IS 608 Final Project Proposal

Jonathan Hernandez

Fall 2017

Visualizing Stop and Frisk Data in NYC from 2003 – 2016

My final project will be based on the visualization of Stop and Frisk Data in NYC from

2003 -2016 across New York City. Stop and Frisk is a common debated issue in NYC and examining this topic and leveraging Data Analytics and Visualization can show where Stop and Frisk is most common, what action was taken by police officers and suspects, reason to perform stop and frisk and if any malicious items was found on the subject. People can look at this group of datasets and visualization to possibly decide if they feel the neighborhood they are in or any NYC neighborhood is police friendly or safe and if their own skin color and race can play a factor.

My data will be from the NYPD public website containing stop, question and frisk Data found here: <http://www1.nyc.gov/site/nypd/stats/reports-analysis/stopfrisk.page>. I will extract yearly data from 2003 – 2016 of stop and frisk data and combine it into a single dataset. With over 100+ features, I plan to use only certain columns as I mentioned in the previous paragraph (location, action, time, suspects features, reason for stop and result of the stop and frisk). For visualization, I plan to use a mixture of ggplot2, R Shiny and possibly the R ggmap package and plotly package. I want to show an interactive map of NYC that lets users look at a particular zipcode or (x,y) coordinate and see how many stop and frisk cases were there. Also I want to add other interactive graphs showing distribution of various features as well.