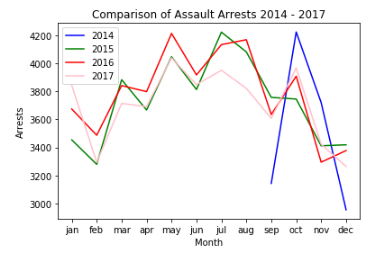
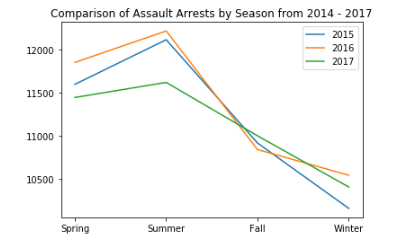
Summary of Data Analysis

Is NYC getting more dangerous by year? By season?

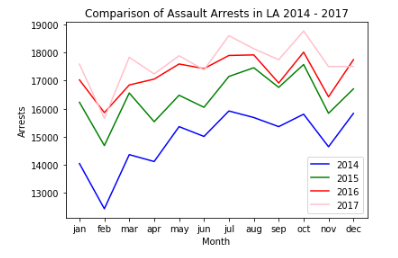
We plotted the number of arrests for assault from Sept 2014 through 2017 and found that assault arrests have remained relatively stable. Based on our ANOVA test, the differences between each year are not statistically significant. We also plotted the number of crimes by season for Sept 2014 through 2017. Although there does appear to be a trend for more arrests in the summer and spring months based on the graphs we plotted, our ANOVA test showed the differences between these seasons at least in 2017 was not statistically significant.

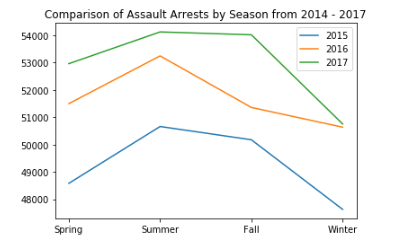




Is LA getting more dangerous by year? By season?

We plotted the number of arrests for assault from 2014 through 2017 and found that assault arrests appear to be steadily increasing based on the graph of the data we plotted. This is consistent without ANOVA test which showed the differences between the years as statistically significant. However, we also plotted the number of crimes by season for 2014 through 2017, although there does appear to also be a somewhat inconsistent trend for more arrests in the summer and spring months based on the graphs we plotted. Our ANOVA test showed the differences between these seasons at least in 2017 was not statistically significant.





Is NYC more dangerous than LA?

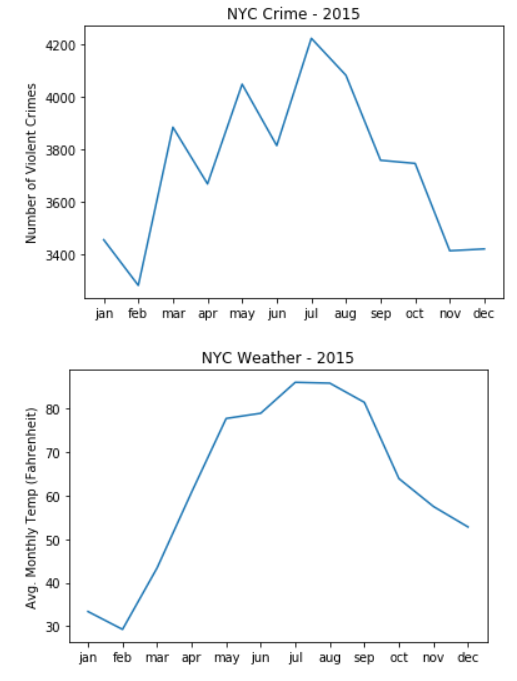
Our independent t-test did not show a statistically significant difference between the number of assault arrests in NYC across 2015 to 2017 and the number of assault arrests in LA across 2015 to 2017.

Are certain restaurants more dangerous in NYC?

We used our assault arrest’s longitudes and latitudes to do a nearby search of restaurants to these incidents to determine if some restaurants have had more incidences of arrests near them than others. We chose the most dangerous month from our dataset, October 2014, to investigate. We were able to call the Google Maps API to do this search and found that some restaurants do have more incidents of arrests than others. Our hypothesis was that McDonalds would have the highest incidences of arrests near them, and while that did turn out to be the case with the highest number of arrests, we also wanted to take into consideration that some restaurants might have more locations than others making them more likely to have an arrest occur nearby. Once we used a ratio to compare the number of unique restaurants that had an arrest occur nearby and the number of arrests that occurred near those locations, we found a particular White Castle to have the highest ratio of arrests to locations.

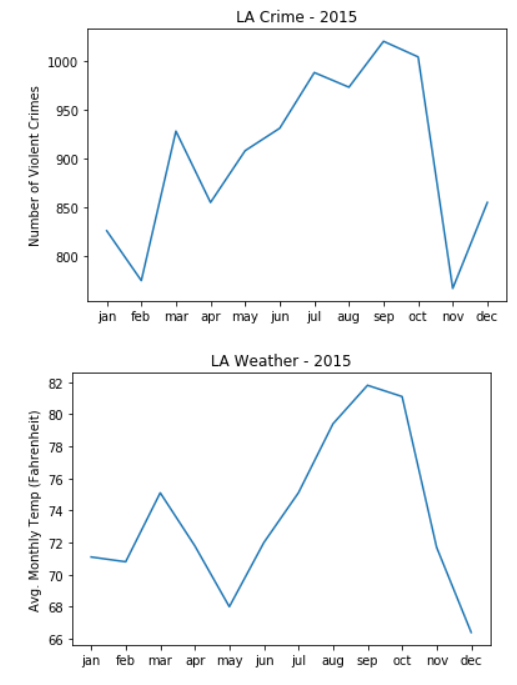
Does the weather in NYC have an impact on violent crimes?

When comparing the two graphs below, one can notice a trend in the data. It appears that as the average monthly temperature goes up, so does the rate at which aggressive crime is committed. With this in mind, it can most likely be determined that the weather has an impact on crime rates.



Do certain weather conditions make a particular day in LA more dangerous? (Mark)

If one were to look at the graphs below, a scary observation can be made. The graphs appear to be very similar, the first graph is arrests in LA, while the second one is the average monthly temperature. Based on the data from these two graphs, it can be said that the weather does have an impact on crime rates in LA.



What problems did we encounter along the way?

The first major problem encountered by the group was deciding what to do our project on. We spent a decent amount of time at the beginning of the project looking through data on Kaggle, Data.gov, and other sites to narrow down our interests and find data that was adequately large and complete. By the second class, we decided on two sets of data. The group chose to do an analysis on arrest data from New York City for assault crimes and we then added Los Angeles too so we could compare our analysis across two major cities. While completing our analysis on these cities, we ran into our main issue during the process of calling the Google Maps API. When attempting to locate the nearest restaurants in NYC, the resulting data frame was not saved in our notebook. This meant none of the data from that run was saved either, resulting in the NYC data needing to be reran. We created a new API key in order to get additional credit and were able to rerun the call. However, a similar issue happened with the much larger LA file, the computer running the call ended the terminal, or restarted it causing us to lose the data. This resulted in the loss of the LA restaurants data after making 13,000 calls to the API. Due to time and credit constraints, we were unable to make this call again.