

# Reported Crimes in New York City

## INFO 3300 Project 2 Write Up

Cari Himel, Ryan Kim, Joaquin Amante

### Description of Our Data

**Reported Crimes in New York City** is a large database of all New York City crimes that occurred from around 1910 to 2015. The focus of this visualization is the most recent, and therefore relevant, year represented in the data: 2015. The data file is so inclusive and large that there are too many crimes to render for a whole year at once, as this would require the code parsing through more than a million rows of data and significantly lagging the system. Our solution to more efficiently looking at this data was to produce a CSV file for each month of our focus year of 2015. The data file is organized in a way that each row represents one crime, and there are extensive columns associated with each crime providing a great deal of column data and insights into each individual crime. Some examples of column data for each crime are complaint number, complaint date, complaint time, crime description, location of crime, how the law categorized the crime, how the department dealt with the crime, and the borough the crime occurred in. We have decided to focus on a select few of these that best represent our overall message of the visualization. We have utilized the data from the columns of complaint date, location of crime, how the law categorized the crime (misdemeanor, felony, or violation), borough of the crime, latitude, longitude, and the more exact description of location (residence, street, etc.).

In order to reduce lagging with such a large data file in our visualization, we created points on the map visualization where we rendered crimes by day, rather than month, so there are infinitely less points to render in the display. In addition, we rendered the points in batches, which was accomplished through essentially re-writing d3. D3 works by rendering all at once, but this was overridden so that we could render points in batches of any desired size.

Our map visualization looks at the total number of reported crimes in each New York City borough by day. In addition, when a borough is clicked on for a specific date, a user can view specific locations of misdemeanors, felonies, and violations in that borough on the selected date. Each circle located on the map tells the user both the type of crime and the location of the crime (street or residence). Our timeline line graph crimes throughout the year for each borough, with a focus on crimes per month. Here, the focus is on locating at where crimes most often happen and at what time of the year, with the hope of guiding users on when the best time to visit certain parts of New York City would be, and which parts of the city to stay in depending on the time of year someone is visiting. In order to allow for these visualizations, we used shapes files of a boroughs map and community districts map.

### Mapping to Visual Elements

Multiples queues are run simultaneously to cut loading time but since this would be asynchronous, we use multiple queue that trigger the next queue as each finishes. This makes the asynchronous queuing become synchronous.

The main visualization is a line graph. We use a queue to gather the data from each borough. Each borough folder has csv files from 1-12 to indicate months. We collect the data through a nest function and it is reorganized so that it is divided by borough and within it is divided by day. Linear scales were used since the data is simply crimes over time. The x-axis is simply day 1 to day 365 and the y scale is the number of crimes happening per day. We compare the five boroughs. Each borough is assigned a specific color from red, light blue, green, orange and blue. The total number of crimes for each day for a borough are represented as circles and paths connect them.

The second visualization uses a linear scale that affects the darkness of the color of the borough. The map is from another dataset but since it small we use geo.mercator to enlarge the graph. Translate is used to make sure the map is in the center of the svg file.

## The Story

The main purpose of our visualization is to guide users and provide assistance in helping individuals know when and where is it most safe to visit NYC. This is mainly intended for non-residents of the city who might want to know what areas are safer during the time they want to visit. The main line graph tracks crimes across boroughs over time so users may compare where is safer depending on the time. Furthermore, clicking on the line graph affects the second visualization. The map of the boroughs are colored with darker areas corresponding to more crime for the day clicked on the line graph. This makes for an easy comparison of safety among boroughs for a specific date. Overall, the users are presented with a general timeline of where and when is safe. From there, users can zoom in to specific day by comparing easily among the boroughs.

There is a notable increase of crime in the middle of the year which decreases on either side of the year. Staten Island also has significantly less crime than the other boroughs.