Background Information

• Revenue for a local government is important to keep that government running

• Around 22.6%, second biggest portion, of a governments income comes from fines and fees, which parking tickets is a part of

• NYC made $993 million off parking fines from fiscal year 2012 to 2016

• How does understanding the data help?
Problem

- How to increase the revenue of NYC parking tickets
- Which tickets could increase the revenue generated by NYC
- Are there safety issues on street that are getting more tickets issued
- Does weather and temperature play a role in the number of tickets issued
Methodology

• Using a Spark cluster to analyze data
• Load in data sets into data frame
• Analyze columns in data to understand what is needed and what can be dropped
• With temperature convert all the Kelvin readings to Fahrenheit
• In the weather and temperature data frame drop all data except for date and data about New York
• Filter out unwanted dates and null values for weather and temperature
Methodology

• Then format date on parking data to match the date format of the weather and temperature creating a data frame with ticket information and weather and temperature columns too
• Start creating csv’s to look at details of the data
• Such as top streets tickets were given on, most common weather when a ticket was issued, and most common temperature when a tickets was issued
• Wrote those data frames out to csv files on an HDFS cluster
Methodology

• Take in csv files produced by our spark cluster
• Use python3 and Jupyter notebooks to analyze data
• Asked 12 different questions like how does weather affect tickets?
• Some data required additional data wrangling to get useable output.
Performance Benchmarks

• 'Sky is clear' is most ticketed weather description
• 40 degrees F is most ticketed temperature, with a nice normal distribution across the range of degrees we measured
• Precinct 19 gave out the most tickets
• 3rd ave with parking violation 37 was the most ticked street and most ticketed violation
• The most violations occurred in precinct 19
• 36 was the most given violation
Performance Benchmarks

- NY plates were the most ticketed, followed by NJ
- Cop with id 18 gave out the most tickets (over 4000)
- The most ticketed violation in the most ticked weather was violation 36 in 'sky is clear'
Insights and Conclusion

- NYC needs to train their law enforcement in better data recording to make further analysis easier.
- We can increase fines for speeding in school zones to easily increase revenue by over $100M a year.
- We can also increase parking in no park zones to achieve the same effect.
- The problem space must be closely monitored in the next 10 years as NYC's public transit system is revamped and with the increasing share of autonomous vehicles in traffic.
References


