Taxis and Uber Ride-Sharing in New York City

Topic Question

How has ride-sharing in New York City changed over time -- particularly, what services are growing and which aren't, and what are the demographics of the customers of the various services?

NON-TECHNICAL EXECUTIVE SUMMARY

To answer the question of whether Uber will overtake all other cab services, a comparative study between various cab types and various demographics was conducted. We found that yellow taxis are declining while Uber is growing the most out of all the ride-sharing services.

Another important question is the demographics of the various services. We found that green cabs seemed to be more popular among lower income customers, while Uber and yellow cabs were more popular among wealthier customers. However, Uber seemed to have broad popularity across different income levels. We did not find any major differences in the typical ages of the different services.

Finally, we found that despite the growing competition, green cab services have not been forced to reduce prices to remain competitive. In fact, they have consistently increased prices.

TECHNICAL EXECUTIVE SUMMARY

Introduction

Since 1897, taxis have been present in New York City. They are a staple in the city, and many people have taken them over the years. Recently however, Uber and other similar ride-sharing apps have begun to encroach on the taxi cab market. In fact, Uber has overtaken green taxis around the beginning of 2015, and (insert info about yellow taxis here). How will the market share for Uber and taxis change in the coming years?

To start, our team decided to approach this problem through two perspectives: one focusing on geographic and demographic trends, and another on the competition between Ubers, yellow cabs, and green cabs over time.

Methods

Throughout the project, our team decided to trim some of the data into large random samples. This cut our computation time significantly, allowing us to focus on making significant analysis without loss of generality.

Through the use of the R package “Spatial Polygons”, we were able to render a visualization of all the boroughs and neighborhoods listed in the geographics dataset. Next, using the pickup or dropoff latitudes and longitudes for each type of taxi service, we plotted each trip on that map. This is shown in the array of four maps below (Figure A).

From there, we were struck with the idea to find a summary of the types of riders who take the service. Although the dataset did not include rider-specific demographic information, we were provided with information about the types of people who lived in each “NTA”, or neighborhood. We decided that to identify which NTA riders were from, we would identify a subsection of them as commuters who were making their way to work. Therefore, we filtered all the ride information according to pickup location and time such that only weekday riders picked up between the hours of 5 and 7 AM were considered. We then looked at their NTAs and created a frequency table to see the distribution of NTAs. The demographic data we prioritized was income and age. The results of these are shown below (Figures B and C).

To assess the competition between different cabs and the effect of Ubers on taxis, a time series plot of number of cab rides aggregated monthly for different types of ride-shares were plotted. Using ARIMA modelling and exponential smoothing, the growth of all 4 types of cab rides were forecasted.

Results

Figure A is a rendering of the maps with the trips taken by riders in the city plotted as per ride service, showing pickup locations. Notable trends here are that the green cab service and yellow cab service service different areas almost completely, with overlap between the two being rare. Uber appears to have encroached mostly

on the yellow cab service areas, but it does seem to also be creeping in on the green cab areas as well. This dynamic has interesting implications regarding how the companies are competing for the average New York City rider.

Figure B shows the distribution of mean income across all the ride services. This was generated by weighting the mean incomes of each NTA according to the frequency that riders from that NTA used that service. For comparison, we also included the current mean income distribution for all of New York City. From this, we find that in general New Yorkers make roughly 50k - 75k per year. Uber customers overtake this statistic by a large margin, holding a median income of 180k. Yellow cab customers have a similar median, but their distribution was much more narrow and concentrated on the central value. Green cab customers are on the other end of the spectrum, and are relatively poorer: their distribution closely mirrors that of the whole city's.

Figure C is similar to figure B, but considers age. These distributions are all roughly similar, and the nuances in age distribution that may be relevant are not able to be gleaned due to the nature of the data given. The only noteworthy takeaways are that green cab customers are usually younger, yellow cab customers are older, and Uber has an evenly split distribution.

Figure D shows the change over time in the various services. Orange represents yellow taxis, green represents green taxis, and red represents Uber. As we can see, Uber has the most growth with yellow taxis being the only service that is declining.

Figure E shows that for a distance of 3 miles, the price of a green cab ride increased in an almost perfectly linear fashion from $14 to $15.50.

Discussion

We found that Uber overtook green taxis around the beginning of 2015. Yellow taxis are the most commonly used, but they are decreasing in use.

Regarding the insights gained by looking at geographic and demographic information, we find that Uber is quickly eating away at the yellow cab and green cab businesses, and is poised to continue making an impact in the space. We expected this, and went looking for further indications that show the yellow cab or green cab service faltering. For instance, we considered trip distance and price for green cabs, expecting that perhaps these would be negatively impacted over time

because of Uber. However, that was not the case: both of these stayed consistent year after year. This may be insightful in itself, indicating that a specific group of customers use green cab for the same kind of trips on a regular basis; they could be considered long-time users who are loyal. However, the old cab companies are certainly feeling the pressure of the new ridesharing services like Uber and Lyft. Price hikes with the green cab service shows that instead of trying to get into a price war with new companies, they are in the phase of trying to recuperate their losses as soon as possible. This seems like a losing strategy in the long run, and will negatively impact their user base in the coming years.

Figures and Graphs

Figure A

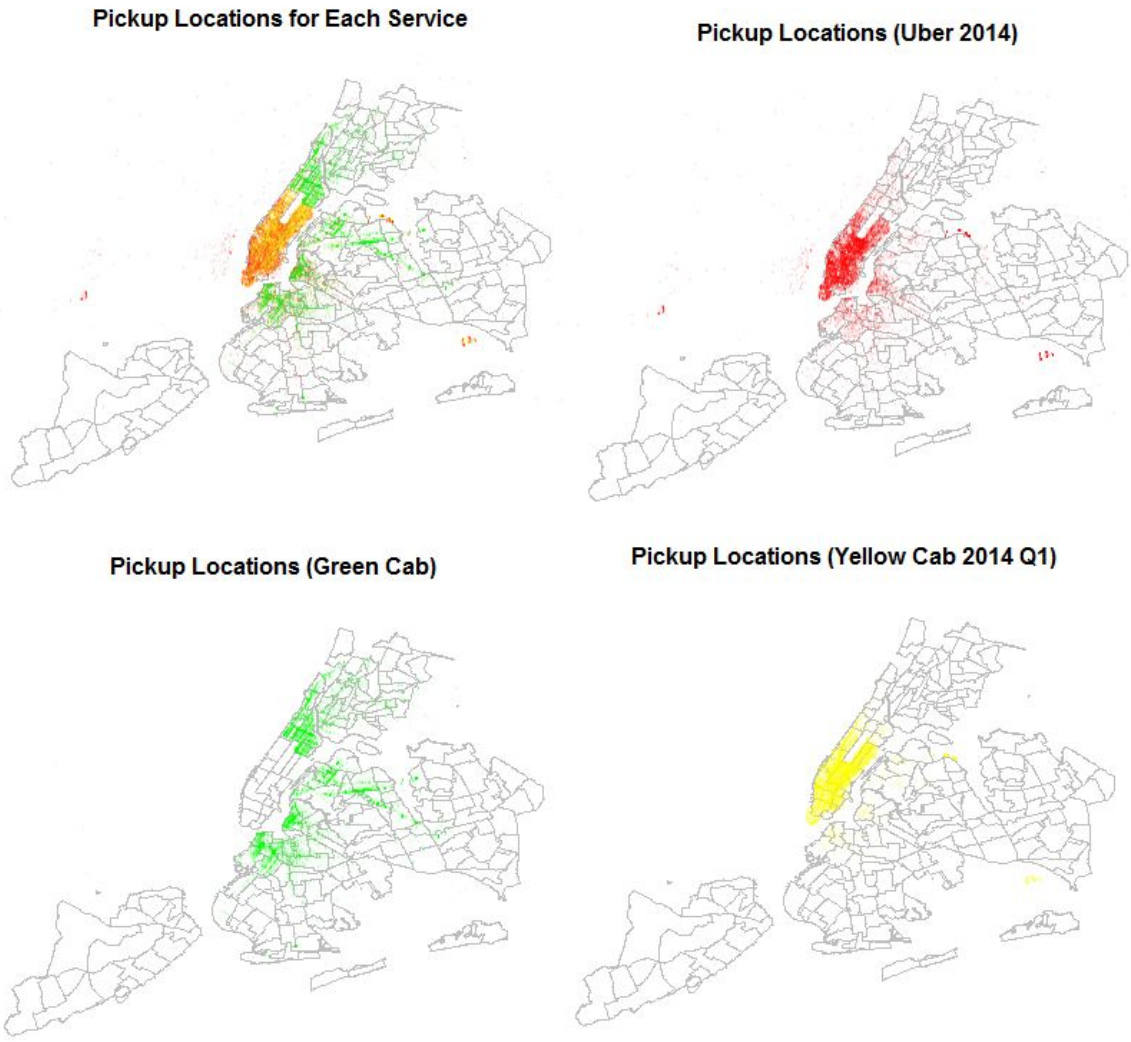


Figure B

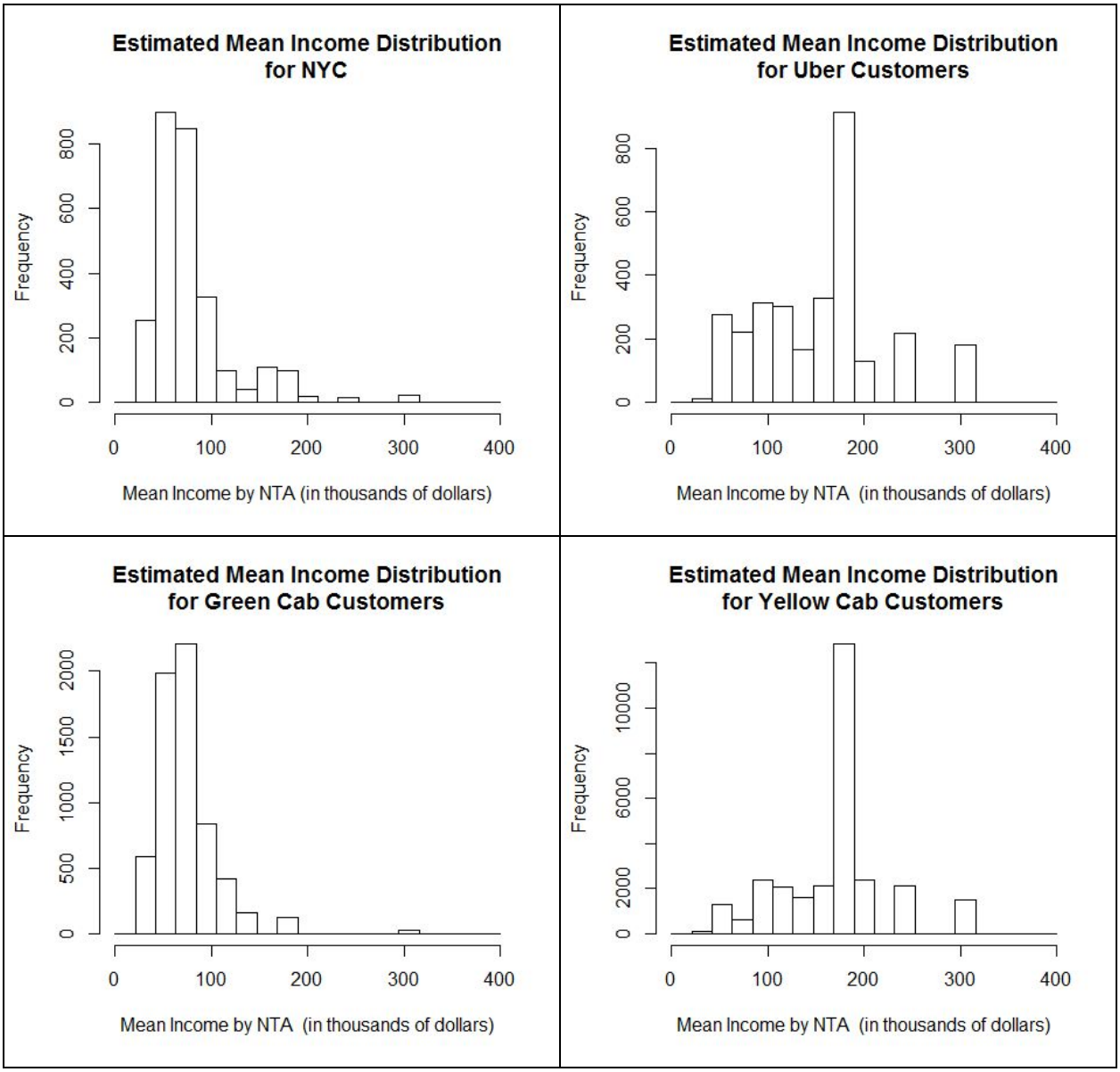

Estimated Mean Income Distribution for NYC
Frequency
0
200
400
600
800
0
100
200
300
400
Mean Income by NTA (in thousands of dollars)
Estimated Mean Income Distribution for Uber Customers
Frequency
0
200
400
600
800
0
100
200
300
400
Mean Income by NTA (in thousands of dollars)
Estimated Mean Income Distribution for Green Cab Customers
Frequency
0
500
1000
1500
2000
0
100
200
300
400
Mean Income by NTA (in thousands of dollars)
Estimated Mean Income Distribution for Yellow Cab Customers
Frequency
0
2000
6000
10000
0
100
200
300
400
Mean Income by NTA (in thousands of dollars)

Figure C

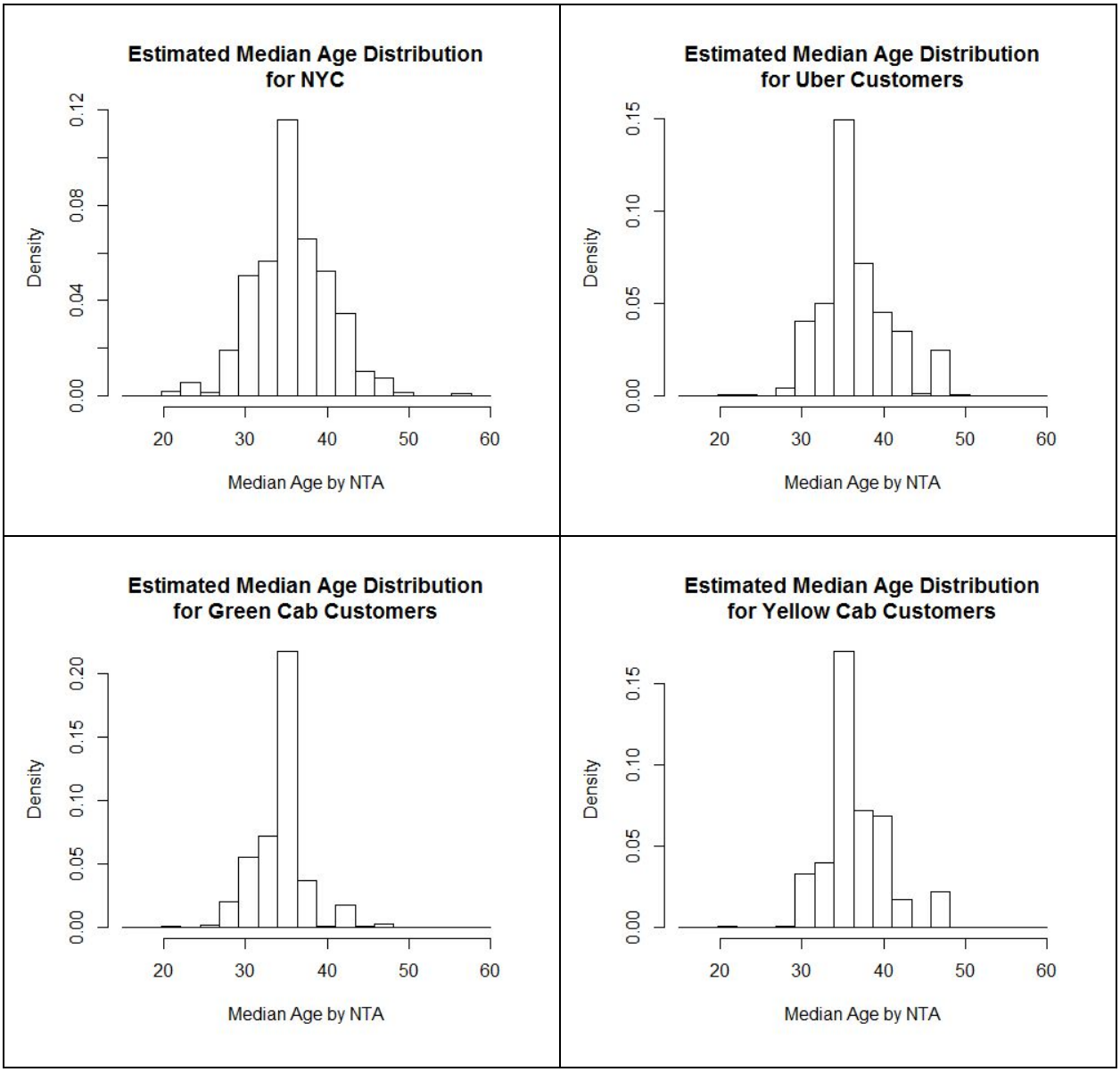

Estimated Median Age Distribution for NYC
Density
0.00
0.04
0.08
0.12
20
30
40
50
60
Median Age by NTA
Estimated Median Age Distribution for Uber Customers
Density
0.00
0.05
0.10
0.15
20
30
40
50
60
Median Age by NTA
Estimated Median Age Distribution for Green Cab Customers
Density
0.00
0.05
0.10
0.15
0.20
20
30
40
50
60
Median Age by NTA
Estimated Median Age Distribution for Yellow Cab Customers
Density
0.00
0.05
0.10
0.15
20
30
40
50
60
Median Age by NTA

Figure D

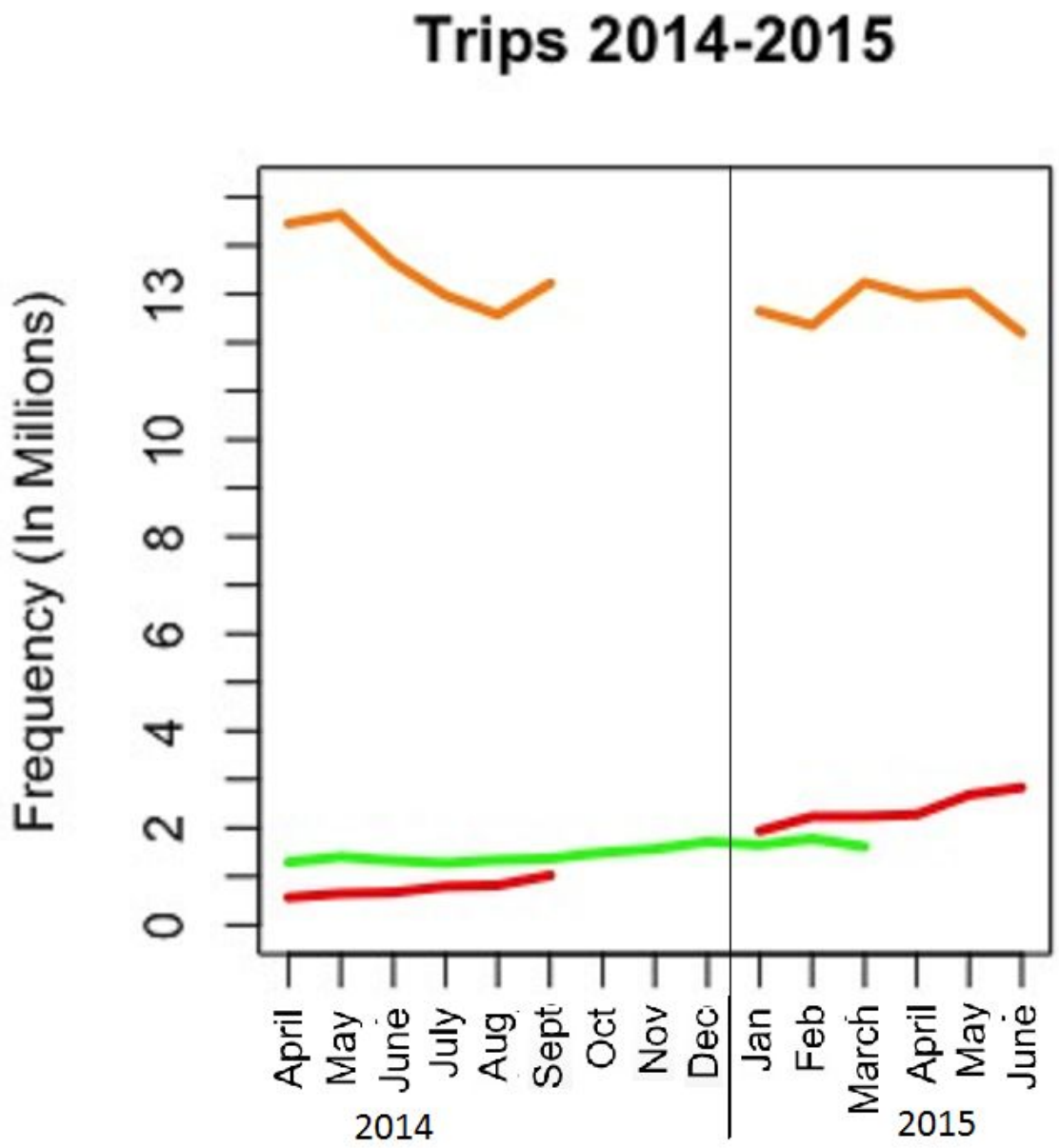

Trips 2014-2015
Frequency (In Millions)
0
2
4
6
8
10
13
April
May
June
July
Aug
Sept
Oct
Nov
Dec
Jan
Feb
March
April
May
June
2014
2015

Figure E

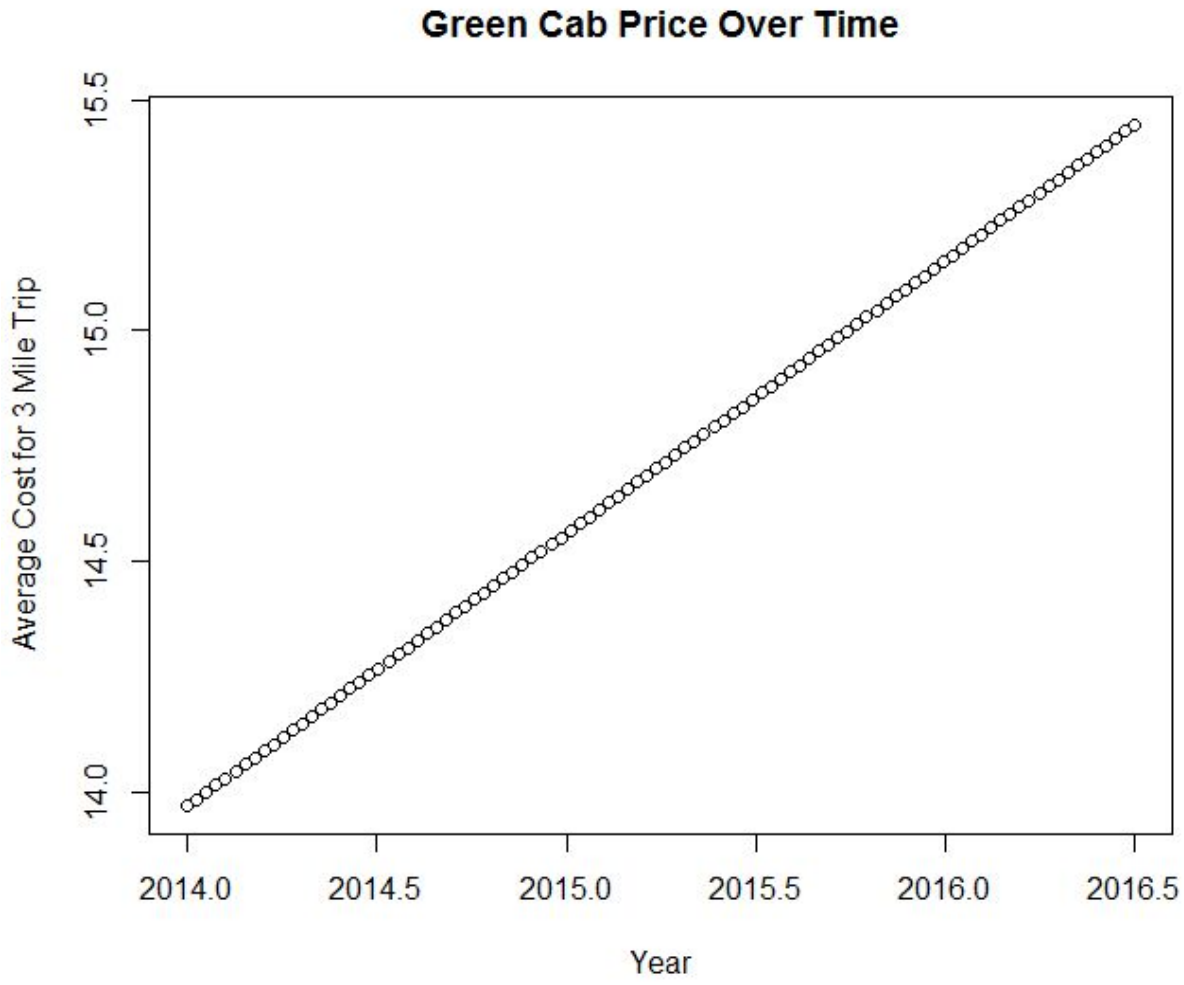
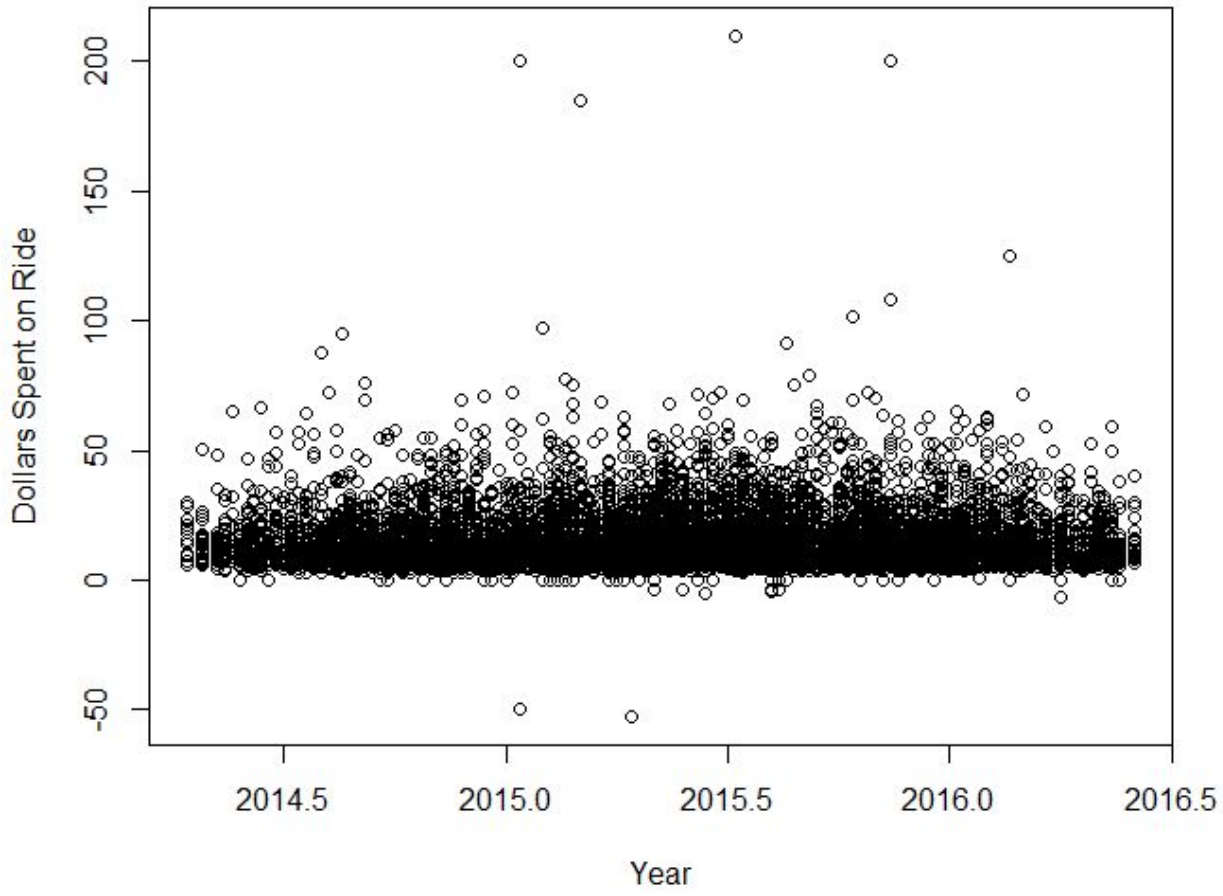

Green Cab Price Over Time
Average Cost for 3 Mile Trip
14.0
14.5
15.0
15.5
2014.0
2014.5
2015.0
2015.5
2016.0
2016.5
Year

Figure F


Dollars Spent on Ride
200
150
100
50
0
-50
2014.5
2015.0
2015.5
2016.0
2016.5
Year