

Chicago Taxi Trips Data Analysis

Gulay Samatli

Introduction

Question: Demand and fare by time and location;
Chicago Taxi Cab Revenue Optimization

- Districts
- Month
- Days of the week
- Competitive market

Motivation:

- to help taxi cab companies to schedule their staff

DATA Acquisition

January-October, 2021

Soda API



CHICAGO
DATA PORTAL



Data Storage

SQL, SQLAlchemy



Data Process

Python



Web App

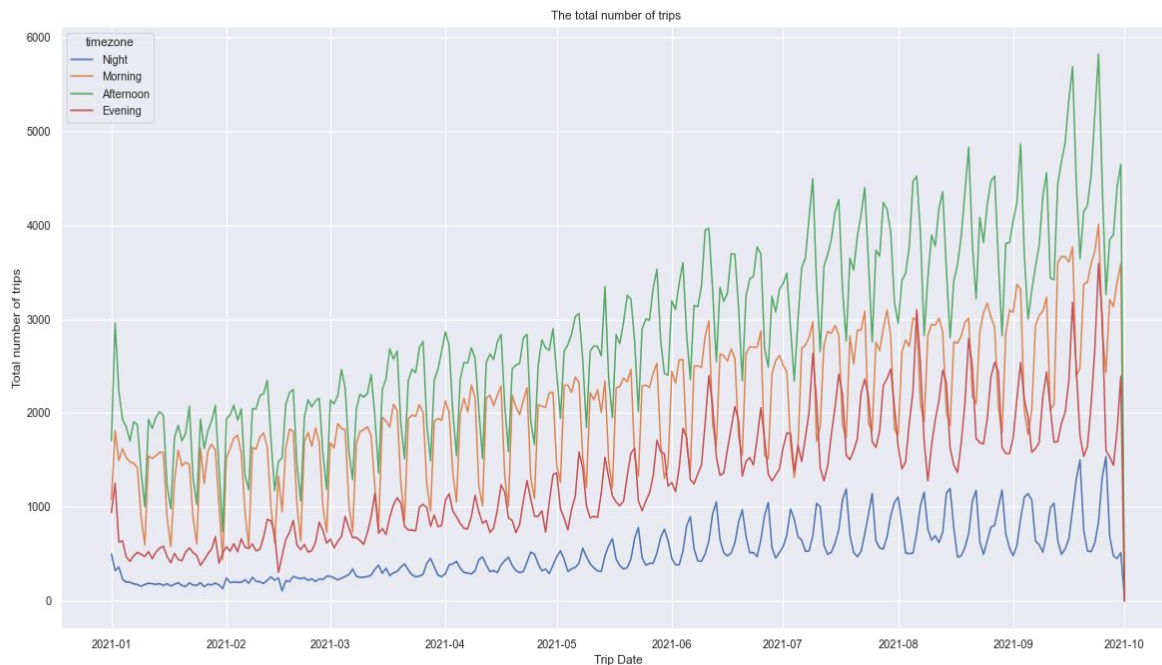
Streamlit



Data, EDA

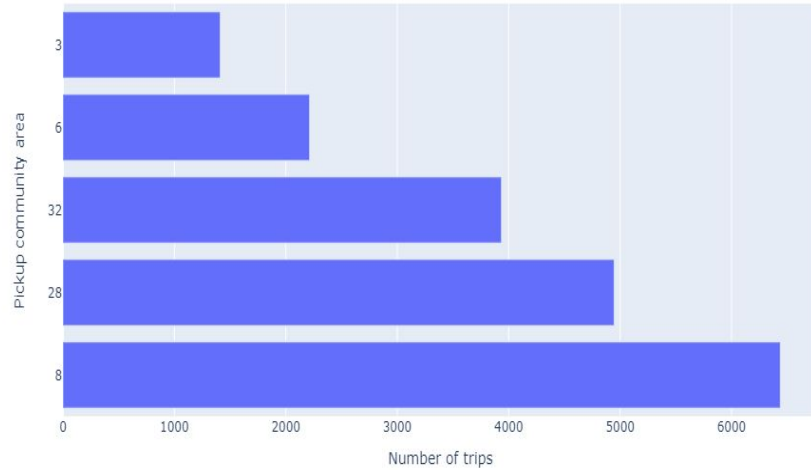
- Features:

trip_id
trip_start_timestamp
trip_seconds
trip_miles
dropoff_community_area
fare
trip_total
payment_type
company
dropoff_centroid_latitude
dropoff_centroid_longitude
pickup_community_area
pickup_centroid_latitude
pickup_centroid_longitude

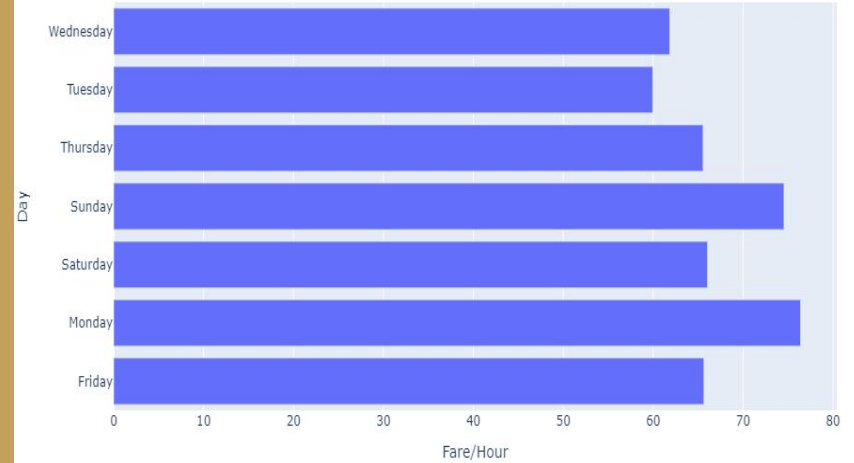


EDA

Total number of trips in the selected time



Average Fare per Hour over the selected period



Chicago Taxi Trips Analysis

Data Taxi trip database can be reached from [City of Chicago Transportation website](#). This data includes trip information between 2021-01-01 and 2021-10-01. By sliding the slider you can explore trips information in different time periods.

Select Trip Info

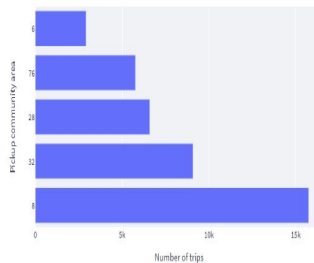
Months:

March

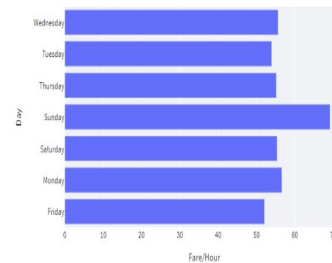
Time of day:

Afternoon

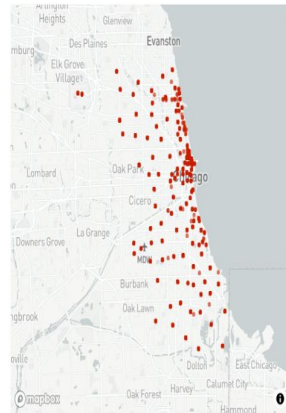
Total number of trips over the selected period



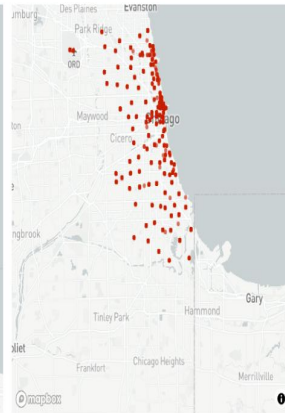
Average Fare per Hour over the selected period



Pickup Location



Drop-off Location



Taxi Company Metrics


Company	Total	Fare/Mile	Mile/Trip
Fresh Cab	24596	3.25	5.66
Taxi Affiliation Services	20019	4.55	4.13
Medallion License	4311	3.43	5.18
Twicab Insurance Agency LLC	3250	3.13	5.48
Sun Taxi	2670	3.03	6.28

- For the selected month and time of day
 - Districts with top 5 demand
 - Popular pickup/Dropoff locations
 - Average fare per hour with respect to day of the week
 - Top 5 best performing companies



Future Works

- Build a predictive model
 - taxi trip duration, fare, demand
- Provide more filters on app
- Provide more graphs/charts on app



Thank you