SURGE PRICING IMPACT OF RIDE-SOURCING SERVICES IN DHAKA CITY

MAHFUZA AKHTER MISHU (170051014)

SIFAT KHAN (170051015)

SHAMS AFRIN ISLAM (170051049)

A THESIS SUBMITTED FOR THE DEGREE OF BACHELOR OF SCIENCE IN CIVIL ENGINEERING

(TRANSPORTATION)

Department of Civil and Environmental Engineering
Islamic University of Technology (IUT)

2022

PROJECT REPORT APPROVAL

This is to certify that the thesis entitled **"Surge Pricing Impact of Ride-Sourcing Services in Dhaka City"**, by Mahfuza Akhter Mishu, Sifat Khan, and Shams Afrin Islam has been approved fulfilling the requirements for the Bachelor of Science Degree in Civil & Environmental Engineering.

SUPERVISOR

Dr. Nazmus Sakib

Assistant Professor

Department of Civil and Environmental Engineering (CEE)

Islamic University of Technology (IUT)

Board Bazar, Gazipur, Bangladesh

DECLARATION OF CANDIDATE

We hereby certify that the undergraduate research work reported in this thesis was completed by us under Dr Nazmus Sakib's supervision, and that this work has not been submitted elsewhere for any reason (except for publication).

Name of the Candidates:

MAHFUZA AKHTER MSHU

Student ID: 170051014

Academic Year: 2020-2021

Date:

SIFAT KHAN

Student ID: 170051015

Academic Year: 2020-2021

Date:

SHAMS AFRIN ISLAM

Student ID: 170051049

Academic Year: 2020-2021

Date:

DEDICATION

We dedicate this thesis to our parents, who for many years gave up important time, money, and effort in order for us to be who we are now. We also want to thank Dr. Nazmus Sakib, our esteemed supervisor.

ACKNOWLEDGEMENTS

"In the name of Allah, the Most Gracious, the Most Merciful."

All gratitude is due to Almighty Allah (SWT) for providing us with the strength and fortitude to complete our Undergraduate Thesis successfully. Our parents have been a constant source of inspiration and support for us, and we want to express our gratitude to them.

We would like to convey our deep gratitude and sincere appreciation to our supervisor, Dr. Nazmus Sakib, Assistant Professor, Department of Civil and Environmental Engineering, Islamic University of Technology, for his courteous direction, astute advice, and never-ending support. His technical and editorial input was critical to the successful completion of this academic project. The paper would not have been completed without his help and instruction.

We'd also like to express our gratitude to all of the faculty members who made meaningful recommendations during our research. For their valuable suggestions and cordial support, we are grateful to our friends, juniors, seniors, and batchmates in our department.

A special mention for our beloved friends Farhan Anjum Badhon and Lubaena Masud for their constant support and feedback throughout our research work. We are very much grateful to them.

Abstract

Surge pricing (or dynamic pricing) has received mixed reactions over the years in case of ride-sourcing services. Different opinions regarding the impact of surge pricing from the stakeholders of ride-sourcing services; especially the drivers and users of these services created a doubt whether it is beneficial or not. In this study, we have tried to explore the impact of surge pricing from both the driver and user's perspective. Our research has been divided into three segments. At first, we wanted to analyze the actual time at which the riders of Dhaka city earn the most using Bayesian Belief Network. Secondly, the investment of a driver (waiting time and fuel cost) for ride-sourcing services has been observed through linear regression in Python. And at the end, we analyzed the responses of an online survey to understand the user characteristics of ride-sourcing services in Dhaka city based on different parameters. Finally, we recommended some policies that can be implemented by the government on ride-sourcing services of Dhaka city. This study gives an idea about the actual impact of surge pricing on the ride-sourcing stakeholders of Dhaka city.

Keywords: Surge pricing, Ride-sourcing services, Peak hour income, Bayesian Belief Network, Waiting time, Fuel cost, Linear Regression.

Table of Contents

Table of Figures

Table of Tables

1. Introduction

1.1 Motivation

Over the years, ride-sourcing platform has gained much popularity worldwide. Ride-sourcing is a practice of sharing automobile trips, which has recently become technology based. Nowadays, websites or mobile apps are being used to connect drivers and those seeking rides. This method is a sustainable and environmentally friendly mode of commute.[1] And so following global trends some of the ride-sharing companies started their operations in Bangladesh like Uber, Pathao, Amarbike, Taxiwala, Garivara, Chalo, Sohoj ride, Obhai, Obon, and others.[2] Gaining market share, ride-sharing companies have given rise to heated discussion over the market's regulation. One such issues is "Surge-pricing or peak-hour pricing", which is a dynamic pricing method or time- based pricing system. This temporarily increases the price with the increasing demand and limited supply.[3] Peak-hour pricing adjusts the time out fare in true time based on the market scenario of a geographic area, the base day out fare is adjusted by using ability of a multiplier automatically generated from the platform's algorithm.[4] This pricing system has received sizeable criticism leaving many people in confusion with questions such as, Is it really good or bad? Who benefits from surge pricing? Some think that surge pricing can hurt the riders as a form of price discrimination, while others think that it can be a drawback for the drivers who do not share rides in peak-hour. But, despite of all the debates, peak-hour pricing has proved to be economically efficient. It has increased the total welfare by 1.59% of gross revenue—or $0.19 per trip—relative to uniform pricing.[5] The peak-hour pricing came to be a holistic metric that measures the relation of vehicle provide and rider demand, every temporally and spatially. As much as riders and drivers would like to forecast peak-hour pricing for trip/business planning, it would benefit ride-sourcing companies to pro-actively manage fleet operation with information of the distribution of vehicles and drivers in advance.[6]

In our study, we want to know the driver and rider's behavioral pattern, their investment in case of peak-hour pricing of ride-sharing services; after predicting the actual peak-hour. This work is mainly done to show a correlation between peak-hour pricing of ride sharing services and the characteristics and lifestyle of the drivers and riders.

1.2 Objectives of the study

The main objectives of the study are to:

- Analyze the time at which the drivers earn most from ride-sourcing services in Dhaka city.
- Analyze a driver's investment for ride-sourcing services in terms of waiting time and fuel cost.
- Finding out behavior pattern and characteristics of riders based on an online survey.

1.3 Primary research questions and challenges

1.3.1 Primary research questions

As this study focuses both on the driver's and rider's perspective, their benefits and detriments; for this purpose, to collect all the necessary data, we have asked some questions to the drivers and riders. So, there were two types of questionnaires. One for the drivers and one for the rides. The primary research questions of this study are discussed below.

Questionnaire for the drivers:

- At which time of the day do you prefer to share rides?
- At which time of the day do you earn the most?
- How much do you earn during that time?
- At which time of the day do you stop sharing rides?
- What is the average number of trips you complete per day?
- What is your average waiting time between the trips?
- At which time of the day do you earn the most?
- What is your average waiting time between the trips at peak period?
- How much money do you spend daily on the fuel of the bike?
- What is your average income?

Questionnaire for the riders:

- What's your regular mode of commuting to and from the workplace?
- If the pricing of ride sharing services decreases by 20% at that time, will you use the services?
- What's the best reason for you to use ride-sharing services in place of public transport?
- If metro rail service is available, will you still use ride-sourcing services?
- At which time of the day do you use ride-sourcing services?
- Are you bound to use ride-sourcing services at that period of the time every day?
- If the pricing of the ride-sourcing services increases by 20% at that time, will you still use the service?

1.3.2 Challenges

Like any other data collection process, we also had to deal with some challenges. For our study, the main challenge was, we have to work with a small-scale data; as we collected our data by one-to-one interview. And our study period was during the covid-19 pandemic. So, maximum time there was lockdown issue. For this reason, we have done our maximum data collection after the lockdown period. Also, in some cases drivers didn't want to share needed information think that, it might hurt their privacy because of sharing that information. Some others were refrained from sharing information thinking this data collection process as time consuming. And, drivers were selected without any demography measure. But, after all those difficulties, we have tried to overcome those challenges and worked to make this study more efficient.

2. Literature Review

In this section, we review the literature from three aspects: (i) peak hour pricing analysis of ride-sharing platforms, (ii) driver's investment in ride-sharing services, (iii) behavior pattern and characteristics of ride-sharing service users. We spotlight our contributions by collating and contrasting our work with previous works and studies.

2.1 Peak hour pricing analysis of ride-sharing platforms

A few related papers analyze how surge-pricing or peak hour pricing helps the ride-sharing (Uber, Pathao, DiDi etc.) drivers in their income.

Surge-pricing or peak hour pricing helps the drivers to make some little more money during the peak hour as the surge multiplier might pay them twice or three times as much for their effort. Li et al. studied impact of dynamic pricing(so-called surge-pricing) on relocation decisions by Uber's driver-partners and corresponding revenue they collected.[7]

Jiao and Junfeng analyzed how ride-sharing(Uber's) peak hour pricing increases when there is a high demand during any special event in the city.[8]

In New York, Uber drivers earn $25 to $30 an hour on average (twice that of cabdrivers and $5 to $10 more on average than Uber drivers earn in other American cities). But there is a disadvantage as for getting a profit, drivers must plan their schedules at the peak period and invest as much as half of their earnings into insurance and car maintenance.[9]

Back in November,2021 Pathao decided to cut commissions on bike rides in line with demands from drivers for potentially increasing the company's competitiveness and pressurizing rival Uber.[10]

It has cut down its commission on bike trips which is allowing the drivers to earn up to 90% of their trip fare during peak hours (8:00 am to 11:00 am and 5:00 pm to 8:00 pm), charging only a 15% commission during the off-peak hours.[11]

Sun et al. worked to determine the optimal pricing strategy for online car hailing platforms, taking both ride details and driver location into account, and assuming that drivers and customers maximize utility.[12]

Chen et al. presented a strategy for avoiding peak hour pricing after analyzing rid-sourcing service's(Uber) surge-pricing algorithm.[13]

Battifarano and Qian used machine learning methods to predict surge prices minutes or hours ahead of time in selected locations in a city, regardless of surge pricing strategies.[6]

2.2 Driver's investment in ride-sharing services

Lu et al. studied whether short-run variation in dynamic prices attracts drivers, the surge heatmap can be seen as an aid that reduces the cost of search for drivers in the ridesharing market.[7]

Coming out of the pandemic, the demand for rides is high but inflation and rising gas prices has made it harder for drivers to earn what they once did.[14]

The goal of surge pricing is to find the "equilibrium price" at which driver supply matches rider demand and riders' wait time is minimized. Studies show that surge pricing achieves what it was designed to do: it brings more drivers online, and it allocates available rides to those who value them more.[15]

Castillo and Juan Camilo showed how Peak hour pricing saves driver's time as it mitigates imbalances between supply and demand, so drivers have to wait less between trips. The value of time to drivers is their average hourly earnings net of driving costs, which is slightly above minimum wages.[5]

Lu et al. showed that Reducing spatial search friction reduces mismatch between the number of riders wishing to take trips and the number of drivers willing to drive, which reduces waiting times, increases the number of trips taken, and improves welfare for both the drivers and users.[7]

So, all the studies showed peak-hour pricing helps the drivers by reducing their investment in the form of waiting time; as well as, reducing waiting time also reduces the unnecessary fuel cost.

2.3 Behavior pattern and characteristics of ride-sourcing service users

From Castillo and Juan Camilo's study we have seen the effects of surge pricing due to a better allocation and time savings (moving from the dashed line to the dotted line) benefit riders across all income levels. High-income riders benefit most: they have a higher value of time and lose more when they are denied a trip. Low-income riders are price sensitive, and thus prefer low prices. High-income riders prefer higher prices that skim other riders and bring in more drivers, resulting in more reliable trips with lower pickup times.[5]

Akbari et al. studied on identifying the factors influencing customer's intention to use ride-sharing services.[16]

Lee et al. studied to find the implications of surge pricing policies for managing congestion and the welfare of both consumer and labor.[17]

Rey et al. characterized ride-sharing service's rides and drivers considering age and gender, and race and showed how different population behave differently.[18]

Junfeng Jiao's collected data shows that the price surges can produce a ride cost that is 6.1 times a non-surge rate, which may make the service prohibitively expensive to some citizens during peak times.[8]

Li et al. showed that ridesharing platforms could reduce traffic congestion by facilitating the use of public transit, which is a more environmentally sustainable mode. On the other hand, ridesharing services could also increase traffic congestion by inducing additional travel demand via ridesharing services switching from public transit.[19]

Sun et al. stated in their literature that the customer only accepts a price and driver combination if the associated total travel cost, including a waiting time cost, is at most equal to the expected total ride cost of the offline taxi service.[12]

3. Chapter 1: Peak hour income analysis of Pathao drivers

In this chapter we are going to discuss the first part of our work, Peak hour income analysis of Pathao drivers. Here we have considered the time period where the drivers earn the most to be the peak hour.

3.1 Sample and Data Collection

To begin the data collection process at first, we decided on some parameters. We selected the following as our parameters:

- Start and End times
- Peak hour earnings
- Total hour of trips
- Total number of trips
- Daily income

After selecting the parameters, we proceeded to data collection. We selected Mirpur 10 to Shewrapara as our location of study. We conducted one on one interviews with forty Pathao drivers from this location. We did not consider any demographic rather chose our candidates on random. However only motorcycle riders were interviewed.

Figure 1: Location of study (Google Maps, n.d.)

In the interview we asked our participants a series of questions. We asked them the following questions:

- At what time do you prefer to start ride sharing?
- At which time of the day do you earn the most?
- What percentage of your income do you earn at the peak hour?
- What is your daily income?
- How many trips do you complete in a day?
- At what time do you stop ride sharing?

From these questions we obtained our required datasets. Afterwards we sorted the datasets and clustered the data. We obtained a total of forty datasets.

3.2 Data

After obtaining the data we sorted and clustered the data. The clustered data is shown below:

Table 1: Clustered data for Bayesian modelling

Highest Earning Time	Trip Time	Daily Income	% Income in peak hour	Total Trips
8 AM - 12 PM	5 - 7 H	500 - 700	40-50	5-7 T
5 PM - 9 PM	8 - 10 H	500 - 700	60-70	8-10 T
5 PM - 9 PM	2 - 4 H	200 - 400	80-90	2-4 T
5 PM - 9 PM	8 - 10 H	500 - 700	40-50	8-10 T
5 PM - 9 PM	2 - 4 H	200 - 400	80-90	2-4 T
8 AM - 12 PM	2 - 4 H	200 - 400	80-90	2-4 T
5 PM - 9 PM	2 - 4 H	200 - 400	60-70	2-4 T
5 PM - 9 PM	8 - 10 H	800 - 1000	40-50	8-10 T
5 PM - 9 PM	8 - 10 H	800 - 1000	60-70	8-10 T
5 PM - 9 PM	5 - 7 H	200 - 400	40-50	5-7 T
5 PM - 9 PM	5 - 7 H	500 - 700	40-50	5-7 T
5 PM - 9 PM	5 - 7 H	500 - 700	60-70	5-7 T
5 PM - 9 PM	5 - 7 H	500 - 700	80-90	5-7 T
8 AM - 12 PM	2 - 4 H	200 - 400	80-90	2-4 T
8 AM - 12 PM	2 - 4 H	200 - 400	80-90	2-4 T
5 PM - 9 PM	2 - 4 H	200 - 400	60-70	2-4 T
5 PM - 9 PM	8 - 10 H	800 - 1000	40-50	8-10 T
8 AM - 12 PM	2 - 4 H	200 - 400	80-90	2-4 T
8 AM - 12 PM	2 - 4 H	200 - 400	80-90	2-4 T
5 PM - 9 PM	5 - 7 H	500 - 700	60-70	5-7 T
9 PM - 11 PM	2 - 4 H	200 - 400	40-50	2-4 T
9 PM - 11 PM	2 - 4 H	200 - 400	40-50	2-4 T
9 PM - 11 PM	2 - 4 H	200 - 400	40-50	2-4 T
9 PM - 11 PM	2 - 4 H	200 - 400	40-50	2-4 T
9 PM - 11 PM	2 - 4 H	200 - 400	40-50	2-4 T
9 PM - 11 PM	2 - 4 H	200 - 400	40-50	2-4 T
9 PM - 11 PM	2 - 4 H	200 - 400	40-50	2-4 T
9 PM - 11 PM	2 - 4 H	200 - 400	40-50	2-4 T
5 PM - 9 PM	8 - 10 H	800 - 1000	60-70	8-10 T
8 AM - 12 PM	2 - 4 H	200 - 400	80-90	2-4 T
5 PM - 9 PM	2 - 4 H	200 - 400	80-90	2-4 T
8 AM - 12 PM	2 - 4 H	200 - 400	60-70	2-4 T
8 AM - 12 PM	5 - 7 H	200 - 400	60-70	2-4 T

5 PM - 9 PM	2 - 4 H	200 - 400	80-90	2-4 T
5 PM - 9 PM	8 - 10 H	500 - 700	40-50	8-10 T
5 PM - 9 PM	2 - 4 H	200 - 400	40-50	2-4 T
5 PM - 9 PM	2 - 4 H	200 - 400	60-70	2-4 T
5 PM - 9 PM	2 - 4 H	200 - 400	80-90	2-4 T
8 AM - 12 PM	2 - 4 H	200 - 400	40-50	2-4 T
5 PM - 9 PM	8 - 10 H	800 - 1000	40-50	8-10 T

3.3 Methodology

In this chapter, we used the Bayesian Belief Network as our main tool for analyzing the obtained data. We used the software Genie which is a Bayesian network inference software to analyze the data. After analyzing we used the software SPSS for determining if our data set was normally distributed. The steps are discussed in the next articles.

3.3.1 Bayesian Belief Network

"A Bayesian network is a representation for probability-based models. It describes the model through direct dependencies between random variables. This eases the computational handling of the model and gives it a simple graphical presentation. Bayesian networks are applied to a wide range of subjects due to their generality and other useful properties."[20] Bayesian belief network has been found to be useful for a wide range of applications. However, one of its significant downfalls is that there is no universal method for applying this technique. This leads to the problem of it being limited by the person programing it. This disadvantage can sometimes also become an advantage as there can be many ways of solving the problem all equally correct. However, since our dataset was relatively simplistic and small in size these problems were minimized.

An advantage of BNN is that adding a new piece of evidence can be done very easily by adding only a few probabilities. Also, the results obtained by analysis using BNN are strictly probabilities.

To carry out the BNN analysis of our data we used the Genie modeler which is a Bayesian network inference software. This software can carry out various types of operations and can analyze dynamic Bayesian networks. We formed a basic model with our assumptions and the software analyzed our data to give us our desired results.

3.3.2 Hypothesis Testing

Hypothesis testing is a statistical inference operation that determines whether the available data can support the assumptions made or not. To determine if our dataset was sufficient, we used the SPSS software which is a statistical software suit. Using SPSS we determined of our data was normally distributed or not.

3.4 Analysis

In this part, we are going to discuss about the analysis process.

3.4.1 Genie modeler

At first step of our analysis, we developed our basic dynamic Bayesian network in Genie.

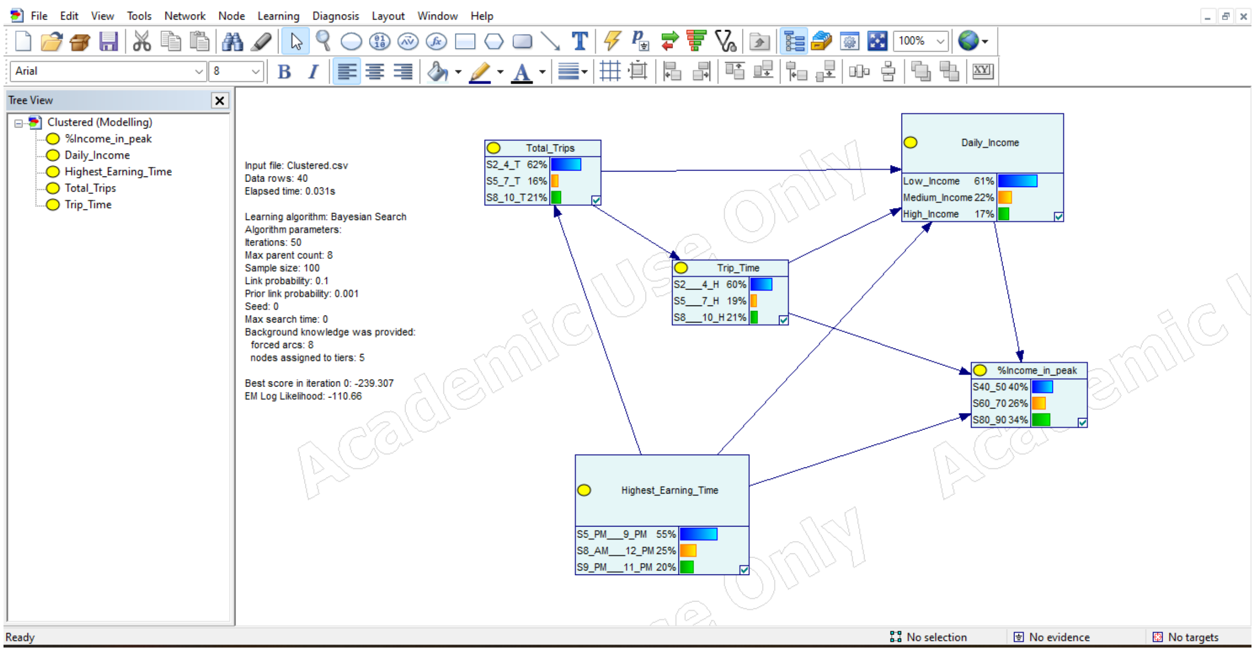

Figure 2: Basic dynamic Bayesian model

At the next step we set our evidence to high income and obtained the following results.

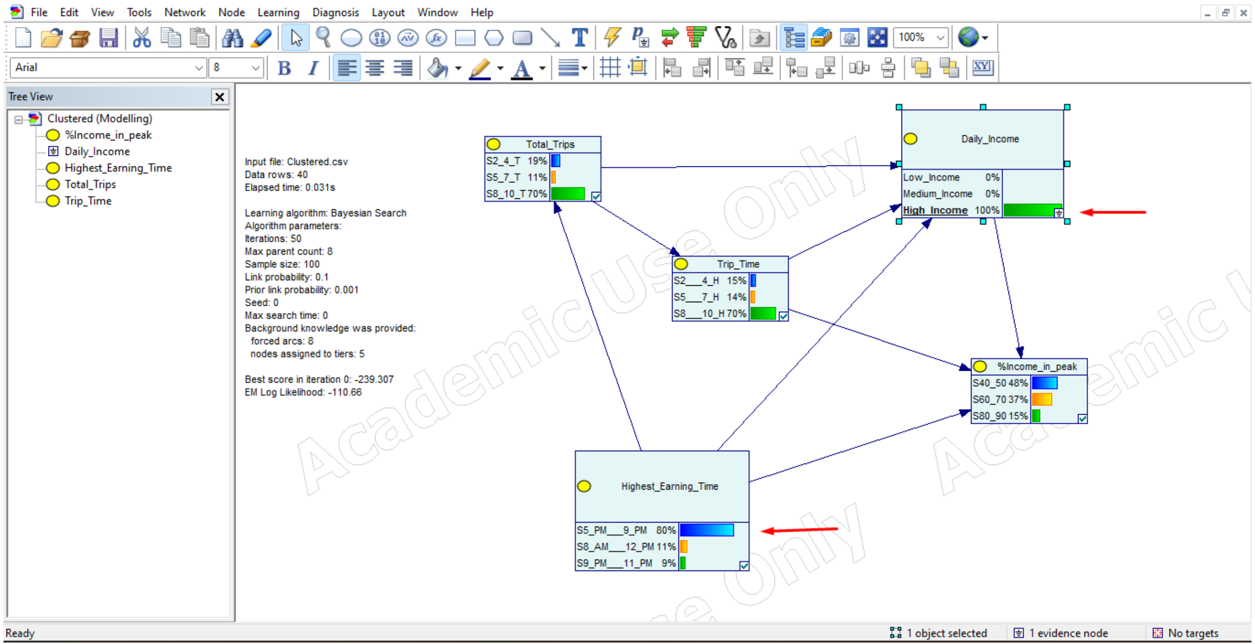

Figure 3: Evidence set to high income

Next, we set the evidence to middle income and obtained its respective results.

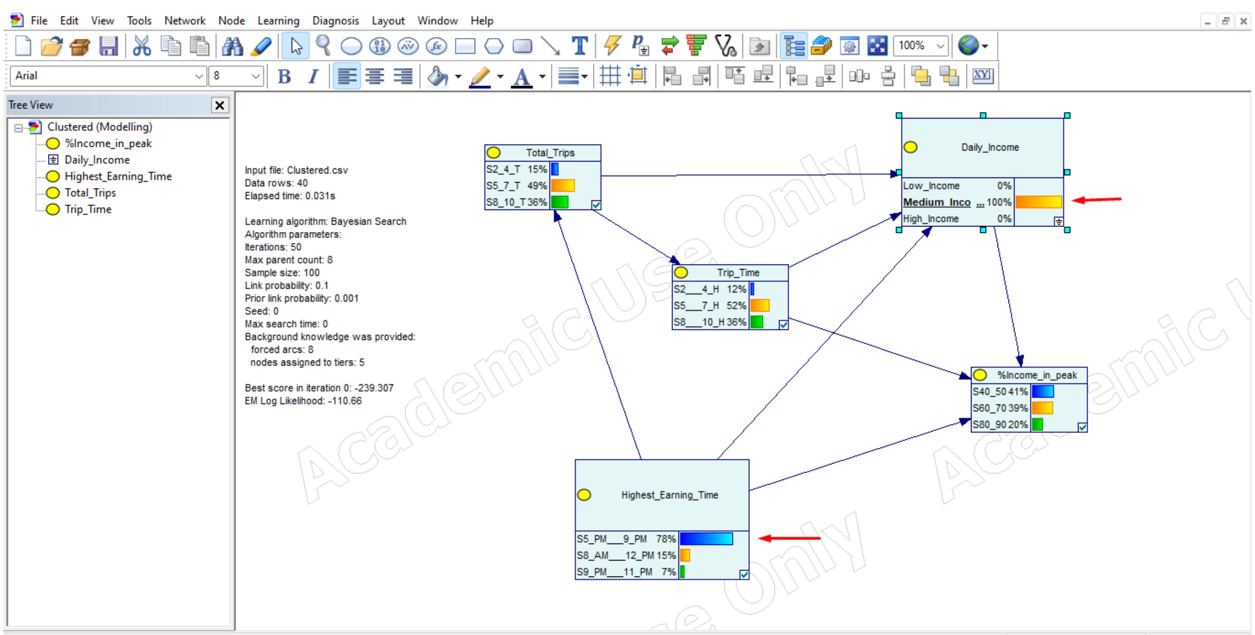

Figure 4: Evidence set to middle income

Then we proceeded to set the evidence to low income and obtain its respective results.

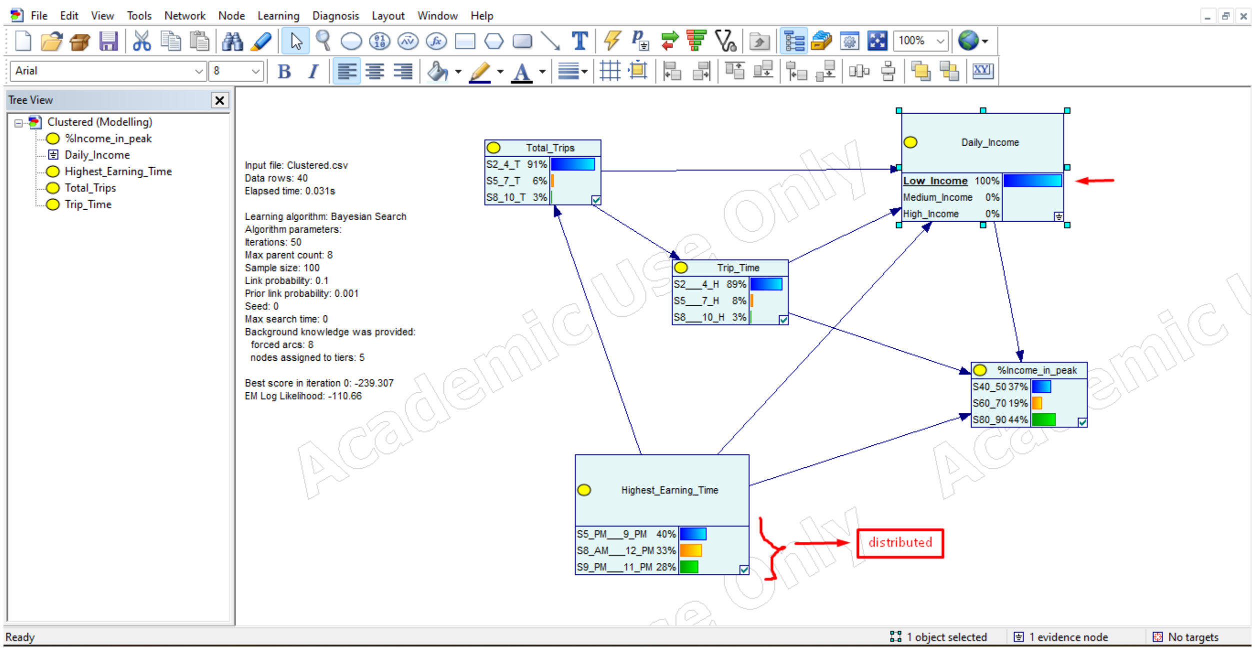

Figure 5: Evidence set to low income

3.4.2 SPSS

We input our data into SPSS to determine if our dataset was normally distributed.

[DataSet1]

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
highest earning time	40	100.0%	0	0.0%	40	100.0%

Descriptives

			Statistic	Std. Error
highest earning time	Mean		1.95	.107
	95% Confidence Interval for Mean	Lower Bound	1.73	
		Upper Bound	2.17	
	5% Trimmed Mean		1.94	
	Median		2.00	
	Variance		.459	
	Std. Deviation		.677	
	Minimum		1	
	Maximum		3	
	Range		2	
	Interquartile Range		1	
	Skewness		.060	.374
	Kurtosis		-.708	.733

Figure 6: SPSS

We obtained the following histogram and graph from the software.

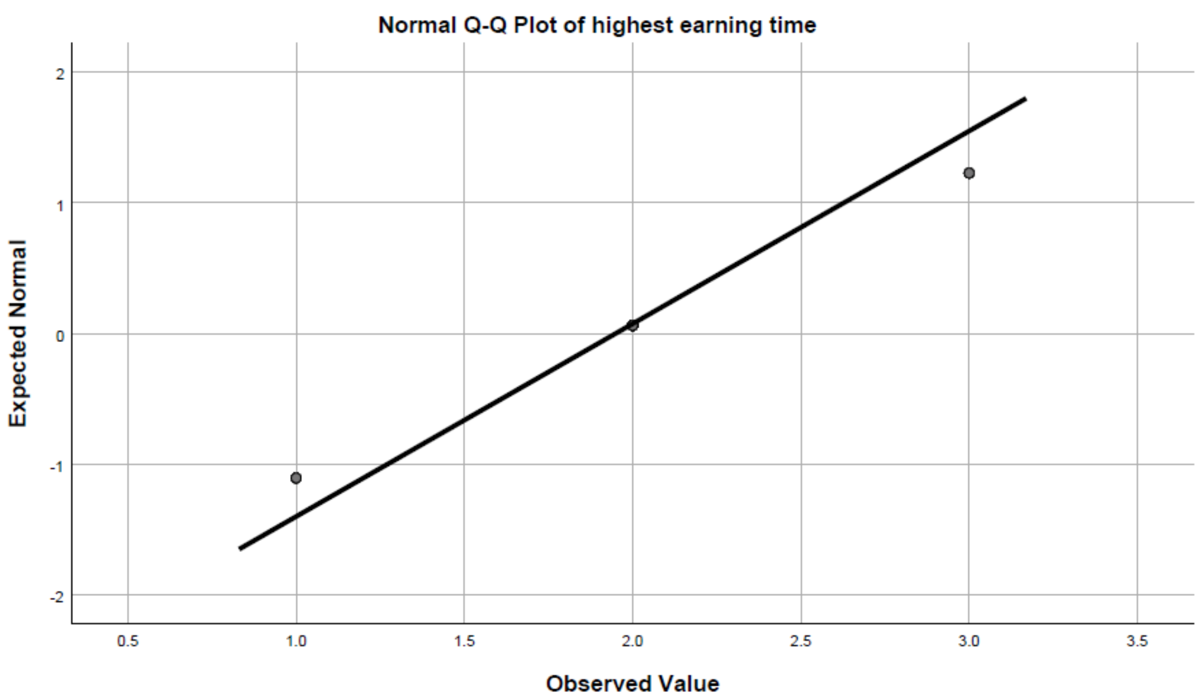

Figure 7: Graph

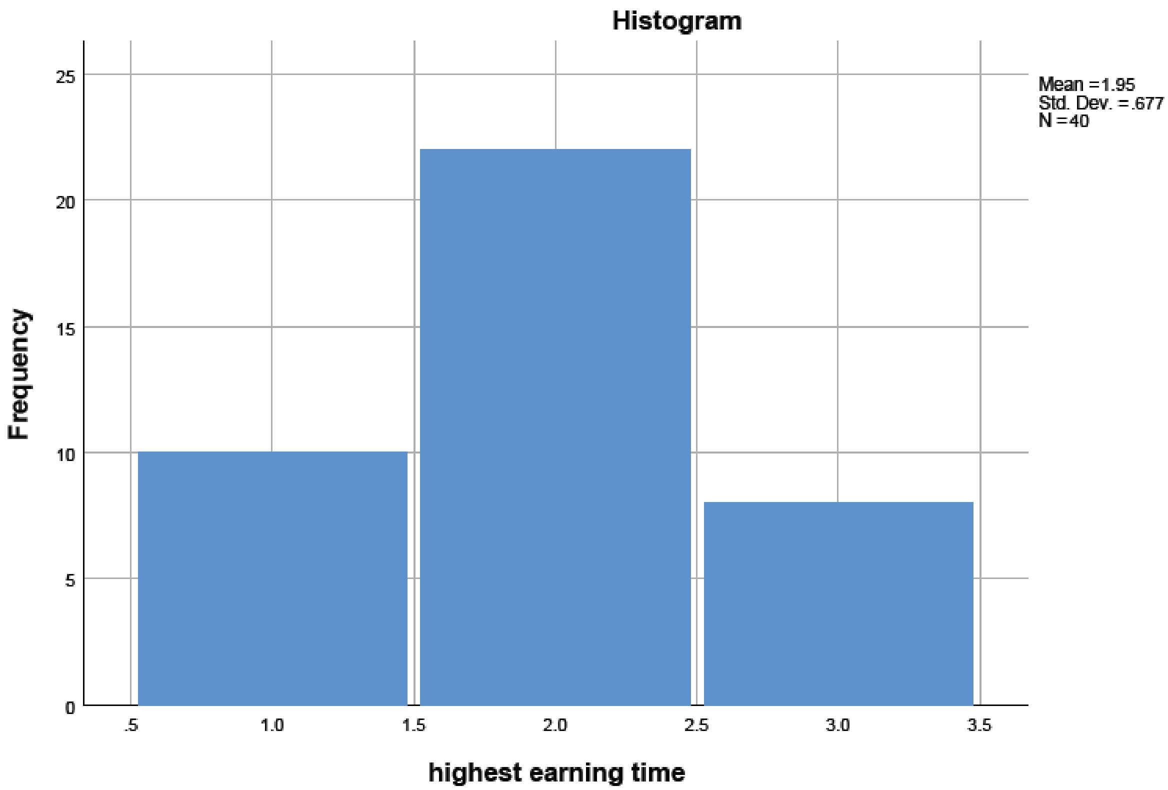

Figure 8:Histogram

3.5 Discussion

From the analysis in Genie, we found that, for the drivers who make high income they prefer to share ride from 5pm to 9pm in which they earn the highest. These types of drivers also share ride for eight to ten hours every day and complete around eight to ten trips. When it comes to drivers with middle income, they also earn the highest during the time frame from 5pm to 9pm. These drivers usually share their rides for 5 to 7 hours daily. Around 49% of these drivers complete around five to seven trips per day. In the next demographic drivers who belong to the low-income category, their preferred time of ride sharing was found to be distributed throughout the day. These people usually share rides for 2 to 4 hours a day and 91% of them share about two to four trips every day. From our analysis we found that the low earners are usually occasional ride sharers. And those who earn medium to highest usually share ride for a significant portion of the day.

Next from analysis carried out on SPSS we found our skewness value to be 0.060 which is very close to one. This indicated our dataset is normally distributed. We also found the kurtosis value to be – 0.708 which is also within the normal range. So, we can conclude that out dataset did not have any significant issues. However, the small size of dataset may have contributed to some error.

4. Chapter-2: Driver's investment in ride-sourcing services

4.1 Sample and Data Collection

In this part of the study, we wanted to understand how much investment a driver while sharing rides using the ride-sourcing platforms and whether it is really worth it or not. In this case we have taken to investment parameters – waiting time and fuel cost.

Castillo et al. stated that, riders are very inelastic in the short run. They are more responsive to prices in the long run, but elasticities are also below one. Riders highly value their time. This is consistent with trips taking place during time sensitive moments: riders need to be in time for an appointment, or they need to get to the airport in time for a flight. High-income riders are even less price sensitive, and time is especially valuable to them. They are more likely to move to the areas with high surge pricing.[5]

We added a few more questions for the bike drivers along with the questionnaire of Chapter-1. We wanted to know their waiting time between trips in regular time and peak time. Also, we asked them about the amount of money they spent daily on fuel while sharing rides too.

According to Pooyan Ehsani, waiting time assigns the exact time that a passenger needs to wait in the passenger's pool before the matching procedure begins. In this work, the waiting time is modeled as an expectation function, and the objective is to minimize the passenger's expected trip cost.[21]

But in our case, we have considered waiting time as a factor which determines how much a driver invests on ride-sourcing services to justify his/her effort and earning along with terms of fuel cost.

4.2 Data

We have done separate analysis for both waiting time and fuel cost. In the first analysis the dependent variable was waiting time. Here we wanted to find out how much time a driver needs to wait between trips during normal time period.

In this case the control variables were – daily number of trips, daily trip time, daily income, and daily average fuel cost. Depending on these variables we wanted to predict the waiting time of drivers while sharing rides in day in the Dhaka city.

On other hand to predict the fuel cost influence we put it as a dependent variable in the analysis and the control variables in this case were - daily number of trips, daily trip time, daily income, and waiting time in normal hours.

Table 2: Data collected for waiting time and fuel cost analysis

SL. NO.	AVG. TRIP	WAITING TIME	EARNING TIME	FUEL COST	DAILY INCOME
1	30	30	3	100-150	500
2	25	25	2	>150	600
3	30	30	2	100	300
4	25	25	1	100-150	700
5	20	20	2	100	300
6	15	15	1	<100	200
7	25	25	2	100	300
8	15	15	2	100-150	800
9	15	15	2	100-150	>1000
10	25	25	3	100-150	400
11	30	30	2	100-150	500
12	30	30	2	100-150	600
13	30	30	1	100-150	500
14	20	20	1	<100	200
15	25	25	2	<100	200
16	20	20	3	100	300
17	15	15	1	100-150	800
18	35	35	2	<100	300
19	30	30	1	100	300
20	25	25	1	100-150	600
21	30	30	1	<100	200
22	40	40	1	<100	200
23	45	45	1	<100	200
24	35	35	1	<100	200
25	45	45	1	<100	200
26	30	30	1	<100	200
27	45	45	1	<100	200
28	30	30	2	<100	200
29	15	15	3	100-150	>1000
30	25	25	1	100	300
31	20	20	3	<100	200
32	20	20	2	100	300
33	25	25	3	100-150	400
34	20	20	3	100	300
35	15	15	3	100-150	700
36	25	25	1	100	400
37	20	20	2	100	300
38	25	25	1	100	300
39	40	40	2	<100	200
40	15	15	3	100-150	800

4.3 Methodology

In this chapter we have done linear regression in Python language to understand the effect of waiting time and fuel cost as investment factors of a driver in ride-sourcing services.

Linear regression using Python has been used in several studies to predict the travel time and forecast demand of ride-sourcing services. For example, in a study done by Shokoohyar et al. we can see the travel time prediction in Ride-Sourcing Networks. This paper explores the applications of machine learning for predicting the travel time in the ride-sourcing networks using the Uber movement dataset. Using the Python programming environment, a case study is presented to analyze the travel time of the ride-sourcing services from the central Washington D.C. to the given specific destinations by considering the distance, railway/subway and street density in different destination zones (areas) and also weather conditions. To this end, in the first step, descriptive analytics is completed to include potential features (attributes) affecting the travel times of Uber (ride-sourcing) services. Then, machine learning techniques such as random forest and robust regressions are applied to identify key attributes (features) for the prediction of the average travel times. The findings and accuracy of the robust regression models are compared with the random forest to select the best model in predicting the mean travel time. This case study provides opportunities in data preparation, descriptive and predictive analytic topics covered in applied machine learning, data science and decision support system courses using data mining programming environments like Python and R. Students are also able to change the study area (city) for this case study based on their interest.[22]

But in this case, we have only wanted to understand the amount of time and money a driver is spending while sharing rides in his bike using the Pathao app in the Dhaka city – whether it's really profitable for him or not.

4.4 Analysis

In case of **waiting time analysis**, we have considered dependent variable or output as waiting time of normal hours. The control variables or inputs were - daily number of trips, daily trip time, daily income, and daily average fuel cost.

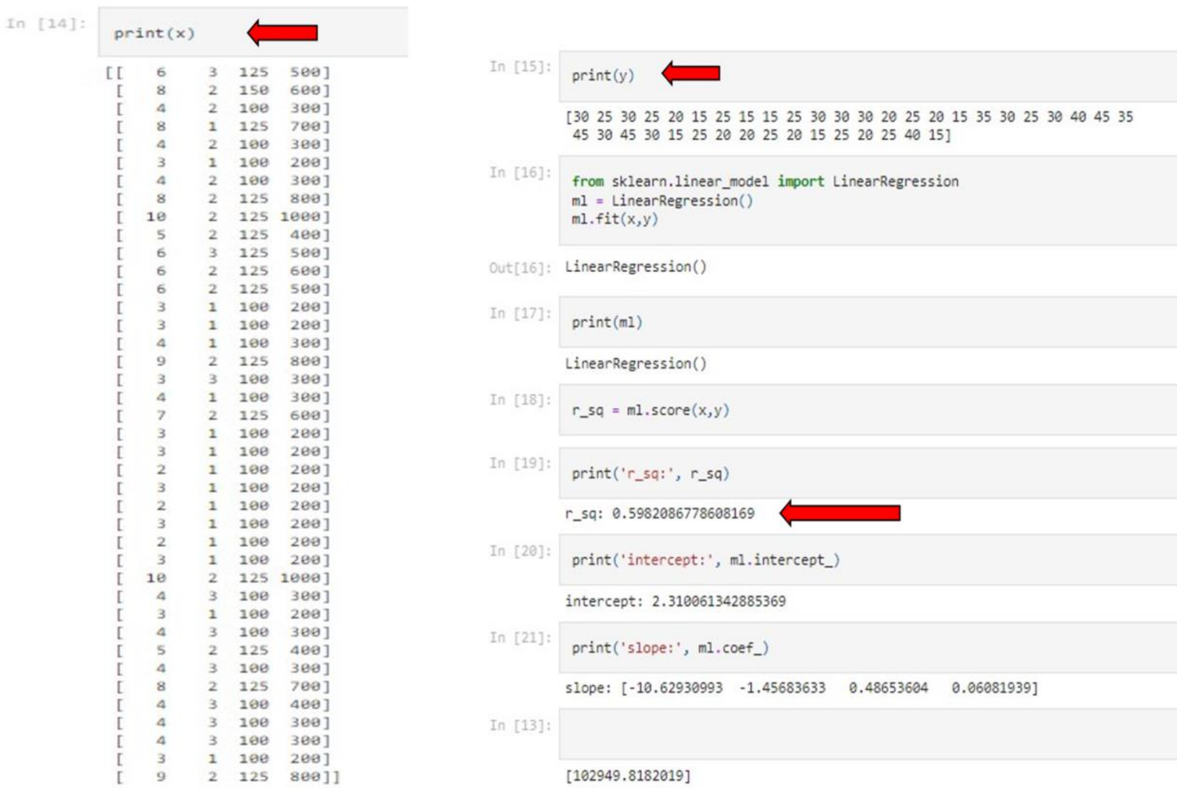

Figure 9: Linear Regression of Waiting Time using Python

In case of **fuel cost analysis**, we have considered dependent variable or output as average daily fuel cost of drivers. The control variables or inputs were - daily number of trips, daily trip time, daily income, and waiting time in normal hours.

```
In [4]: x=df.drop(['Fuel_Money','Waiting_Time_Between_Trips(Peak)'],axis=1).values
        y=df['Fuel_Money'].values

In [5]: print(x)

[[   6   30    3  500]
 [   8   25    2  600]
 [   4   30    2  300]
 [   8   25    1  700]
 [   4   20    2  300]
 [   3   15    1  200]
 [   4   25    2  300]
 [   8   15    2  800]
 [  10   15    2 1000]
 [   5   25    2  400]
 [   6   30    3  500]
 [   6   30    2  600]
 [   6   30    2  500]
 [   3   20    1  200]
 [   3   25    1  200]
 [   4   20    1  300]
 [   9   15    2  800]
 [   3   35    3  300]
 [   4   30    1  300]
 [   7   25    2  600]
 [   3   30    1  200]
 [   3   40    1  200]
 [   2   45    1  200]
 [   3   35    1  200]
 [   2   45    1  200]
 [   3   30    1  200]
 [   2   45    1  200]
 [   3   30    1  200]
 [  10   15    2 1000]
 [   4   25    3  300]
 [   3   20    1  200]
 [   4   20    3  300]
 [   5   25    2  400]
 [   4   20    3  300]
 [   8   15    2  700]
 [   4   25    3  400]
 [   4   20    3  300]
 [   4   25    3  300]
 [   3   40    1  200]
 [   9   15    2  800]]
```

```
In [6]: print(y)

[125 150 100 125 100 100 100 125 125 125 125 125 125 100 100 100 125 100
 100 125 100 100 100 100 100 100 100 100 125 100 100 100 125 100 125 100
 100 100 100 125]

In [7]: from sklearn.linear_model import LinearRegression
        ml = LinearRegression()
        ml.fit(x,y)

Out[7]: LinearRegression()

In [8]: print(ml)

LinearRegression()

In [9]: r_sq = ml.score(x,y)

In [10]: print('r_sq:', r_sq)

r_sq: 0.8068904299937216

In [11]: print('intercept:', ml.intercept_)

intercept: 58.160850492812216

In [12]: print('slope:', ml.coef_)

slope: [12.46489701  0.58713     0.5814194  -0.06194779]
```

Figure 10:Linear regression of Fuel Cost using Python

4.5 Discussion

From the linear regression analysis of waiting time, we got to know that, the average waiting time during normal hours of a Pathao driver in Dhaka city is **26.25 minutes**. But during peak hour the waiting time reduces to **8-20 minutes**. One of the most significant observations here is that, the occasional drivers face more waiting time due to less trip time and trip numbers than the regular drivers. The waiting time being lower during peak hours, the trip generation is very high in this period of time. The R square value of this analysis is **0.59** which means the analysis is valid and satisfactory.

From the linear regression analysis of fuel cost, we got to know that, the average money spent on fuel everyday by a Pathao driver in Dhaka city is **BDT 100**. Also, irrespective of occasional and regular drivers, the average daily income of a Pathao driver here is **BDT 273**. Significant observation is fuel cost is higher when trip generation is high – which is obvious. The R square value of this analysis is **0.80** which means the analysis is valid and satisfactory.

The key takeaway from this analysis in terms of both waiting time and fuel cost it is seen that only the regular or full time Pathao drivers are benefitted sharing rides using the ride-sourcing platform in the Dhaka city.

5. Chapter-3: Behavior pattern and user characteristics of ride-sourcing service users

5.1 Sample and Data Collection

In this chapter we mainly want to understand the user perspective in terms of using ride-sourcing services. We have prepared a questionnaire based on the transportation modes of the users. We wanted to see their urge to use ride-sourcing services based on the pricing, peak hour, metro rail, timing, etc.

We want to utilize the understandings of this chapter to give some recommendations about surge pricing policy in the next chapter.

5.2 Data

We have conducted an online survey through social media groups and collected almost 160+ responses. But while collecting these responses we ignored the demographic division.

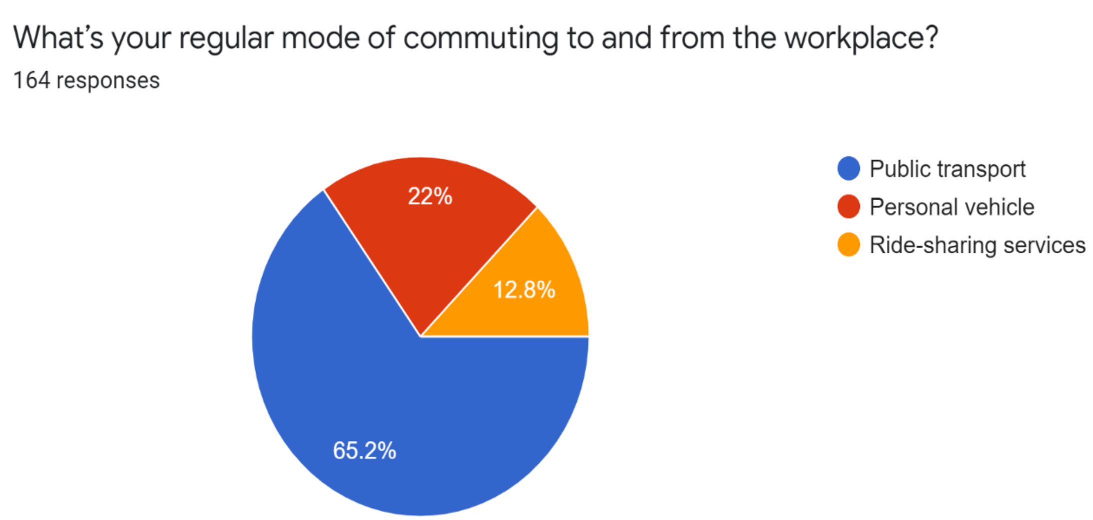

Figure 11: Mode of commute

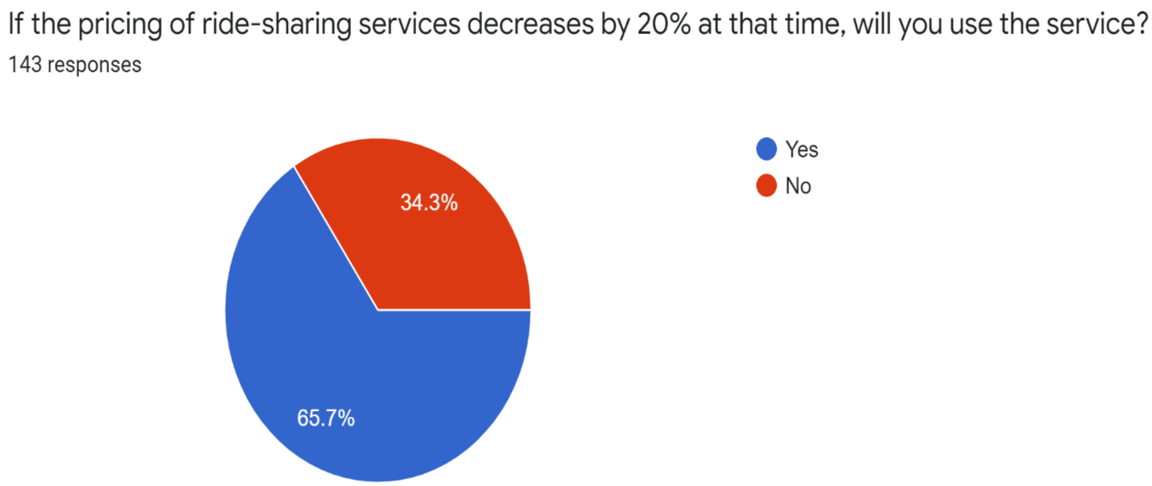

Figure 12:If pricing of ride sharing decreases, will you use the service?

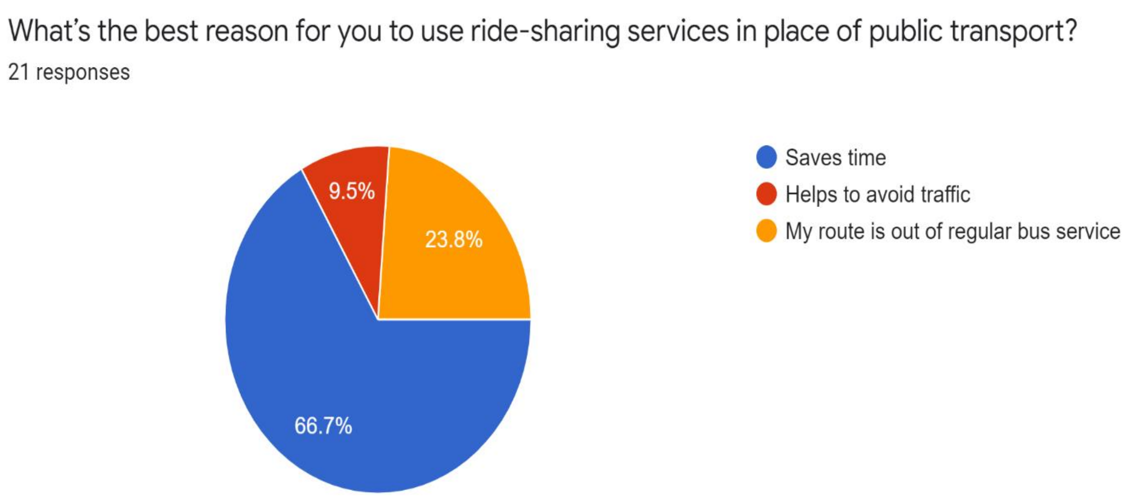

Figure 13: Reason for using ride sharing services

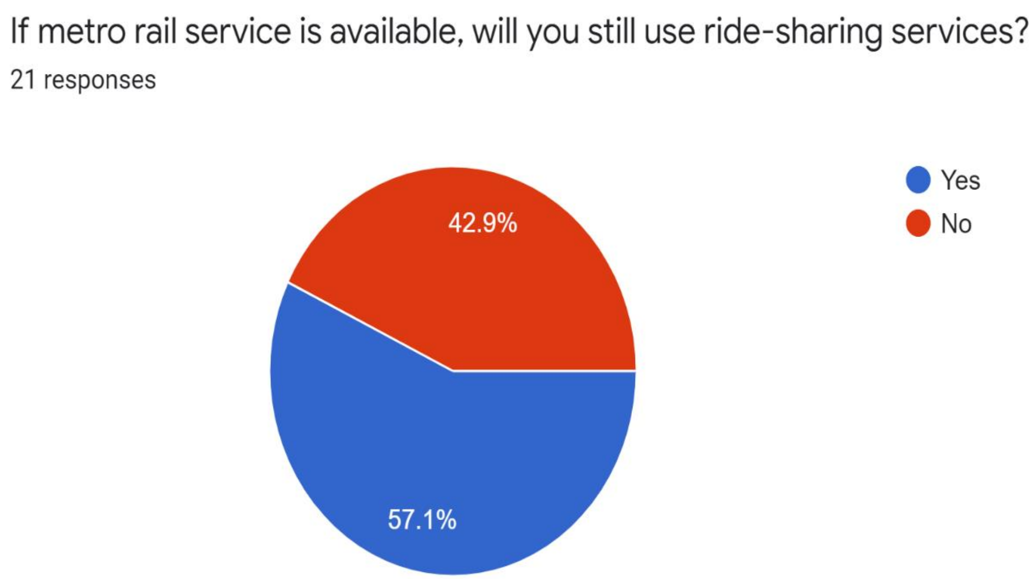

Figure 14:If metro rail is available, will you still use ride sharing services?

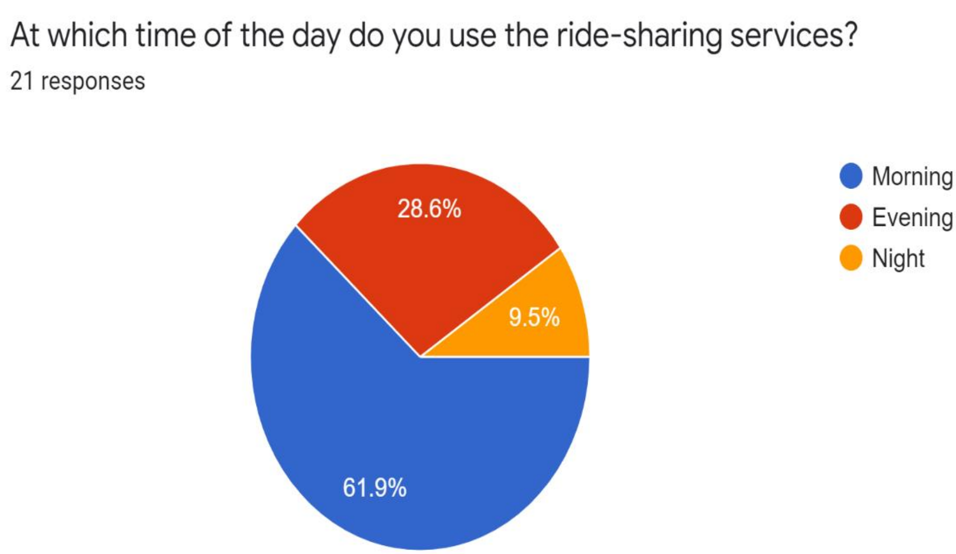

Figure 15: At which time of day do you use ride sharing services?

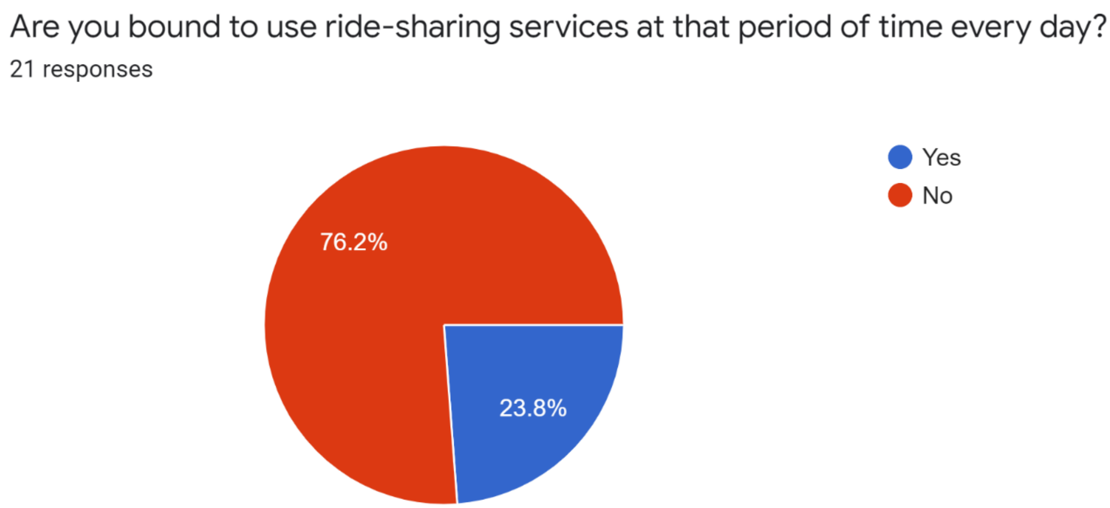

Figure 16: Are you bound to use ride sharing services?

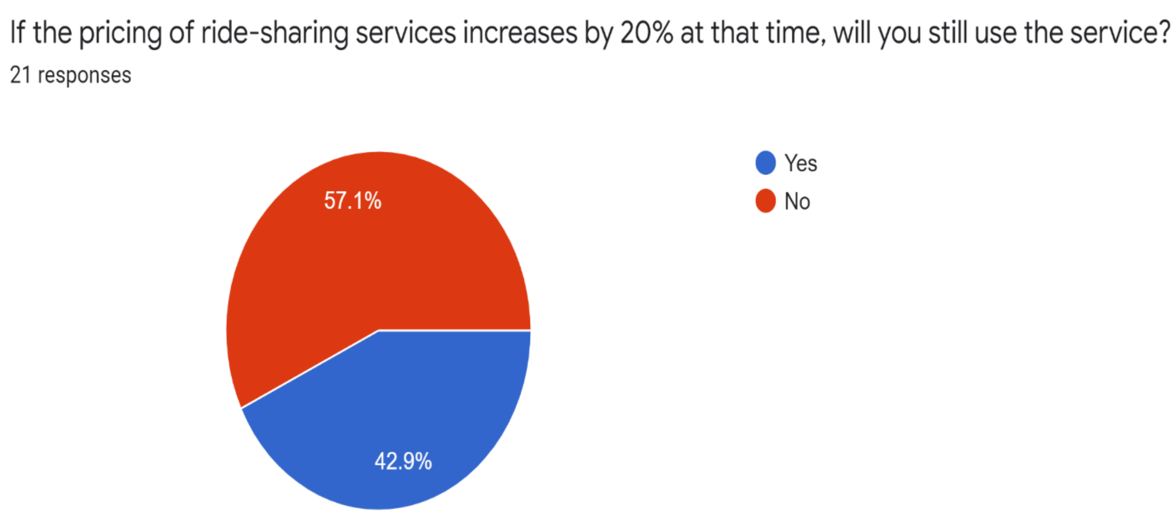

Figure 17: If pricing of ride sharing services increase, will you still use it?

5.3 Findings

The main thing that we wanted to see how the users respond when the pricing of rides is increased and decreased. Also, how much they are bound to use ride sharing services in different time periods of the day. In addition, we wanted to see whether any change will come in the user characteristics of when Metro Rail in Dhaka city is full on operation. Based on the findings of this chapter we want to put on some recommendations in the next section.

6. Recommendations

Based on the fuel cost and waiting time analysis of our study it is evident that only full-time drivers are financially benefitted from the surge pricing of ride-sourcing platforms. But the burning question is – was the core objective of ride-sourcing services to create full time drivers?

According to V Armant [4], In general, the aims of ridesharing are to improve mobility by providing rides to users without cars, to reduce travel costs, to ad- dress environmental concerns by reducing the total driven distance on roads, and to alleviate congestion by reducing vehicle numbers.

From this study it is observed that, in Dhaka city the Pathao bike drivers normally earn highest within the time period of 5 to 9 PM of a day. Keeping this in mind and based on our research findings and user characteristics - we want to recommend the government to apply a policy in which an extra tax will be applied on surge pricing between 5 to 9 Pm period of the day on all the ride-sourcing services.

This will make common people use public transport more and emergence of metro rail in near future will add extra value to this policy.

7. Limitations and Future Scope of Study

Due to lack of resources and Covid-19 situation we had to select a limited range of study area and number of drivers interviewed in the process gave us a small data set. We also only interviewed the bike drivers and ignored other ride-sourcing transportation modes like car, CNG, etc. In case of the online survey, demographic division was ignored and no text mining was done prior to questionnaire set up.

In future, we want researchers to work with a larger study area and data set to get a clear idea about justification of our findings and policy recommended as well.

8. Conclusion

At the end we just want to summarize the findings of our analytical study:

1. Drivers earn most during 5 to 9 PM in Dhaka city while sharing rides in their bikes.
2. Fuel cost is higher and waiting time lower for drivers whose trip time is higher and always active during the peak period of 5 to 9 PM.
3. Only the full-time drivers are financially benefitted from surge pricing.
4. If 20% pricing increases, people will tend to avoid TNCs and prefer using public transport more.

9. References

[1] R. S. Thangaraj, K. Mukherjee, G. Raravi, A. Metrewar, N. Annamaneni, and K. Chattopadhyay, "Xhare-A-Ride: A search optimized dynamic ride sharing system with approximation guarantee," *Proc. - Int. Conf. Data Eng.*, pp. 1117–1128, 2017, doi: 10.1109/ICDE.2017.156.

[2] S. Islam, E. Huda, and F. Nasrin, "Ride-sharing Service in Bangladesh: Contemporary States and Prospects," *Int. J. Bus. Manag.*, vol. 14, no. 9, p. 65, 2019, doi: 10.5539/ijbm.v14n9p65.

[3] UBER, "Surge Pricing." 2019, [Online]. Available: https://www.uber.com/en-GB/drive/partner-app/how-surge-works/.

[4] L. Zha, Y. Yin, and Y. Du, "Surge Pricing and Labor Supply in the Ride-Sourcing Market," *Transp. Res. Procedia*, vol. 23, pp. 2–21, 2017, doi: 10.1016/j.trpro.2017.05.002.

[5] J. C. Castillo, "Who Benefits from Surge Pricing?," *SSRN Electron. J.*, 2018, doi: 10.2139/ssrn.3245533.

[6] M. Battifarano and Z. Sean, "Predicting real-time surge pricing of ride-sourcing companies," pp. 1–27, 2019.

[7] A. Lu, P. I. Frazier, and O. Kislev, "Surge Pricing Moves Uber's Driver-Partners," vol. i, pp. 3–3, 2018, doi: 10.1145/3219166.3219192.

[8] J. Jiao, "Investigating Uber price surges during a special event in Austin, TX," *Res. Transp. Bus. Manag.*, vol. 29, no. September 2017, pp. 101–107, 2018, doi: 10.1016/j.rtbm.2018.02.008.

[9] M. Goncharova, "Ride-Hailing Drivers Are Slaves to the Surge," *The New York Times*. 2017, [Online]. Available: https://www.nytimes.com/2017/01/12/nyregion/uber-lyft-juno-ride-hailing.html.

[10] "Pathao cuts commission on bike trips _ The Daily Star." .

[11] “Pathao Reduces Commission for Bike Rides During Peak Hours - Business Inspection BD.” .

[12] L. Sun, R. H. Teunter, M. Z. Babai, and G. Hua, “Optimal pricing for ride-sourcing platforms,” *Eur. J. Oper. Res.*, vol. 278, no. 3, pp. 783–795, 2019, doi: 10.1016/j.ejor.2019.04.044.

[13] L. Chen, A. Mislove, and C. Wilson, “Peeking beneath the hood of uber,” *Proc. ACM SIGCOMM Internet Meas. Conf. IMC*, vol. 2015-Octob, pp. 495–508, 2015, doi: 10.1145/2815675.2815681.

[14] “Soaring gas prices are forcing some Uber, Lyft drivers off the road.” .

[15] U. M. Dholakia, “Everyone Hates Uber’s Surge Pricing – Here’s How to Fix It,” *Harvard Business Review*. pp. 1–6, 2015, [Online]. Available: https://hbr.org/2015/12/everyone-hates-ubers-surge-pricing-heres-how-to-fix-it.

[16] M. Akbari, A. Moradi, N. SeyyedAmiri, M. Á. Zúñiga, Z. Rahmani, and H. Padash, “Consumers’ intentions to use ridesharing services in Iran,” *Res. Transp. Bus. Manag.*, vol. 41, no. February, 2021, doi: 10.1016/j.rtbm.2020.100616.

[17] K. (Brad) Lee, M. A. Bellamy, N. Joglekar, C. Wilson, and S. Jiang, “Surge Pricing on A Service Platform under Spatial Spillovers: Evidence from Uber,” *Acad. Manag. Proc.*, vol. 2019, no. 1, p. 16279, 2019, doi: 10.5465/ambpp.2019.16279abstract.

[18] M. Rey, L. M. Aiello, M. Rey, and K. Lerman, “2017 Analyzing Uber ’ s Ride-sharing Economy,” *Www*, no. January, pp. 574–582, 2017.

[19] Z. Li, C. Liang, Y. Hong, and Z. Zhang, “How Do On-demand Ridesharing Services Affect Traffic Congestion? The Moderating Role of Urban Compactness,” *Prod. Oper. Manag.*, vol. 31, no. 1, pp. 239–258, 2022, doi: 10.1111/poms.13530.

[20] J. Nikkarila, I. Kangasniemi, and J. Valtonen, “Bayesian networks- an example of software and some defence applications,” *Conf. (R)evolution War*, no. October 2015, pp. 1–18, 2016.

[21] “Pooyan Ehsani (0000-0002-0155-3593).” .

[22] S. Shokoohyar, A. Sobhani, R. Malhotra, and W. Liang, “Travel Time Prediction in Ride-Sourcing Networks: A Case Study for Machine Learning Applications,” *J. Bus. Cases Appl.*, vol. 26, p. 1, [Online]. Available: http://www.aabri.com/copyright.html.