**Chicago Taxi Data Analysis**

**Advait Kulkarni**

**Suhit Datta**

**Varad Tupe**

**Abstract**

The space of public transport has exploded in the recent years with the advent of many a ride-hailing services. A major concern for users has been about the pricing of rides dependent on timings, areas and surges. This project to perform exploratory analysis on Chicago taxi data.

**Introduction**

The data comprises of all the Chicago taxi records between April-July 2017 extracted from the Chicago Data Portal. It has a total of 113 million records containing features like trip duration, miles, area codes and pricing.

**Data**

The data has the following features:

Trip\_Start\_Timestamp - Boarding time

Trip\_End\_Timestamp - Departing time

Trip\_Seconds - Duration of the trip

Trip\_Miles - Distance of the trip

Pickup\_Community\_Area - Unique area code for each each area of boarding

Dropoff\_Community\_Area - Unique area code for each each area of departure

Fare - Cost of the ride

Company - Company providing the taxi services

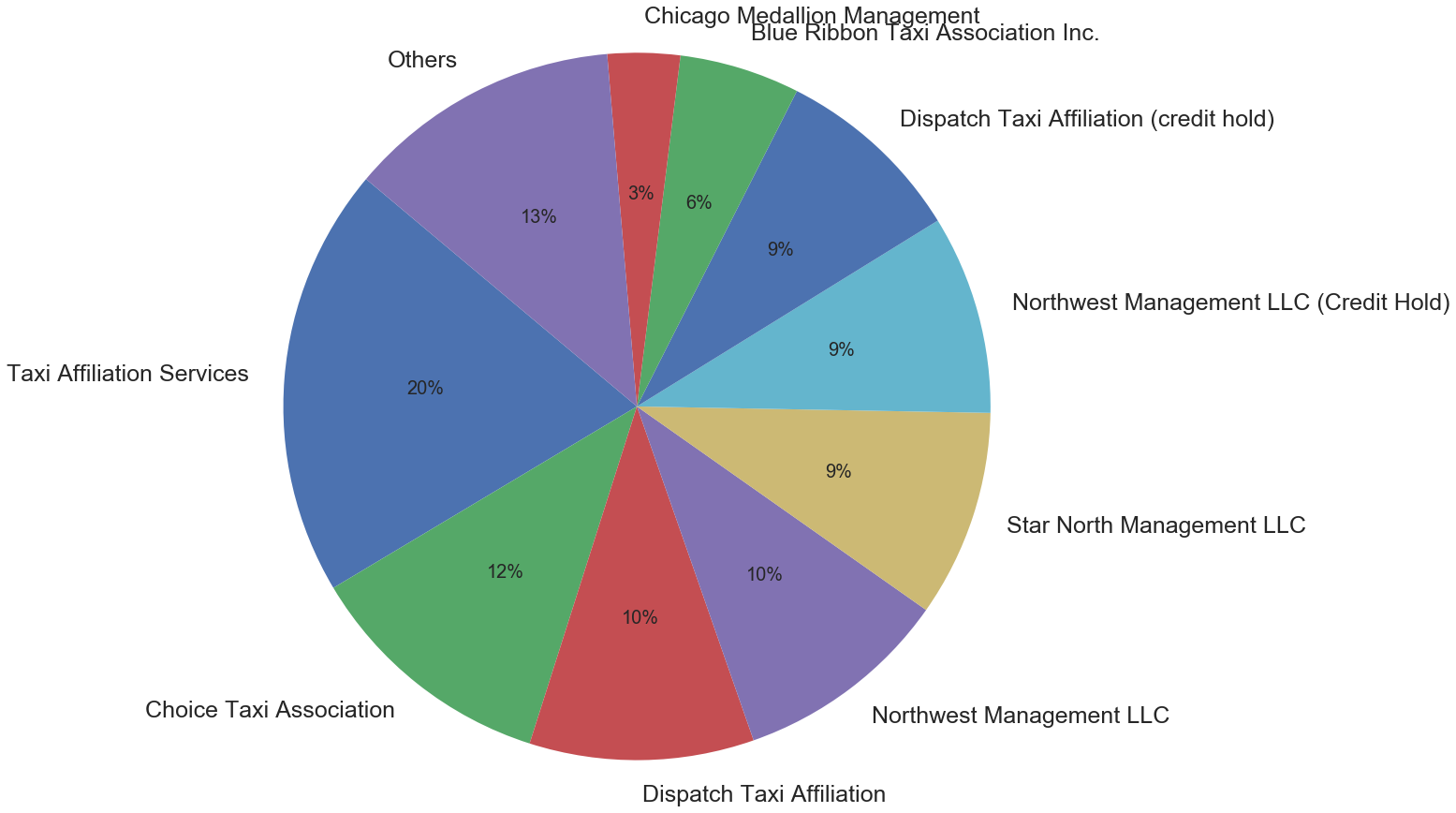
Pickup\_Centroid\_Latitude - Pickup Latitude

Pickup\_Centroid\_Longitude - Pickup Longitude

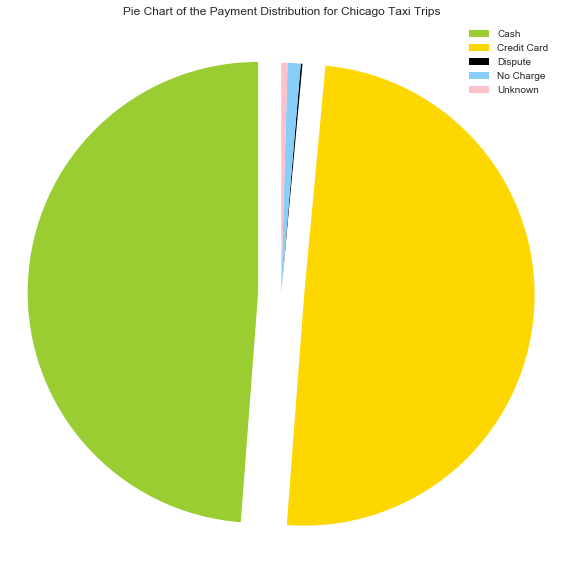
Dropoff\_Centroid\_Latitude - Drop-off Lattitude

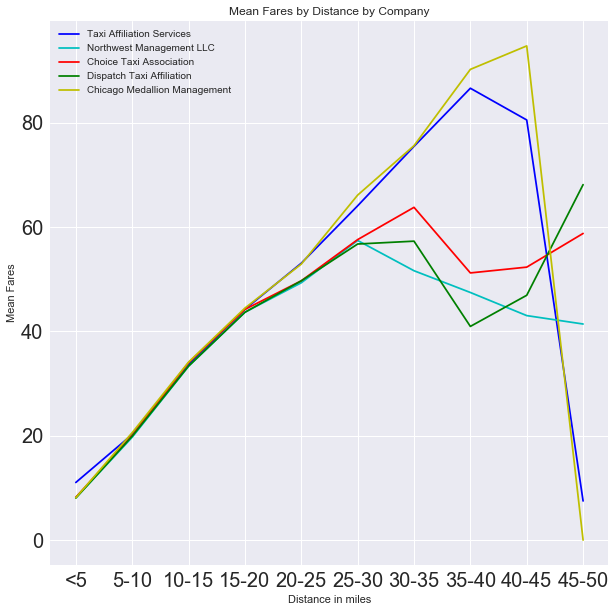
Dropoff\_Centroid\_Longitude - Drop-off Longitude

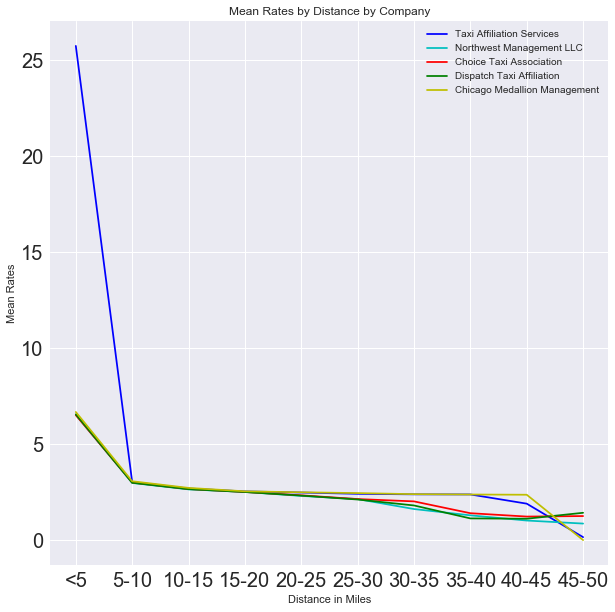
**Analysis**

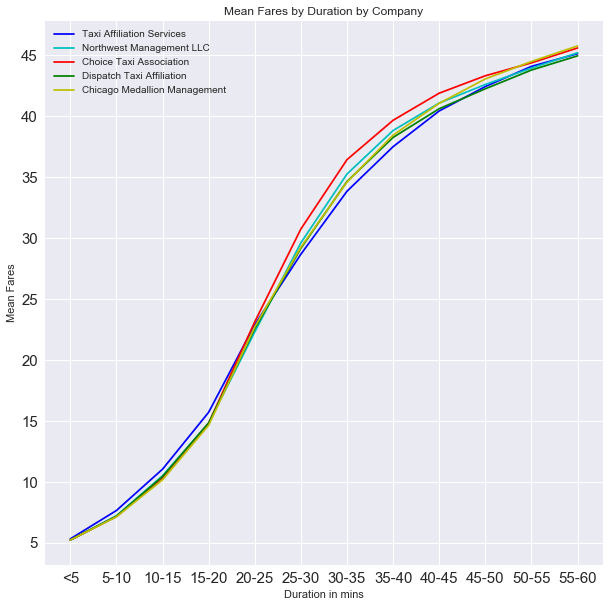
The following pie chart represents the share of each company in the Chicago taxi space.

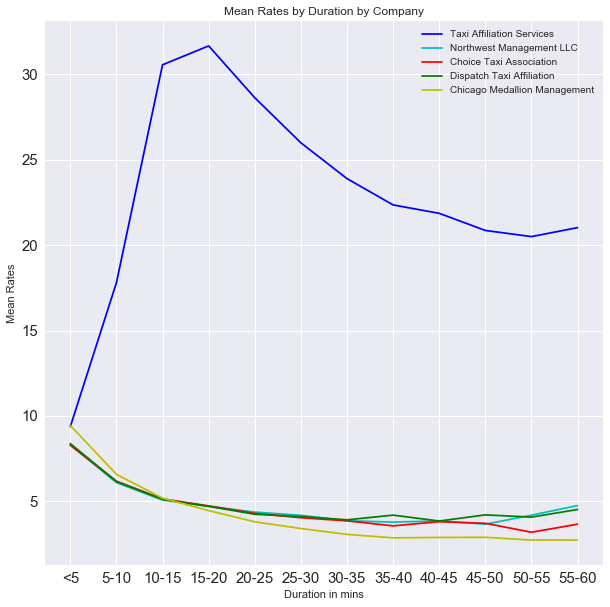
The distribution of payment types is as follows-

****

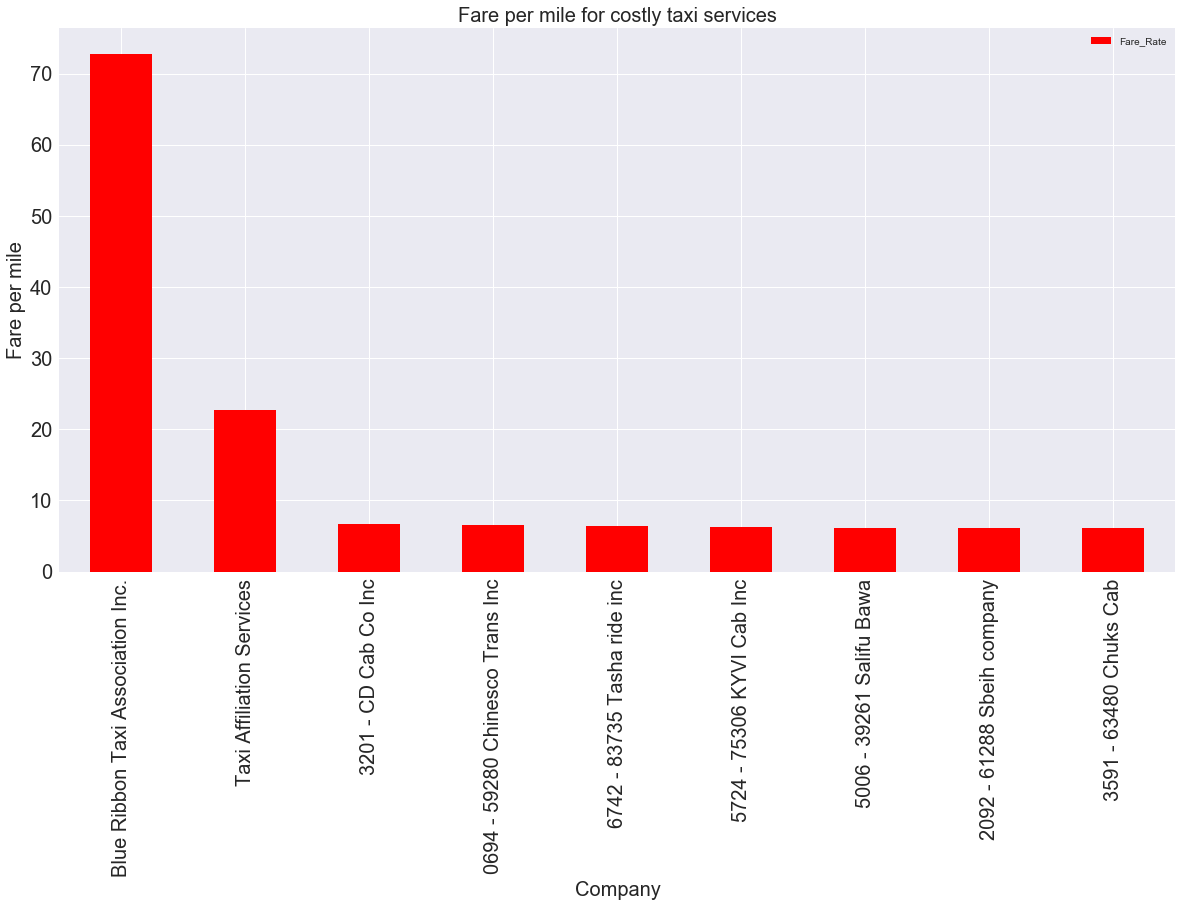
The following plots display the variation of costs and rates with distance.

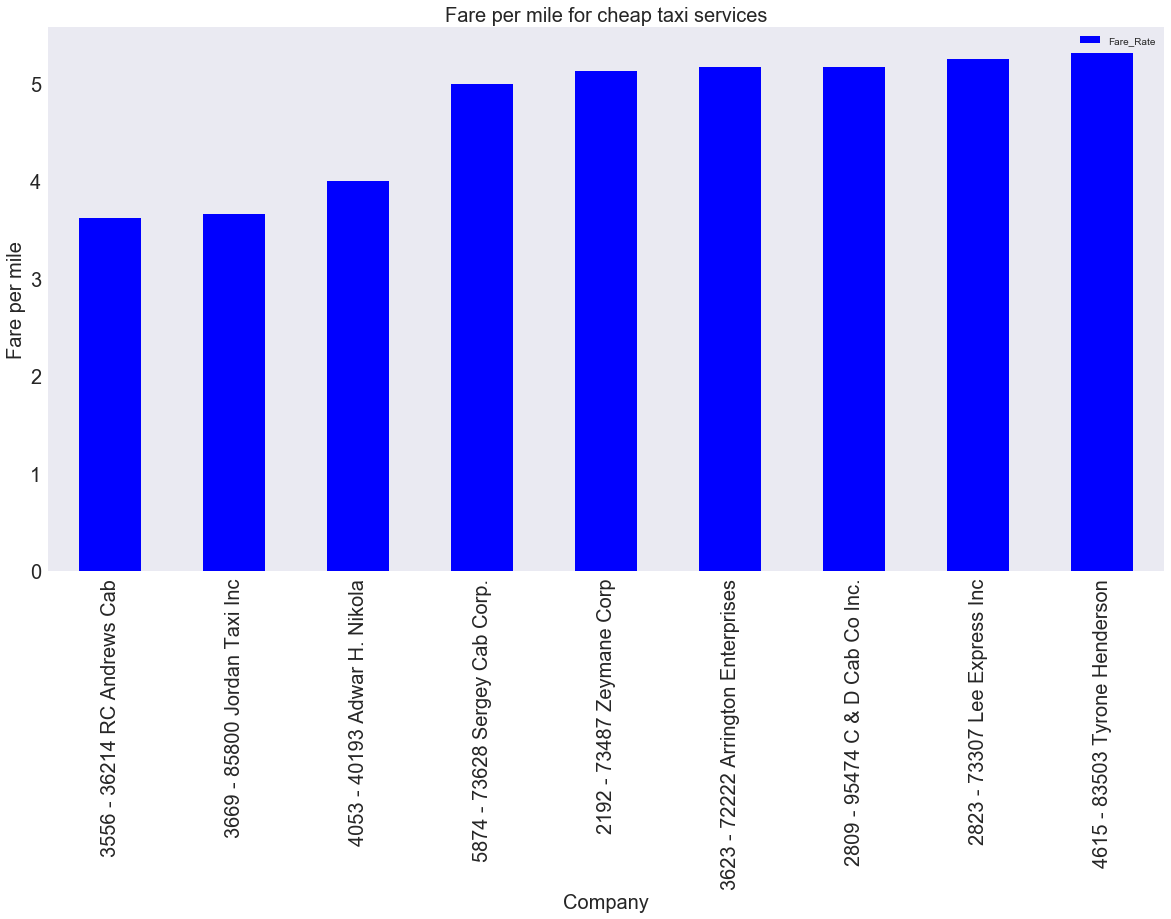
****

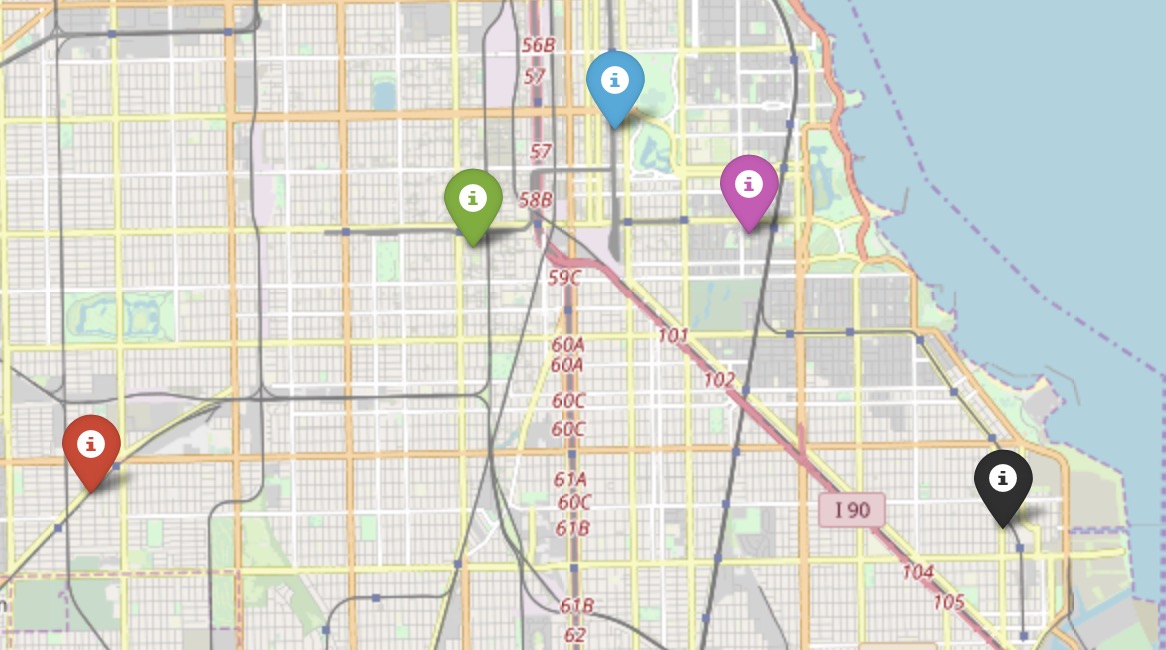
The following plots show the variation of fares and rates with duration.

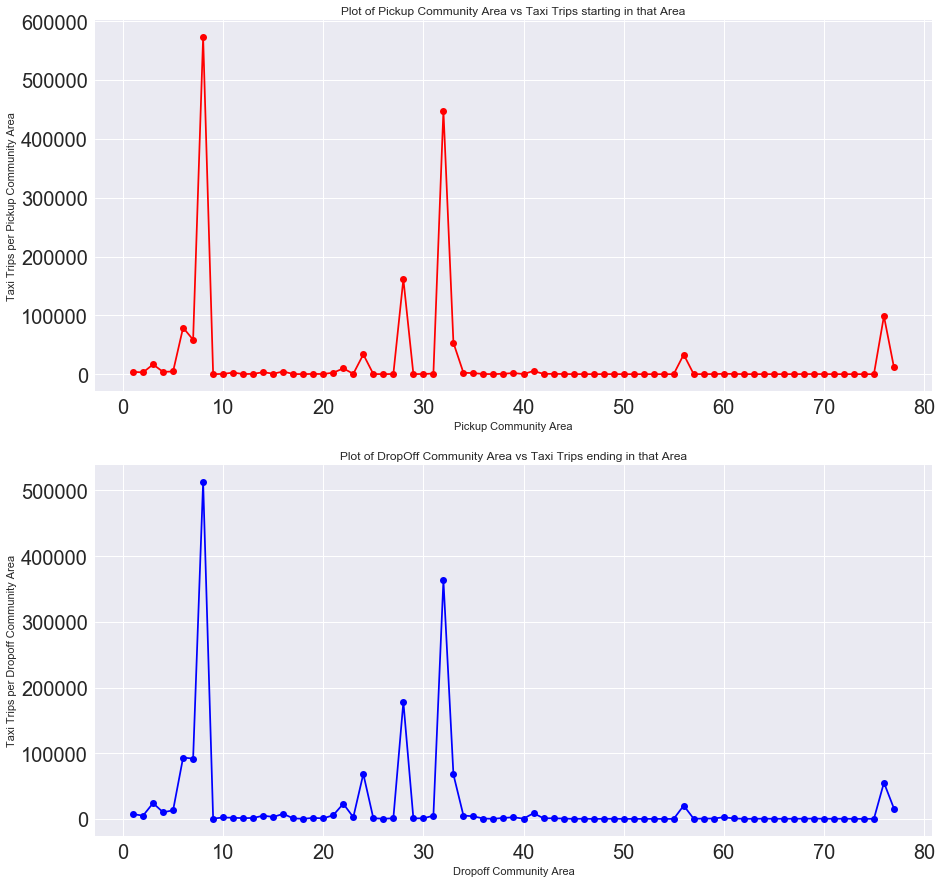
****

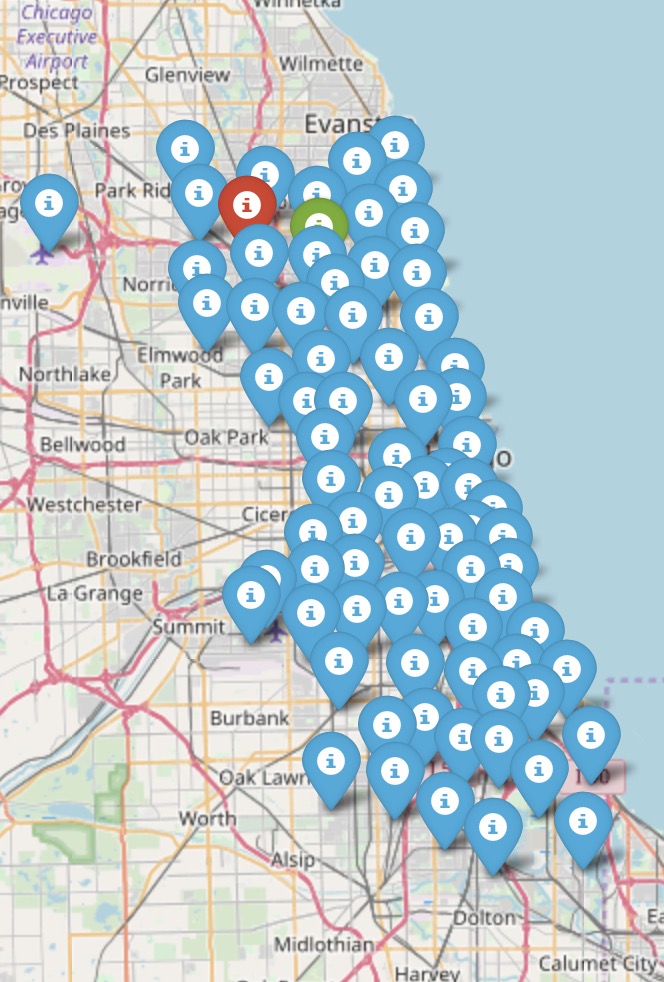
These plots show us the relative costs of the costliest and the cheapest taxi services.

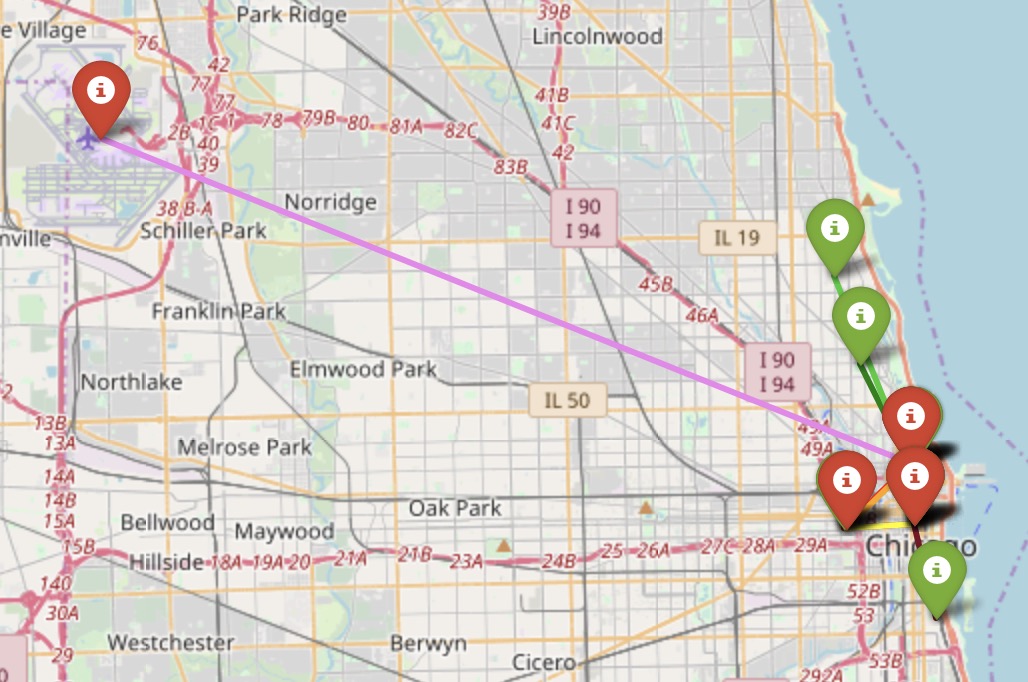


****

****

****

****

****

**Conclusion**

The project provided an elaborate understanding of the Chicago taxi business. It can be useful for newer players like Uber to move into bigger markets. Conversely, it can also be useful for traditional companies to maintain their businesses in their territories. It provides an idea of how competitive pricing can be used both at the user and the company’s end.