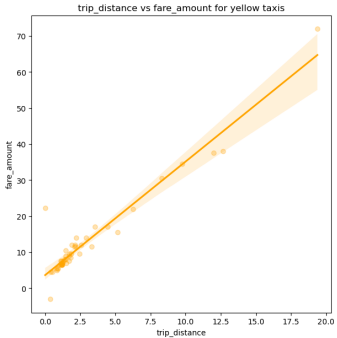
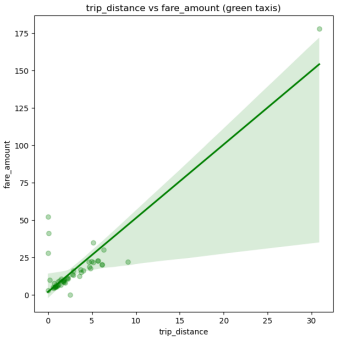

DTU

Muaz Ahmed - S205430
Ahmad Sandhu - S215813
Sohaib Sajjad - S215784
Milad Ibrahimi - S215796
Noorulein Jamil - S215810

Group 11

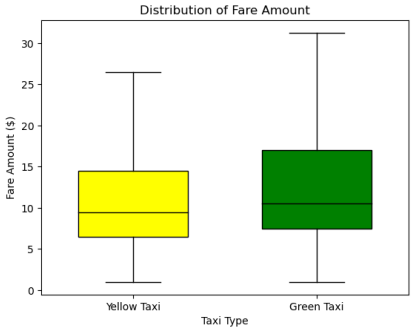
Project 1 - Analyzing NYC Taxi Rides





While both yellow and green taxis show that longer trips generally cost more, yellow taxis display a more consistent pricing pattern as indicated by the tighter clustering. Green taxis have more variability in fare costs and potentially some outliers with unusually high fares.

Fare cost for green and yellow taxis



Yellow fare
mean: 13.55
max: 401092.32

Green fare
mean: 14.25
max: 604.5

Green taxis, on average, charge slightly more than yellow taxis. The green taxis also display some fluctuation in their fares. The provided maximum fare for yellow taxis is insanely high and should be investigated for accuracy.

Trip distance for green and yellow taxis

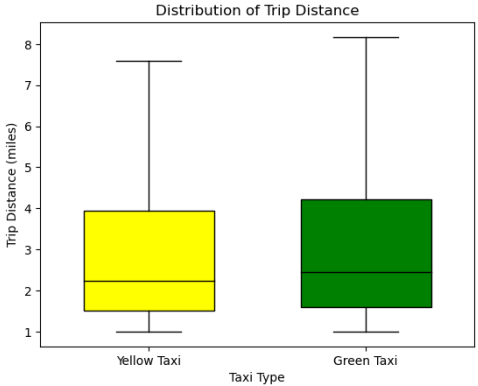


Figure: While the median trip distance is similar for both yellow and green taxis, green taxis have a wider range of trip distances, suggesting that they are more likely to have both shorter and longer trips compared to yellow taxis.

Passenger count

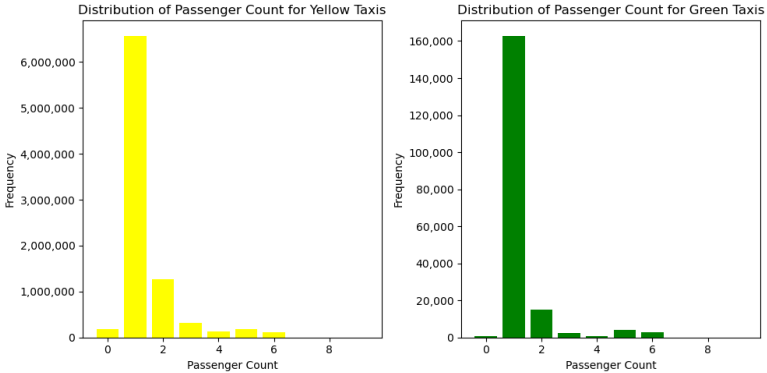


Figure: The vast majority of trips for both yellow and green taxis are made with just one passenger. Yellow taxis have a higher overall number of trips across all passenger counts.

Spatial Analysis

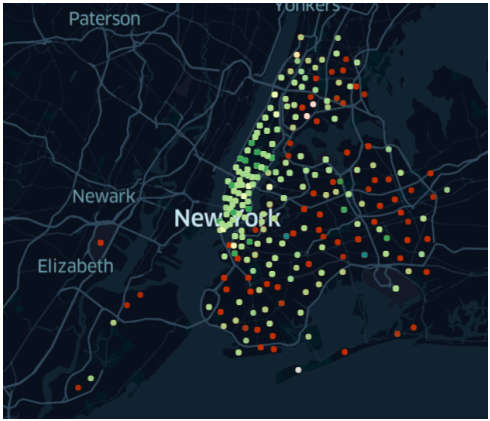


Figure: Spartial map with kepler. Pickup(PU) is green, Dropoff(DO) is red. Drop off and pickup is overlapping, that's why there is more green than red.

Temporal Analysis - Yellow Taxi

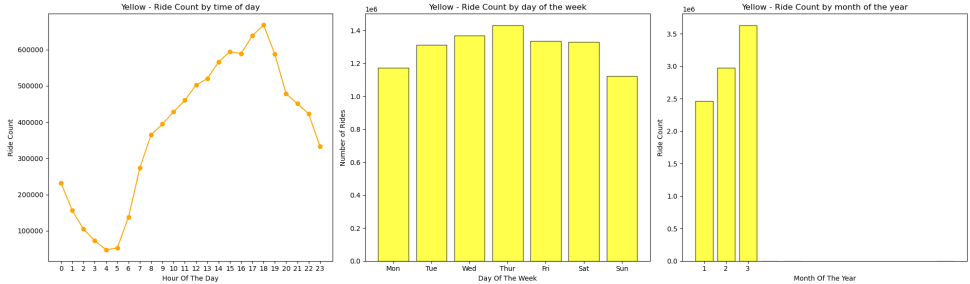


Figure: Yellow taxis are busiest in the evening hours of the day, particularly on weekdays, with Friday being slightly busier than other days. There's a decrease in rides over the weekend. Also, there seems to be a monthly variation in ride counts, with a increase in rides per month.

Temporal Analysis - Green Taxi

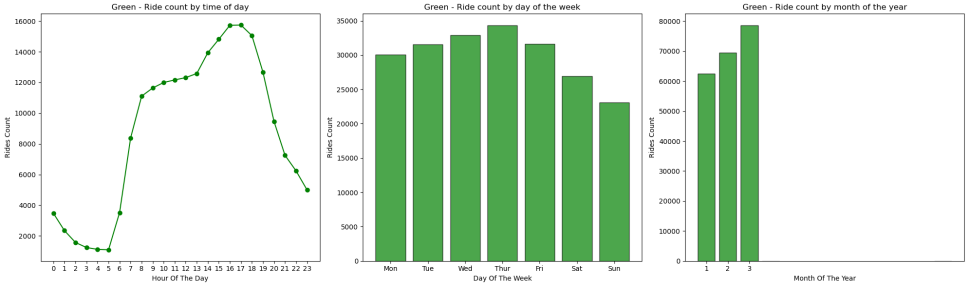
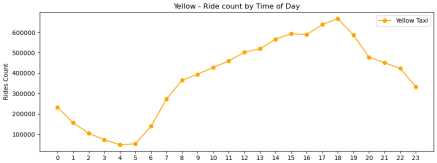
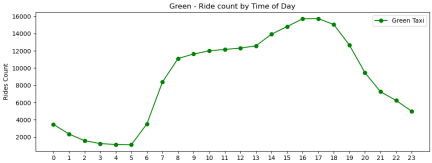
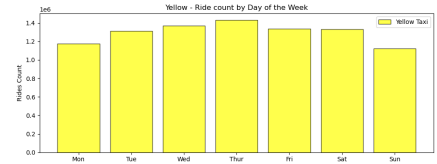


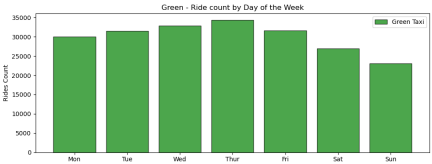
Figure: Green taxis see the most activity during weekday evenings, particularly in the middle of the week, with a decline in rides on weekends. There is also a monthly variation in ride counts with a increase from the first to the second to the third month.

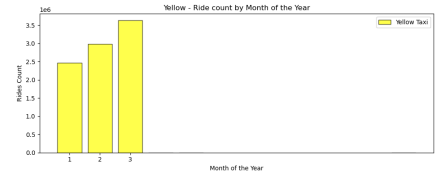
Temporal Analysis - Comparison

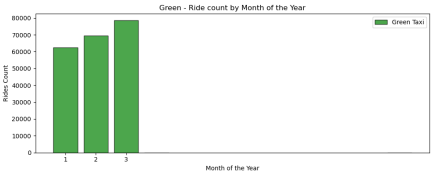












Temporal Analysis - Ride Count Conclusion

In summary, both yellow and green taxis have the highest demand in the late afternoon and early evening hours. Yellow taxis maintain a steadier demand throughout the week with a slight uptick on Fridays, while green taxis are doing well on weekdays with a decline during the weekends. Both yellow and green taxis experience a substantial increase in demand from the first to the third month of the year.

Temporal Analysis - Average Fare

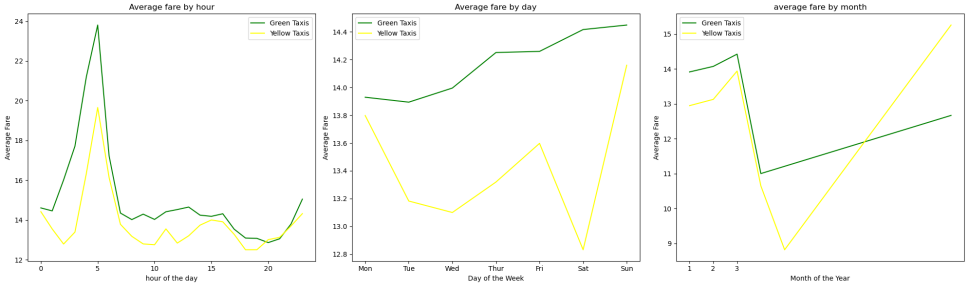


Figure: Green taxis have a notable early morning fare spike and a general increase in fares towards the weekend, while yellow taxis have more variability in average fares across different days of the week and show a significant increase in average fares over the months.

Temporal Analysis - Trip Duration

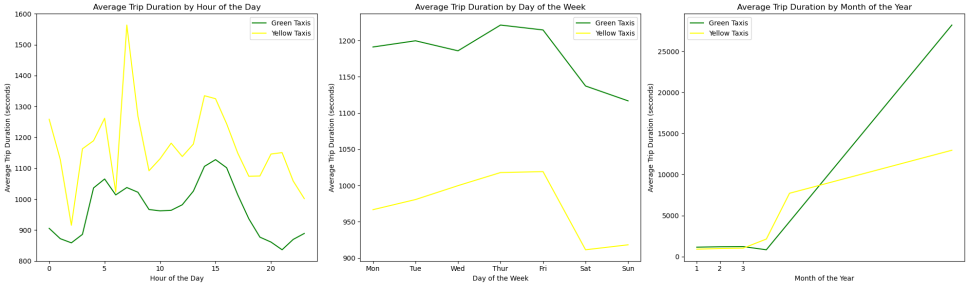


Figure: Average trip durations for both green and yellow taxis peak in the early morning hours and change throughout the week, with longer trips typically occurring on weekdays. Trip durations for both types of taxis increase markedly over the first quarter of the year, with green taxis starting their increase earlier than yellow taxis.

Forecasting - Yellow Taxi Passengers

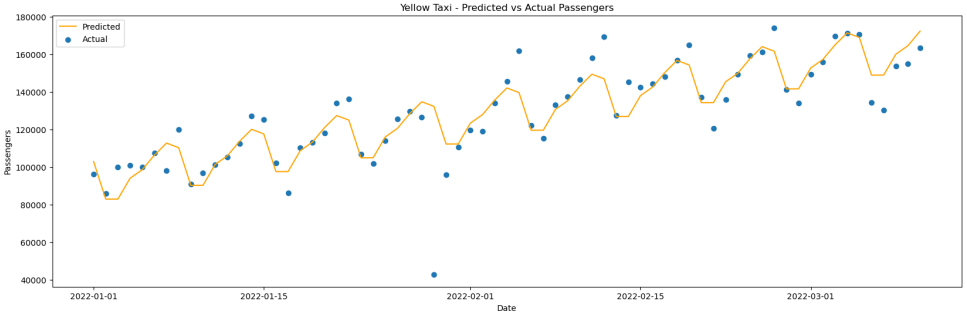


Figure: The predictive model for the number of passengers in yellow taxis appears to be generally effective, with some periods of underestimation/overestimation. Improvement of the model will be necessary to improve its accuracy.

Forecasting - Green Taxi Passengers

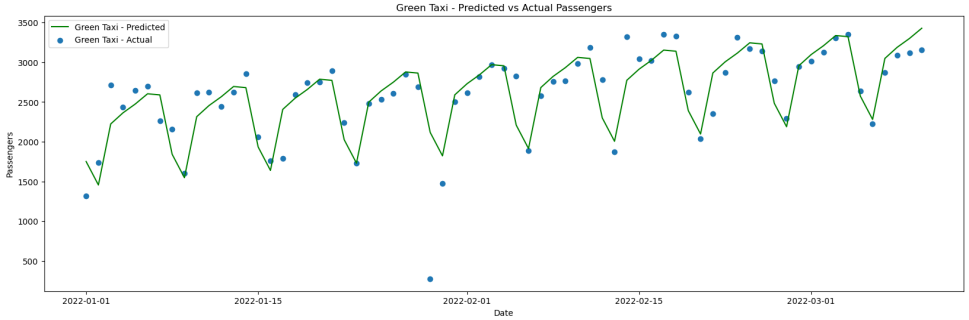


Figure: While the prediction model for green taxi passengers captures the overall trend, there are specific periods where it does not align well with the actual data, indicating potential areas for model refinement/improvement.

Project 2 - NASA Data, Visualization and Analysis

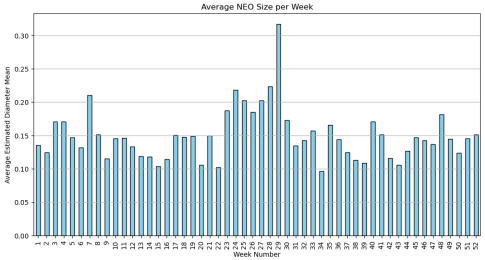


Figure: While the average daily size of NEOs is variable, most NEOs observed on any given day are relatively small, with occasional days where larger NEOs are recorded. These outliers could be of particular interest for further investigation to understand why certain days have much larger average NEO sizes.

NEOs per week

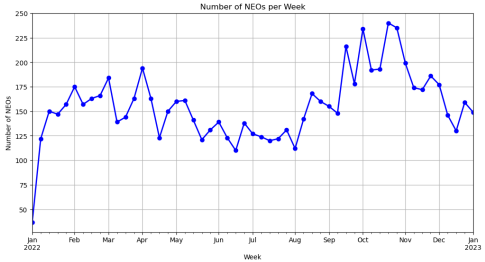


Figure: The detection of NEOs varies from week to week. The data does not show a consistent seasonal pattern, but rather periods of higher and lower detections, which could be influenced by a variety of factors such as observational campaigns or the position of Earth.

Hazardous NEOs

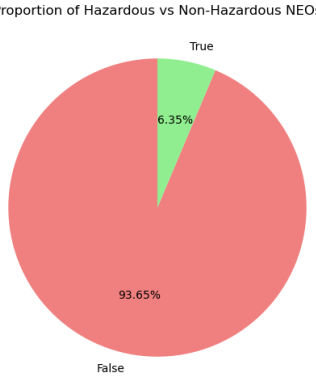


Figure: While there is a significant number of NEOs near Earth, only a small fraction of them pose a potential hazard.

Correlation between NEO size and close approach distance

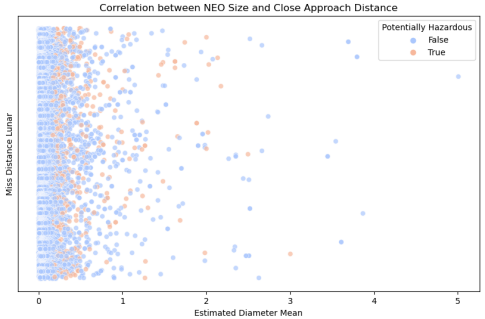


Figure: The data indicates that potentially hazardous NEOs are not exclusively large in size; they can also be relatively small. The determination of an NEO being potentially hazardous likely includes other factors such as velocity, not just size and close approach distance. The lack of a pattern in the scatter plot suggests that there is no simple direct relationship between an NEO's size and how closely it passes by Earth.

Recommendations

- **Special Focus Weeks:** The heatmap indicated specific weeks where larger average sizes of PHAs were detected. These periods could be considered special weeks, where monitoring efforts are increased and resources are allocated to ensure comprehensive coverage and analysis.
- **Increase Monitoring of Large NEOs:** Since larger NEOs are more likely to be classified as potentially hazardous, it's important to prioritize the monitoring of larger NEOs. Telescopic resources could be allocated to track their trajectories more closely, especially those above a certain size.