**BEM/Ec 150 Final Assignment**

Dataset: NYPD Motor Vehicle Collisions

Author: Bianca Yang

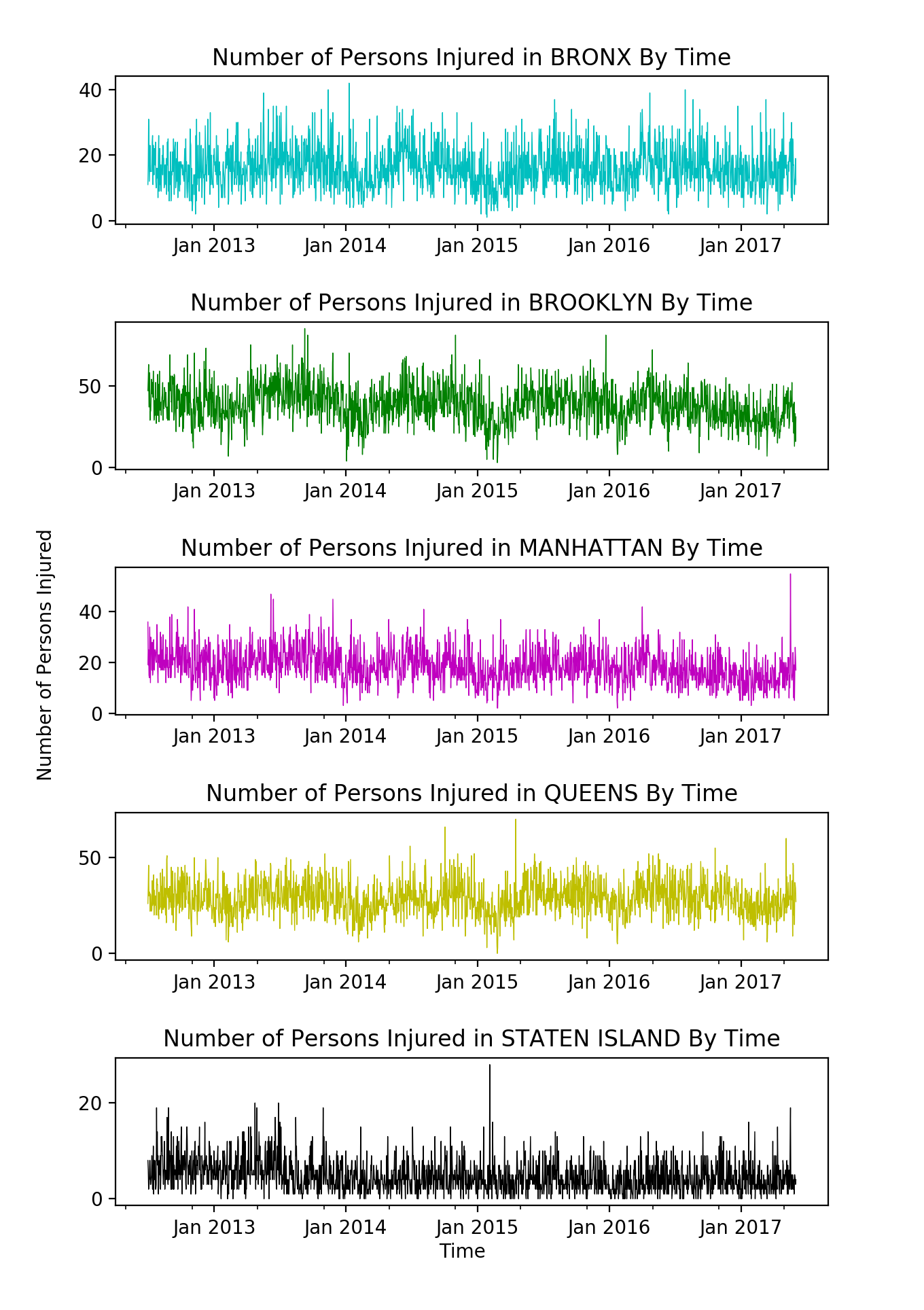
Date: 6/8/2017

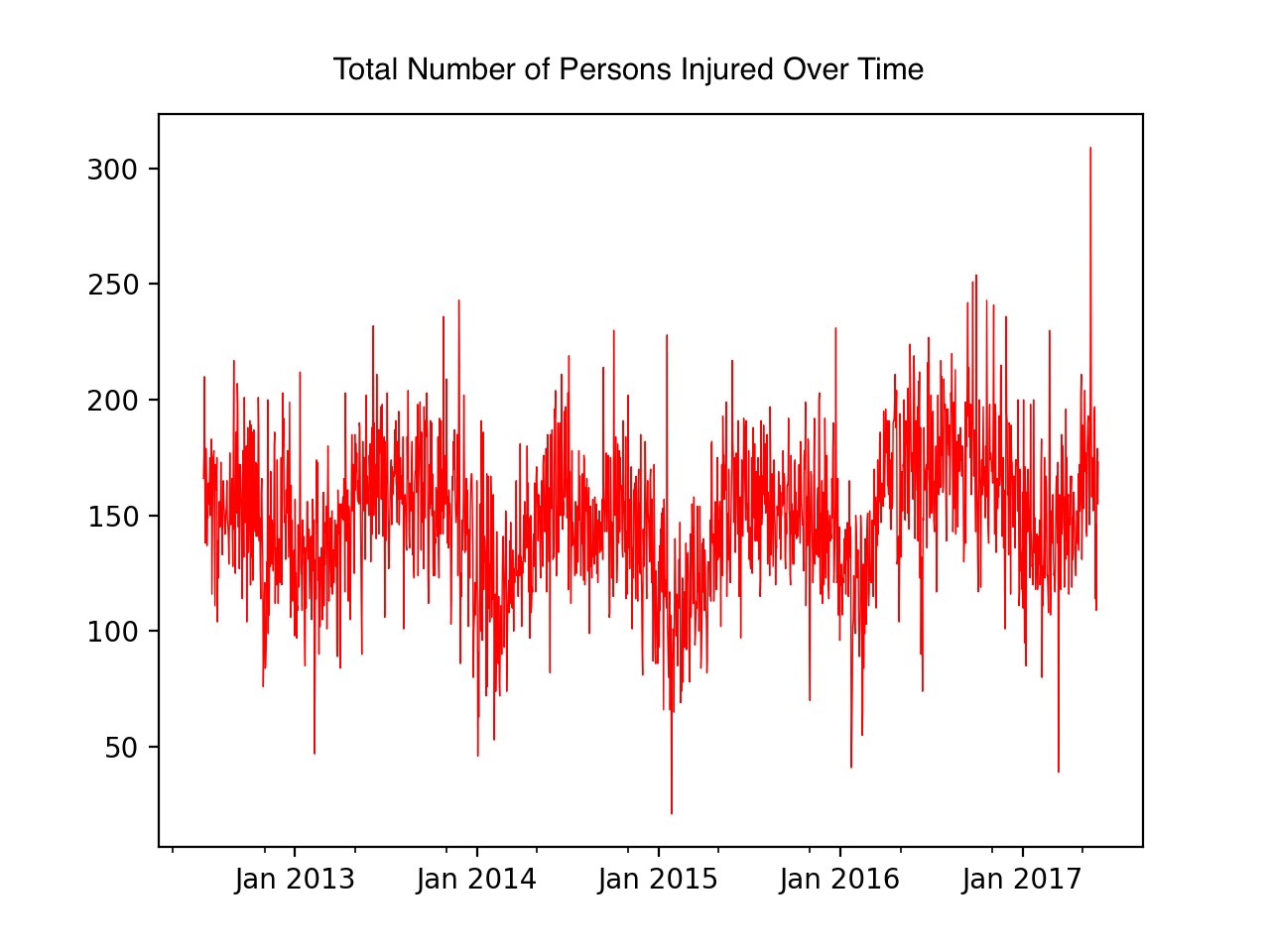
The purpose of this series of data analysis is to try to identify trends in motor vehicle collisions from July 2012 - June 2017 in New York City. The dataset has over 1 million instances and 29 features. The features includes: type of vehicle involved, contributing factor to accident, borough, and number of people injured. I have created three time series graphs which show trends in total number of people injured by borough, trends in contributing factors to number of accidents, and trends in the type of motor vehicles causing accidents.

Below are some details explaining the insights from each graph.

Graph 1: Number of people injured per borough per time

There is some cyclicality in the data, with peak number of people injured in the summer. This could be due to increased tourist traffic in the summer months. Tourists may not be familiar with traffic patterns in New York City and cause disruptions to normal traffic flow which confuse other drivers. This trend could also be explained by increased travel frequency over the summer. Families tend to travel more over the summer, since kids are out of school. This trend appears to be the strongest in Brooklyn.





Graph 2: Contributing factors by time

The contributing factors are ranked at time 1. The top 5 factors’ frequencies are plotted over the rest of the time series.

Graph 3: Incidents by vehicle type over time.

It is commonly assumed that SUVs are safer than sedans because they are able to absorb more force than SUVs. This graph \*confirms\* that assumption. We can clearly see that passenger vehicles have resulted in the most injuries. If one is going to travel through NYC by automobile, taxis are the safest option. This is intuitive, since taxi drivers have the most experience driving. This graph could simply be a reflection of the relative proportions of cars in the city.

Conclusion:

These graphs help to confirm the intuition that increased traffic in summer due to tourism results in increased motor vehicle collisions. The graphs also demonstrate that driver inattention is the biggest cause of collisions.