**Group 5:**

Rebecca Kwon

Neil Tipton

Turgut Ozsirkinti

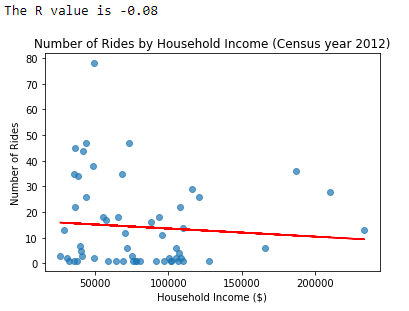
Daniel Garza

**Project 1 Write Up**

1. **Are income & number of rides correlated?**

No, based on the low R value, income and number of rides are not correlated.

The initial assumption was that people who make more money would take more uber rides. However, the data shows that this is not necessarily the case.

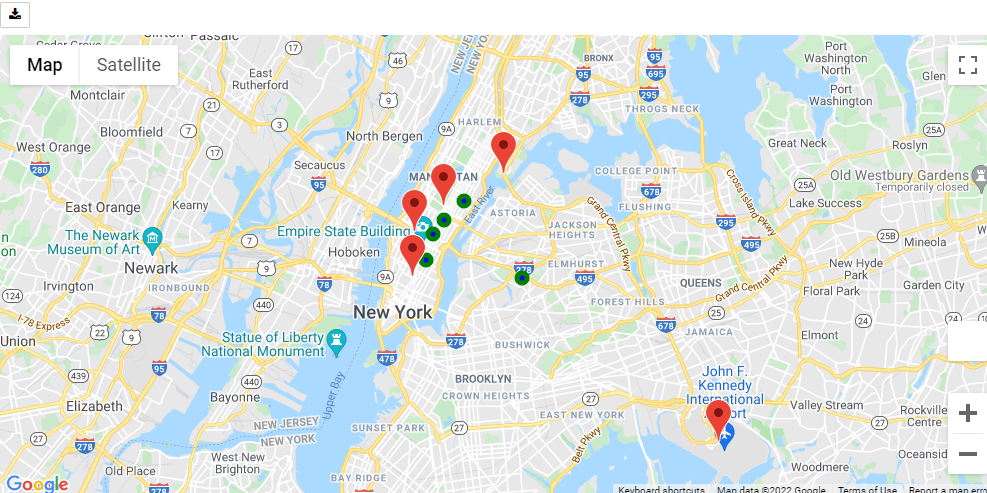


1. **What are the locations for the most popular drop off and pick up places**

The most popular locations are in New York, Guttenberg, Weekawken, Edgewater, and Woodmere.

Following the data analysis, the data set appears to be central to New York.

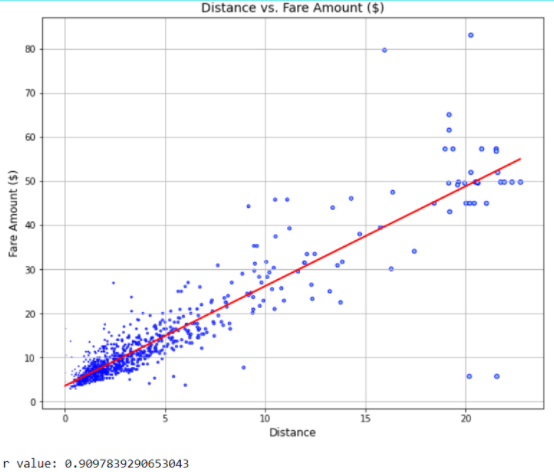
A notable pick up location is JFK international airport.



1. **Are distance and fare amount correlated?**

The high R value of 0.9 indicates high positive correlation– the longer the distance, the more it costs.

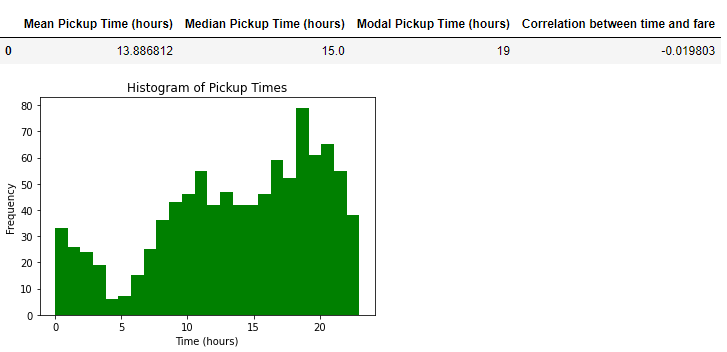
Clustering is also shown below the ~5 mile mark, indicating most trips happen for distances between 0 and ~5 miles.

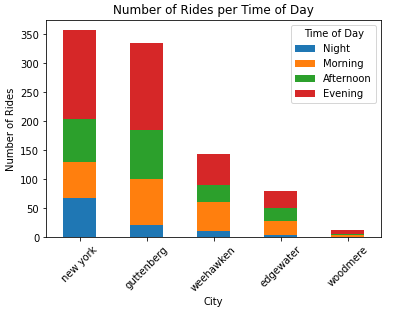


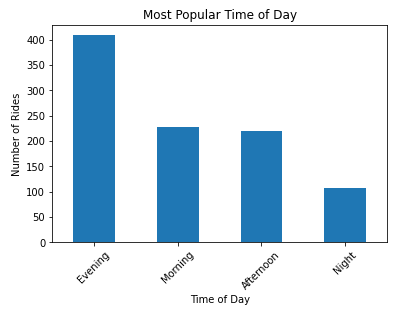
1. **What are the most popular times of travel?**

Modal pick up for this random sample (1000) is at the 19 hour mark, or 7 PM.

For the top 5 cities, and overall data set, Evening times (17-24) from 5 PM to midnight were the most popular.

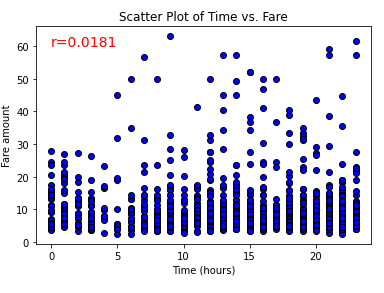
****

****

****

1. **Are Pick up times and Fare correlated?**

Initial assumptions were that fares would spike at popular times of the day. However, based on the low R value, there is no significant correlation between pick up times and fare.



1. **Does the number of passengers affect the fare amount?**

There is no significant correlation between the number of passengers and fare amount based on the low p value.

