Study the Trip Attraction Rates and Developing the Trip Attraction Models for Mega Shopping Centers of Pakistan

By

Muhammad Hassam Jadoon

(00000328518)

A thesis submitted in partial fulfillment of

The requirements for the degree of

Master of Science

In

Transportation Engineering

NATIONAL INSTITUTE OF TRANSPORTATION (NIT) RISALPUR

DEPARTMENT OF CIVIL ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY

ISLAMABAD, PAKISTAN.

November, 2023

Study the Trip Attraction Rates and Developing the Trip Attraction Models for Mega Shopping Centers of Pakistan

Author

Muhammad Hassam Jadoon

Registration Number

(00000328518)

A thesis submitted in partial fulfillment of the requirements for the degree of

MS Transportation Engineering

Thesis Supervisor:

Dr. Wasim Irshad Kayani

Thesis Supervisor's Signature: ____________________

NATIONAL INSTITUTE OF TRANSPORTATION (NIT) RISALPUR

DEPARTMENT OF CIVIL ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY

ISLAMABAD, PAKISTAN.

November 2023.

Thesis Acceptance Certificate

Certified that final copy of MS thesis written by **ENGR. MUHAMMAD HASSAM JADOON** Reg. No. 00000328518 of National Institute of Transportation, National University of Sciences & Technology, Risalpur Campus, has been vetted by the undersigned, found complete in all respects as per NUST Statutes / Regulations, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for the award of MS degree in Transportation Engineering. It is further certified that necessary amendments, as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature:

Name of Supervisor: Dr. Wasim Irshad Kayani

Date: 13/12/2023

Signature (HOD):

Date: 12/12/2023

Signature (Dean/Principal):

Date: 13/12/2023

Declaration

I certify that this research work titled **"Study the Trip Attraction Rates and Developing the Trip Attraction Models for Mega Shopping Centers of Pakistan"** is my work. The work has not been presented elsewhere for assessment. The material that has been used from other sources has been properly acknowledged/referred to.

Signature of Student

Muhammad Hassam Jadoon

(00000326823)

Plagiarism Certificate (Turnitin Report)

This thesis has been checked for Plagiarism. The Turnitin report endorsed by Supervisor is attached.

Signature of Student

Muhammad Hassam Jadoon

(00000328518)

Signature of Supervisor

DEDICATION

My beloved Parents.

For their support and cooperation, For their Inspiration and courage.

My kind Teachers.

For their parental guidance,

&

My Friends.

For sharing their insight……

ACKNOWLEDGEMENTS

I am thankful to ALLAH (SWT) for granting me courage and strength to complete my master's degree.

I am deeply indebted to my research advisor **Brig Dr. Wasim Irshad Kayani** for his support and guidance throughout the research period and for providing such a conductive environment for research at National Institute of Transportation (NIT), NUST. I would like to express my deepest gratitude to the committee members **Lt Col Dr. Muhammad Jawed Iqbal** and **Dr. Inamullah Khan**. I would like to express my deepest appreciation to all my fellows and friends who helped me throughout the course work and research work especially **Engr. Kamran Shafqat** who helped me a lot in the data collection phase. This endeavor would not have been possible without the help of management of the **Centaurus Mall, Packages Mall, Emporium Mall, Amanah Mall, Mall of Lahore and Fortress Square Mall**. In the end I pay my earnest gratitude to my loving **Parents** for their unending support, encouragement, prayers and patience.

(Engr. Muhammad Hassam Jadoon)

ABSTRACT

Shopping centers play an important role in the traffic impact of any city. The proliferation of shopping malls, particularly multi-purpose shopping centers (trans-marts) has caused a significant change in traffic patterns in the cities. To study the travel patterns of these shopping centers, there is a need for proper data and models for the travel demand forecasting. This will ultimately help in the utilization of optimum resources with maximum benefits. The modern revolutionized world requires planning, which necessitates the use of real data, models, travel patterns, and trip generation models.

This study bridged the gap by developing the trip attraction rates of two metropolitan cities. First, shopping centers with multiple facilities such as shopping, dining, restaurants, play areas, and cinemas are chosen for research in Islamabad and Lahore. Six shopping centers are selected, and data is collected every 15 minutes during the three peak hours on weekdays and weekends. For the collection of data, two days on both weekends and weekdays are selected. The data collected includes the number of people and vehicles entering and leaving the shopping centers for every 15 minutes intervals. The data related to physical features like gross floor area, shopping area, playing area, watching area, dining area, number of shops, number of parking spaces and number of stories is collected from the management of shopping centers.

Different statistical techniques like multiple linear regression followed by non- linear regression, partial least square regression, and artificial neural networks are used. All the models gave significant results. Pearson Correlations showed that all the explanatory variables have significant correlations with response variables except for the number of stories of shopping centers. The data is observed to be nonlinear and multicollinear. Logarithmic

regression is used to solve the problem of non-linearity and partial least square regression is used tackle the multicollinearity.

The logarithmic regression explained that gross floor area and number of shops are the most important variables for the shopping trips with R-square = 0.89 on weekends and R-square = 0.76 on weekdays. Partial least square regression showed R-square = 0.98 on weekdays and R-square = 0.87 on weekdays using all the explanatory variables except the number of stories. ANN models have the same results as partial least square regression with gross floor area, shopping area, playing area and no of shops as independent variables.

Several studies have been done on the trip attraction rates of shopping centers around the globe. There is a dire need for the transportation planners of Pakistan to plan effectively by using spatial models because every country has its own demographic, social and economic characteristic. The computed trip attraction models will be helpful for prediction, planning and managing the traffic around the existing and new shopping centers.

LIST OF ABBREVIATIONS

GFA - Gross Floor Area

TA - Trip Attraction

TAZ - Traffic Analysis Zone

SC - Shopping Center

SA - Shopping Area

MLRA - Multiple Linear Regression Analysis

ANN – Artificial Neural Network

PCA – Principal Component Analysis

PLSR – Partial Least Square Regression

ML – Machine Learning

LM Algorithm - Levenberg-Marquardt Algorithm

UNDP – United Nations Development Programme

TABLE OF CONTENTS

Contents

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

1.1 Background:

With the advancement of modern technology, adjustments have been made to every area of engineering as well as our social lives. The world population is growing every day, so adjustments must be made while taking the future into consideration. The transportation sector should likewise undergo these adjustments.

Pakistan is a developing country, and every developing country needs proper planning to utilize its limited resources. Transportation sector is one of the largest sectors of Pakistan and most of the freight is transported across the country using the road network. People also travel from rural areas to the cities due to greater opportunities in education, jobs, health, social life etc. This process is called urbanization. Urbanization has increased in the past two decades.

According to UNDP "In South Asia, Pakistan ranks highest in urbanization. According to the 2017 Population Census, 36.4% of the population lives in urban areas. The UN Population Division estimates that, by 2025, nearly half the country's population will be living in cities. Urbanization is generally considered to be closely related to economic growth, particularly in developed countries. Globally, it is estimated that cities generate more than 80% of the global GDP. The more urbanized areas indicate higher per capita income and more employment opportunities. Urbanization has positive impacts on technological innovation and economic progress."

Due to urbanization the traffic also increases in the metropolitan cities like Islamabad, Lahore, Peshawar, and Karachi. These cities are the main hub of business and economic activities. With the sprawl of cities, people prefer to travel in private vehicles which in turn causes an increase in the traffic. The supply of transportation infrastructure and services has not kept pace with this surge in demand. As a result, traffic congestion has become a common problem, causing delays, air pollution, and increased travel times.

Transportation planning includes the traditional four step travel demand forecasting method, which are trip generation, trip distribution, mode split and trip assignment. The first step i.e., trip generation has trip productions and trip attractions. Trip productions are related to residential areas while trip attractions related to non-residential trip ends like offices, shopping centers and schools which we called them as work trips, shopping trips and educational trips respectively.

One key aspect of transportation planning that traffic engineers and planners must consider is trip attraction. An area's traffic patterns and the number of vehicles on the road can be drastically altered when a new office building, shopping mall, or housing complex is built there. The number of trips created and their distribution can be affected by the development's location, size, and other characteristics.

The construction of a large office complex, for instance, near a major transit hub can increase demand for public transportation in the area. It may be essential to build additional roads and parking lots in order to accommodate the increased influx of cars if a large commercial complex attracts a large amount of vehicular traffic. Changes in traffic patterns may result from the construction of new residential areas, since inhabitants may decide to increase their car use or alter their commuting patterns.

Hasen et al showed that to effectively manage transportation planning issues in Asian cities, it is important for traffic engineers and planners to carefully consider the impact of trip attraction when evaluating new development proposals. This can help ensure that transportation infrastructure and services are designed to meet the evolving needs of the community while minimizing the negative impacts of increased travel demand.

The first step of travel demand forecasting is the basic and most important step in planning. Any error in this step can replicate into next steps and make the whole planning process wrong. In Pakistan, no study has been done before for the production and attraction rates. The transportation planning process in Pakistan uses the trip generation rates of US, which are not suitable for a developing country. This study mainly focuses on studying and developing the trips attraction model for the shopping centers of two metropolitan cities. One is Islamabad, the capital city of Pakistan, and the second is Lahore, the capital city of Punjab province.

1.2 Important Definitions:

Total Covered Area: It is the area that is covered by the shopping center which includes the floor areas, open parking, basements etc.

Gross Floor Area: This includes the floor area but not the basements for parking.

Shopping Area: All the areas of shops and corridors inside the shopping mall which people use for the sake of shopping activity.

Watching area: Cinemas are included in this area.

Dining Area: This area comprises of restaurants, food courts and area for sitting and eating.

Playing Area: This is the area for the children to play.

Vehicular Trips: Number of vehicles coming and leaving the shopping mall.

Personal Trips: Number of people entering and leaving the mall.

Trip Attraction: Number of trip ends connected to the non-residential areas.

Trip Production: Number of trip ends associated with the residential land use.

1.3 Problem Statement:

Trans-marts have a significant impact on the traffic impact of any city. The increase in shopping malls has caused the change in travel patterns of the cities. To study these travel patterns real data and models play a crucial role.

Transportation planning and travel demand forecasting in Pakistan uses the models of developed countries. These models are not suitable for the actual conditions of this country. No study has been carried out to find the parameters and to develop the trip attraction models for Pakistan. This research is an attempt to develop the trip attraction rates for the shopping centers of metropolitan cities.

1.4 Objectives:

The main objective is to develop trip attraction model for the multi-purpose shopping centers of Pakistan, mainly two metro-Politian cities i.e., Islamabad and Lahore.

1.5 Outcome:

Personal and vehicular trip attraction models will be developed for the weekdays and weekends using statistical methods and machine learning techniques.

1.6 Organization of Thesis:

Chapter 2: This chapter is related to the past studies and techniques which have been done in other countries.

Chapter 3: In this chapter, methodology which is adopted throughout the research is discussed.

Chapter 4: This chapter shows how and which data is collected from different shopping centers.

Chapter 5: This chapter explains the results of different statistical and machine learning techniques in the form of models.

Chapter 6: Conclusion and different recommendations are made in this chapter.

Chapter 2

LITERATURE REVIEW

2.1 General:

The accuracy in trip generation model is equally important since subsequent stages of transport modeling are highly dependent on the trip generation output. Trip attraction and trip production are the two main components that play an important role in trip generation. Shopping centers are known to generate significant amounts of vehicular traffic, which can have a significant impact on the surrounding transportation network. Accurately estimating the trip attraction rates for shopping centers is crucial for transportation planners and engineers to properly assess this impact and design appropriate transportation infrastructure and services. In the past, there has been growing interest in using statistical and machine learning methods to develop and model trip attraction rates for shopping centers. This literature review provides an overview of the past and current state of research for the development of trip generation rates.

Travel demand forecasting is crucial in designing transportation facilities and services, as well as in planning, investing, and developing policies. The initial stage in the conventional four-step process of travel demand forecasting is trip generation, which is of utmost importance as the accuracy of the values produced in this stage serve as the foundation for the subsequent steps. Any errors in this stage can propagate throughout the entire estimation process, making it critical to ensure its precision. To forecast travel demand accurately, it is necessary to divide the study area into smaller units known as traffic analysis zones (TAZ). These TAZs serve as the building blocks for analyzing travel movements within, into, and out of the urban area. For trip attraction models the multiple regression technique was preferred based on the variables used (Sosslau, Hassam, Carter, & Wickstrom, 1978). Equations based on regression analysis represent 'best fits' through the data points (Manual, 2000).

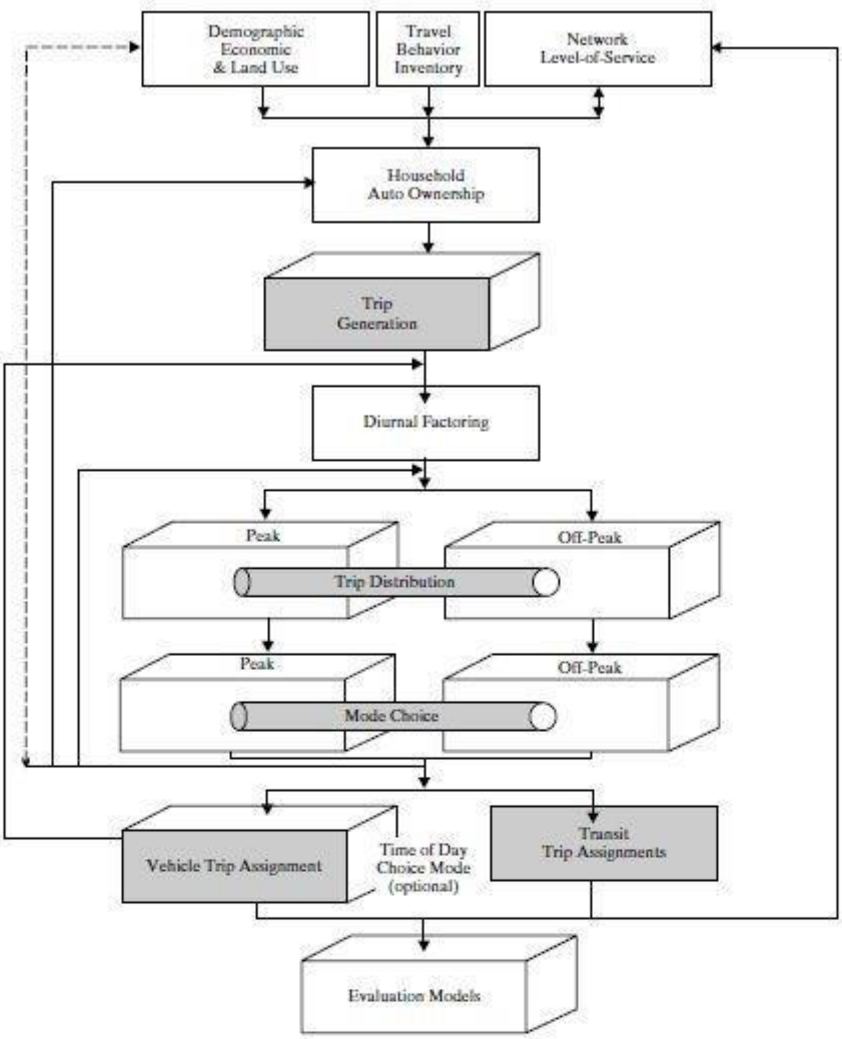


Figure 2-1 Travel Demand Model Flow Chart
(SAMEER, ISLAM, & LIANA, 2015)

2.2 Past Studies Related to Trip Attraction Rates:

(Shamim Al Razib, 2017) estimated trip attraction rates using the trip rate analysis method. The study involved surveying six shopping centers to count the number of vehicles and people entering the centers during peak hours at 15-minute intervals. The data collected from the survey is used to establish a relationship between the trip attraction rates of the shopping centers. Several physical features of the shopping centers, such as total parking spaces, gross floor area, and the number of stores, were considered in the study. The results estimated the traffic volume of newly constructed shopping centers and their impact on the geometric design of roadways in the surrounding area.

(PASRA, RAMLI, ADISASMITA, & PALLU, 2013) conducted a study that builds a relationship model between the number of visitors and the number of kiosks on the traditional markets in Makassar, Indonesia. Additionally, the research investigated the elements that greatly influenced persons in attaining their shopping journey frequency. This study used data of the number of customers and stalls at six different traditional marketplaces. It also considered the socio-demographic features of the households, the particulars of a trip to a conventional market for shopping, and the travel characteristics of duration and mode. The research has used regression analysis as well as the multinomial logit model to create the phenomenon of trip attractiveness as well as the phenomenon of trip frequency in relation to traditional marketplaces. There is a significant correlation between the number of kiosks selling accessories and household goods and the total number of trip attractions. In addition, the state of the house, the number of members in the family, level of income, amount of money spent, and length of time spent at the market are the key elements that influence trip frequencies.

To create a travel attraction model for the education area in Gunung Pangilun, Padang City, data was collected through random sampling and a questionnaire containing 17 variables from 200 respondents. The researchers utilized crosstabs and multiple linear regression analysis to comprehend trip attraction patterns in the area. The results of the crosstabs analysis revealed a correlation between various variables such as gender, destination, mode of transportation used, reasons for mode selection, distance, travel time, and congestion to the overall number of trips. Multiple linear regression techniques were used to explain the trip attraction amount through the Building Floor, Trip Distance, and Age variables. (Roza, Yusnita, & Rusli, 2021)

The trip attraction model for commercial land use in medium-sized towns of Kerala was developed by (George, Kattor, & technology, 2013) using multiple regression methods. In the three medium-sized towns of Kerala—Tirur, Perinthalmanna, and Ponnani, eight commercial nodes are randomly selected to participate in a questionnaire survey. The origin-destination data was obtained from socioeconomic surveys in the selected municipalities. These survey results serve as the basis for the correlation and regression analyses. Using the number of businesses, the proportion of stores in the commercial node, and the proportion of banks in the commercial node as independent variables, the study found a multiple regression model with an R-squared value of 0.99. In addition to formulating the model, we then examined the multiple regression assumptions to ensure its accuracy.

Sustainable development plans often include looking forward to identifying potential impacts. Future difficulties can be reduced via careful planning. (Rahayu, Basuki, & Gritanarum, 2020) implemented the concept of sustainable development and developed model that is used to predict the effects of growth. The purpose of the study was to identify must-see destinations near the Trans-mart four-in-one shopping centers. Trip rate analysis and a multiple linear regression model were employed for this research. He developed a trip attraction model using shopping area, dining area, playing area and watching area. The model predicted the amount of traffic congestion that would result from the development of multiple purpose shopping malls.

Another study with objectives to develop models to estimate trip attraction rates for person trips and vehicle trips was carried out by (Suthakaran, Dharmakeerthi, & Sathyaprasad). In addition to this, parking regulations for three-wheelers, motorbikes, vehicles, and vans were devised and implemented for supermarkets in the Kandy region. During peak hours, a survey will be conducted using an Android application designed for use on smart phones to count the number of people and cars entering and exiting the supermarkets at regular intervals of fifteen minutes. The data was analyzed using a method known as multiple linear regression analysis. The central limit theorem was utilized in the process of formulating guidelines for parking. A more complex model was designed for supermarket designers, while a simpler model was constructed for local authorities. Both models were created for each category. According to the findings, the wine shop is the primary factor that determines the rate of trip attractiveness as well as the parking standards in the Kandy region.

The objective of the study conducted by (Waloejo & Surjono, 2012) is to discover the characteristics of trip attraction in relation to market land usage, as well as the elements that influence it. The researchers intend to develop a model that accurately represents the trip attraction at the Dinoyo market area. In conclusion, this study aims to examine the correlation between the attraction of Dinoyo Market as a tourist destination and the level of service provided by the Mayjen Haryono Street. The study employed Pearson correlation analysis, followed by regression analysis using the stepwise technique and ANOVA, to assess and develop a model that examines the impact of various factors on the level of service of road. The study revealed that the primary source of the market visitations originated predominantly from residential regions (90%), with a particular emphasis on the Lowokwaru District. Public transport was the prevailing means of transportation. The present investigation developed a model to assess the factors influencing trip attraction at Dinoyo market. The dependent variable in this model is trip attraction, while the independent variables include the floor area of the meat and fish zone, daily basic requirement products, fruits, vegetables, and parking area. The Dinoyo market accounted for 7.18% of the daily vehicle volume on Mayjen Haryono Street. The impact of road level of service (LOS) on volume per capacity has been observed to decline. This suggests a direct correlation between trip attraction and land use, resulting in an effect on the LOS of the road network system.

A study conducted by (Uddin et al., 2012) in Bangladesh employed the trip rate analysis approach to assess the trip attractiveness rates of shopping centers located on Mirpur Road in the Dhanmondi region of Dhaka. As part of the research conducted, a comprehensive survey was conducted on a total of six shopping malls. The purpose of this survey was to record the number of vehicles (specifically cars) and individuals entering the shopping centers during the peak hour, with data collected at 15-minute intervals. The study examined the relationship between the trip attraction of a shopping center and its physical characteristics, such as the total number of parking spaces, the gross floor area, and the number of retailers within the shopping center.

(Majeed & Qasim, 2021) studied the trip attraction rates of selected zones of Baghdad city. The selected zone had educational centers, state institutions, shopping centers and health centers. They developed six types of attraction rates including educational, state institutions, shopping, health, others, and total trips of the selected zones. Regression analysis was performed in SPSS and models was developed. the important variable for educational trips is number of students and for shopping trips number of shops was most significant variable. Health trips showed that the number of beds and number of patients in the consulting and emergency clinics are the most important variables. The total trip attraction of the study area was described best using number of employees, number of health staff, number of visitors and number of students.

(Yulianto, Purnomo, & Prasetyo, 2020) studied the trip attraction rates for hotels and minimarket, in Surakarta City. 3-star, 4-star and 5-star hotels were selected for the analysis. Trip attraction models are developed using regression analysis and number of rooms as an independent variable. Trip attraction model for the mini markets is developed with building floor area as an explanatory variable using regression analysis.

Models based on multiple regressions (Fillone, Tecson, Sia, & Viray, 2003) found that floor area was associated with trip attraction. Thirty condominiums were chosen at random in Metro Manila, and a survey questionnaire was used to collect data on the condos in order to learn about their features and use that information in a modeling exercise. Information collected includes residential square footage, retail space, parking spots, commercial space, occupancy rates, building age, number of stories, and maintenance costs.

Chapter 3

METHODOLOGY

3.1 General:

The overall framework of this research methodology is given in Figure 3-1**.** The trip data is collected for each shopping center for 15 minutes time interval during the three peak hours on weekdays and weekends. Data related to physical features of shopping centers is taken from the respective management. The statistical analysis has been done by SPSS while the ANN model is made using MATLAB.

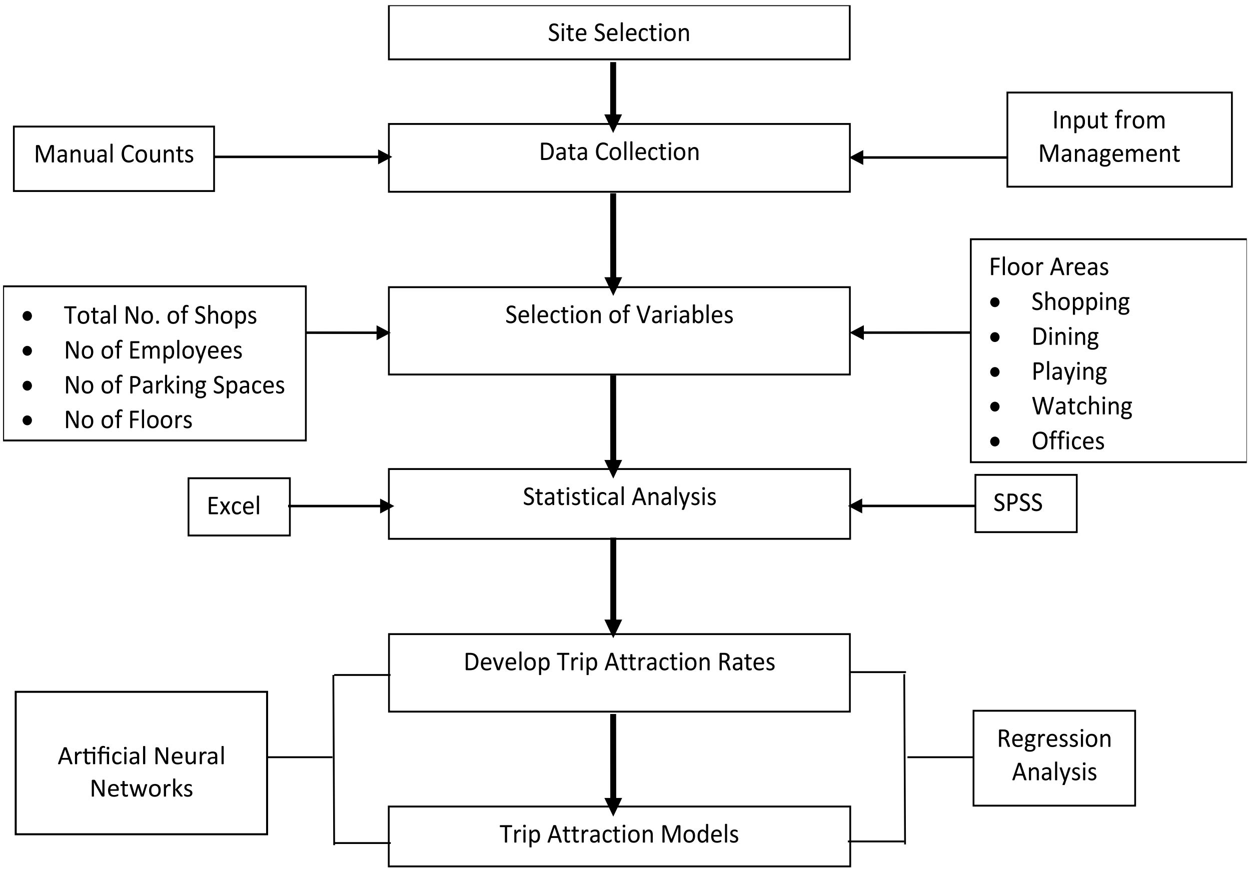


Figure 3-1 Study Flow Chart

3.2 Study Area:

Shopping Centers are selected across the different places so we can get the actual picture of the trips. One shopping center i.e., The Centaurus Mall is selected from the Islamabad and the five shopping centers are selected from the Lahore. The location and pictures of each shopping mall is given:

3.2.1 The Centaurus Mall:

The Centaurus Mall is in Sector F-8 near the Blue Area which is the Central Business District of Islamabad along Jinnah Avenue. Multi story buildings are present in this area and the road leads to the government buildings. This shopping center is selected because it is the largest shopping mall in the area, and it better represents the trip attraction of the whole city due to its size and location. The Centaurus mall has three towers with 36 floors. There are five floors for shopping while the remaining floors include residential flats and offices. The data is collected only for the shopping center because of the study scope. Separate parking areas are dedicated for vehicles coming for shopping.

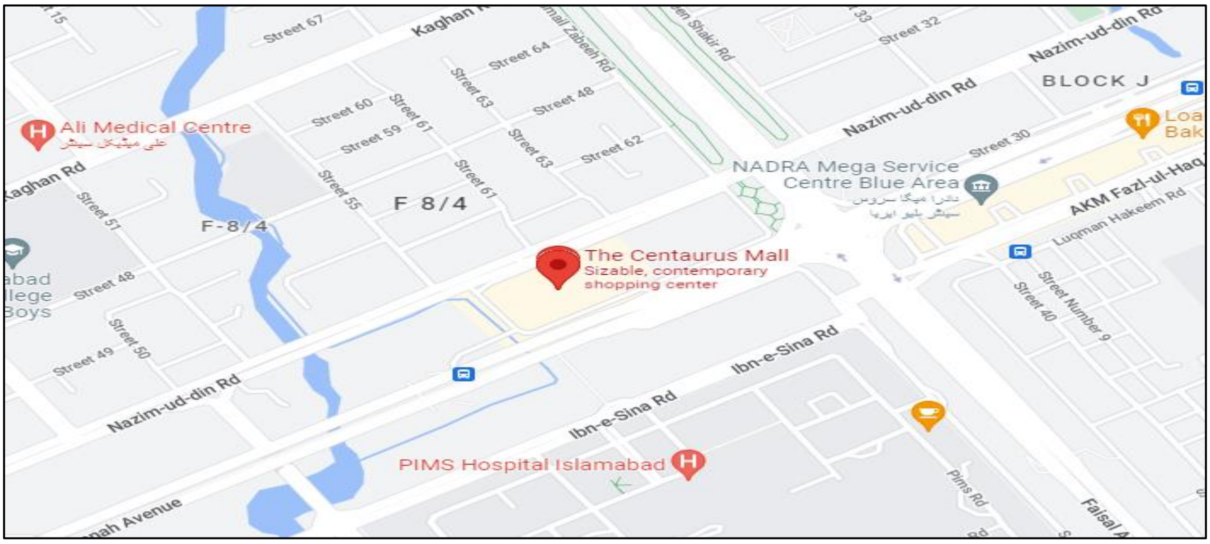


Figure 3-2 Map Location of Centaurus Mall

Figure 3-3 Pictures of Centaurus Mall

3.2.2 Packages Mall:

Packages Mall is one of the biggest shopping malls which is located on Walton Road, Lahore. Packages Mall is dedicated to shopping and does not have residential flats, guest rooms or offices. It has a big cinema, food courts, a playing area and branded shops. The packages mall has three stories and a gross floor area of 1200000 square feet. Indoor and outdoor parking facilities are available with 1700 parking spaces. There are more than 200 shops selling cosmetics, clothing, footwear and many more. This mall has the largest footfall (number of people coming to mall) as compared to other shopping centers of Lahore.

Figure 3-4 Map Location of Packages Mall

Figure 3-5 Pictures of Packages Mall

3.2.3 Emporium Mall:

Emporium Mall is another major mall of Lahore and located near the Expo-Center in Johar Town, Lahore. It is also the busiest mall like Packages Mall, and it also comes under the category of Mega Mall. It has 4 floors dedicated to shopping, playing, food and cinema. Other floors have guest rooms and halls. Parking is in the basement with 1500 parking spaces. The data is collected only for the shopping center as per the scope of study. There are more than 200 branded shops.

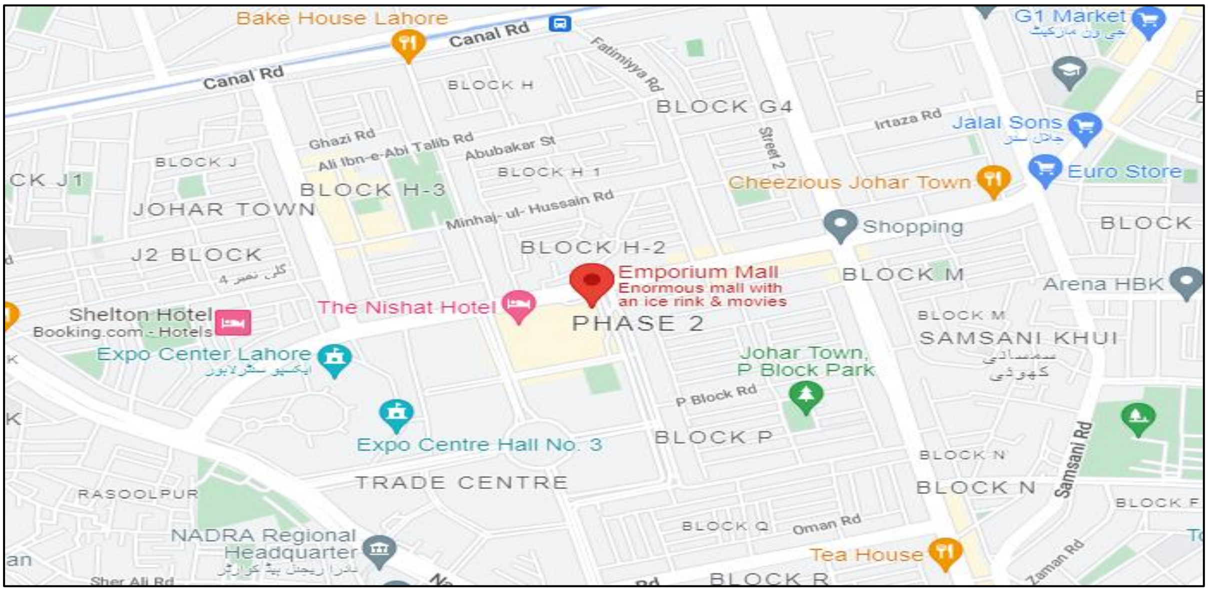


Figure 3-6 Map Location of Emporium Mall

Figure 3-7 Pictures of Emporium Mall

3.2.4 Amanah Mall:

Amanah Mall is somewhat smaller mall than Packages and Emporium Mall. It is located on Model Town Link Road, Lahore. In this mall 4 floors have been dedicated to shopping with 40 branded stores and one floor for food and one floor for playing and cinema. So, there are total of 6 floors having gross floor area of 300000 square feet. There are 400 parking spaces in the basement. This mall comes under the category of US shopping centers classification.

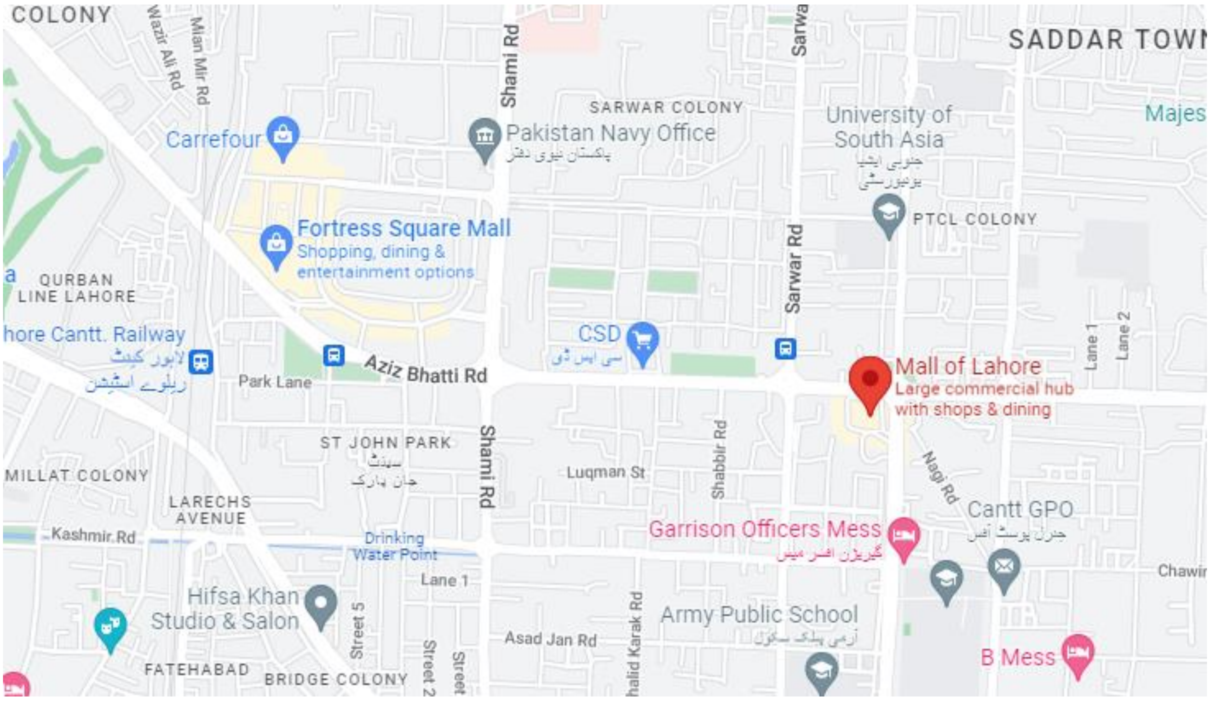


Figure 3-8 Map Location of Amanah Mall

Figure 3-9 Pictures of Amanah Mall

3.2.5 Mall of Lahore:

Mall of Lahore is located at 172-Tufail Road, Lahore Cantt. This mall has five floors and has shopping, playing, and watching areas along with food courts. The mall has gross floor area of 275000 square feet. It has a separate building for parking, having 200 parking spaces.

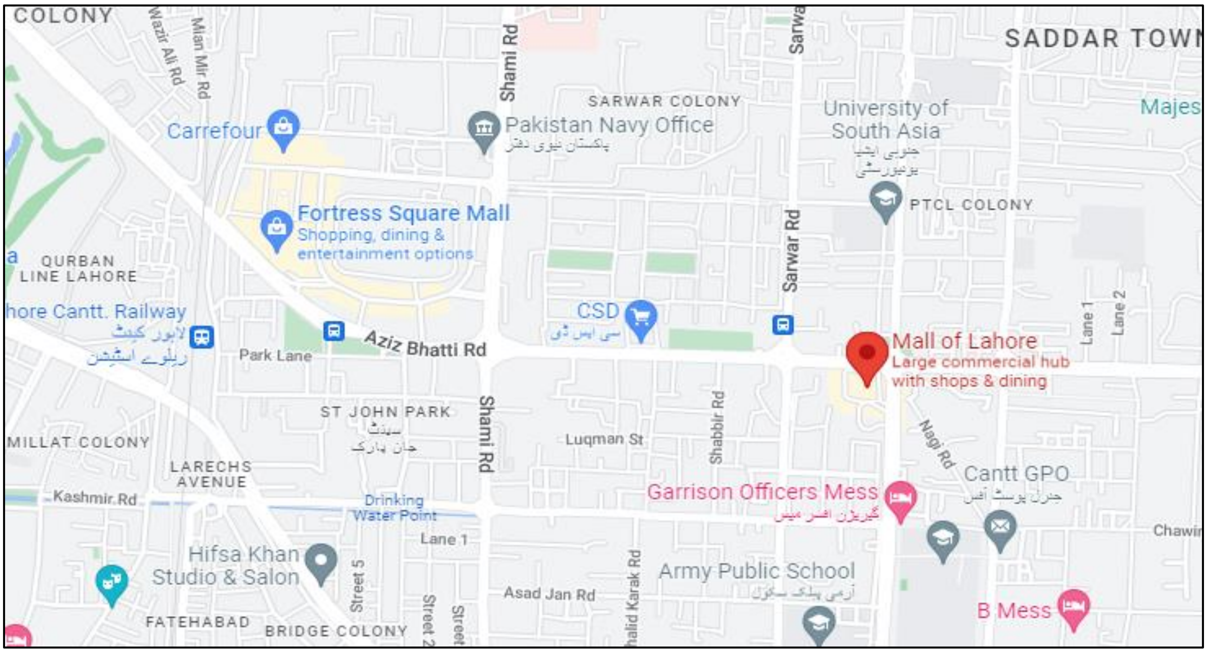


Figure 3-10 Map Location of Mall of Lahore

Figure 3-11 Pictures of Mall of Lahore

3.2.6 Fortress Square Mall:

Fortress Square Mall is a popular shopping mall on Fortress Stadium Road in the Cantonment area, Lahore. The gross floor area of the mall is 135000 square feet having 3 stories. The mall has shopping, playing and dining facilities with 60 shops. There are 400 parking spaces.

Figure 3-12 Map Location of Fortress Square Mall

Figure 3-13 Pictures of Fortress Square Mall

3.3 Data Collection:

The data collection process consists of searching for shopping malls that have multiple activities like shopping, watching, playing, dining etc. The process is divided into two stages. First collection of data from Islamabad and then Lahore.

In Islamabad, Centaurus Mall was selected. In Centaurus Mall, there are residential flats, offices, and a shopping mall. As our scope of study is only related to shopping trips, data is collected only for the shopping as Centaurus mall has separate parking spaces for residential, offices and shopping. for the three peak hours for two weekdays and 2 weekends. The peak hours are 07:00 PM to 10:00 PM. Data is collected for every 15 minutes interval for both People and Vehicles coming and going from the shopping mall.

In Lahore five shopping malls are selected which are Packages Mall, Emporium Mall, Amanah Mall, Mall of Lahore, and Fortress Square Mall. The US shopping centers classification can not be used for the Pakistani shopping centers because the shopping centers of US have much more larger areas than Pakistan. So, the shopping centers are of Pakistan is classified based on Gross Floor Area. We got three types of shopping centers large, medium and small. Shopping centers with GFA greater than 1 million square feet is considered as large, GFA between 0.5 to 1 million square feet is considered as medium and GFA with less than less than 0.5 million square feet is considered as small. Based on these classification Packages mall and Emporium mall is in the category of large malls, Centaurus mall is a medium mall while Amanah mall. Mall of Lahore and Fortress Square mall are small malls.

Data is collected only for shopping malls for 3 peak hours and 2 days on weekend and weekdays. Surveyors were present at each gate of the mall, and they recorded the number of people and vehicles coming and leaving from the shopping mall for each 15-min interval.

3.3.1 Traffic Data:

The peak hour trips are given as:

Table 3-1 Peak hour Trips for Weekend and Weekday

	Weekend		Weekday	
	Peak hour Volume of People	Peak hour Volume Vehicles	Peak hour Volume of People	Peak hour Volume Vehicles
Packages Mall	7499	2169	4702	1398
Emporium Mall	7225	2024	4361	1499
Centaurus Mall	7319	2052	4459	1127
Amanah Mall	4803	1330	3047	949
Mall of Lahore	1955	555	1548	382
Fortress Square Mall	2479	754	1996	497

Results show that Packages Mall has the greatest peak hour volume 7499 and 2169 for both people and vehicles respectively on weekend hile 4702 and 1398 for people and vehicles respectively on weekday. Centaurus Mall and then Emporium Mall have greater peak hour volumes. Mall of Lahore has the lowest peak hour volume. The average occupancy is calculated using this data and it is 3.5.

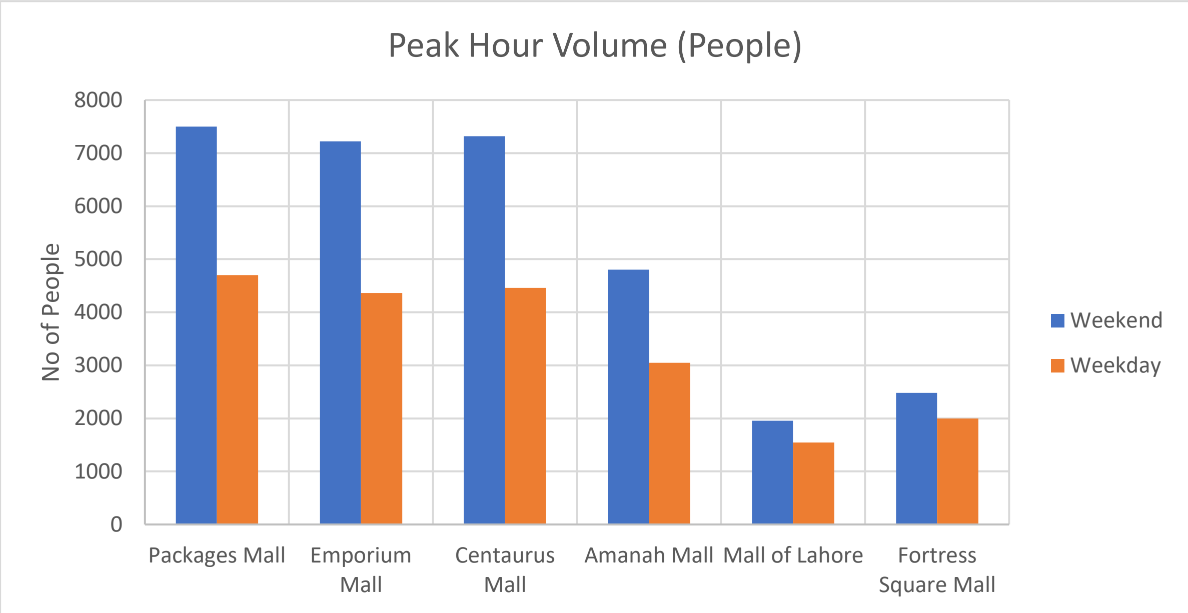


Figure 3-14 Personal Trips on Weekday and Weekend

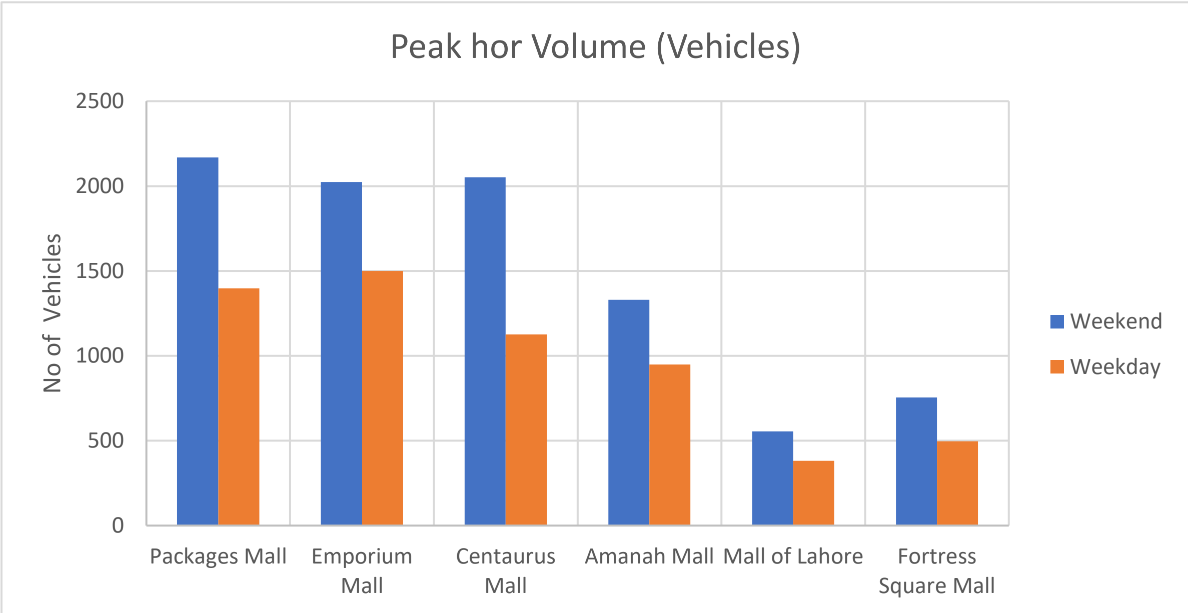


Figure 3-15 Vehicular Trips on Weekday and Weekend

3.3.2 Physical Features of Shopping Malls:

Data related to the physical features of shopping malls is taken from the management of the shopping malls. The features which are included in the study are Gross Floor Area (GFA), Shopping Area, Playing Area, Dining Area, Watching Area, Number of Shops, Number of Parking Spaces and No of Stories.

Table 3-2 Data of Physical features of shopping malls

	Gross Floor Area (sq.ft)	Shopping Area (sq.ft)	Playing Area (sq.ft)	Watching Area (sq.ft)	Dining Area (sq.ft)	No of Shops	Parking Spaces	Number of Stories
Packages Mall	1200000	948000	60000	100000	80000	210	1700	3
Emporium Mall	1200000	800000	42000	60000	85000	200	1500	4
Centaurus Mall	800000	537492	22446	15420	22735	250	1111	5
Amanah Mall	300000	184000	13500	13500	16000	40	400	6
Mall of Lahore	275000	136000	12000	14487	29000	94	200	5
Fortress Square Mall	135000	90000	11000	0	12000	60	400	3

Data showed that Packages Mall and the Emporium Mall have the largest GFA and shopping areas while Fortress Square Mall has the smallest areas. The comparison of GFA and shopping areas are shown in Figure 3-16.

Figure 3-16 GFA and Shopping Areas

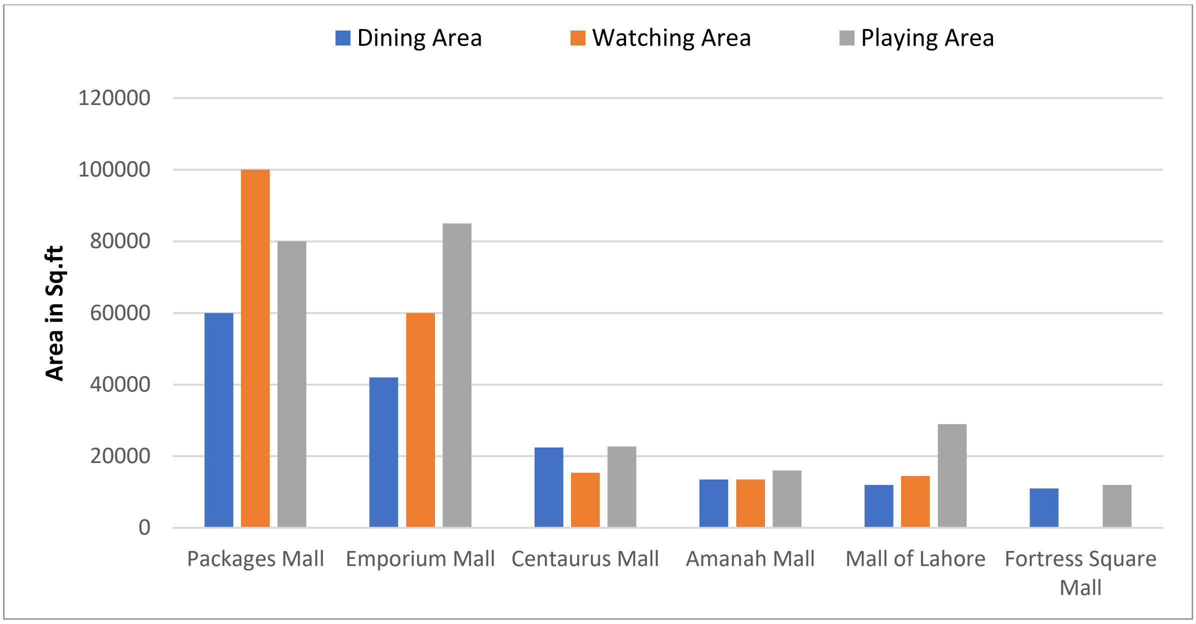


Figure 3-17 Dining, Watching and Playing Area of Shopping Malls

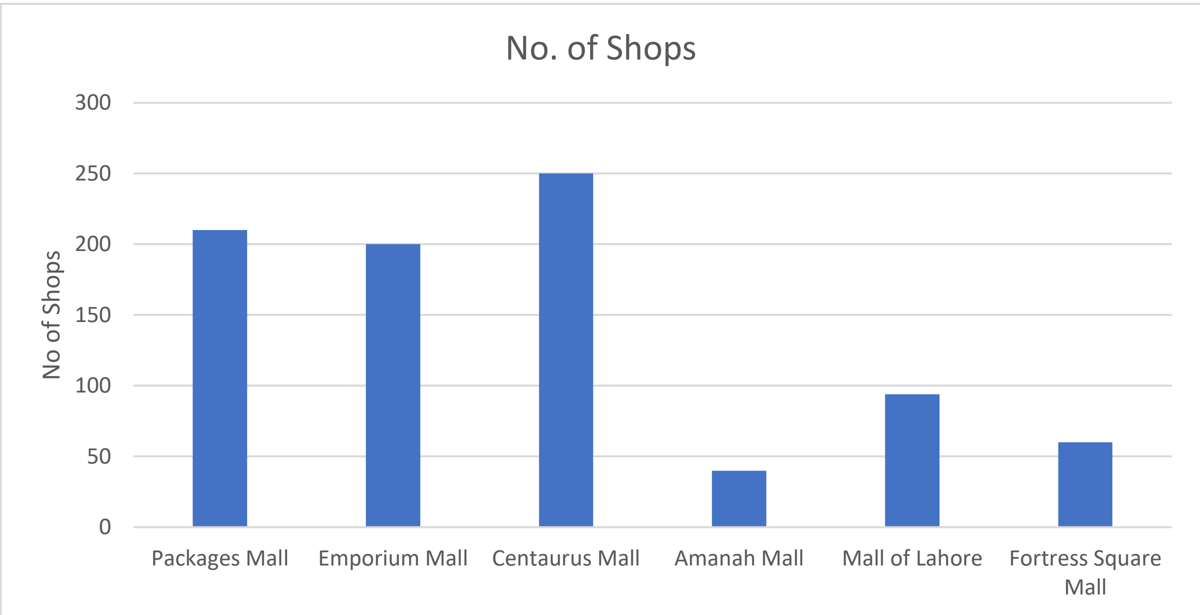


Figure 3-18 No of Shops for Shopping Malls

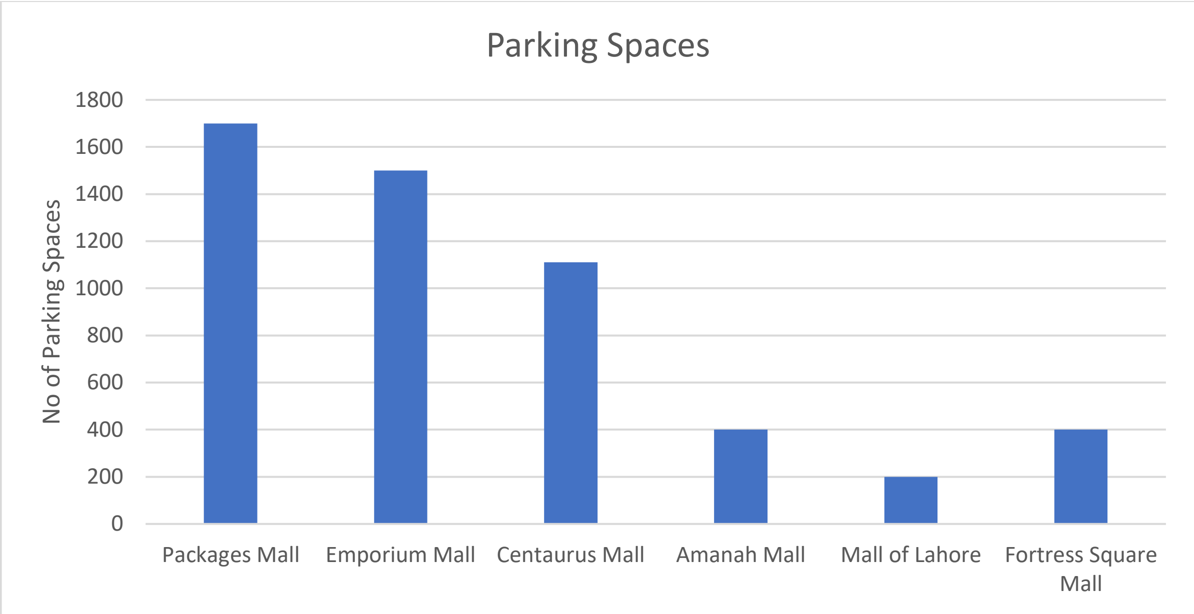


Figure 3-19 No of Parking Spaces for Shopping Malls

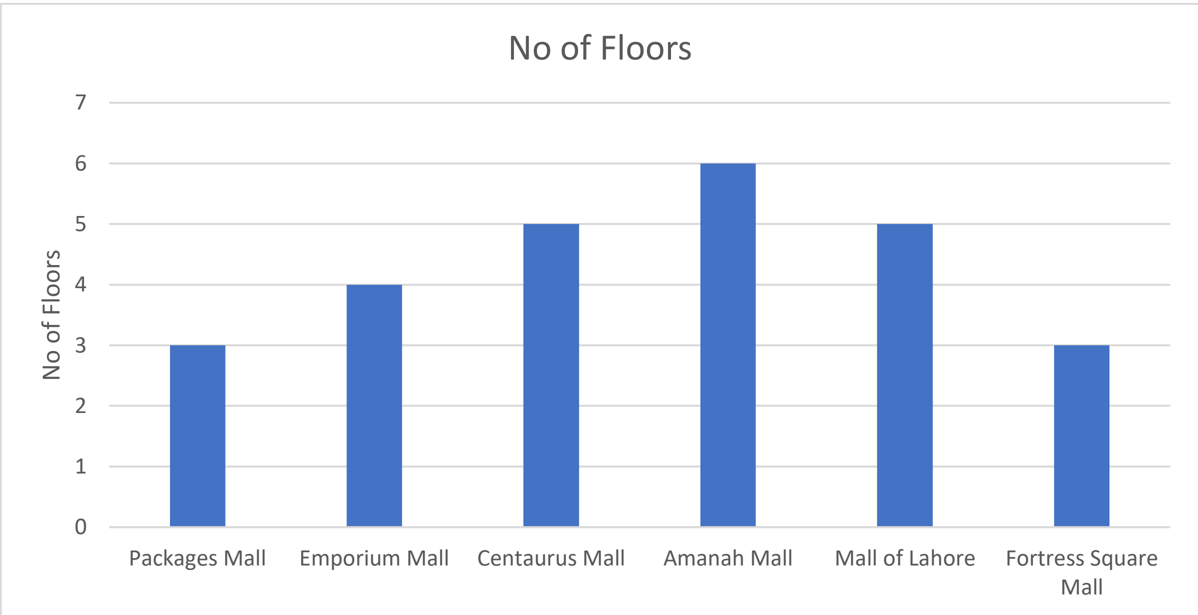


Figure 3-20 No of Floors for Shopping Malls

Traffic data is used as the response (dependent) variable while the features of shopping centers are used as the explanatory (independent) variables.

3.4 Regression Analysis:

Regression analysis is a statistical method used to analyze the relationship between a dependent variable and one or more independent variables. In the context of trip attraction rates for shopping centers, regression analysis is used to model the relationship between the number of trips generated by a shopping center and various factors such as the size of the shopping center, the number of stores, and the surrounding land uses.

Trip Generation Manual analyzed both the independent variable and the total number of trips to produce a regression curve, a regression equation, and a coefficient of determination (R2). The computer program chooses the equation for regression that has the highest R2 value and then plots that equation.

3.5 SPSS Analysis:

SPSS is used for developing the regression models. Pearson Correlations are used to check how the response (dependent) variables are related to explanatory (independent) variables and how explanatory variables are related to each other. This analysis is used for deciding the variables which should be included in the model.

Step wise linear regression analysis is done, and four models are developed two for weekdays and two for weekends. Personal Trip Attraction Model and Vehicular Trip Attraction Model are made for both weekdays and weekends.

3.6 Artificial Neural Networks (ANN):

Artificial neural networks (ANNs) are machine learning models inspired by the structure and function of biological neural networks. ANNs are composed of interconnected nodes or neurons that process information and make predictions based on input data. In the context of trip attraction rates for shopping centers.

3.7 MATLAB Analysis:

Four ANN models are made in total. Two for weekdays and two for weekends. Levenberg-Marquardt (LM) Algorithm is used in the analysis and optimum number of neuros in hidden layer is found using RMSE of training and validation model.

Because of its many features and extensive toolkits tailored for data analysis and machine learning, Matlab is crucial to Artificial Neural Networks (ANN) research. Because it offers an environment that is easy to use for prototyping and testing with multiple architectures, Matlab is widely used by academics to develop, implement, and fine-tune ANN models. The extensive array of pre-built functions expedites the whole research process by making feature extraction, performance evaluation, and data preprocessing easier. Moreover, researchers may assess and clarify the intricate patterns learned by neural networks thanks to Matlab's visualization features, which improves comprehension of model behavior.

CHAPTER 4

ANALYSIS AND RESULTS

The analysis and results are mainly divided into two categories, one is for weekdays, and the other is for weekends. Then each category is further divided into sub-categories i.e., Regression Analysis and ANN. Data is analyzed, and results are developed for People and Vehicles for every 15 minutes during three peak hours. The Statistical Analysis is done using SPSS software while the Artificial Neural Network model is trained and validated using MATLAB.

4.1 Weekend Results:

Weekends are most important to study the trip attraction rates. First scattered plot is formed using SPSS software so we can get the exact picture of dependent and independent variables.

4.1.1 Scattered Plot Matrix:

The scattered plot showed that the variables have variable degree of correlations.

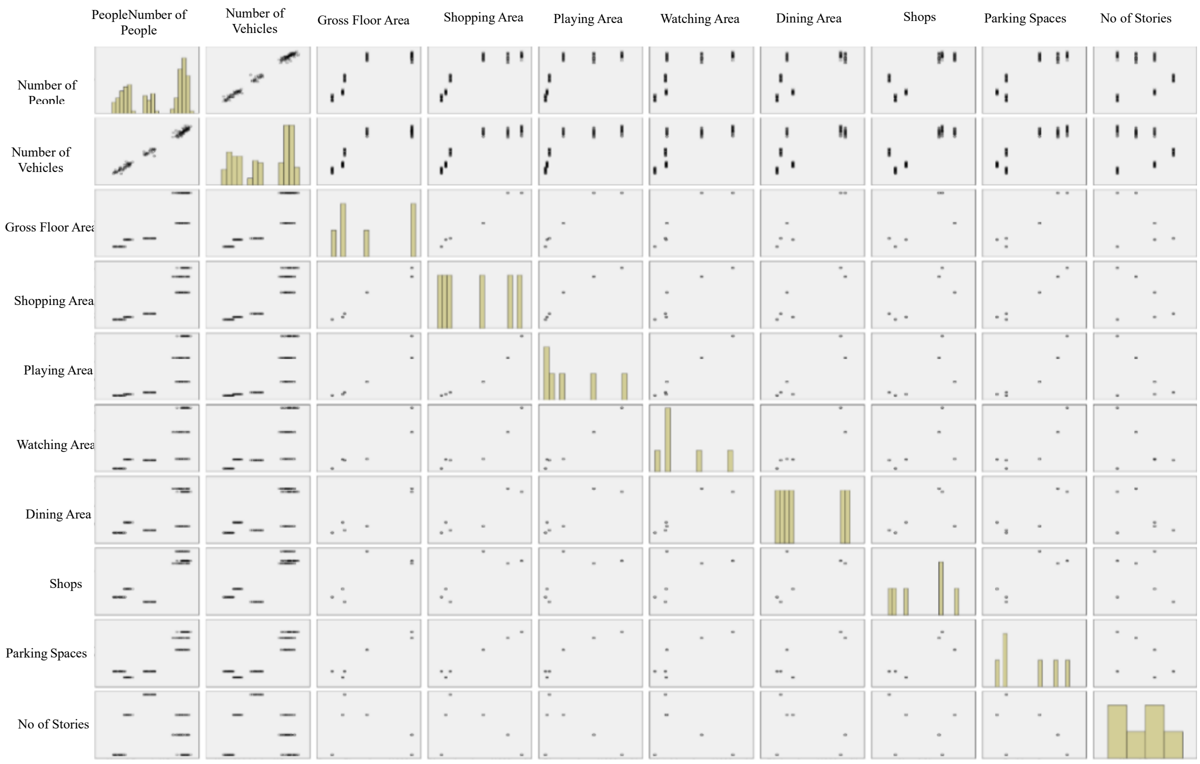


Figure 4-1 Scattered Plot for Weekend Data

It can be seen in the scattered plot that all independent variables except number of stories have almost linear or logarithmic relation with the dependent variable.

4.1.2 Pearson Correlation:

Pearson correlation is the ratio of covariance between two variables and the product of their standard deviation. Pearson correlations show that explanatory variables are highly correlated to the response variable. Only the number of stories have negative and non-significant value of correlation with the response variable. So, all the variables will be used in the analysis except the number of stories. One thing also important in the scattered plot matrix is that the independent variables are also highly correlated with each other, so there will be a chance of multi-collinearity in the model.

Table 4-1 Correlation Matrix

Correlations									
	No of People/15 min	Gross Floor Area	Shopping Area	Playing Area	Watching Area	Dining Area	No of Shops	Parking Spaces	No of Stories
No of People/15 min	1								
Gross Floor Area	.856	1							
Shopping Area	.902	.981	1						
Playing Area	.776	.951	.961	1					
Watching Area	.702	.917	.910	.984	1				
Dining Area	.680	.956	.894	.917	.920	1			
No of Shops	.861	.761	.832	.683	.579	.601	1		
Parking Spaces	.899	.958	.988	.936	.863	.851	.843	1	
No of Stories	-.051	-.381	-.402	-.522	-.478	-.451	-.244	-.456	1

4.1.3 Regression Analysis:

Stepwise Ordinary Least Square Regression is performed in SPSS, the main advantage of this Stepwise regression is that this starts from one independent variable and then adding independent variables one by one. Seven models have been made, and the last model has the greatest R-Square value of 0.984.

4.1.3.1 Number of Personal Trips:

4.1.3.1.1 Model 1

Table 4-2 Model Summary of Regression

Model Summary[h]			
Model	R Square	Adjusted R Square	Std. Error of the Estimate
1	.813	.811	253.099
2	.919	.918	166.933
3	.941	.940	142.972
4	.968	.967	106.219
5	.967	.966	106.749
6	.983	.983	77.027
7	.984	.984	73.903
a. Predictors: (Constant), Shopping Area (1000sq.ft)			
b. Predictors: (Constant), Shopping Area (1000sq.ft), Playing Area (1000sq.ft)			
c. Predictors: (Constant), Shopping Area (1000sq.ft), Playing Area (1000sq.ft), Dining Area (1000sq.ft)			
d. Predictors: (Constant), Shopping Area (1000sq.ft), Playing Area (1000sq.ft), Dining Area (1000sq.ft), Gross Floor Area (1000sq.ft)			
e. Predictors: (Constant), Playing Area (1000sq.ft), Dining Area (1000sq.ft), Gross Floor Area (1000sq.ft)			
f. Predictors: (Constant), Playing Area (1000sq.ft), Dining Area (1000sq.ft), Gross Floor Area (1000sq.ft), Watching Area (1000sq.ft)			
g. Predictors: (Constant), Playing Area (1000sq.ft), Dining Area (1000sq.ft), Gross Floor Area (1000sq.ft), Watching Area (1000sq.ft), No of Shops			
h. Dependent Variable: No of People/15 min			

R square is coefficient of determination, and it considers all the variables in the model while adjusted R square considers only those variables which have significant effect on the model and it eliminates the impact of non-significant variables in the model.

Table 4-3 Regression Coefficients

Coefficients[a]								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	525.743	35.220		14.927	.000		
	Shopping Area	1.559	.063	.902	24.825	.000	1.000	1.000

	(1000sq.ft)							
2	(Constant)	652.393	25.022		26.073	.000		
	Shopping Area (1000sq.ft)	3.519	.150	2.035	23.493	.000	.076	13.082
	Playing Area (1000sq.ft)	-37.547	2.757	-1.180	-13.617	.000	.076	13.082
3	(Constant)	657.800	21.444		30.676	.000		
	Shopping Area (1000sq.ft)	3.631	.129	2.100	28.099	.000	.075	13.273
	Playing Area (1000sq.ft)	-28.617	2.665	-.899	-10.737	.000	.060	16.664
	Dining Area (1000sq.ft)	-7.240	1.002	-.374	-7.226	.000	.157	6.364
4	(Constant)	511.351	20.997		24.353	.000		
	Shopping Area (1000sq.ft)	-.635	.410	-.367	-1.550	.123	.004	241.899
	Playing Area (1000sq.ft)	-4.926	2.969	-.155	-1.659	.099	.027	37.469
	Dining Area (1000sq.ft)	-34.991	2.697	-1.807	-12.976	.000	.012	83.531
	Gross Floor Area (1000sq.ft)	4.139	.387	3.093	10.707	.000	.003	359.149
5	(Constant)	532.854	15.839		33.642	.000		

	Playing Area (1000sq.ft)	-8.828	1.583	-.277	-5.578	.000	.095	10.539
	Dining Area (1000sq.ft)	-31.118	1.019	-1.607	-30.544	.000	.085	11.805
	Gross Floor Area (1000sq.ft)	3.557	.091	2.657	39.082	.000	.051	19.707
6	(Constant)	813.274	27.130		29.977	.000		
	Playing Area (1000sq.ft)	-46.208	3.473	-1.452	-13.305	.000	.010	97.474
	Dining Area (1000sq.ft)	-39.188	1.021	-2.024	-38.395	.000	.044	22.757
	Gross Floor Area (1000sq.ft)	4.315	.093	3.224	46.160	.000	.025	39.932
	Watching Area (1000sq.ft)	17.186	1.508	1.035	11.397	.000	.015	67.463
7	(Constant)	811.307	26.035		31.162	.000		
	Playing Area (1000sq.ft)	-49.852	3.482	-1.566	-14.317	.000	.009	106.441
	Dining Area (1000sq.ft)	-37.713	1.061	-1.948	-35.537	.000	.037	26.727
	Gross Floor Area (1000sq.ft)	4.120	.105	3.079	39.379	.000	.018	54.359

	Watching Area (1000sq.ft)	19.269	1.558	1.160	12.369	.000	.013	78.218
	No of Shops	.626	.174	.087	3.606	.000	.193	5.180
a. Dependent Variable: No of People/15 min								

An important thing to notice here is the Variance Influence Factor (VIF). VIF shows whether there is a multicollinearity present in a model or not. VIF should be less than 10 for a model without multicollinearity. If the value of VIF is greater than 10 then it means that there is an issue of multicollinearity in the model.

The issue of multi collinearity can be solved by removing the least important independent variable from the model one by one and then checking the VIF values of the remaining variables. If the values of VIF is less than 10 for all the variables, then we can say that now model has no multicollinearity. If VIF value is still greater than 10, then we must remove another independent variable from the model until we get VIF < 10 for remaining variables.

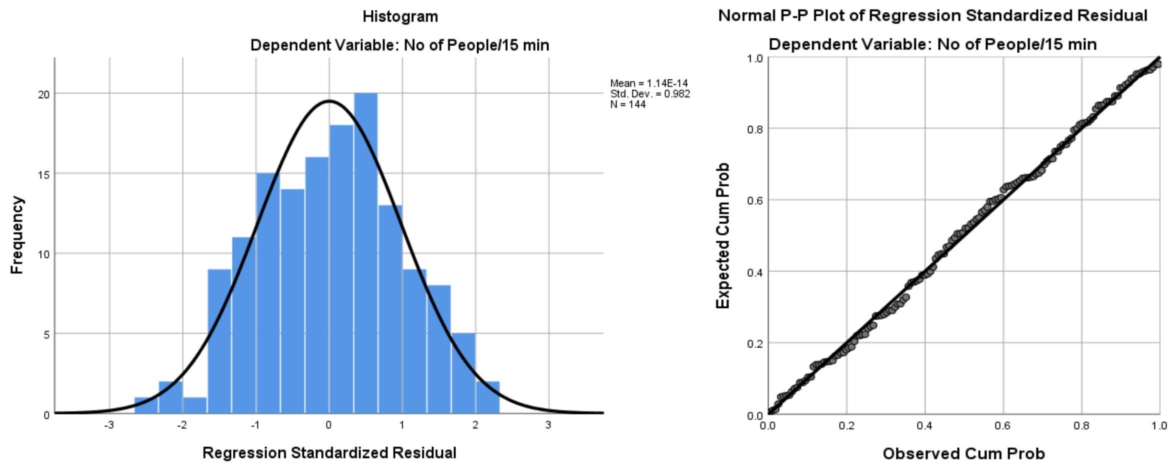


Figure 4-2 Histogram and P-P Plot of the Regression Residuals

One of the assumptions of the regression analysis is that the residuals should be normally distributed. Histogram and P-P plot shows the normality of the regression residuals. In Figure 4-3, the histogram of the residuals of the regression is along the normal line and P-P plot have the points along the 45-degree line which means that the residuals are normally distributed. To check the normality of the residuals clearly, there are two tests. One test is Kolmogorov-Smirnov (KS) Test and the other is Shapiro-Wilk (SW) Test. These tests are also performed in SPSS and the results of these tests are shown in Table 4-4:

Table 4-4 Kolmogorov-Smirnov and Shapiro-Wilk Test

Tests of Normality		
	Kolmogorov-Smirnov[a]	Shapiro-Wilk

	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.035	144	.200*	.991	144	.522
*. This is a lower bound of the true significance.						
a. Lilliefors Significance Correction						

The null hypothesis for these tests states that the data follows the normal distribution. If the significant value is less than 0.05 then the null hypothesis is rejected. Kolmogorov-Smirnov and Shapiro-Wilk test shows that residuals are normally distributed as the significance value is greater than 0.05 for both the tests.

It can be concluded that there is a problem of multicollinearity in the model which can be seen in Table 4-3. To overcome this problem, variables having VIF > 10 will be removed one by one. It is important that those independent variables should be removed that are least important in the model.

After several trails, we got the best model without multicollinearity and VIF<10.

4.1.3.1.2 Model 2:

Table 4-5 Model Summary of Regression

Model Summary[b]			
Model	R Square	Adjusted R Square	Std. Error of the Estimate
1	.837	.835	236.847
a. Predictors: (Constant), No of Shops, Gross Floor Area (1000sq.ft)			
b. Dependent Variable: No of People/15 min			

Table 4-6 Regression Coefficients

Coefficients[a]								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	321.209	40.094		8.011	.000		
	Gross Floor Area (1000sq.ft)	.640	.070	.478	9.125	.000	.421	2.375
	No of Shops	3.578	.377	.497	9.493	.000	.421	2.375
a. Dependent Variable: No of People/15 min								

Now we get a model with only two explanatory variables having positive coefficients and VIF < 10. The R-square of the model is 0.837. There is a problem of normality of the regression residuals. Histogram shows that that the residuals are skewed which means that residuals are not normally distributed. Further KS and SW tests will be conducted to check the normality.

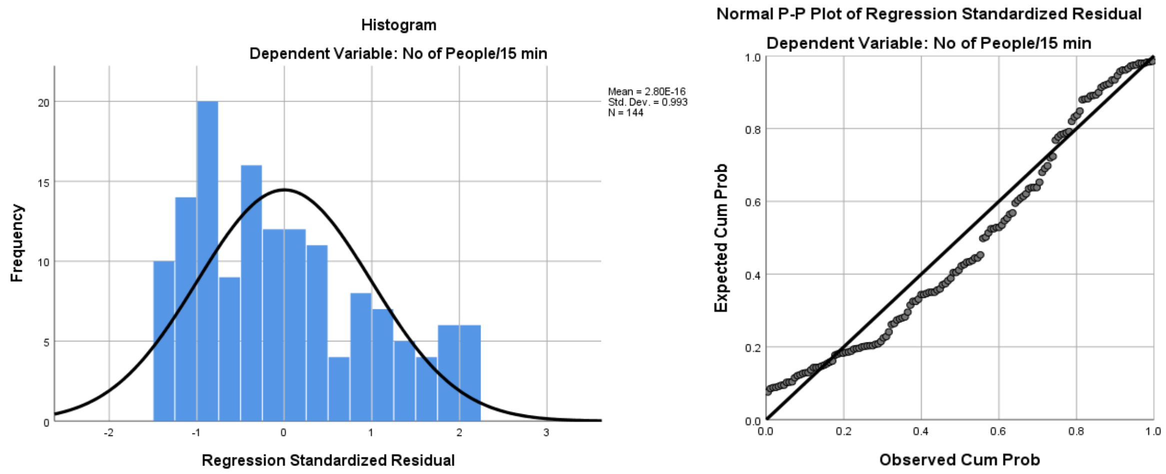


Figure 4-3 Histogram and P-P Plot of the Regression Residuals

Table 4-7 Kolmogorov-Smirnov and Shapiro-Wilk Test

Tests of Normality						
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.105	144	.001	.932	144	.000
a. Lilliefors Significance Correction						

As significance value for KS and SW tests is less than 0.05 so the residuals are not normally distributed. As Residuals are not normal so we have to go for data transformation or nonlinear model.

Transforming data into Logrithmic Function:

Data is transformed into logrithmic function because in the scattered plot matrix it was concluded that the data must be linear or logrithmic. If the data was linear then the residulas of regression had to be normally distributed. In this case, these are not normally distributed, so logrithmic transformation may solve this problem. As we got two variable having VIF < 10. One is gross floor area and other is number of shops. From the scattered plot it can be seen that gross floor area follows the logrithmic didtribution while number of shops follows linear distribution. So, gross floor area is only the varible which is transformed.

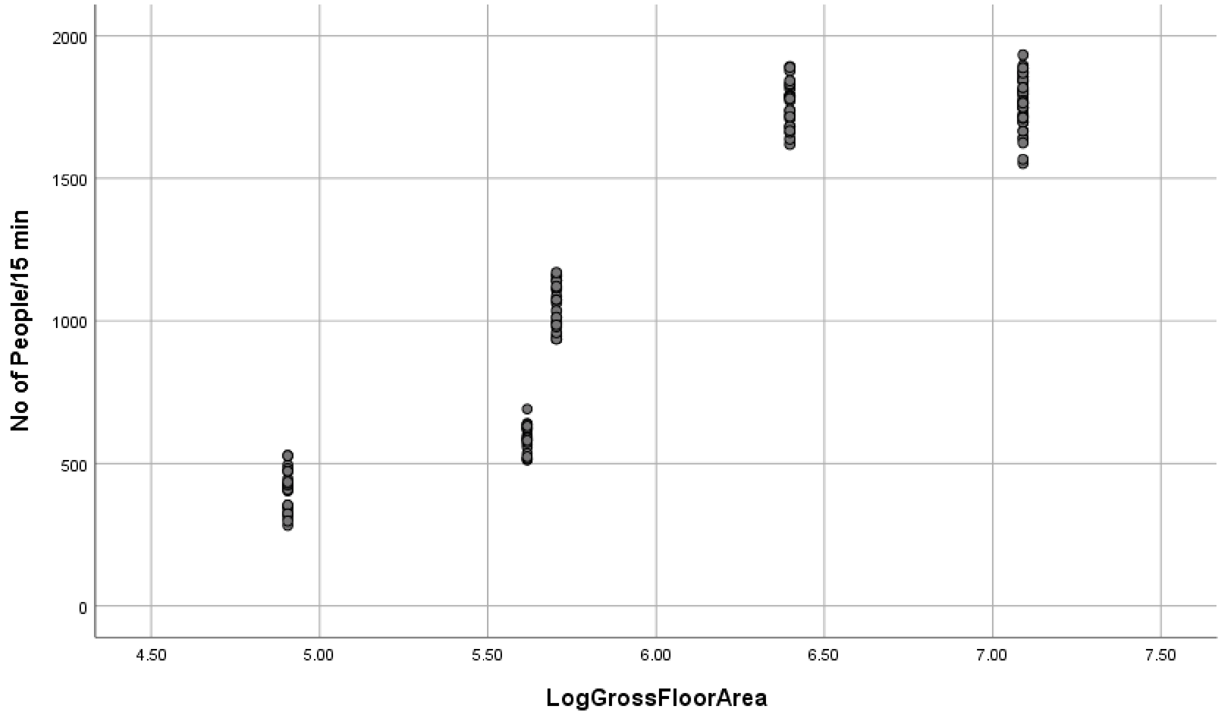


Figure 4-4 Log(GFA) vs No of People / 15 min

Pearson Correlation of No of People/15min with Ln (Gross Floor Area) is 0.929.

4.1.3.1.3 Model 3:

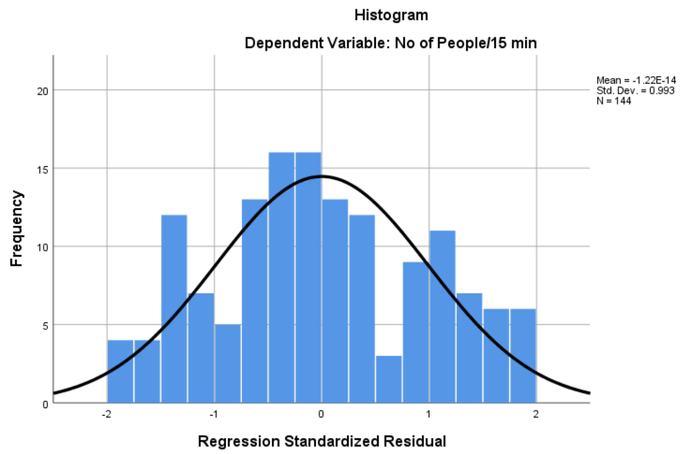
Table 4-8 Model Summary of Regression

Model Summary[b]			
Model	R Square	Adjusted R Square	Std. Error of the Estimate
1	.893	.891	192.056
a. Predictors: (Constant), No of Shops, log(Gross Floor Area 1000 sq.ft)			
b. Dependent Variable: No of People/15 min			

Table 4-9 Regression Coefficients

Coefficients[a]								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-2106.415	176.270		-11.950	.000		

	Ln(Gross Floor Area)	492.643	34.829	.680	14.145	.000	.328	3.046
	No of Shops	2.182	.346	.303	6.305	.000	.328	3.046
a. Dependent Variable: No of People/15 min								



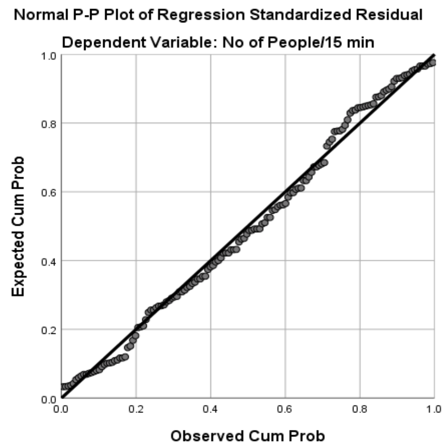


Figure 4-5 Histogram and P-P Plot of the Regression Residuals

Table 4-10 Kolmogorov-Smirnov and Shapiro-Wilk Test

Tests of Normality						
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.060	144	.200*	.975	144	.10
a. Lilliefors Significance Correction						

Kolmogorov-Smirnov and Shapiro-Wilk test shows that residuals are normally distributed.

Conclusion:

It can be concluded by comparing all the above models, Model 3 gives the best result having R-Square of 0.893 and the equation we get from the Model is:

$$\boldsymbol{Y = 492.64[ln(X1)] + 2.183(X2) - 2106.4} \qquad \textbf{(R-square = 0.893)}$$

Where, Y is the Number of Personal Trips/15 min.

X1 is the Gross Floor Area in 1000 sq.ft.

X2 is the No of Shop.

4.1.3.2 Number of Vehicular Trips

Same procedure is followed as for the above (No of People on Weekend) Model. Step wise regression is run and there is multicollinearity, and the residuals are not normally distributed. So first transforming data into log and then performing regression we get the best model as follows:

Table 4-11 Model Summary of Regression

Model Summary[b]			
Model	R Square	Adjusted R Square	Std. Error of the Estimate
1	.894	.892	53.924
a. Predictors: (Constant), No of Shops, Log(Gross Floor Area)			
b. Dependent Variable: No of Vehicles/15 min			

Table 4-12 Regression Coefficients

Coefficients[a]								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-603.499	49.491		-12.194	.000		
	Log(Gross Floor Area)	140.584	9.779	.688	14.376	.000	.328	3.046
	No of Shops	.600	.097	.295	6.173	.000	.328	3.046
a. Dependent Variable: No of Vehicles/15 min								

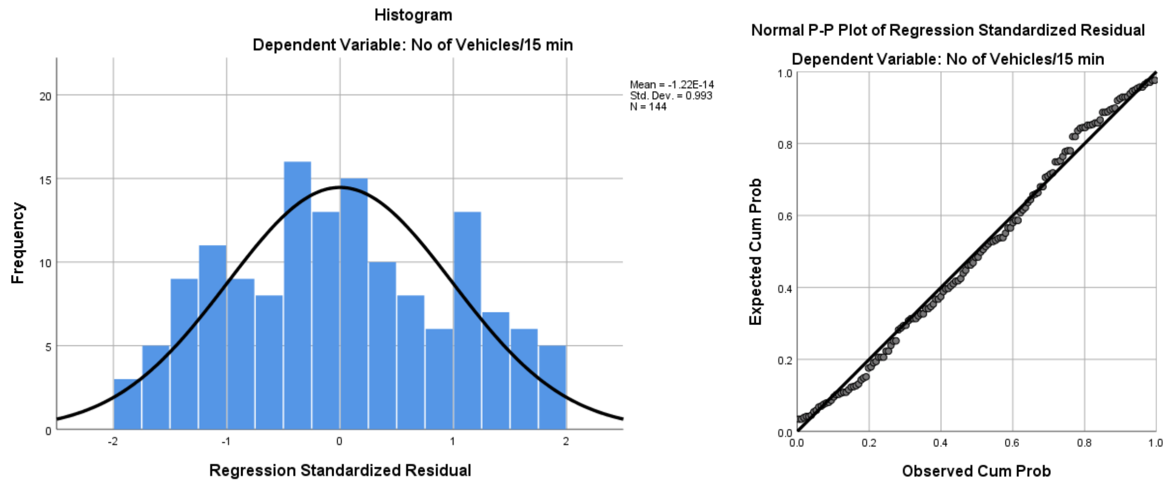


Figure 4-6 Histogram and P-P Plot of the Regression Residuals

Table 4-13 Kolmogorov-Smirnov and Shapiro-Wilk Test

Tests of Normality						
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.035	144	.200*	.991	144	.522
a. Lilliefors Significance Correction						

Model has R-square value .894 and the Regression equation is given as:

$$Y = 140.58[ln(X1)] + 0.6(X2) - 603.49 \qquad \textbf{(R-square 0.894)}$$

Where Y is the Number of Vehicular Trips/15 min.

X1 is the Gross Floor Area in 1000 sq.ft.

X2 is the Number of Shops.

4.1.4 Principal Component Analysis:

Principal component analysis, or PCA, is a dimensionality reduction method that is often used to reduce the dimensionality of large data sets, by transforming a large set of variables into a smaller one that still contains most of the information in the large set. PCA is used in those cases where the independent variables are highly correlated to each other. PCA is performed in SPSS and the results are as shown in Table 4-14.

4.1.4.1 Eigen analysis of the Correlation Matrix

Table 4-14 Eigen Analysis

Eigenvalue	6.2528	0.5598	0.1219	0.0559	0.0096	0.0000	0.0000
Proportion	0.893	0.080	0.017	0.008	0.001	0.000	0.000
Cumulative	0.893	0.973	0.991	0.999	1.000	1.000	1.000

All the independent variables are used in the analysis and a total of seven components are evaluated. Two principal components capture 97.3% of the variation from the data set while five principal components capture 100% of the variation. Same results can clearly be seen in the scree plot in Figure 4-7 Scree Plot.

4.1.4.2 Eigenvectors

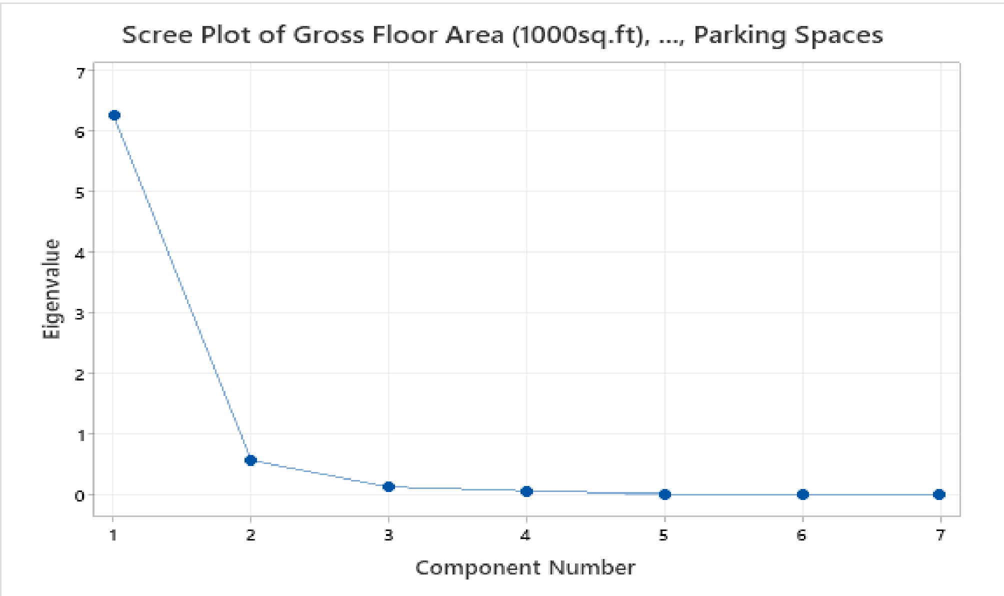


Figure 4-7 Scree Plot

4.1.5 Partial Least Square Regression:

PLS regression is performed because of multicollinearity in the linear regression model. Another reason for applying PLS regression is that it will include all the explanatory variables in the model without the issue of multicollinearity. In PLSR, we seek to obtain a set of new variables, called latent variables, that maximize covariance between variables and that are uncorrelated to each other. So the problem of multicollinearity is solved using PLS regression.

4.1.5.1 Number of Personal Trips/15 min:

Table 4-15 PLSR Components Evaluation

Cross-validation	Leave-one-out
Components to evaluate	Adjusted
Number of components evaluated	5
Number of components selected	5

Cross validation technique of Leave-one-out is used to find the components or latent variables and then the coefficients are calculated. The number of components that are evaluated and selected is 5.

Table 4-16 Analysis of Variance PLSR

Source	DF	SS	MS	F	P
Regression	5	47820944	9564189	1751.17	0.000
Residual Error	138	753700	5462		
Total	143	48574644			

Table 4-17 Model Selection and Validation for No of People/15 min

Components	X Variance	Error	R-Sq	PRESS	R-Sq (pred)
1	0.89222	11941851	0.754155	12232544	0.748170
2	0.97293	6809469	0.859814	7053391	0.854793
3	0.98296	4142540	0.914718	4403991	0.909336
4	0.99494	2540489	0.947699	2776969	0.942831
5	1.00000	753700	0.984484	820664	0.983105

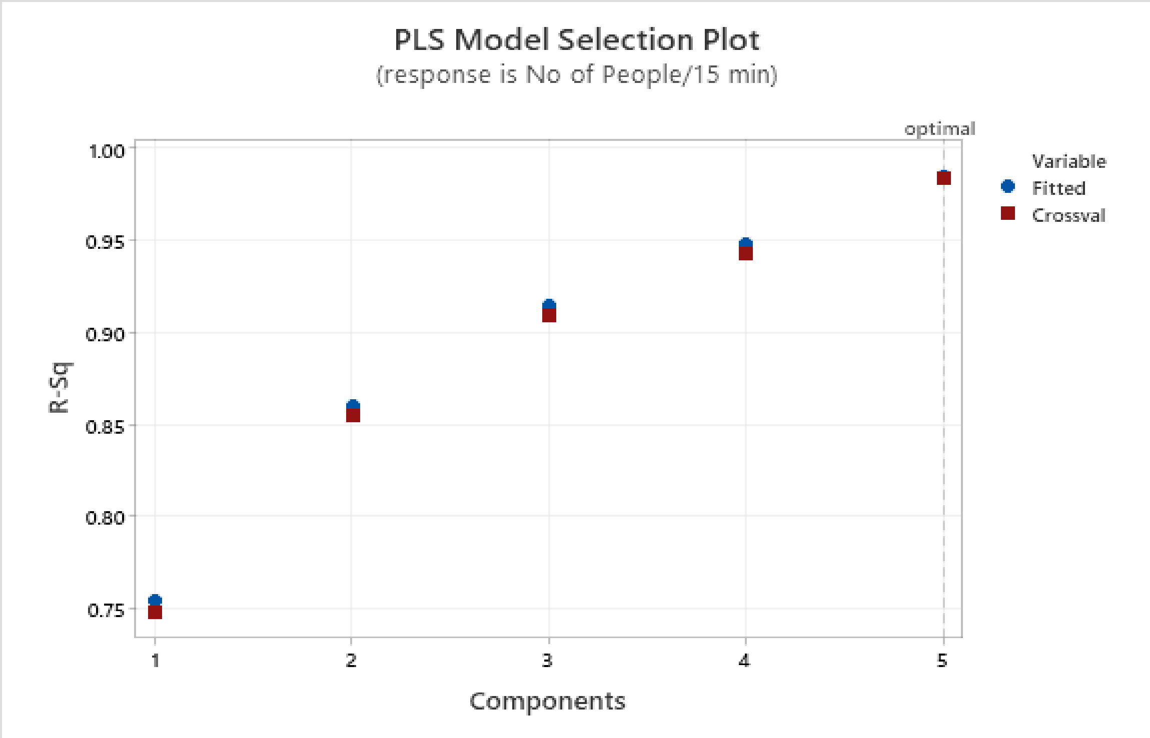


Figure 4-8 PLS Model Selection Plot

Table 4-18 PLSR Cofficients

	No of People/15 min	No of People/15 min standardized
Constant	858.494	0.00000
Gross Floor Area (1000sq.ft) (X1)	2.830	2.11468
Shopping Area (1000sq.ft) (X2)	2.966	1.71575
Playing Area (1000sq.ft) (X3)	-38.633	-1.21377
Watching Area (1000sq.ft) (X4)	7.770	0.46771
Dining Area (1000sq.ft) (X5)	-30.345	-1.56749
No of Shops (X6)	-0.962	-0.13362
Parking Spaces (X7)	-0.641	-0.64284

$$\boldsymbol{Y = 858.49 + 2.83X1 + 2.699X2 - 38.633X3 + 7.77X4 - 30.345X5 - 0.962X6 - 0.641X7}$$

R-Square = 0.98

4.1.5.2 Number of Vehicular Trips/15 min:

Same procedure has been done for the vehicular attraction rates.

Table 4-19 Analysis of Components of PLSR

Cross-validation	Leave-one-out
Components to evaluate	Adjusted

Number of components evaluated	5
Number of components selected	5

Table 4-20 Analysis of Variance for No of Vehicles/15 min

Source	DF	SS	MS	F	P
Regression	5	3792020	758404	1394.68	0.000
Residual Error	138	75042	544		
Total	143	3867062			

Table 4-21 Model Selection and Validation for No of Vehicles/15 min

Components	X Variance	Error	R-Sq	PRESS	R-Sq (pred)
1	0.89227	937886	0.757468	960986	0.751495
2	0.97294	547008	0.858547	567091	0.853354
3	0.98314	343013	0.911299	365393	0.905512
4	0.99489	212683	0.945001	232974	0.939754
5	1.00000	75042	0.980595	81709	0.978870

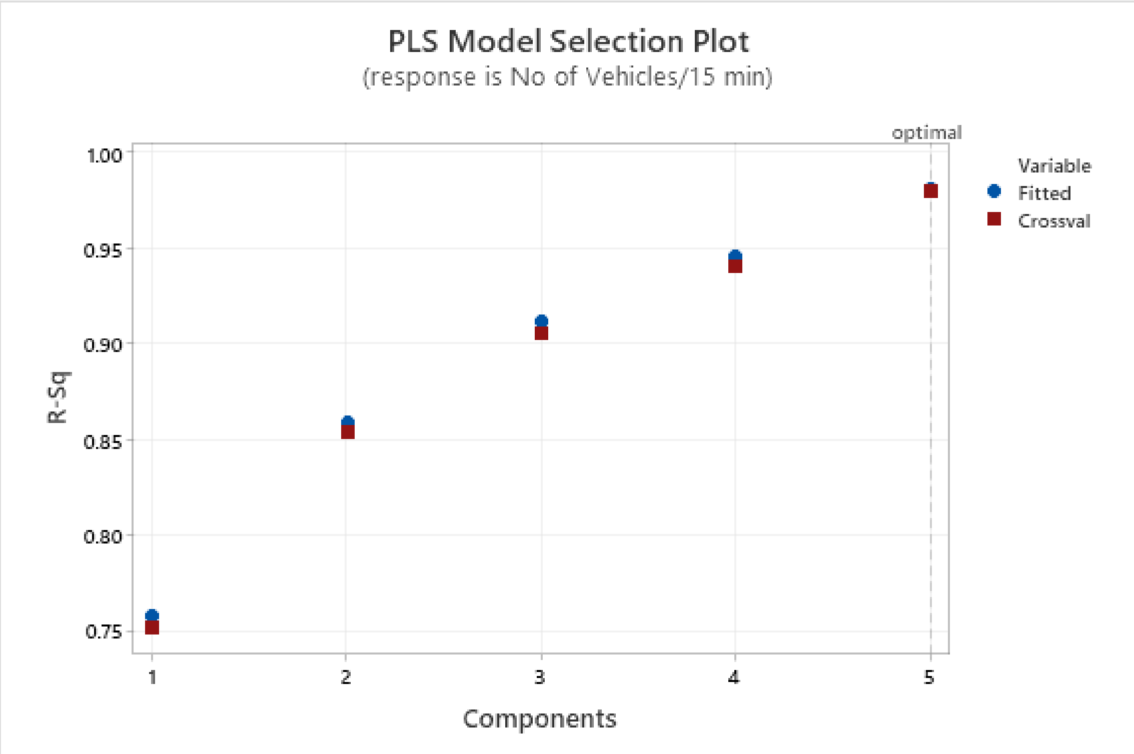


Figure 4-9 PLS Model Selection Plot

Table 4-22 Coefficients of PLSR

	No of Vehicles/15 min	No of Vehicles/15 min standardized
Constant	240.129	0.00000
Gross Floor Area (1000sq.ft) (X1)	0.793	2.10070
Shopping Area (1000sq.ft) (X2)	0.831	1.70445
Playing Area (1000sq.ft) (X3)	-10.814	-1.20419

Watching Area (1000sq.ft) (X4)	2.282	0.48692
Dining Area (1000sq.ft) (X5)	-8.493	-1.55487
No of Shops (X6)	-0.249	-0.12260
Parking Spaces (X7)	-0.187	-0.66378

$$\boldsymbol{Y = 240.12 + 0.793X1 + 0.831X2 - 10.814X3 + 2.82X4 - 8.493X5 - 0.249X6 - 0.187X7}$$

R-Square = 0.98

4.1.6 MATLAB Results:

The ANN model with Levenberg-Marquardt (LM) Algorithm is used for analysis in MATLAB software. Two models were trained for weekend and weekdays each, one for number of personal trips per 15 minutes and second is number of vehicular trips per 15 minutes. In this analysis, first all the variables are used but removing some variables we get no effect on the results.

Total Data points for the ANN Model are 144 of which 70% data (100 data points) is used for Training the model and remaining 30% of the data (44 data points) is used for Validation. It is recommended to use the 70-30 ratio for smaller data set and 80-20 ratio for larger data set. (Nguyen et al., 2021) studied the effect of data splitting on machine learning models for the prediction of shear strength of soil with 538 data set sample and found that 70-30 ratio have the best performance. Sigmoid (log) function is used as a transfer function.

4.1.6.1 No of People:

Response Variable = No of Personal Trips/15-min

Explanatory Variables = Gross Floor Area, Shopping Area, Playing Area, and No of Shops. (All areas are in 1000 sq. ft)

70% Data is used for training 30% Data is used for validation.

RMSE for Training = 73.38 RMSE for Validation = 71.35

R-Square=0.984 1 hidden layer with 2 neurons.

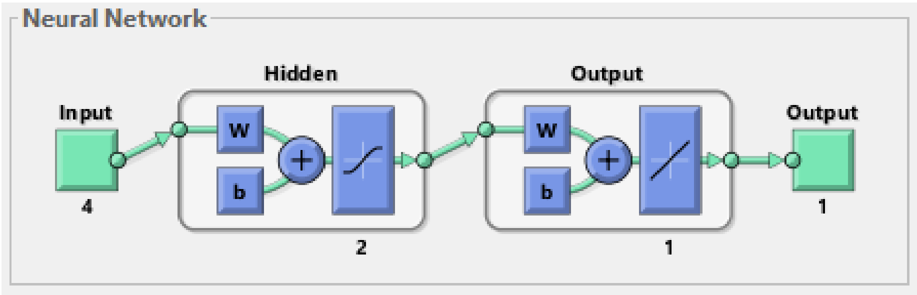


Figure 4-10 Neural Network

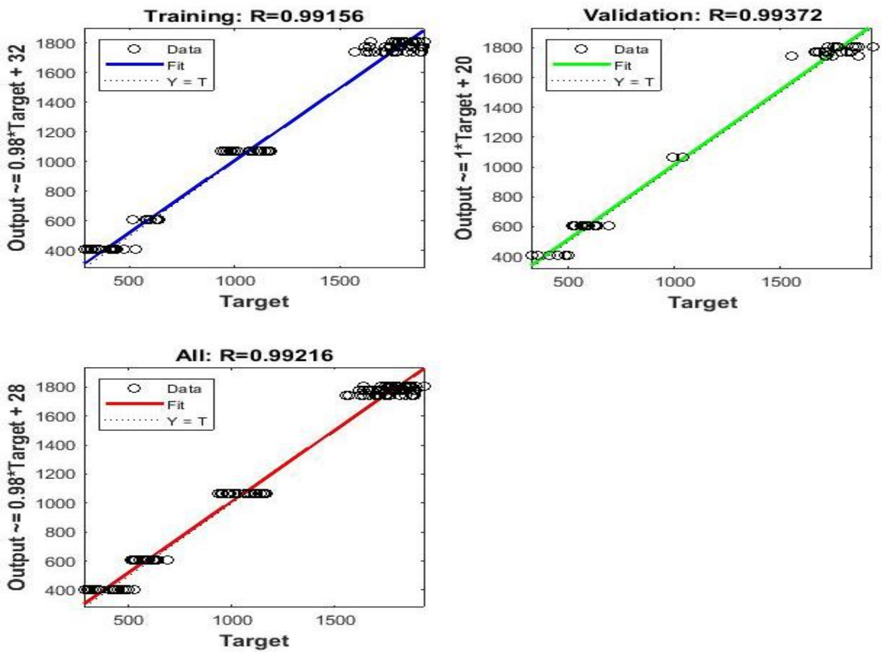


Figure 4-11 Regression Plots

For optimum number of neurons, RMSE of training and validation is plotted against number of neurons. This is done by running a loop for the number of neurons starting from 1 to 20. Then the optimum number of neurons is selected for the minimum value of RMSE for both training and validation and from where the slope of the RMSE of training or validation line changes from negative to positive or RMSE starts increasing. From Figure 4-12 Number of Neurons vs RMSE can be seen that RMSE starts to increase at number of neurons are 2. Thumb rules are also available for selecting the optimum number of neurons.

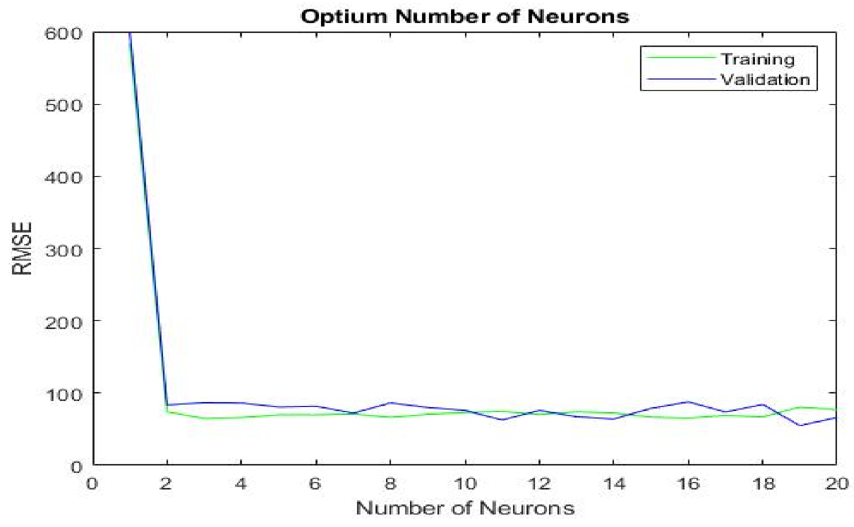


Figure 4-12 Number of Neurons vs RMSE

4.1.6.2 No of Vehicles:

Response Variable = No of Vehicular Trips/15-min

Explanatory Variables = Gross Floor Area, Shopping Area, Playing Area and No of Shops.

(All areas are in 1000 sq. ft)

RMSE for Training = 23.01 RMSE for Validation = 22.95

R-Square = 0.98

1 hidden layer and 2 neurons

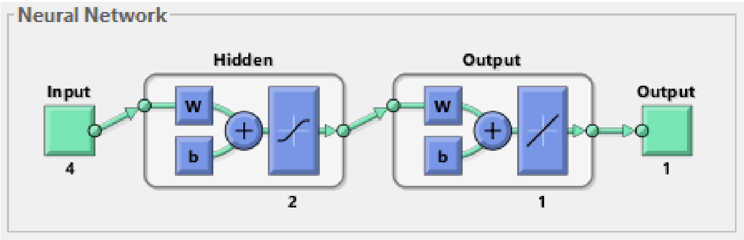


Figure 4-13 Neural Network

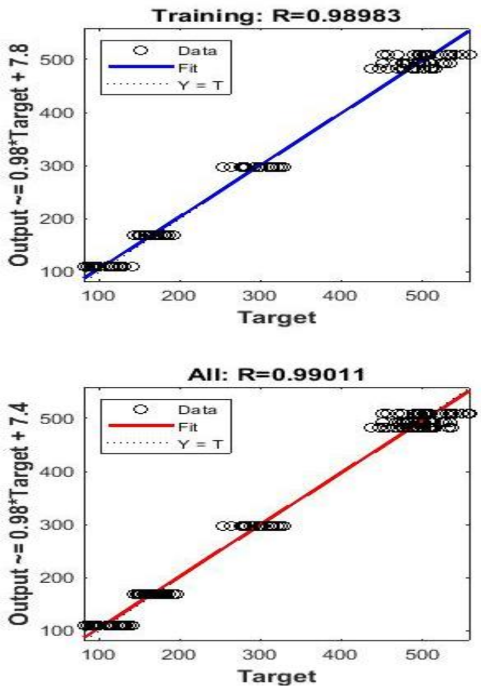


Figure 4-14 Regression Plot

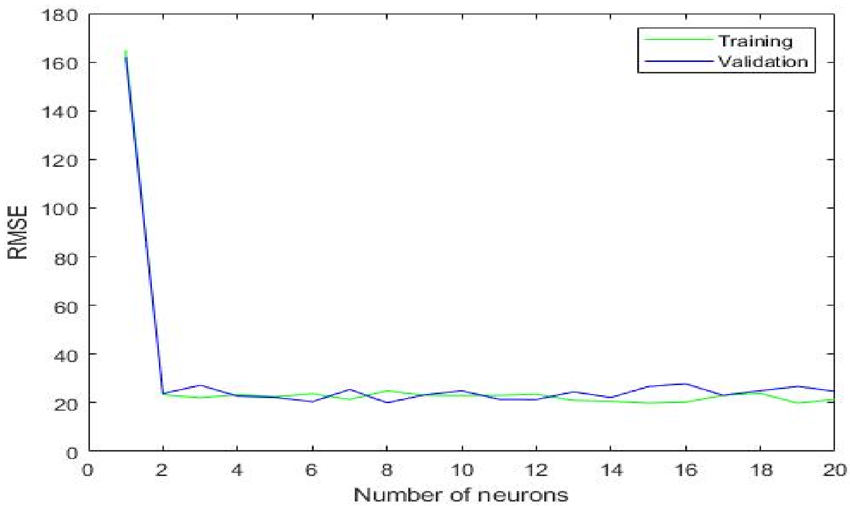


Figure 4-15 Number of Neurons vs RMSE

Optimum number of neurons is 2.

4.2 Weekday Results:

4.2.1 Scattered Plot:

The scattered plot showed that there are some variables which have positive correlations, and some have less correlations. It can also be seen in the scattered plot that some independent variables have almost linear or logarithmic relation with dependent variable.

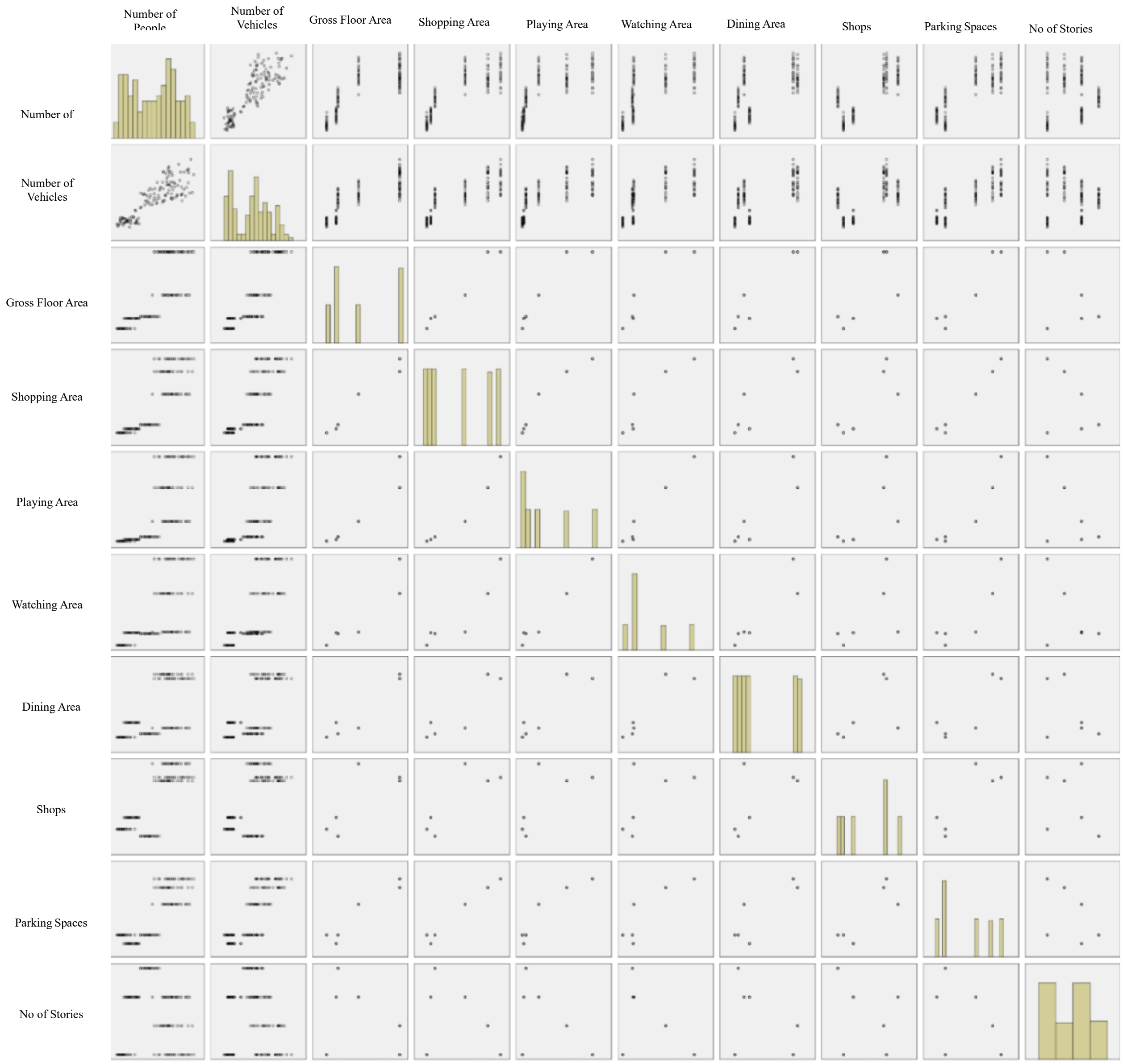


Figure 4-16 Scattered Plot for Weekdays

4.2.2 Pearson Correlations:

Pearson Correlations show that explanatory variables are highly correlated to the response variable. Only the number of stories have negative but vary less correlation with the response variable. So, we will use all the variables in the analysis except the number of stories. One thing also important in this matrix is that the independent variables are also highly correlated with each other, so there will be a chance of multi-collinearity in the model.

Table 4-23 Pearson Correlations for Weekday Data

Correlations									
	No of People/15 min	Gross Floor Area	Shopping Area	Playing Area	Watching Area	Dining Area	No of Shops	Parking Spaces	No of Stories
No of People/15 min	1								
Gross Floor Area	.790	1							
Shopping Area	.840	.981	1						
Playing Area	.728	.951	.961	1					
Watching Area	.663	.917	.910	.984	1				
Dining Area	.619	.956	.894	.917	.920	1			
No of Shops	.804	.761	.832	.683	.579	.601	1		
Parking Spaces	.833	.958	.988	.936	.863	.851	.843	1	
No of Stories	-.020	-.381	-.402	-.522	-.478	-.451	-.244	-.456	1

Number of Stories is not used in the Regression Model.

4.2.3 Regression Analysis:

4.2.3.1 Number Of People/15min:

Stepwise Ordinary Least Square Regression is performed in SPSS, the main advantage of this Stepwise regression is that this starts from one independent variable and then adding independent variables one by one. Six models have been made, and the last model has the greatest R-Square value of 0.871.

4.2.3.1.1 Model 1:

Table 4-24 Model Summary of Regression

Model Summary[g]				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.840[a]	.706	.704	166.572165140233270
2	.891[b]	.794	.791	139.867857571195600
3	.920[c]	.847	.844	121.062734186957630
4	.919[d]	.845	.843	121.505513189782650
5	.933[e]	.871	.868	111.183317658399290
6	.937[f]	.878	.875	108.376927422985800
a. Predictors: (Constant), Shopping Area (1000sq.ft)				
b. Predictors: (Constant), Shopping Area (1000sq.ft), Dining Area (1000sq.ft)				

c. Predictors: (Constant), Shopping Area (1000sq.ft), Dining Area (1000sq.ft), Gross Floor Area (1000sq.ft)
d. Predictors: (Constant), Dining Area (1000sq.ft), Gross Floor Area (1000sq.ft)
e. Predictors: (Constant), Dining Area (1000sq.ft), Gross Floor Area (1000sq.ft), Parking Spaces
f. Predictors: (Constant), Dining Area (1000sq.ft), Gross Floor Area (1000sq.ft), Parking Spaces, Shopping Area (1000sq.ft)
g. Dependent Variable: No of People/15 min

Table 4-25 Regression Coefficients

Coefficients[a]								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	415.349	23.179		17.919	.000		
	Shopping Area (1000sq.ft)	.763	.041	.840	18.472	.000	1.000	1.000
2	(Constant)	448.470	19.924		22.509	.000		
	Shopping Area (1000sq.ft)	1.302	.078	1.434	16.792	.000	.200	4.996
	Dining Area (1000sq.ft)	-6.749	.868	-.664	-7.772	.000	.200	4.996
3	(Constant)	405.688	18.313		22.153	.000		
	Shopping Area (1000sq.ft)	-.353	.248	-.389	-1.426	.156	.015	68.038
	Dining Area (1000sq.ft)	-18.392	1.838	-1.808	-10.008	.000	.033	29.863
	Gross Floor Area	2.040	.294	2.901	6.943	.000	.006	159.729

	(1000sq.ft)							
4	(Constant)	413.088	17.627		23.435	.000		
	Dining Area (1000sq.ft)	-16.350	1.156	-1.608	-14.145	.000	.085	11.730
	Gross Floor Area (1000sq.ft)	1.637	.080	2.328	20.481	.000	.085	11.730
5	(Constant)	455.138	17.956		25.347	.000		
	Dining Area (1000sq.ft)	-23.367	1.689	-2.298	-13.835	.000	.033	29.912
	Gross Floor Area (1000sq.ft)	2.706	.214	3.848	12.671	.000	.010	100.036
	Parking Spaces	-.471	.088	-.899	-5.329	.000	.032	30.841
6	(Constant)	493.984	22.072		22.380	.000		
	Shopping Area (1000sq.ft)	.873	.302	.961	2.889	.004	.008	126.301
	Dining Area (1000sq.ft)	-21.753	1.739	-2.139	-12.512	.000	.030	33.357
	Gross Floor Area (1000sq.ft)	2.232	.265	3.175	8.424	.000	.006	162.118
	Parking Spaces	-.701	.117	-1.338	-5.974	.000	.017	57.251
a. Dependent Variable: No of People/15 min								

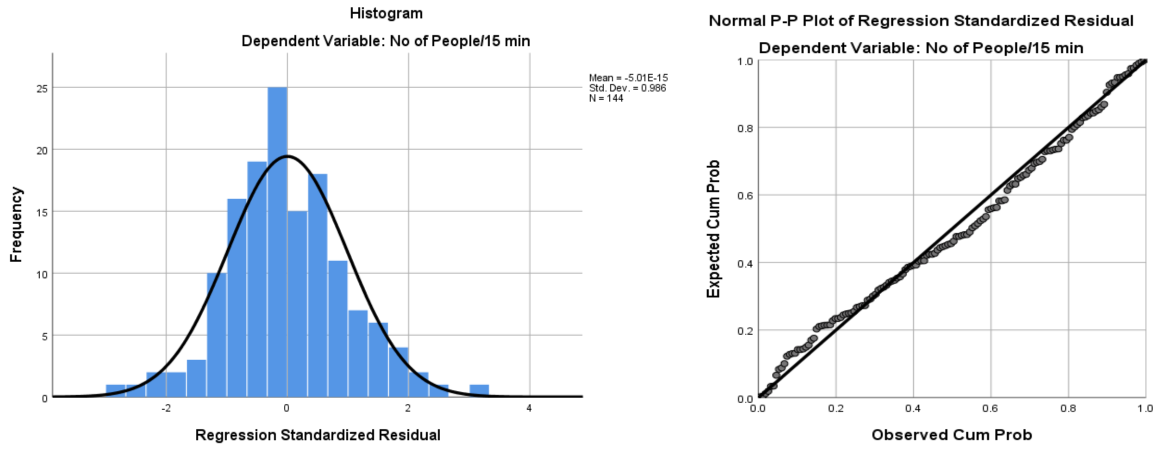


Figure 4-17 Histogram and P-P Plot of the Regression Residuals

Table 4-26 Kolmogorov-Smirnov and Shapiro-Wilk Test

Tests of Normality						
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.059	144	.200*	.987	144	.202
a. Lilliefors Significance Correction						

The analysis shows that there is multicollinearity in the model, same procedure is applied to remove the multicollinearity (by removing the variables having VIF > 10. We have two variables one is gross floor area and the other is number of shops. When the normality is checked for the regression residual, it was not normally distributed. So, data transformation is done just like the previous analysis.

Table 4-27 Correlation for Log(Gross Floor Area)

Correlations			
		No of People/15 min	Ln(Gross Floor Area 1000 sq.ft)
No of People/15 min		1	.869

4.2.3.1.2 Model 2:

Table 4-28 Model Summary of Regression

Model Summary[b]				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.884[a]	.781	.778	144.16
a. Predictors: (Constant), No of Shops, GrossFloorArea_Log				
b. Dependent Variable: No of People/15 min				

Table 4-29 Regression Coefficients

Coefficients[a]								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-887.460	132.316		-6.707	.000		
	Ln(Gross Floor Area 1000 sq.ft)	243.844	26.144	.641	9.327	.000	.328	3.046
	No of Shops	1.054	.260	.279	4.057	.000	.328	3.046
a. Dependent Variable: No of People/15 min								

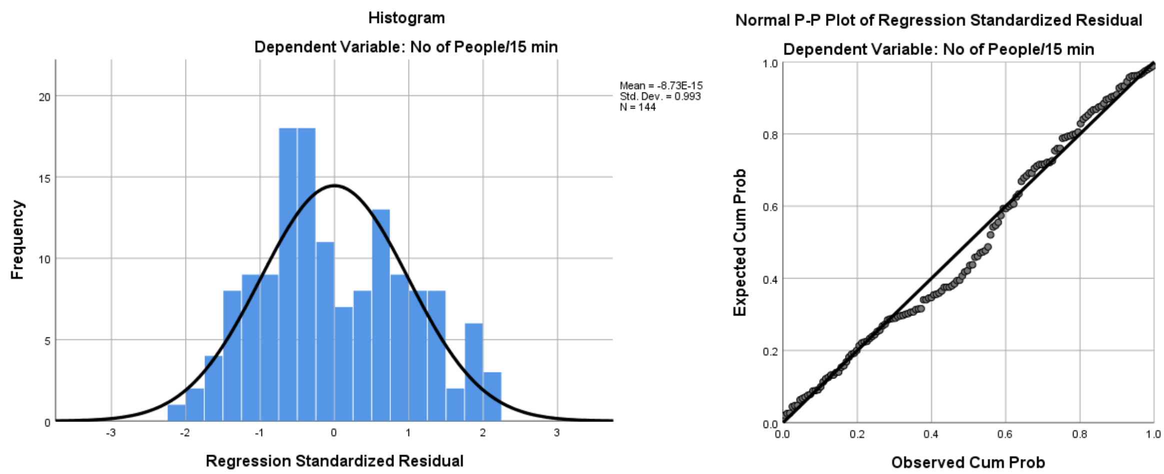


Figure 4-18 Histogram and P-P Plot of the Regression Residuals

Table 4-30 Kolmogorov-Smirnov and Shapiro-Wilk Test

Tests of Normality						
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.085	144	.012	.981	144	.046
a. Lilliefors Significance Correction						

4.2.3.1.3 Model 3:

Table 4-31 Model Summary of Regression

Model Summary[b]				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.869[a]	.756	.754	151.808989210227300
a. Predictors: (Constant), GrossFloorArea_Log				
b. Dependent Variable: No of People/15 min				

Table 4-32 Regression Coefficients

Coefficients[a]								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-1270.623	97.573		-13.022	.000		
	GrossFloorArea_Log	330.768	15.773	.869	20.970	.000	1.000	1.000
a. Dependent Variable: No of People/15 min								

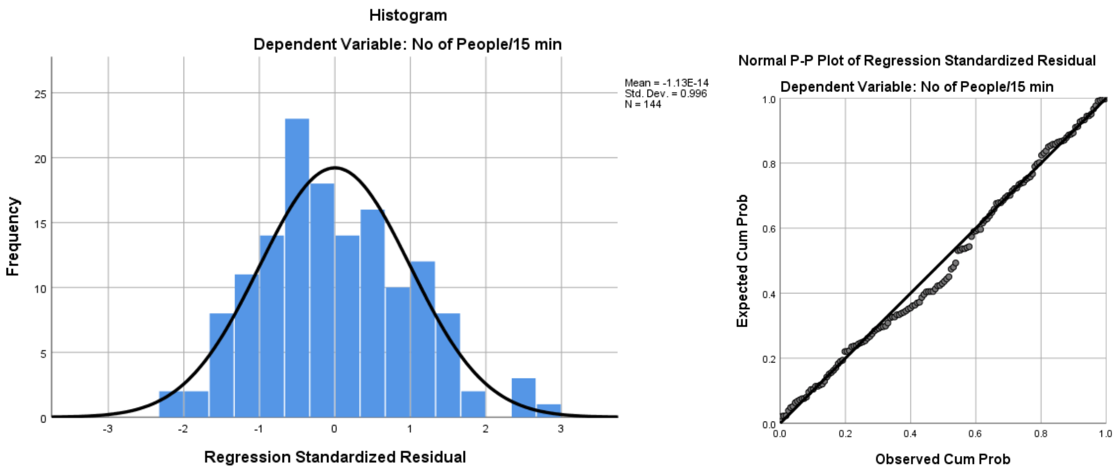


Figure 4-19 Histogram and P-P Plot of the Regression Residuals

Table 4-33 Kolmogorov-Smirnov and Shapiro-Wilk Test

Tests of Normality						
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.

Standardized Residual	.071	144	.075	.986	144	.161
a. Lilliefors Significance Correction						

Best Model for the weekdays is:

$\boldsymbol{Y = 330.7\,[ln(X)] - 1270.6}$ **(R-square 0.756)**

Where,

Y = No of People/ 15-min

X = Gross Floor Area (1000 sq.ft)

4.2.3.2 Vehicles/15-min:

Same procedure is followed as for the above (No of People on Weekend) Model. Step wise regression is run and there is multicollinearity, and the residuals are not normally distributed. So first transforming data into log and then performing regression we get the best model as follows:

4.2.3.2.1 Pearson Correlations:

Table 4-34 Pearson Correlations for vehicular data

Correlations									
	No of Vehicles/15 min	Gross Floor Area	Shopping Area	Playing Area	Watching Area	Dining Area	No of Shops	Parking Spaces	No of Stories
No of Vehicles/15 min	1								
Gross Floor Area	.819	1							
Shopping Area	.830	.981	1						
Playing Area	.754	.951	.961	1					
Watching Area	.708	.917	.910	.984	1				
Dining Area	.691	.956	.894	.917	.920	1			
No of Shops	.655	.761	.832	.683	.579	.601	1		
Parking Spaces	.822	.958	.988	.936	.863	.851	.843	1	
No of Stories	-.059	-.381	-.402	-.522	-.478	-.451	-.244	-.456	1

4.2.3.2.2 Model 1:

Table 4-35 Model Summary of Regression

Model Summary[f]			
Model	R Square	Adjusted R Square	Std. Error of the Estimate
1	.689	.687	58.443507485165230
2	.713	.709	56.305859106176650
3	.793	.789	48.006366129075910
4	.820	.814	44.966469419283830
5	.835	.829	43.152377253702520
a. Predictors: (Constant), Shopping Area (1000sq.ft)			
b. Predictors: (Constant), Shopping Area (1000sq.ft), Playing Area (1000sq.ft)			
c. Predictors: (Constant), Shopping Area (1000sq.ft), Playing Area (1000sq.ft), No of Shops			

d. Predictors: (Constant), Shopping Area (1000sq.ft), Playing Area (1000sq.ft), No of Shops, Dining Area (1000sq.ft)
e. Predictors: (Constant), Shopping Area (1000sq.ft), Playing Area (1000sq.ft), No of Shops, Dining Area (1000sq.ft), Parking Spaces
f. Dependent Variable: No of Vehicles/15 min

Table 4-36 Regression Coefficients

Coefficients[a]								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	97.015	8.134		11.927	.000		
	Shopping Area (1000sq.ft)	.257	.015	.830	17.665	.000	1.000	1.000
2	(Constant)	107.827	8.440		12.776	.000		
	Shopping Area (1000sq.ft)	.425	.051	1.372	8.391	.000	.077	13.054
	Playing Area (1000sq.ft)	-3.214	.931	-.564	-3.451	.001	.077	13.054
3	(Constant)	184.344	12.690		14.527	.000		
	Shopping Area (1000sq.ft)	.989	.088	3.192	11.199	.000	.018	54.564
	Playing Area (1000sq.ft)	-10.113	1.232	-1.776	-8.206	.000	.032	31.444
	No of Shops	-1.015	.139	-.789	-7.321	.000	.128	7.800
4	(Constant)	203.581	12.625		16.125	.000		

	Shopping Area (1000sq.ft)	1.144	.090	3.693	12.775	.000	.016	63.967
	Playing Area (1000sq.ft)	-9.799	1.156	-1.721	-8.474	.000	.032	31.558
	No of Shops	-1.255	.140	-.975	-8.944	.000	.110	9.103
	Dining Area (1000sq.ft)	-1.554	.344	-.446	-4.520	.000	.134	7.464
5	(Constant)	233.693	14.744		15.851	.000		
	Shopping Area (1000sq.ft)	1.544	.141	4.984	10.964	.000	.006	171.713
	Playing Area (1000sq.ft)	-10.636	1.134	-1.868	-9.379	.000	.030	32.953
	No of Shops	-1.344	.137	-1.044	-9.815	.000	.106	9.410
	Dining Area (1000sq.ft)	-2.135	.368	-.613	-5.807	.000	.108	9.262
	Parking Spaces	-.171	.048	-.954	-3.584	.000	.017	58.919

a. Dependent Variable: No of Vehicles/15 min

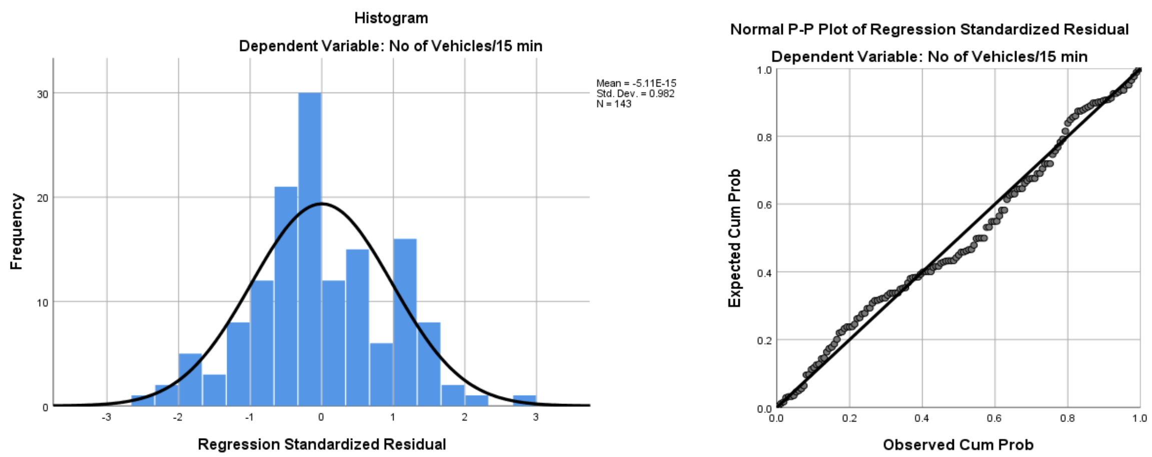


Figure 4-20 Histogram and P-P Plot of the Regression Residuals

Table 4-37 Kolmogorov-Smirnov and Shapiro-Wilk Test

Tests of Normality						
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.074	143	.052	.989	143	.339
a. Lilliefors Significance Correction						

4.2.3.2.3 Model 2:

Table 4-38 Model Summary of Regression

Model Summary[b]				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.855[a]	.731	.729	54.35
a. Predictors: (Constant), GrossFloorArea_Log				
b. Dependent Variable: No of Vehicles/15 min				

Table 4-39 Regression Coefficients

Coefficients[a]					
Model	Unstandardized Coefficients	Standardized Coefficients	t	Sig.	Collinearity Statistics

		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-468.489	35.075		-13.357	.000		
	GrossFloorArea_Log	111.042	5.676	.855	19.563	.000	1.000	1.000
a. Dependent Variable: No of Vehicles/15 min								

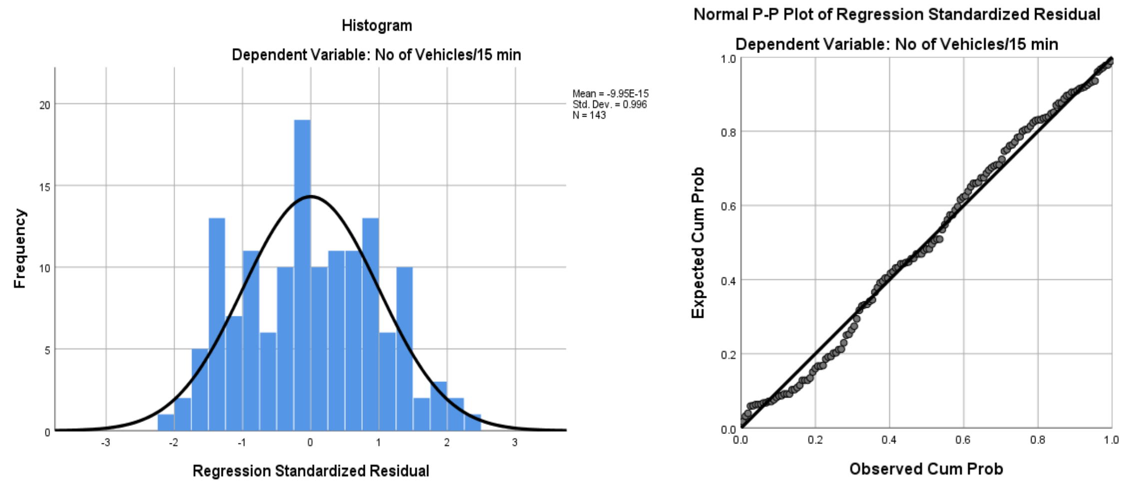


Figure 4-21 Histogram and P-P Plot of the Regression Residuals

Table 4-40 Normality Tests

Tests of Normality						
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.061	143	.200*	.980	143	.036
*. This is a lower bound of the true significance.						
a. Lilliefors Significance Correction						

$$\boldsymbol{Y = 111.04\,[ln(X)] - 468.4}$$ **(R-square 0.731)**

Where, Y is the Vehicular Trips/ 15-min

X is the Gross Floor Area (1000 sq. ft)

4.2.4 Partial Least Square Regression:

4.2.4.1 Number of Personal Trips/15 min:

Table 4-41 PLSR Components Evaluation

Cross-validation	Leave-one-out
Components to evaluate	Adjusted
Number of components evaluated	5
Number of components selected	5

Table 4-42 Analysis of Variance for No of People/15 min

Source	DF	SS	MS	F	P
Regression	5	11775150	2355030	199.14	0.000
Residual Error	138	1632012	11826		
Total	143	13407162			

Table 4-43 Model Selection and Validation for No of People/15 min

Components	X Variance	Error	R-Sq	PRESS	R-Sq (pred)
1	0.89220	4660500	0.652387	4786894	0.642960
2	0.97275	3390563	0.747108	3542587	0.735769
3	0.98706	2806088	0.790702	3000937	0.776169
4	0.99495	2070424	0.845573	2274145	0.830378
5	1.00000	1632012	0.878273	1777011	0.867458

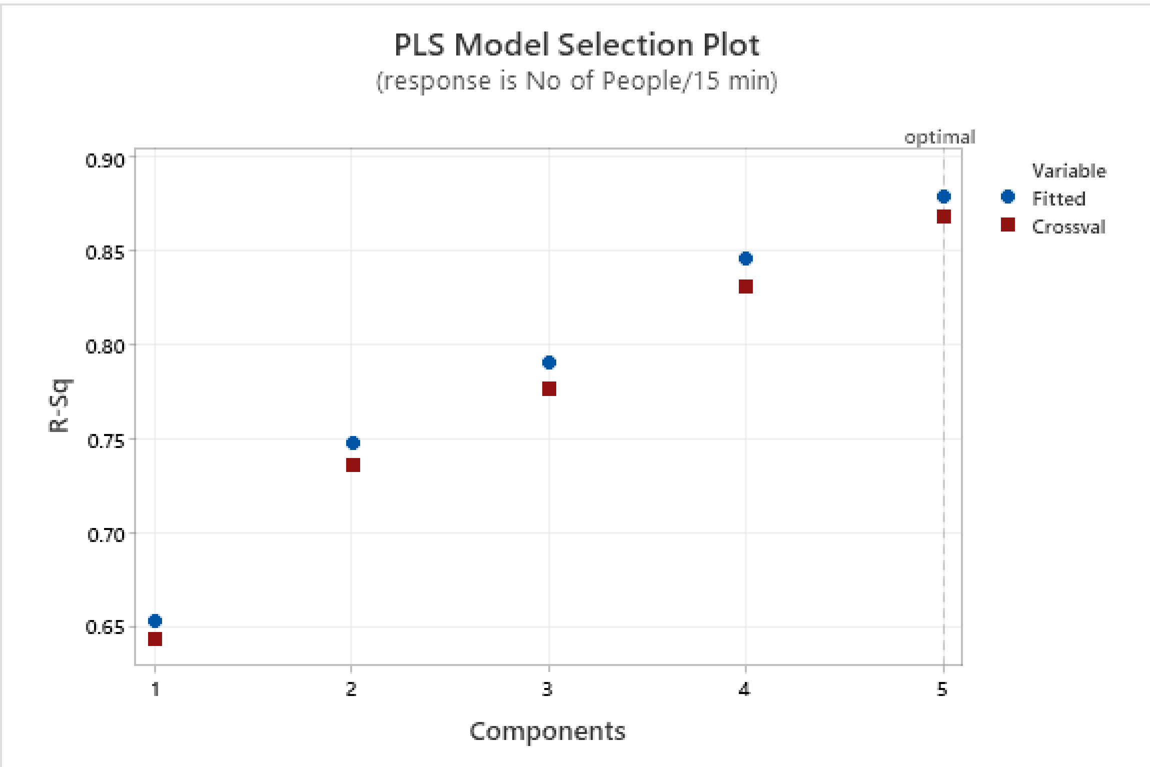


Figure 4-22 PLS Model Selection Plot

Table 4-44 PLSR Coefficients

	No of People/15 min	No of People/15 min standardized
Constant	614.969	0.00000
Gross Floor Area (1000sq.ft) (X1)	1.491	2.12005
Shopping Area (1000sq.ft) (X2)	1.598	1.75908
Playing Area (1000sq.ft) (X3)	-20.355	-1.21725
Watching Area (1000sq.ft) (X4)	5.726	0.65612
Dining Area (1000sq.ft) (X5)	-17.327	-1.70366
No of Shops (X6)	-0.355	-0.09394
Parking Spaces (X7)	-0.437	-0.83430

$$\boldsymbol{Y = 614.9 + 1.49X1 + 1.59X2 - 20.3550X3 + 5.726X4 - 17.3271X5 - 0.355X6 - 0.437X7}$$

R-Square = 0.87

4.2.4.2 Number of Vehicular Trips/15 Min:

Table 4-45 PLSR Components Evaluation

Cross-validation	Leave-one-out
Components to evaluate	Adjusted
Number of components evaluated	5
Number of components selected	5

Table 4-46 Analysis of Variance for No of Vehicles/15 min

Source	DF	SS	MS	F	P
Regression	5	1303127	260625	140.98	0.000
Residual Error	138	255111	1849		
Total	143	1558238			

Table 4-47 Model Selection and Validation for No of Vehicles/15 min

Components	X Variance	Error	R-Sq	PRESS	R-Sq (pred)
1	0.89316	551812	0.645875	568510	0.635159
2	0.96610	500906	0.678543	523569	0.663999
3	0.98122	369887	0.762625	407336	0.738592
4	0.99219	310045	0.801029	337586	0.783354
5	1.00000	255111	0.836282	277777	0.821736

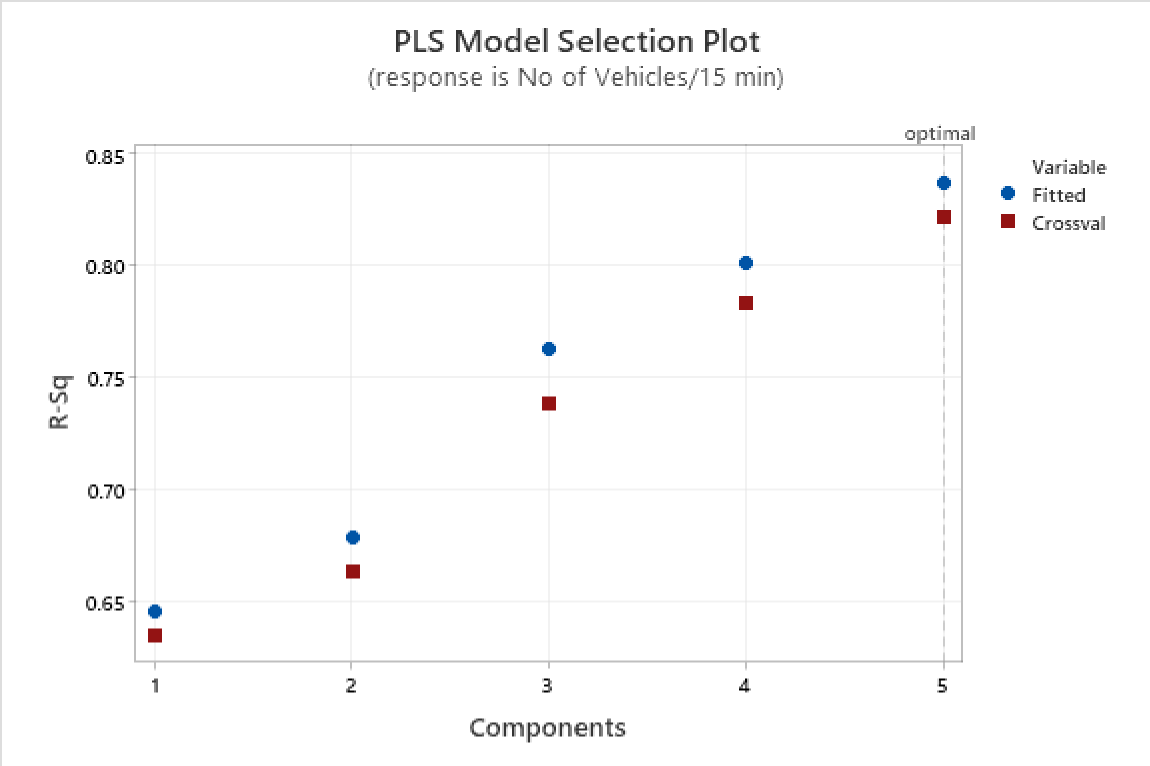


Figure 4-23 PLS Model Selection Plot

	No of Vehicles/15 min	No of Vehicles/15 min standardized
Constant	183.168	0.00000
Gross Floor Area (1000sq.ft) (X1)	0.557	2.32463
Shopping Area (1000sq.ft) (X2)	0.557	1.79935
Playing Area (1000sq.ft) (X3)	-7.210	-1.26479
Watching Area (1000sq.ft) (X4)	1.163	0.39098
Dining Area (1000sq.ft) (X5)	-5.341	-1.54038
No of Shops (X6)	-0.812	-0.62990
Parking Spaces (X7)	-0.088	-0.49350

$$\boldsymbol{Y = 183 + 0.557X1 + .557X2 - 7.210X3 + 1.163X4 - 5.341X5 - 0.812X6 - 0.088X7}$$

R-Square = 0.83

4.2.5 MATLAB results:

The ANN model is made using MATLAB software. Levenberg-Marquardt (LM) Algorithm has been used for the analysis. Two Models have been made for the weekend, one for number of personal trips per 15 minutes and second is number of vehicular trips per 15 minutes. In this analysis, first all the variables are used but removing some of the variables we get no effect on the results. By the hit and trail method, minimum number of variables are selected with maximum R value and minimum RMSE (Root Mean Squared Error) value.

4.2.5.1 No of People:

Response Variable = No of Personal Trips/15-min

Explanatory Variables = Gross Floor Area, Shopping Area, Playing Area, and No of Shops. (All areas are in 1000 sq. ft)

RMSE Training = 109.9

MSE Validation = 111.3

R-Square = 0.87

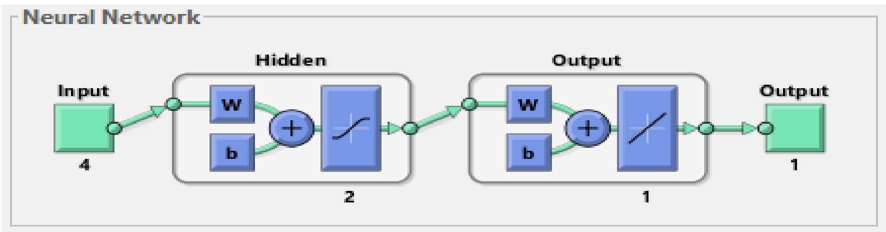


Figure 4-24 Structure of Neural Network

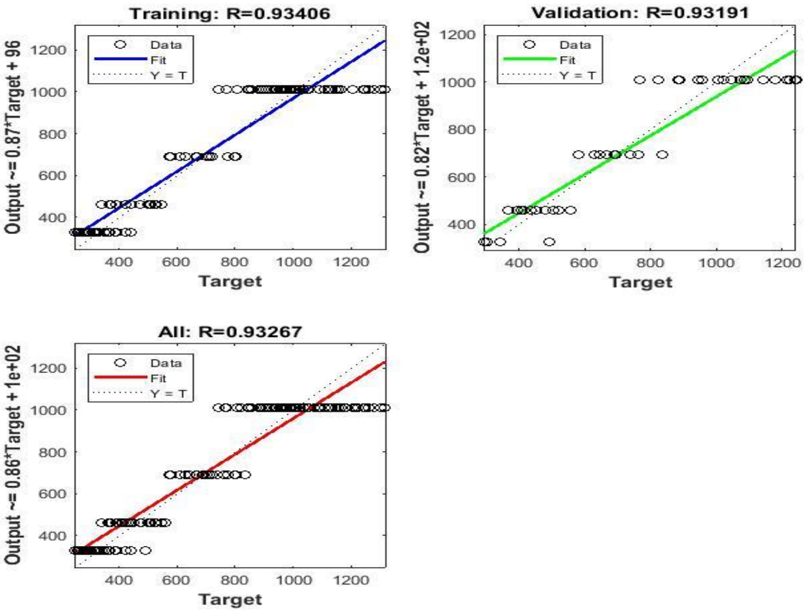


Figure 4-25 Regression Plots

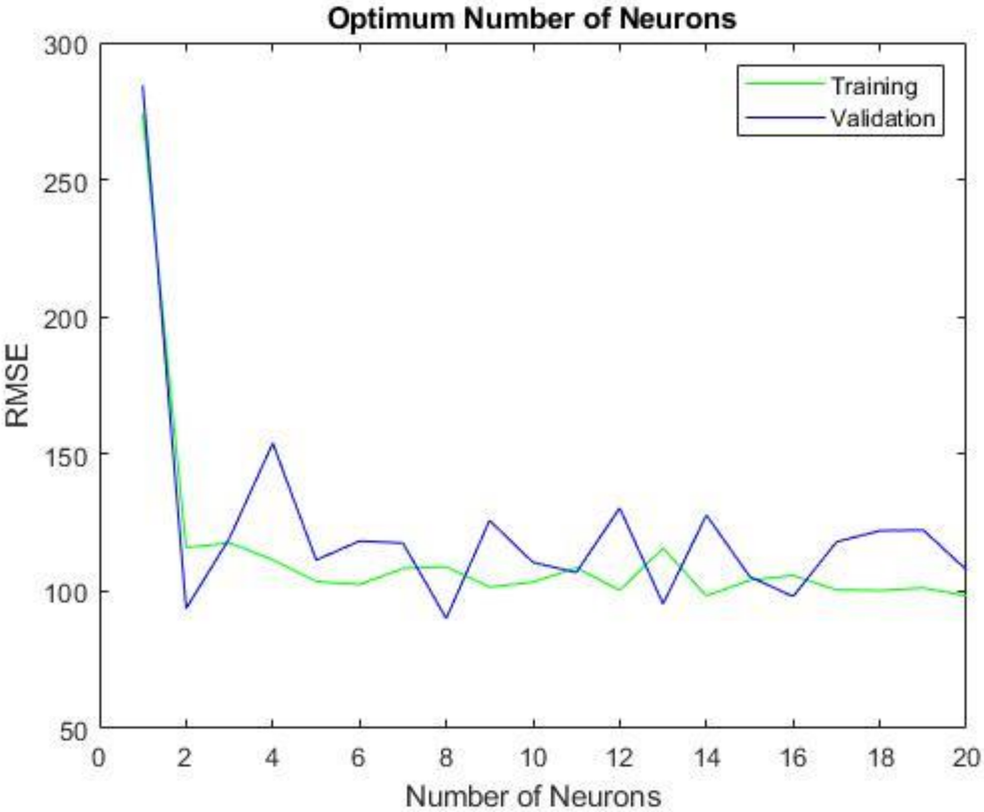


Figure 4-26 Optimum Number of Neurons

4.2.5.2 No of Vehicles:

Response Variable = No of Personal Trips/15-min

Explanatory Variables = Gross Floor Area, Shopping Area, Playing Area, and No of Shops.

(All areas are in 1000 sq. ft)

RMSE Training = 43

RMSE Validation = 40

R-Square = 0.83

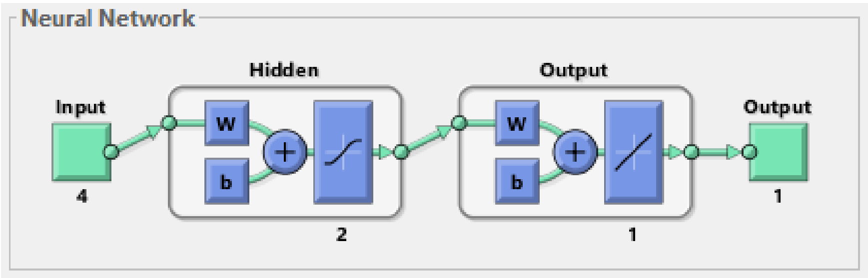


Figure 4-27 Structure of Neural Network

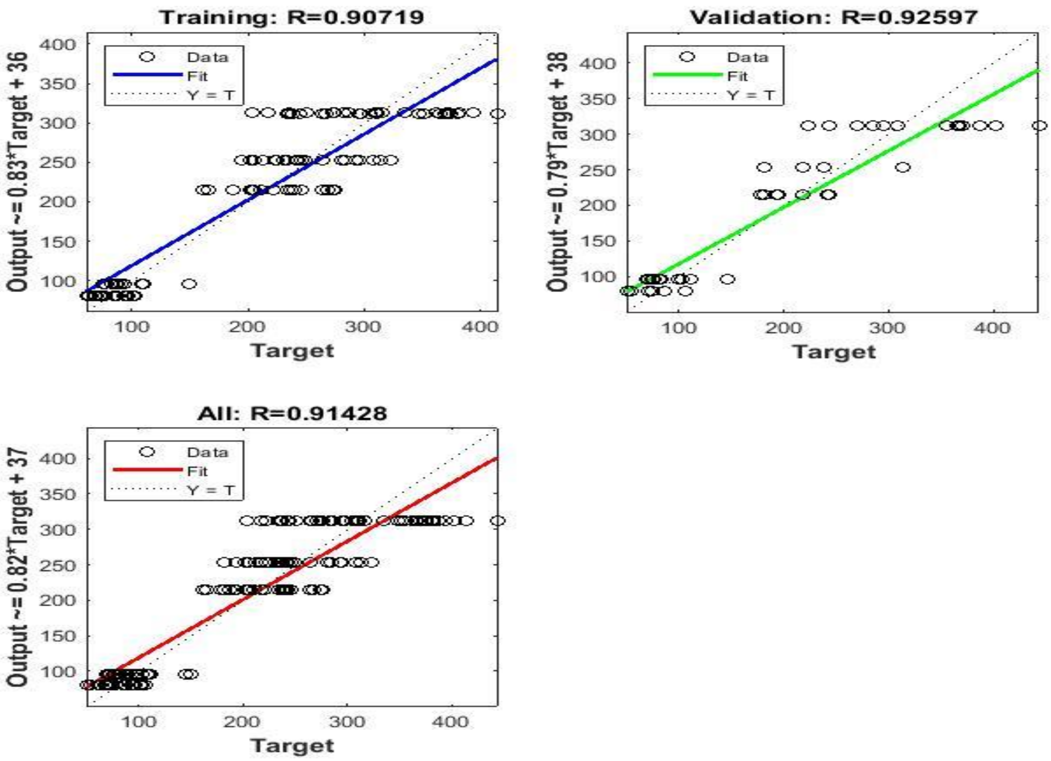


Figure 4-28 Regression Plots

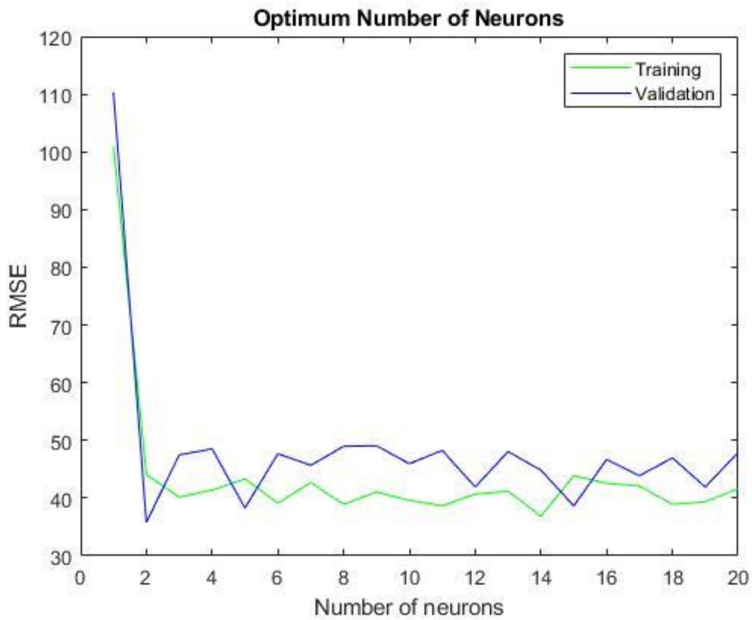


Figure 4-29 Optimum Number of Neurons

Optimum number of neurons is 2.

4.3 The final models:(Multiple Linear Regression)

4.3.1 Weekend Trips:

4.3.1.1 Number of Personal Trips/15 Min:

$\boldsymbol{Y = 292.64[ln(X1)] + 2.183(X2) - 2106.4}$ **(R-square = 0.893)**

Where, Y is the Number of Personal Trips/15 min

X1 is the Gross Floor Area in 1000 sq.ft

X2 is the No of Shop.

4.3.1.2 Number of Vehicular Trips/15 min:

$\boldsymbol{Y = 140.58[ln(X1)] + 0.6(X2) - 603.49}$ **(R-square = 0.894)**

Where Y is the Number of Vehicular Trips/15 min.

X1 is the Gross Floor Area (1000 sq.ft)

X2 is the Number of Shops

4.3.2 Weekday Trips:

4.3.2.1 Number of Personal Trips/15 Min:

$\boldsymbol{Y = 330.7\ [ln(X)] - 1270.6}$ **(R-square = 0.756)**

Where, Y = No of People/ 15-min

X = Gross Floor Area (1000 sq.ft)

4.3.2.2 Number of Vehicular Trips/15 min:

$\boldsymbol{Y = 111.04\ [ln(X)] - 468.4}$ **(R-square = 0.731)**

Where, Y is the Vehicular Trips/ 15-min.

X is the Gross Floor Area (1000 sq. ft)

4.4 The final models:(Partial Least Square Regression)

4.4.1 Weekend Trips:

4.4.1.1 Number of Personal Trips/15 Min:

$$Y = 858.49 + 2.83X1 + 2.699X2 - 38.633X3 + 7.77X4 - 30.345X5 - 0.962X6 - 0.641X7$$

R-Square = 0.98

Where, Y is the Number of Personal Trips/15 min.

4.4.1.2 Number of Vehicular Trips/15 min:

$$Y = 240.12 + 0.793X1 + 0.831X2 - 10.814X3 + 2.82X4 - 8.493X5 - 0.249X6 - 0.187X7$$

R-Square = 0.98

Where Y is the Number of Vehicular Trips/15 min.

4.4.2 Weekday Trips:

4.4.2.1 Number of Personal Trips/15 Min:

$$Y = 614.9 + 1.49X1 + 1.59X2 - 20.3550X3 + 5.726X4 - 17.3271X5 - 0.355X6 - 0.437X7$$

R-Square = 0.87

Where, Y = No of People/ 15-min

4.4.2.2 Number of Vehicular Trips/15 min:

$$Y = 183 + 0.557X1 + .557X2 - 7.210X3 + 1.163X4 - 5.341X5 - 0.812X6 - 0.088X7$$

R-Square = 0.83

Where, Y is the Vehicular Trips/ 15-minutes.

Gross Floor Area (1000sq.ft) = (X1)

Shopping Area (1000sq.ft) = (X2)

Playing Area (1000sq.ft) = (X3)

Watching Area (1000sq.ft) = (X4)

Dining Area (1000sq.ft) = (X5)

No of Shops = (X6)

Parking Spaces = (X7)

Figure 4-30 Comparison of R-Square

Comparison of R- Square				
	Weekend Trips		**Weekday Trips**	
Method	Personal Trips	Vehicular Trips	Personal Trips	Vehicular Trips
Multiple Linear Regression	0.893	0.894	0.756	0.731
Partial Least Square Regression	0.98	0.98	0.87	0.83
Artificial Neural Networks	0.984	0.98	0.87	0.83

Chapter 05

Conclusion and Recommendations

5.1 Conclusion:

This study has covered the gap by developing the Trip Attraction rates of the shopping centers of two metropolitan cities. Real data collection enabled this study to be more efficient and its use in actual real-world problems. Statistical techniques are used for the analysis of the data i.e., Multiple Linear Regression followed by logarithmic regression, partial least square regression. Artificial neural networks are used for training and validation of the model. All the models gave significant results.

A total of eight models have been made. Among them four models are made using logarithmic regression and four using partial least square regression. The logarithmic models have gross floor area and number of shops as independent variables. These models will be used where we have only two variables for the prediction of variables (Gross Floor Area and Number of Shops). Partial least square regression has gross floor area, shopping area, playing area, dining area, watching area, number of shops and parking spaces as independent variables.

Several studies have been done on the trip attraction rates of shopping centers characteristics but there is a need for the transportation planners of Pakistan to plan effectively by using our own models because every country has its own demographic, social and economic characteristic. The computed trip attraction rates will be helpful for planning traffic around the shopping centers and for the parking. The information would be helpful in determining how much traffic a new shopping mall will generate both locally and across the entire region.

5.2 Recommendations:

By considering the scope and work of this study, the recommendations can be summarized as follows:

- The data collection in Pakistan has a lot of problems and it is very difficult to collect the data. All the stakeholders should be educated about the importance of research for the country.
- Traffic Problems like congestion around the shopping centers can be solved as these models predict the trip attraction rates.
- This study has great use on implementing proper transportation facilities for Islamabad and Lahore.
- As it is the first step for developing the trip attraction models, more study should be done by including hospitals, educational centers, recreational centers and offices.
- There was no study in the past, so the change in travel change pattern can’t be studied but in future it will help to identify the changes in the travel pattern.
- This study forms the base and first step for the development of local database in the form of trip generation manual or guidelines for the country. Trip generation manual can be made on national level if more study has been done.

References

Fillone, A. M., Tecson, M. R., Sia, R., & Viray, P. J. P., Eastern Asia Society for Transp. Studies. (2003). Trip attraction of mixed-use development in Metropolitan Manila. *4*, 860-868.

George, P., Kattor, G. J. J. I. J. o. R. i. E., & technology. (2013). Forecasting trip attraction based on commercial land use characteristic. *2*(9).

Majeed, N., & Qasim, G. (2021). *Trip attraction model of selected zones in Baghdad city.* Paper presented at the Journal of Physics: Conference Series.

Manual, I. T. G. J. W., DC, USA, ISBN-10. (2000). Institute of Transportation Engineers. 1-9333452.

Nguyen, Q. H., Ly, H.-B., Ho, L. S., Al-Ansari, N., Le, H. V., Tran, V. Q., . . . Pham, B. T. J. M. P. i. E. (2021). Influence of data splitting on performance of machine learning models in prediction of shear strength of soil. *2021*, 1-15.

PASRA, M., RAMLI, M. I., ADISASMITA, S. A., & PALLU, M. S. (2013). *The shopping trips for traditional markets in Makassar City: Trip attraction and frequency analysis using regression and logit model.* Paper presented at the Proceedings of the Eastern Asia Society for Transportation Studies.

Rahayu, S., Basuki, Y., & Gritanarum, M. (2020). *Trip Attraction Model of 4 in 1 Shopping Center Concept for Sustainable Development in Semarang City.* Paper presented at the IOP Conference Series: Earth and Environmental Science.

Roza, A., Yusnita, Y., & Rusli, A. M. (2021). *Trip Attraction Model on the Education Center on Pangilun Mountain Area, Padang City.* Paper presented at the 4th International Conference on Sustainable Innovation 2020–Technology, Engineering and Agriculture (ICoSITEA 2020).

SAMEER, M., ISLAM, M. S., & LIANA, F. K. (2015). DETERMINATION OF TRIP ATTRACTION RATES OF DIFFERENT SIZE HOSPITALS IN DHAKA CITY.

Shamim Al Razib, M. (2017). Determination of Trip Attraction Rates of Shopping Centers in Uttara Area, Dhaka. *American Journal of Management Science and Engineering, 2*(5). doi:10.11648/j.ajmse.20170205.19

Sosslau, A. B., Hassam, A. B., Carter, M. M., & Wickstrom, G. V. (1978). *QUICK-RESPONSE URBAN TRAVEL ESTIMATION TECHNIQUES AND TRANSFERABLE PARAMETERS. USER'S GUIDE* (0077-5614). Retrieved from

Suthakaran, S., Dharmakeerthi, U., & Sathyaprasad, I. Development of Trip Attraction Rates and Parking Standards for Supermarkets in Kandy Area.

Uddin, M., Hasan, M. R., Ahmed, I., Das, P., Uddin, M. A., Hasan, T. J. I. J. o. C., & Engineering, E. (2012). A comprehensive study on trip attraction rates of shopping centers in dhanmondi area. *12*(4), 12-16.

Waloejo, B. S., & Surjono, H. S. J. J. A. E. B. S. (2012). The influence of trip attraction on the road's level of service (LOS) at traditional market land use. *2*(1), 92-96.

Yulianto, B., Purnomo, S., & Prasetyo, R. (2020). *Study of standard trip attraction models of various land use in the Surakarta City.* Paper presented at the Journal of Physics: Conference Series.

Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	Muhammad Hassam Jadoon
Assignment title:	Quick Submit
Submission title:	Study the Trip Attraction Rates and Developing the Trip Attr...
File name:	Hassam_Jadoon_Plagiarism.docx
File size:	3.88M
Page count:	71
Word count:	11,681
Character count:	63,519
Submission date:	29-Nov-2023 01:40AM (UTC-0800)
Submission ID:	2241820899

Study the Trip Attraction Rates and Developing the Trip Attraction Models for Mega Shopping Centers of Pakistan

Final Thesis Dissertation

By

Muhammad Hassam Jadoon

(00000328618)

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY

ISLAMABAD, PAKISTAN

NOVEMBER 2023

1

Study the Trip Attraction Rates and Developing the Trip Attraction Models for Mega Shopping Centers of Pakistan

ORIGINALITY REPORT

17%	14%	10%	8%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.researchgate.net Internet Source	2%
2	home.nps.gov Internet Source	2%
3	Submitted to Multimedia University Student Paper	1%
4	www.diva-portal.org Internet Source	1%
5	Submitted to University of Essex Student Paper	1%
6	download.atlantis-press.com Internet Source	1%
7	edoc.tips Internet Source	1%
8	Chang Gao, Li Wang, Xu Tian, Guo-Min Song. "Sleep quality and the associated factors among in-hospital nursing assistants in	<1%

	general hospital: A cross-sectional study", Heliyon, 2022 Publication	
9	eprints.utar.edu.my Internet Source	<1%
10	cse.anits.edu.in Internet Source	<1%
11	www.scribd.com Internet Source	<1%
12	N H Majeed, G J Qasim. "Trip attraction model of selected zones in Baghdad city", Journal of Physics: Conference Series, 2021 Publication	<1%
13	Mojtaba Khatami, Afsana Jerin Shayerey, Santosh Kumar Behera, Deyuan He, Noor Azam Haji-Othman. "Exploring Bangladeshi EFL Learners’ Mindfulness and Engagement in Online Learning Context", Research Square Platform LLC, 2023 Publication	<1%
14	Submitted to Université Saint-Esprit Kaslik Student Paper	<1%
15	www.mdpi.com Internet Source	<1%
16	sportdocbox.com Internet Source	<1%

17	pure.tue.nl Internet Source	<1 %
18	media.neliti.com Internet Source	<1 %
19	www.dcee.ruh.ac.lk Internet Source	<1 %
20	core.ac.uk Internet Source	<1 %
21	elibrary.matf.bg.ac.rs Internet Source	<1 %
22	eurchembull.com Internet Source	<1 %
23	repository.unhas.ac.id Internet Source	<1 %
24	Noor Moutaz Asmael, Noor Kadhim. "Estimate Attraction Rate for Shopping Centers", Transport technic and technology, 2020 Publication	<1 %
25	"Recent Advances in Civil Engineering", Springer Science and Business Media LLC, 2023 Publication	<1 %
26	Tony Shun-Te Yuo, Colin Lizieri. "Tenant Placement Strategies within Multi-Level	<1 %

Large-Scale Shopping Centers", Journal of Real Estate Research, 2020
Publication

27	www-sre.wu-wien.ac.at Internet Source	<1 %
28	apessay.elementfx.com Internet Source	<1 %
29	www.ijitee.org Internet Source	<1 %
30	civilejournal.org Internet Source	<1 %
31	etda.libraries.psu.edu Internet Source	<1 %
32	wiki.commres.org Internet Source	<1 %
33	Submitted to Capella University Student Paper	<1 %
34	vdocuments.pub Internet Source	<1 %
35	www.testmagzine.biz Internet Source	<1 %
36	article.sciencepublishinggroup.com Internet Source	<1 %
37	erepository.uonbi.ac.ke Internet Source	<1 %

38	Submitted to University of Hong Kong Student Paper	<1 %
39	statistics.ats.ucla.edu Internet Source	<1 %
40	Md. Shamim Al Razib. "Determination of Trip Attraction Rates of Shopping Centers in Uttara Area, Dhaka", American Journal of Management Science and Engineering, 2017 Publication	<1 %
41	Submitted to Open University Malaysia Student Paper	<1 %
42	www.vre.org Internet Source	<1 %
43	repository.up.ac.za Internet Source	<1 %
44	jurnalmahasiswa.uma.ac.id Internet Source	<1 %
45	Submitted to Asia Pacific University College of Technology and Innovation (UCTI) Student Paper	<1 %
46	Submitted to Flinders University Student Paper	<1 %
47	Submitted to Malaviya National Institute of Technology Student Paper	<1 %

48	ejournal.unp.ac.id Internet Source	<1 %
49	www.coursehero.com Internet Source	<1 %
50	www.researchsquare.com Internet Source	<1 %
51	Debebe Dangiso, Addisu Abebe, Kanbiro Orkaido. "Determinants of Employee Recruitment in Sidama National Regional State, Ethiopia", Qeios Ltd, 2023 Publication	<1 %
52	Submitted to Visvesvaraya Technological University Student Paper	<1 %
53	Y Basuki, S Rahayu, A F Andrika. "Complexity of small-scale commercial area on Menoreh Street in Semarang and its contribution to trip attraction model", IOP Conference Series: Earth and Environmental Science, 2021 Publication	<1 %
54	acikbilim.yok.gov.tr Internet Source	<1 %
55	eopwebsvr.blob.core.windows.net Internet Source	<1 %
56	eprints.usm.my Internet Source	<1 %

57	etd.uum.edu.my Internet Source	<1 %
58	faculty.salisbury.edu Internet Source	<1 %
59	ikesra.kra.go.ke Internet Source	<1 %
60	irc.itum.mrt.ac.lk Internet Source	<1 %
61	researchbank.swinburne.edu.au Internet Source	<1 %
62	sciendo.com Internet Source	<1 %
63	"Recent Advances in Modeling and Forecasting Kaiyu", Springer Science and Business Media LLC, 2023 Publication	<1 %
64	B Yulianto, Setiono, Sugiyarto, S Purnomo, R A Prasetyo. "Study of Standard Trip Attraction Models of Various Land Use in the Surakarta City", Journal of Physics: Conference Series, 2020 Publication	<1 %
65	DAGIMWORK ASELE MANUKA, TAREKEGN SHIRKO LACHORE. "Modeling Trip Generation Pattern by Adopting Multiple Linear	<1 %

Regression in at Existing Traffic Condition", Research Square Platform LLC, 2022
Publication

66	Meyer, Michael D.. "Travel Demand and Network Modeling", Transportation Planning Handbook, 2016. Publication	<1%
67	Submitted to The Robert Gordon University Student Paper	<1%
68	hdl.handle.net Internet Source	<1%
69	repository.nida.ac.th Internet Source	<1%
70	worldwidescience.org Internet Source	<1%
71	Arjun Kathayat. "Leadership Style as a Predictor of Employee Safety Performance in the Oil and Gas Industry.docx", SAGE Publications, 2023 Publication	<1%
72	S Rahayu, Y Basuki, M Gritanarum. "Trip Attraction Model of 4 in 1 Shopping Center Concept for Sustainable Development in Semarang City", IOP Conference Series: Earth and Environmental Science, 2020 Publication	<1%

73	scholarcommons.usf.edu Internet Source	<1 %
74	www.science.gov Internet Source	<1 %

Exclude quotes On

Exclude bibliography On

Exclude matches Off