# Welcome to New York

Members: Dan Manzo
E-Mail: dvmanzo@wpi.edu
GitHub: dvmanzo
Link: dvmanzo.github.io/DataVisFinal

## Background/Motivation

Having lived in New Jersey my entire life, I live about 20 minutes away from New York City. I've been aware of "Stop and Frisk" since I can remember, but being a white male have yet to be on the receiving end. Recently, Stop and Frisk has found its way back into public discussion due to the Presidential debates and it's highly controversial and sensitive subject. It was referred to as something that was no longer in effect, which is not true. The subsequent discussion thereafter demonstrated how most people are unaware of whether or not stop and frisk is actually effective, what this practice yielded, and exactly how this activity ran and still continues to run on a micro-scale even now.

## Objective

The goal of this project is to highlight who is being targeted, where one is most likely to be stopped, and shed light into what is resulting from an ideology that may or may not be as effective as people would assume. For those viewing the visualization in its final state, I aim to have it answer these questions:

- In a given location, how likely am I to be stopped?

- When and where are most people stopped?

- Is there any general indication of wrongdoing before being stopped?

- How often are the police using force when stopping an individual and how often is it unwarranted?

- Have crime rates gone down in recent years since stop and frisk?

These questions are important as the notion of stop and frisk itself is controversial to say the least. As citizens of the United States, we are protected from search and seizure without probable cause by the 4th Amendment of the Constitution. Stop and frisk, much like the PATRIOT ACT, feeds into the fear and paranoia of suspecting

that anyone and everyone is likely breaking the law or plotting some attack that we should be countering by signing away our rights and freedoms through constant surveillance. The ideology that being able to search someone without cause is completely unconstitutional. The PATRIOT ACT, as we recently found out thanks to whistleblower Edward Snowden, is being abused as we are actively spying on every citizen and recording all data transmissions regardless of whether or not they are under investigation of criminal activity. Is stop and frisk another type of policy that is overextending the reach of the law? If so, is it even effective in lowering the rate of crime?

## Strategy

My data source for this project is the New York Police Department's "The Stop, Question and Frisk Data" report. They offer, as public record, the .csv files of all Stop and Frisk activity by year from the year 2003 through 2015. The type of data that is recorded in these reports include:

- Date and time of stop
- The race, age, and gender of the person stopped
- The location of the stop
- Was the person arrested?
- Was the person carrying a weapon or other contraband?
- Did the officer explain the reason for stopping?
- Did the person run away?
- How long was each stop?
- Was physical force used?
- The reason for stopping
- The reason for using physical force

The data itself is encoded using timestamps, coordinates, police codes, and binary values of Y and N for most of the Boolean questions so there will be some cleaning of the data in order to effectively use it in my visualizations. There is also a difference in some of the data from 2003 to 2015 where some fields aren't recorded for various technical or logistical reasons. Overall, though, I should not have an issue with being able to get it effectively running in my project.

Because this project is data-rich, I plan to have multiple illustrations, which effectively highlight the narratives that the data reflects. For example, one will show the geographical data of each stop and map it to the geographical map of New York City with a timeline scrubber that will allow users to scrub through a day to see the time and locations of where most stops occurred. They will also be able to click on an area to of that map at that time to see what demographics make up a stop at that particular time and place. Another visualization will include an interactive timeline of crime activity through the years with data point lines such as race, gender, age

and type of crime as toggled values that can be added and removed from the timeline. A third type of visualization will include comparable graphs of each of the demographic criterion compared. For example, given a period of time what is the portion of whites being stopped in comparison to blacks? Male to female? Low-income areas to high-income areas?

I plan to replicate the structure of the MTA project demonstrated in class in that they were effectively able to give a micro-scaled view of data that is presented on a macro-scale. I am separating the data into several different visualizations because I believe that having all of these different aspects and perspectives of the data will lessen the user experience and overcomplicate the visualization itself. Furthermore, the narrative of each visualization is complex and must give room for the user to explore succinct data and decide their own conclusions based on the information presented. Having too many options for toggling and displaying will overcomplicate the presentation and overload the user thus preventing them from absorbing the narrative that each visualization will be expressing. Doing so would be strictly counterintuitive of this course.

At the very least, I plan to have visualizations for these narratives:

- An interactive map of New York City that shows the geographic data of each stop with a timeline scrubber that will allow users to scrub through a day to see the time and locations of where most stops occurred. I will use heatmaps to show the occurrences of stops by region and will be able to show the demographics of the stops when clicked on.

- An interactive timeline view of crime activity through the years with data point lines such as race, gender, age and type of crime as toggled values that can be added and removed from the timeline. This will highlight any trends that might have simultaneously occurred with other controversial events involving hate crimes, terrorist attacks, mass shootings, etc.

- A graph that will clearly breakdown the demographics of each stop in comparison to the other demographics. For example, what is the portion of whites being stopped in comparison to blacks? Male to female? Low-income areas to high-income areas?

- A graph that shows the use of police force against each demographic

If time allows, I would like to have additional features in my project that will investigate the correlations between other components related to stop and frisk such as:

- The correlation of weapons and contraband found in stops

- How many stops actually lead to arrests?

- Where and when you would most likely be stopped and frisked based on the user's inputted information

Here is my schedule upon approval of this project:

November 1 – Submit proposal - continue work assuming that it will be approved

November 8 – Have data cleaned and ready for import into project

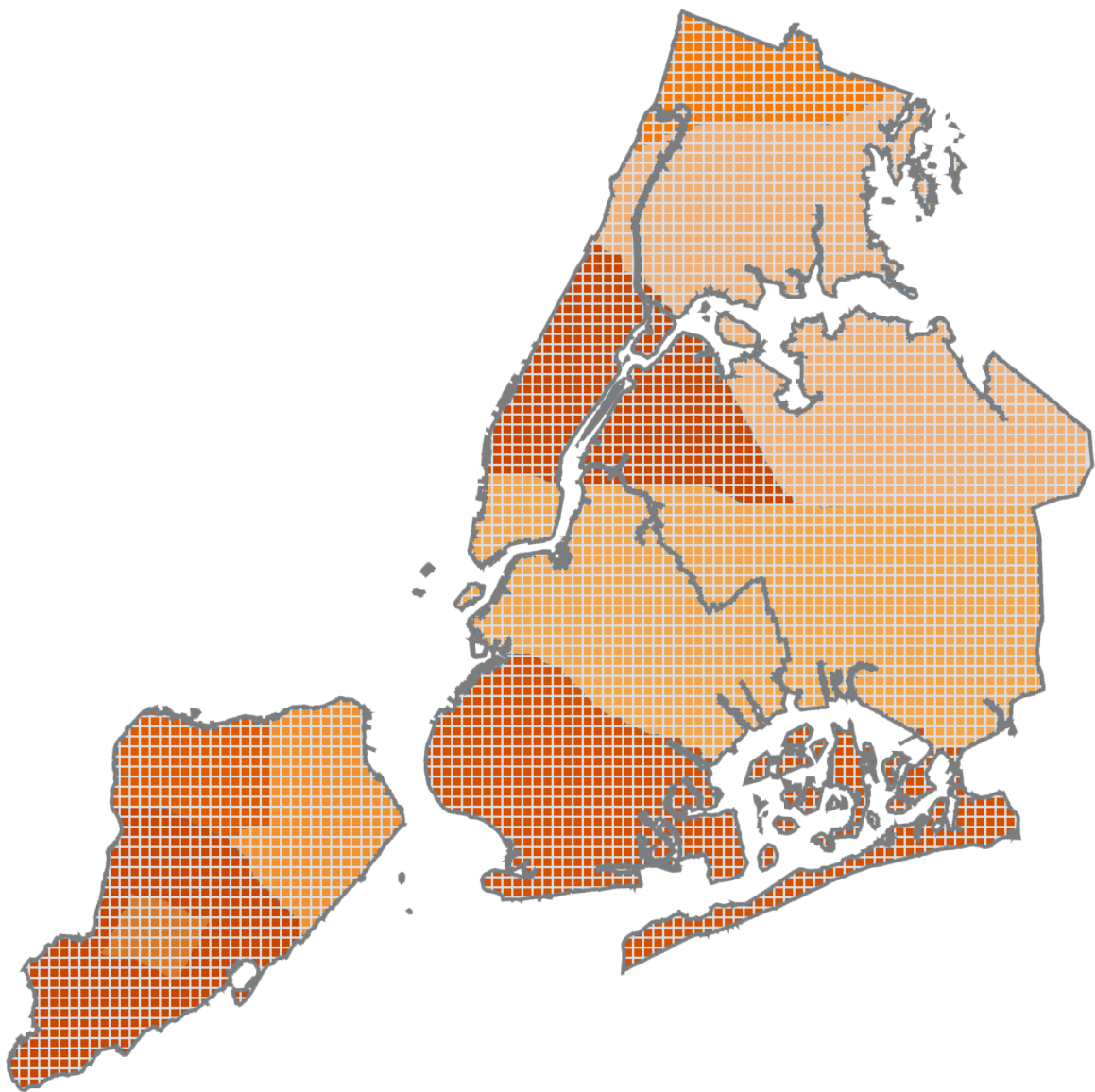November 15 – Have data imported into geographic map

November 22 – Have two of the four narratives written and the visualizations completed for each with the expectation of having the remaining two ready for the prototype on the 29th.
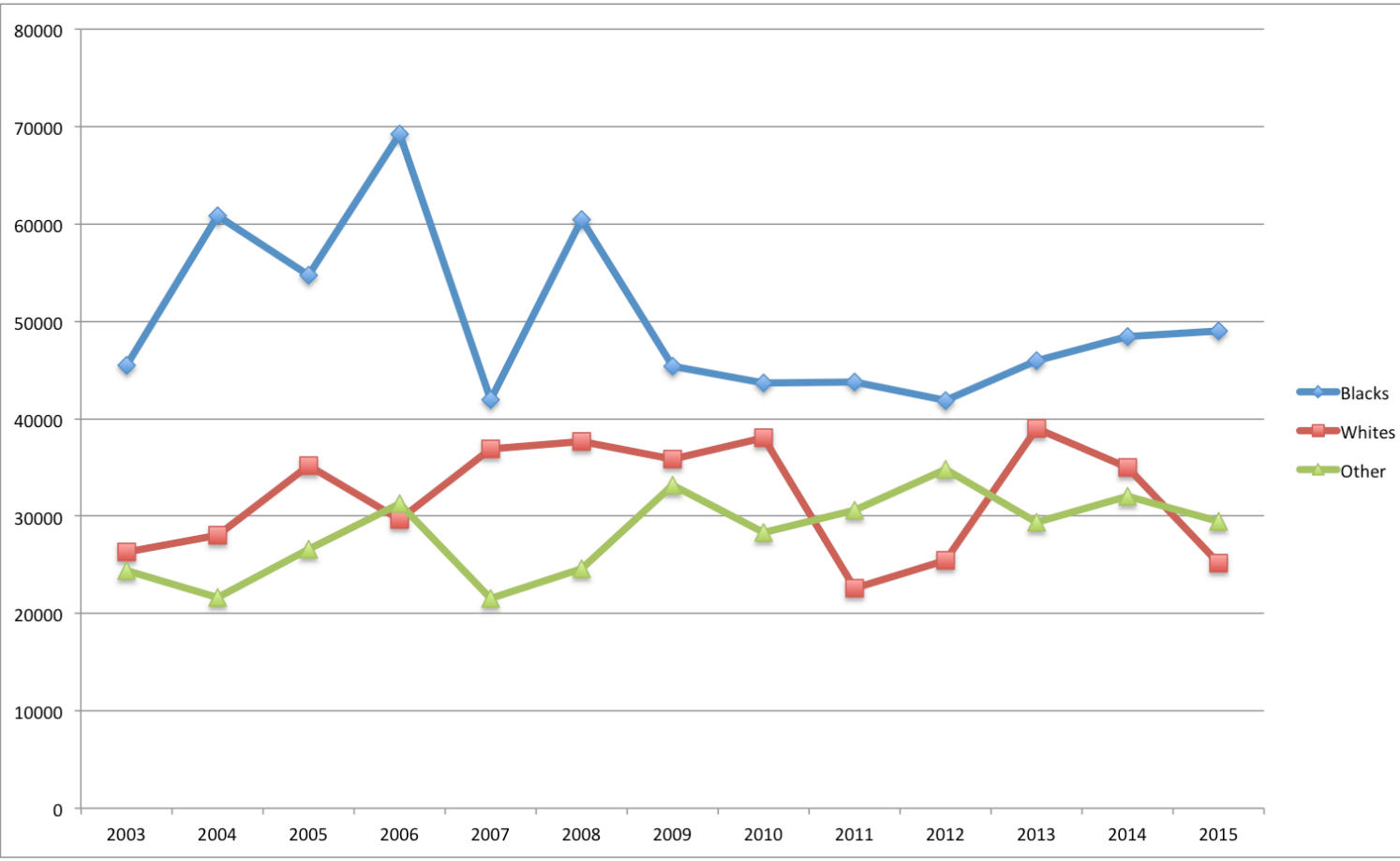
December 6 – Have all of the narratives and visualizations completed. Begin to work on final touches for color and begin user-testing.
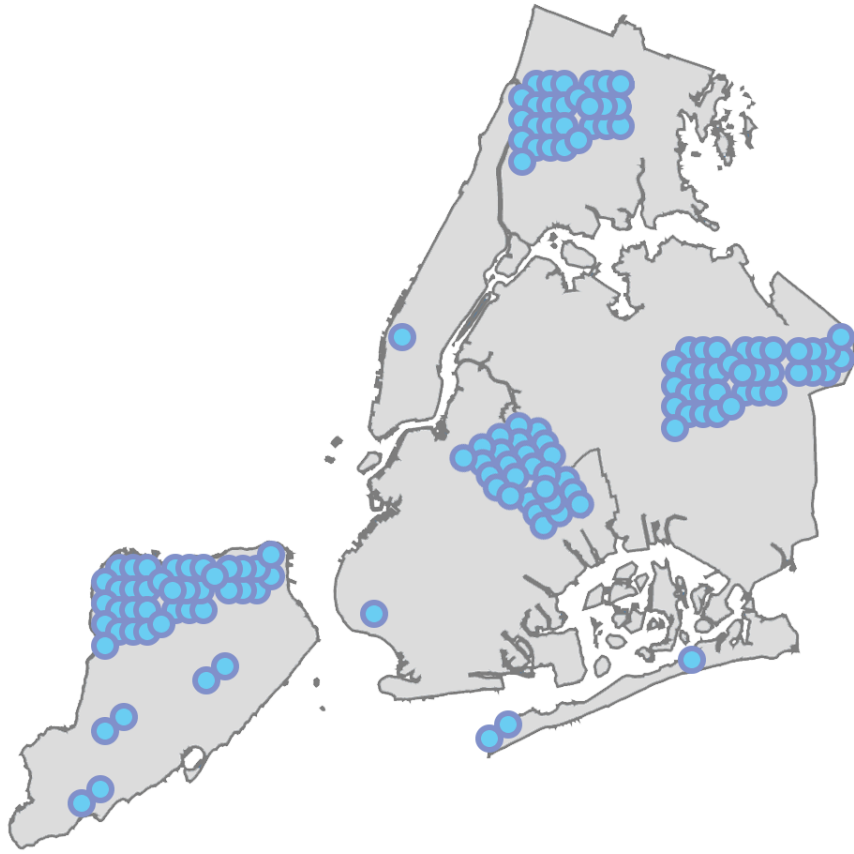
December 10 – Record final presentation video explaining each visualization within two minutes

December 13 – Present the presentation video in class

## Sketches/Illustrations

7:30am