Problem: Analyzing and Predicting Traffic Violations in NYC

It is very frequent to see illegal or double parkers in urban areas and this can cause a lot of frustration for those affected by this group. Often a 311 service call/inquiry with a follow-up by NYPD is needed to resolve this situation. I would like to analyze the 311 service requests dealing with blocked driveways and illegal parking and attempt to create a prediction of areas and times of when it would be best for NYPD traffic to patrol.

Client: NYPD, Department of Traffic, private towing companies

Through this analysis, the NYPD and DOT can task patrols around specific areas during select time periods and this would improve ticket violations being submitted and thus higher revenue to the city. This will also help the allocation of resources in a more efficient manner unlike the times when you see patrol cars go around the same block multiple times. Additionally, private towing companies would be able to advertise their services towards areas of higher incidents.

Sample Data:
https://data.cityofnewyork.us/Social-Services/311-Service-Requests-from-2010-to-Present/erm2-nwe9

This dataset is actively updated daily with current 311 service requests.

Methodology:
1. Wrangle data from NYC Open Data on 311 service requests
2. Exploratory Data Analysis
   a. Is it possible to perform a times series data analysis?
   b. How does the density of violations differ over time?
3. Machine Learning
   a. Perform time series forecasting techniques such as ARIMA and moving average. Compare results to actual data.
   b. Create a binary classification model to predict the probability of incoming service request to be that of traffic violation.
4. Record results and offer any insights or future suggestions

Deliverables:
- Powerpoint
- Paper to summarize the results
- Github Repository
   ○ Jupyter Notebook containing code and explanations