

Group Member: Yifan Leng, Yunwen Cai, Gurneet Kaur, Sarthak Rajvanshi

Topic: What features affect Uber daily rides?

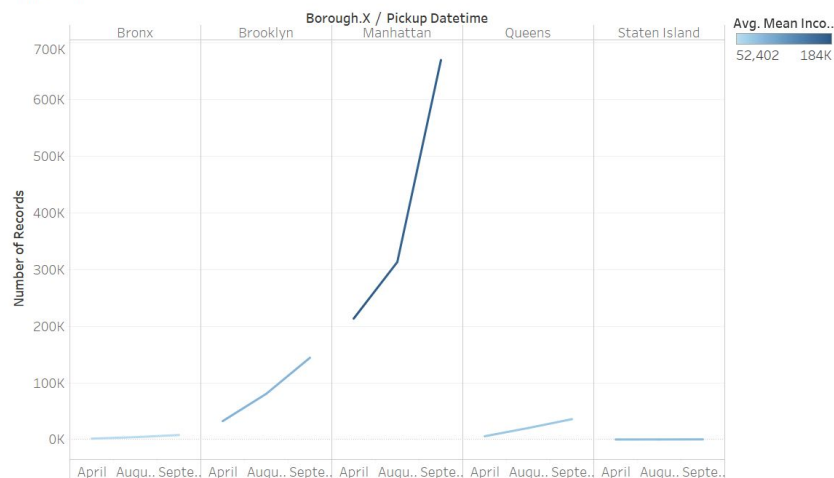
### Intriduction:

As a new means of transportation in the industry, Uber has a large influence on how people in New York live. Managers at Uber would be interested to understand the potential of the booming market in New York. We will explore features that affect the number of Uber rides a day in 2014 and 2015, so that managers can make more efficient marketing strategies in New York. The datasets we chose to use are uber\_trips\_2014, uber\_trips\_2015 and geographic. We mapped the latitude, longitude or location code of each pick up to the NTAs in New York.

### EDA:

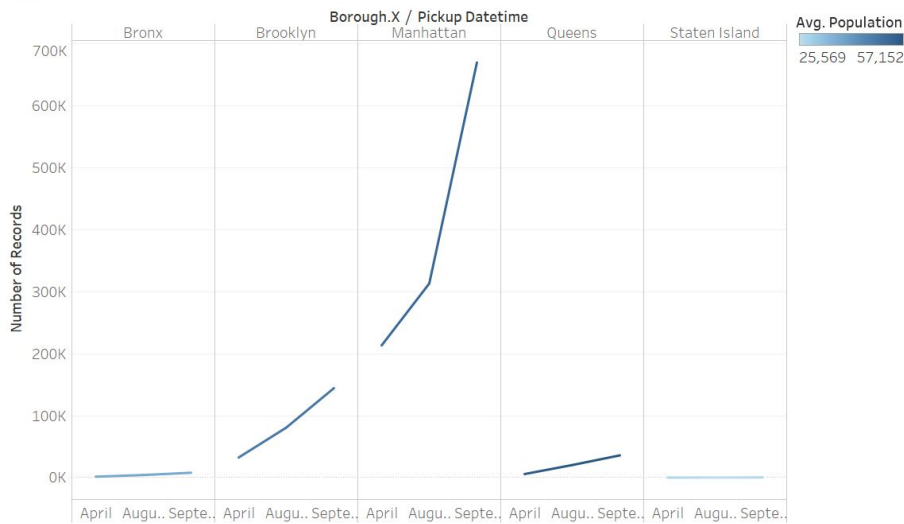
Uber ride trends for year 2014:

Sheet 1



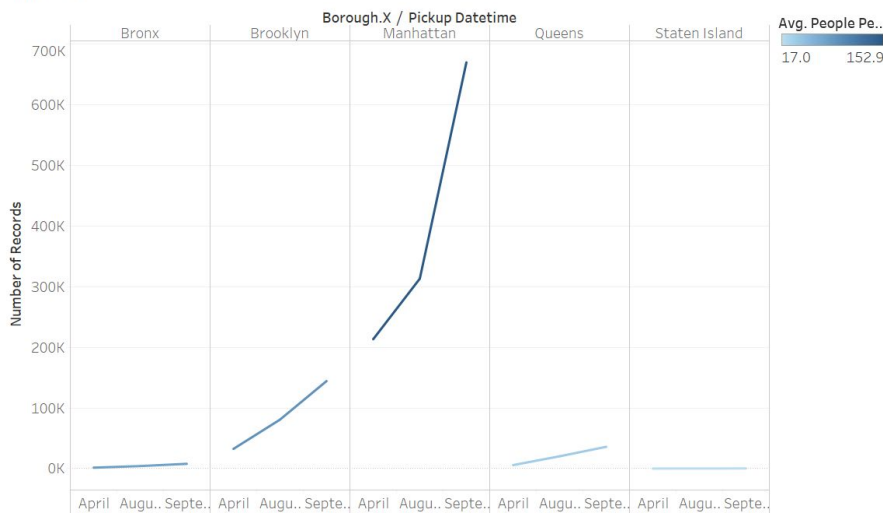
The above graph shows the trends of number of rides with respect to mean income of each borough. We can see that as the boroughs with higher mean income have more number of riders.

Sheet 4



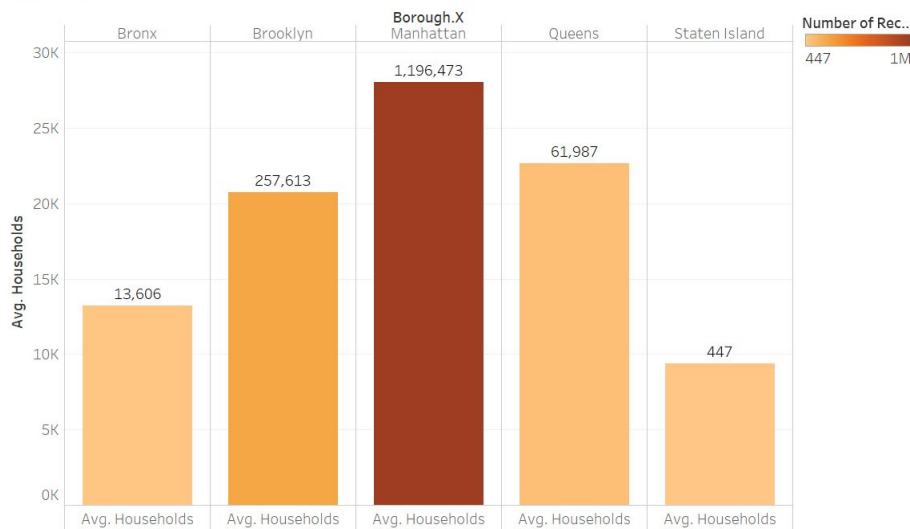
The above graph shows the trends of number of rides with respect to population of each borough. We can see that as the boroughs with higher the population have more number of riders.

Sheet 5



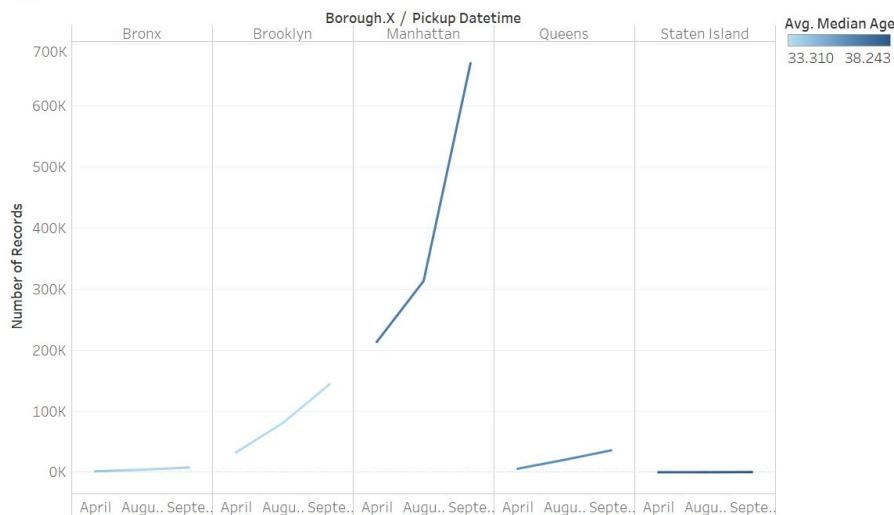
The above graph shows the trends of number of rides with respect to people per acre of each borough. We can see that as the boroughs with higher number of people per acre have more number of riders.

Sheet 6



The above graph shows the trends of number of rides with respect to number of households of each borough. We can see that as the boroughs with higher number of households have more number of riders.

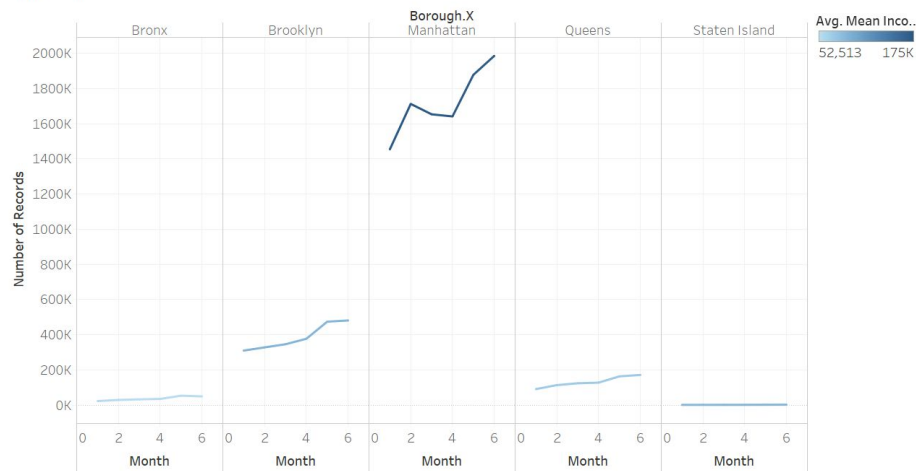
Sheet 7



The above graph shows the trends of number of rides with respect to median age of each borough. We can see that as the boroughs with people of higher median age have more number of riders.

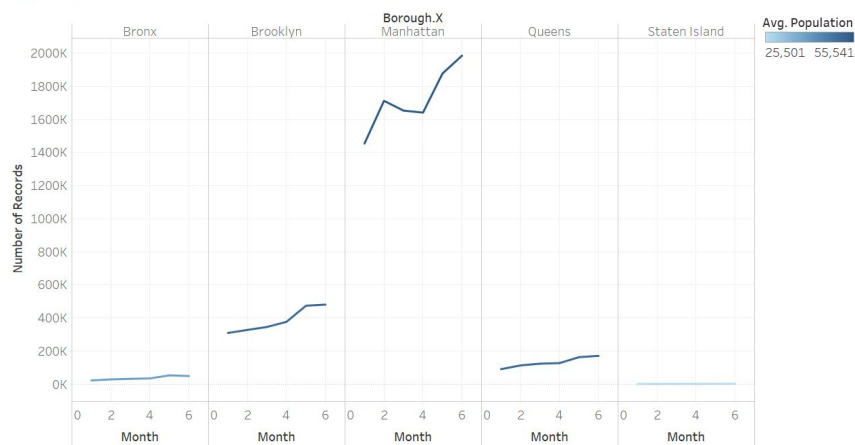
Uber ride trends for year 2015:

Sheet 1



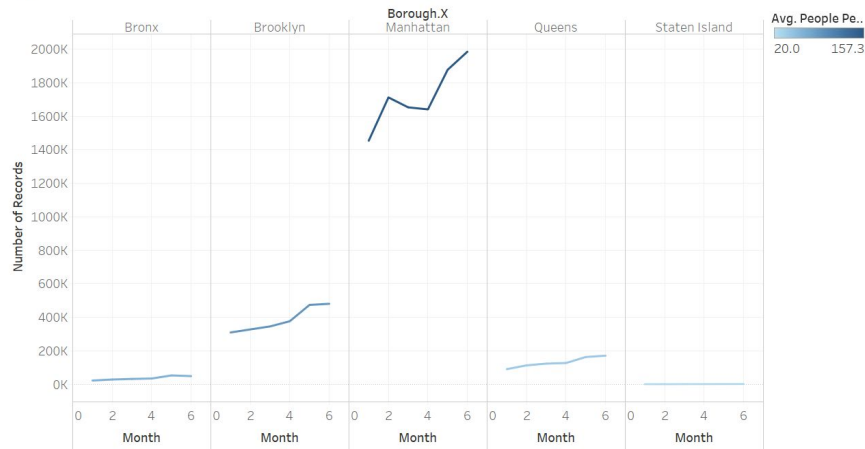
The above graph shows the trends of number of rides with respect to mean income of each borough. We can see that as the boroughs with higher mean income have more number of riders.

Sheet 2



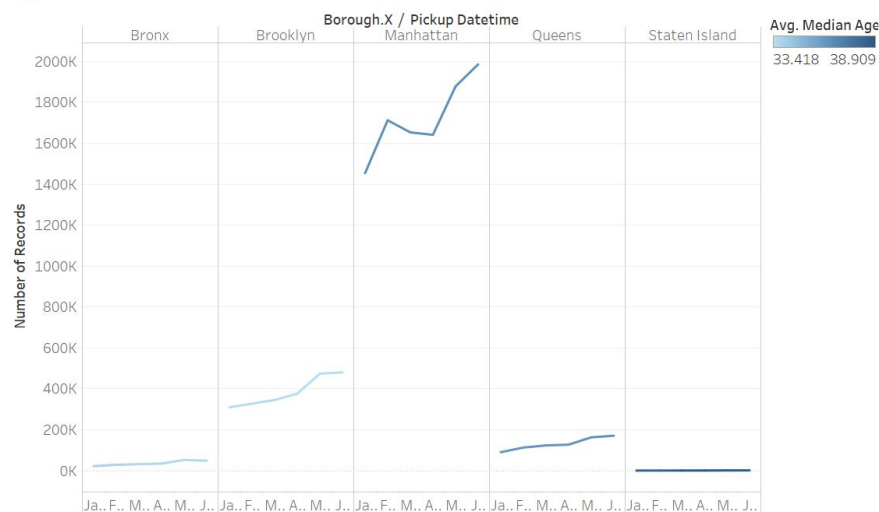
The above graph shows the trends of number of rides with respect to population of each borough. We can see that as the boroughs with higher the population have more number of riders.

Sheet 3



The above graph shows the trends of number of rides with respect to people per acre of each borough. We can see that as the boroughs with higher number of people per acre have more number of riders.

Sheet 4



The above graph shows the trends of number of rides with respect to median age of each borough. We can see that as the borough with people of highest median age, i.e. Manhattan, has largest number of riders. The other boroughs are not showing specific trend for number of riders with change in median age.

Overall, it has been observed that Uber has performed much better in 2015 cumulatively than 2014. The number of riders have increased from 700,000 to 2,000,000 in Manhattan, which is the borough with most number of riders.

## Random Forest Regression

We applied random forest regression to predict the number of daily uber rides with demographic features including population, population density, mean income, median income, and the age distributions (i.e. the population in each age category from under 5 to over 65). We also include the geographic information of borough and the time information of month.

To rank these features by importance, we find the mean decrease in Gini index, which measures the difference between two levels of trees. If classification based on a feature results in a large difference of misclassification rate, then this feature will correspond to a greater decrease in Gini index. Thus we will consider this feature to be important. The following ranks the top 10 features by their importance.

	MeanDecreaseGini
base	53.13496
Month	51.22228
borough	40.73417
population	18.46647
people_per_acre	17.93740
mean_income	16.34232
median_income	16.19067
X20.24_years	14.95154
X15.19_years	14.90822
X25.29_years	14.05982

To visualize the effect on the

