the difference between the logarithm of the odds of having poor or fair health minus the logarithm of having poor health is the same regardless of *x*; similarly, the logarithm of the odds of having poor, fair, or good health minus the logarithm of having poor or fair health is the same regardless of *x*; etc.

Ordered logit can be derived from a latent-variable model, similar to the one from which binary logistic regression can be derived.

- Assumptions

When you choose to analyse your data using ordinal regression, part of the process involves checking to make sure that the data you want to analyse can actually be analysed using ordinal regression. You need to do this because it is only appropriate to use ordinal regression if your data "passes" four assumptions that are required for ordinal regression to give you a valid result.

- ✓ Assumption 1: Your dependent variable should be measured at the ordinal level. Examples of ordinal variables include Likert items (e.g., a 7-point scale from "strongly agree" through to "strongly disagree"), amongst other ways of ranking categories (e.g., a 3-point scale explaining how much a customer liked a product, ranging from "Not very much", to "It is OK", to "Yes, a lot")
- ✓ Assumption 2: one or more independent variables that are continuous, ordinal or categorical. However, ordinal independent variables must be treated as being either continuous or categorical.
- ✓ Assumption 3: There is no multicollinearity. Multicollinearity occurs when you have two or more independent variables that are highly correlated with each other. This leads to problems with understanding which variable contributes to the explanation of the dependent variable and technical issues in calculating an ordinal regression
- ✓ Assumption 4: You have proportional odds, which is a fundamental assumption of this type of ordinal regression model; that is, the type of ordinal regression that we are using in this guide (i.e., cumulative odds ordinal regression with proportional odds). The assumption of proportional odds means that each independent variable has an identical effect at each cumulative split of the ordinal dependent variable

Measuring strength of association could also be done Calculating the pseudo-R-Square. There are several $R^2$ like statistics that can be used to measure the strength of the association between

the dependent variable and the predictor variables. They are not as useful as the statistic in multiple regression, since their interpretation is not straightforward.

# APPENDIX D

## Multinomial Logistic Model

(Source: bookdown.org, en.wikipedia.org)

For Binary logistic regression, the number of dependent variables is two, whereas the number of dependent variables for multinomial logistic regression is more than two.

In statistics, multinomial logistic regression is a classification method that generalizes logistic regression to multiclass problems, i.e. with more than two possible discrete outcomes. That is, it is a model that is used to predict the probabilities of the different possible outcomes of a categorically distributed dependent variable, given a set of independent variables (which may be real-valued, binary-valued, categorical-valued, etc.).

Examples: Consumers make a decision to buy or not to buy, a product may pass or fail quality control, there are good or poor credit risks, and employee may be promoted or not. Multinomial logistic regression is used when the dependent variable in question is nominal (equivalently categorical, meaning that it falls into any one of a set of categories that cannot be ordered in any meaningful way) and for which there are more than two categories. Some examples would be:

- Which major will a college student choose, given their grades, stated likes and dislikes, etc.?
- Which blood type does a person have, given the results of various diagnostic tests?
- In a hands-free mobile phone dialling application, which person's name was spoken, given various properties of the speech signal?
- Which candidate will a person vote for, given particular demographic characteristics?
- Which country will a firm locate an office in, given the characteristics of the firm and of the various candidate countries?

The multinomial logistic model assumes that data are case-specific; that is, each independent variable has a single value for each case. The multinomial logistic model also assumes that the dependent variable cannot be perfectly predicted from the independent variables for any case. As with other types of regression, there is no need for the independent variables to be statistically independent from each other (unlike, for example, in a naive Bayes classifier); however, collinearity is assumed to be relatively low, as it becomes difficult to differentiate between the impact of several variables if this is not the case. If the multinomial logit is used

to model choices, it relies on the assumption of independence of irrelevant alternatives (IIA), which is not always desirable. This assumption states that the odds of preferring one class over another do not depend on the presence or absence of other "irrelevant" alternatives.

- Logits or Log Odds

Odds value can range from 0 to infinity and tell you how much more likely it is that an observation is a member of the target group rather than a member of the other group.

$$\text{Odds} = \frac{p}{1-p}$$

If the probability is 0.80, the odds are 4 to 1 (0.80/0.20); if the probability is 0.25, the odds are 0.33 (0.25/0.75).

The odds ratio (OR), estimates the change in the odds of membership in the target group for a one unit increase in the predictor. It is calculated by using the regression coefficient of the predictor as the exponent or exp.

Assume in the example earlier where we were predicting accountancy success by a math's competency predictor that b = 2.69. Thus, the odds ratio is exp (2.69) or 14.73. Therefore, the odds of passing are 14.73 times greater for a student for example who had a pre-test score of 5 than for a student whose pre-test score was 4.

- Hypothesis Test of Coefficients

In logistic regression, hypotheses are of interest:

- ✓ The null hypothesis, which is when all the coefficients in the regression equation take the value zero, and
- ✓ The alternate hypothesis that the model currently under consideration is accurate and differs significantly from the null of zero, i.e., gives significantly better than the chance or random prediction level of the null hypothesis.

Evaluation of Hypothesis

We then work out the likelihood of observing the data we actually did observe under each of these hypotheses. The result is usually a very small number, and to make it easier to handle, the natural logarithm is used, producing a log-likelihood (LL). Probabilities are always less than one, so LLs are always negative. Log-likelihood is the basis for tests of a logistic model.

- Likelihood Ratio Test

The likelihood ratio test is based on the -2LL ratio. It is a test of the significance of the difference between the likelihood ratio (-2LL) for the researcher's model with predictors (called model chi-square) minus the likelihood ratio for the baseline model with only a constant in it.

Significance at the .05 level or lower means the researcher's model with the predictors is significantly different from the one with the constant only (all 'b' coefficients being zero). It measures the improvement in fit that the explanatory variables make compared to the null model.

Chi-square is used to assess the significance of this ratio.

- $H_0$: There is no difference between the null model and the final model.
- $H_1$: There is a difference between the null model and the final mode