
DTU

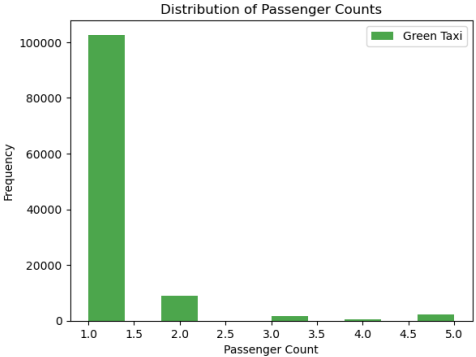
DTU

Course 62444, Group 18

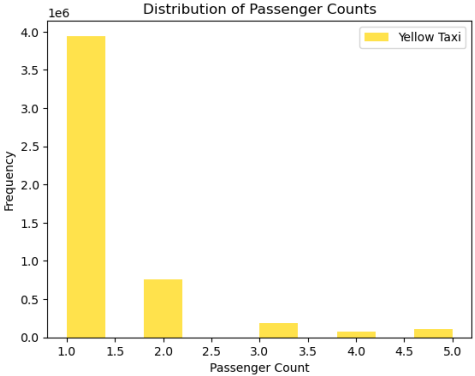
Data Visualization and Analysis

Project 1: Analysis and Forecasting of NYC Taxi Rides

Exploratory Data Analysis



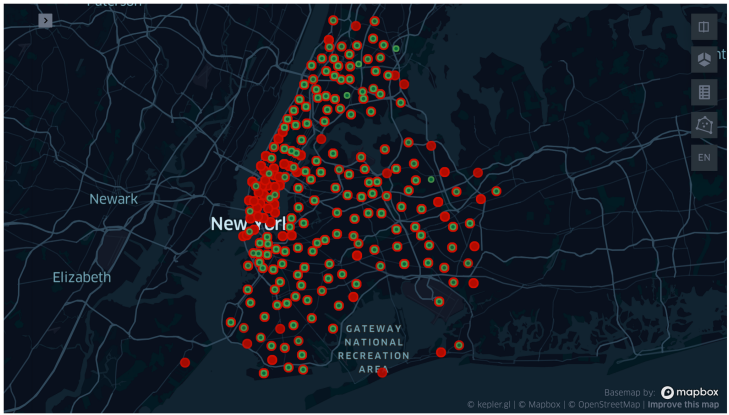
Histogram



Histogram

- The histograms display the distribution of passenger counts for yellow taxis and green taxis

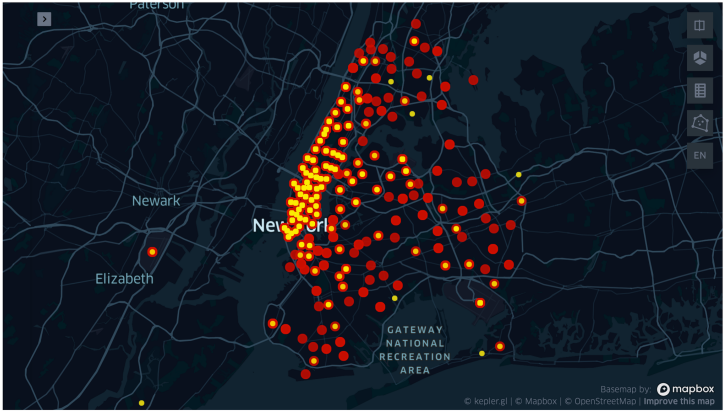
Spatial Analysis - Green Taxis



Heatmap over pickup and dropoff

- Central Park hotspot
- Pickup (green) vs Dropoff (red)

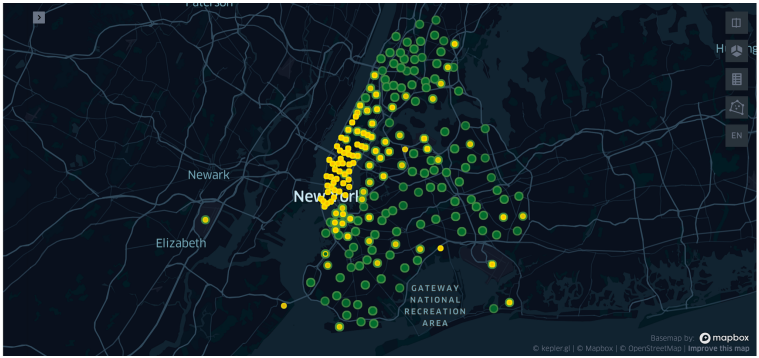
Spatial Analysis - Yellow Taxis



Heatmap over pickup and dropoff

- Central Park hotspot
- Pickup (yellow) vs Dropoff (red)

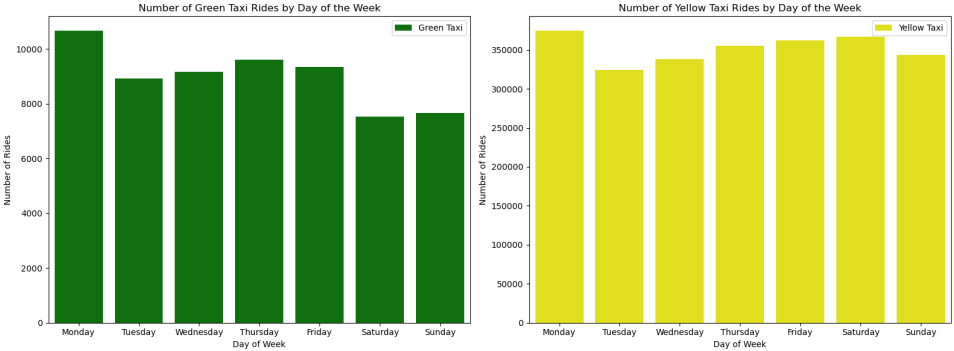
Spatial Analysis - Yellow and Green taxis



Heatmap - Pickup Hotspots

- Green VS Yellow
- Differences in hotspots

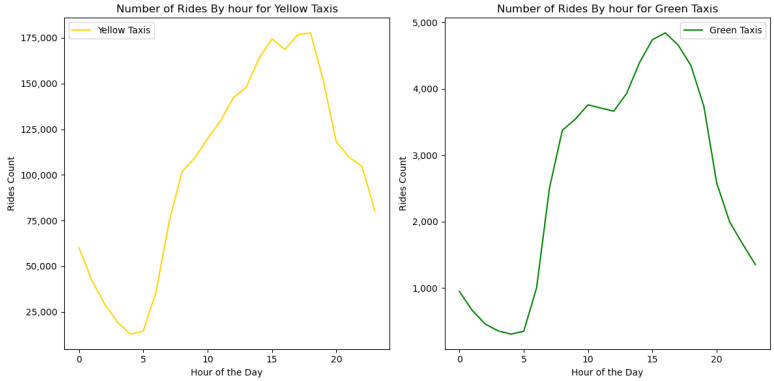
Temporal Analysis



(a) Bar Chart

- The number of rides by day of the week - For the yellow taxis the highest number of rides occurs on Saturday. For green taxis the highest number of rides occurs on Thursday.

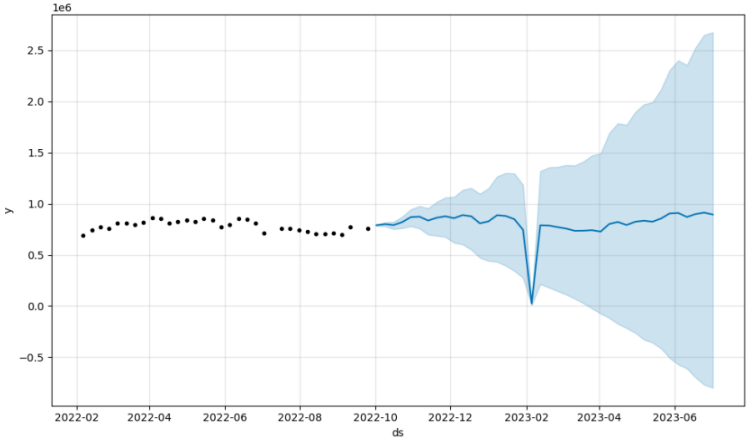
Temporal Analysis



(b) Line-plot

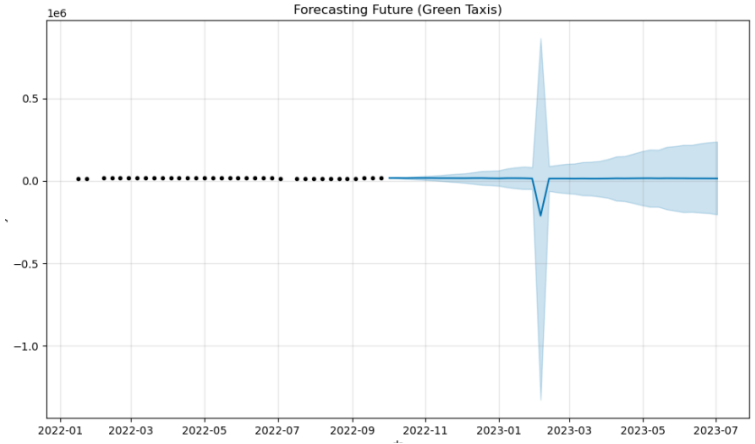
- The line plot compares the number of trips by hour of the day.

Forecasting of NYC Taxi Rides - Yellow Taxis



- This graph shows the predicted values for future trends along with the uncertainty range

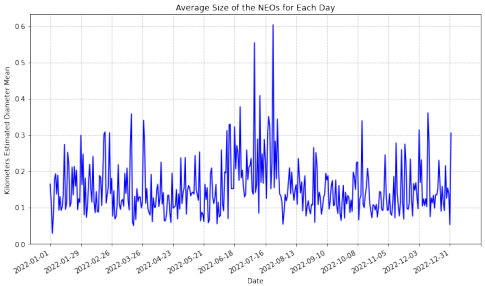
Forecasting of NYC Taxi Rides - Yellow Taxis



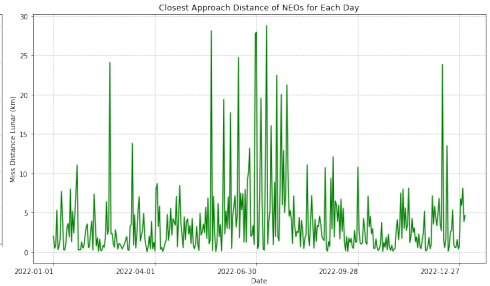
- This graph gives us insight on how the future of taxi rides will look for the green taxis

Project 2:NASA Data Acquisition, Visualization, and Analysis

Data Analysis



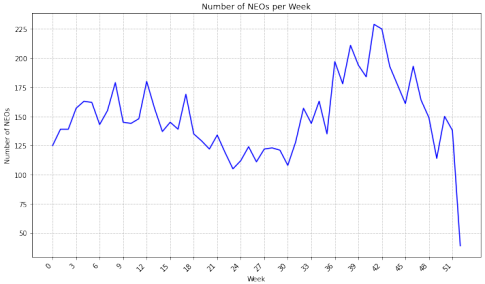
Average Size of the NEOs for each day



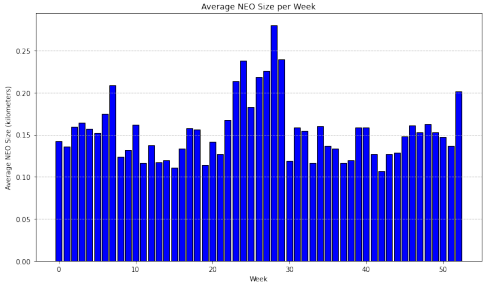
Closest Approach Distance of NEOS for each day

- In the summer the size and distance is increased a lot

Data Visualization Part A



Number of NEOS per week

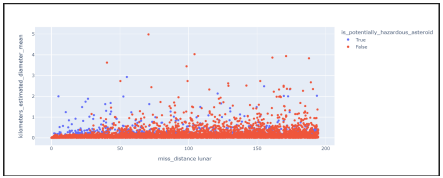


Average NEO Size per week

- Week 21-30 = Number is low
- Week 40-45 = Number is high
- Week 24-30 = Very large sizes

Data Visualization Part B

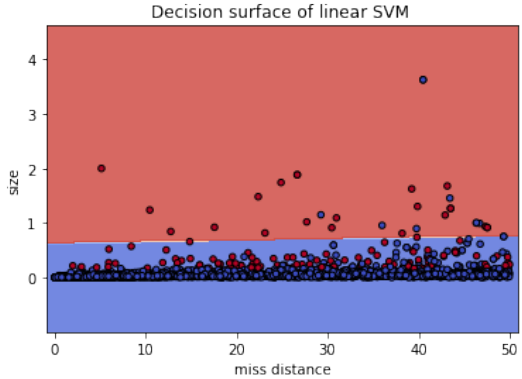
Distribution of hazardous vs non-hazardous NEOs



Potentially Hazardous

- The proportion of hazardous Near-Earth Objects (NEOs) is noticeably smaller in comparison to non-hazardous ones.
- Almost no correlation

Support Vector Machine



- The results are inconclusive as the data points are densely packed and closely situated, making it challenging to reliably distinguish between classes.