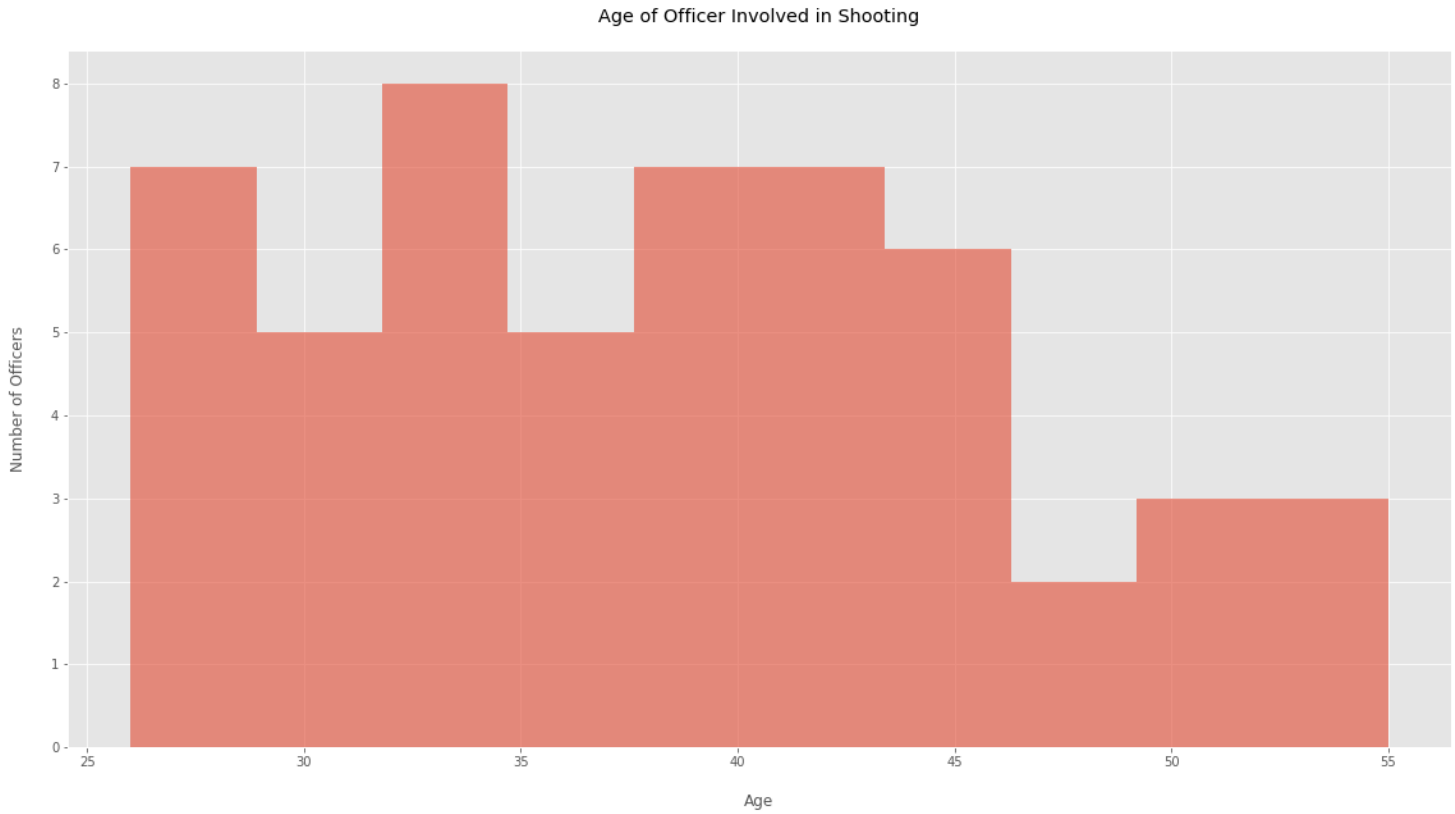


### **Introduction**

I am originally from Denver, Colorado and am planning on moving back there after this program is over, so I thought it would be cool to analyze data about the city. Thankfully, the City and County of Denver has a great website with lots of publicly available datasets. For the final project, I am combining two data sets together. The first dataset has information about “all incidents in which one or more Denver police officers discharged a firearm beginning January 1<sup>st</sup>, 2015.” With all the issues around police brutality right now, I thought this data would be very interesting to dive into. I was also a Criminal Justice major for a while in undergrad, so I am always excited to integrate that passion of mine with data science whenever possible. While this dataset is incredibly interesting, it (thankfully) contains only 90 rows. Therefore, in order to make my visualizations more interesting, I decided to pull in a second dataset which contains all reported crimes in Denver since January 1<sup>st</sup>, 2014. Combining these two datasets together, I’m hoping to answer the following question: “Is there a relationship between the type/location of crimes in Denver and police shootings?” For example, are police officers more likely to shoot their guns in if they’re in an area that typically has a lot of robberies?

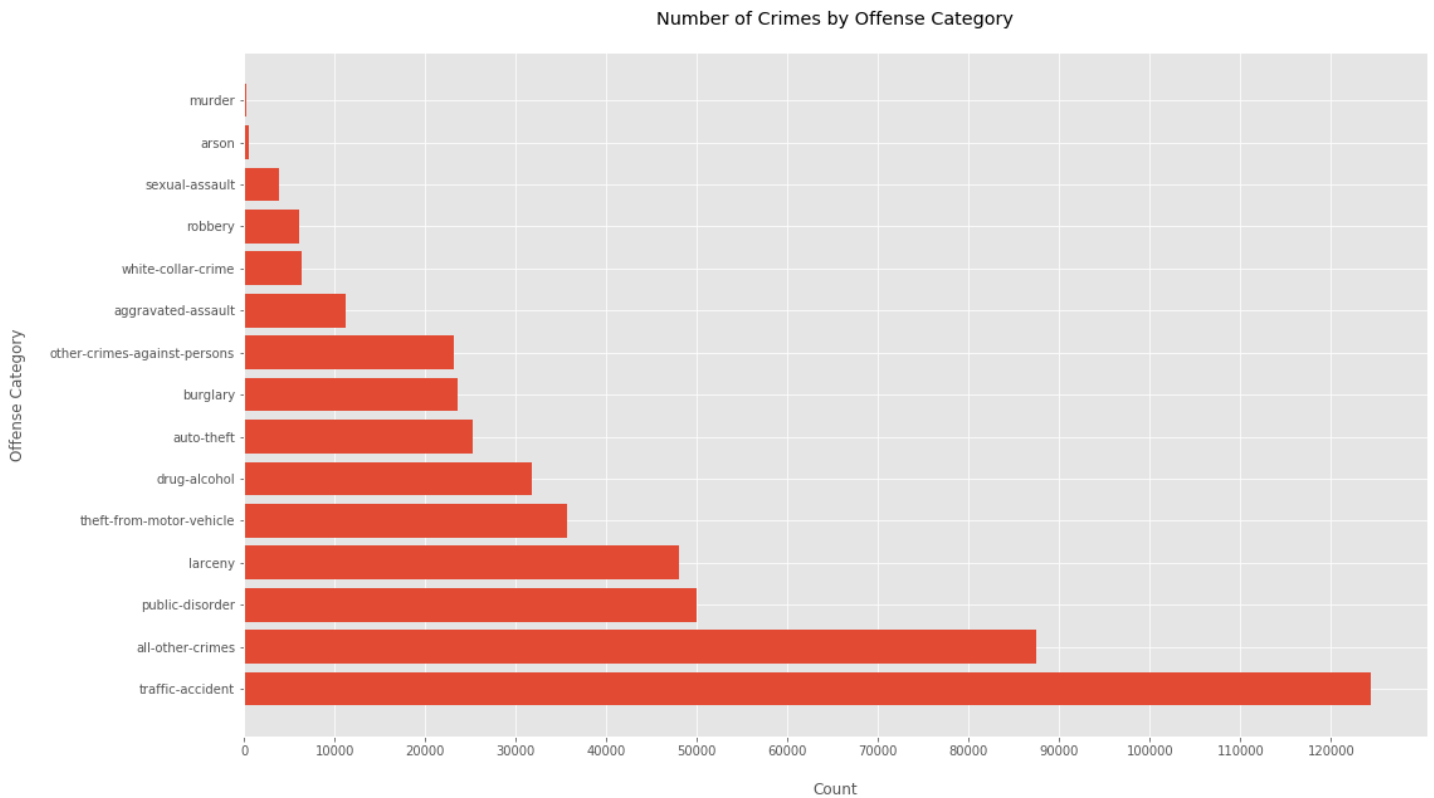
## Summary of Data

### Histogram



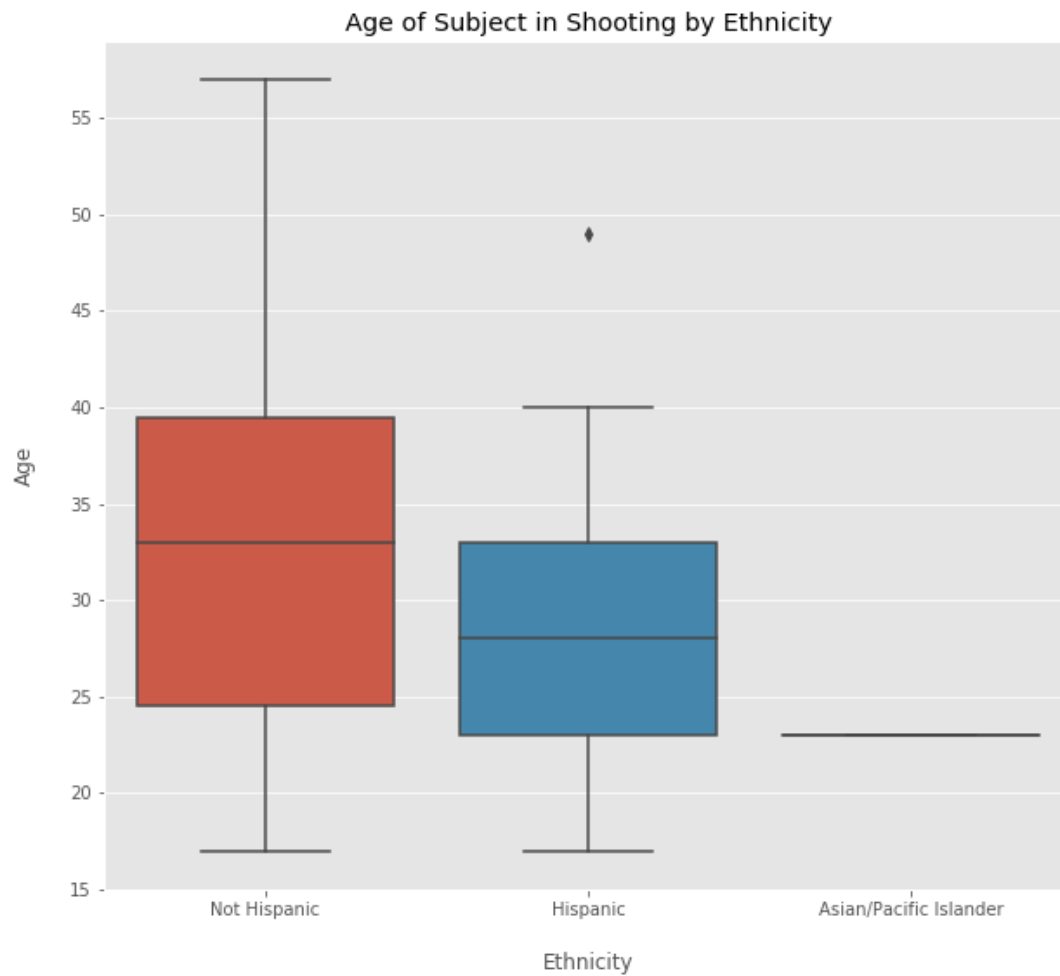
Above is a histogram of the officers' age from the shooting dataset. The officers range from 26 to 55 years old, with the majority of the officers being younger than 46 years old.

## Barplot



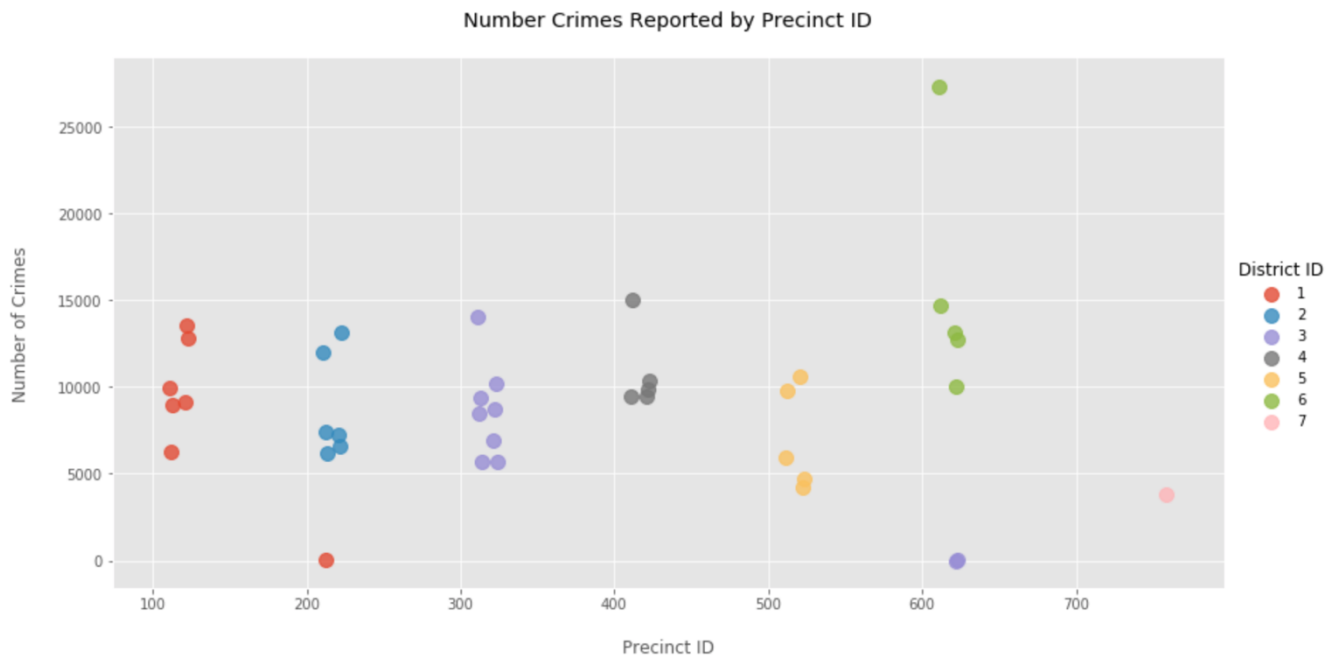
I plotted the counts of a categorical column called “Offense\_Category\_ID”. This column contains the high-level categories of the crimes committed. It is clear from the above graph that the majority of Denver’s crimes are traffic accidents. Thankfully, murder is the offense category with the least number of crimes committed.

## Boxplot



For my boxplot, I looked at the age of the subjects involved in police shootings by their ethnicities. Hispanic subjects' ages are more tightly distributed and are about 5 years younger on average than not Hispanic subjects. There was only one subject whose ethnicity was Asian/Pacific Islander which is why there is just a flat line for that ethnicity.

## Scatterplot

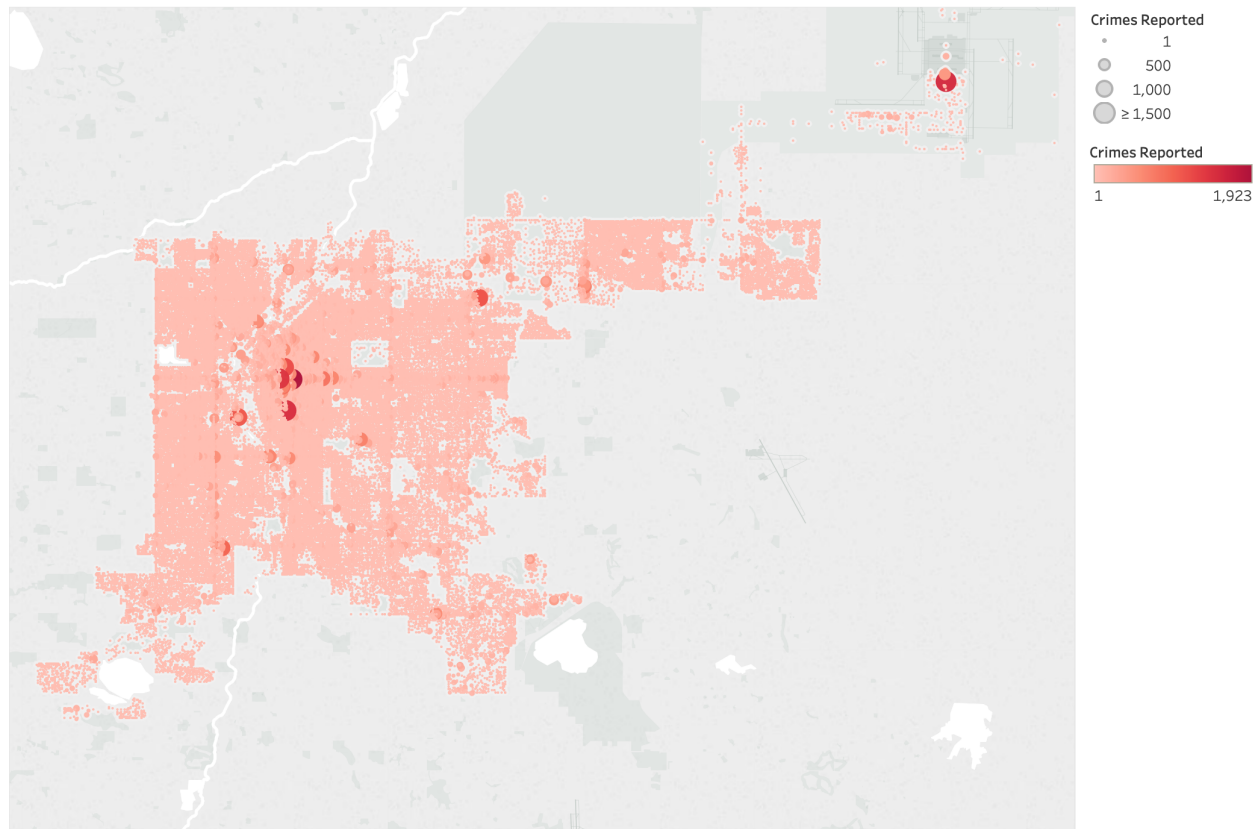


I plotted the number of crimes (not including traffic) by Precinct ID and colored by District ID.

You can see that one of the precincts – 611 from district 6 – has way more crimes reported than any of the others. Also, districts 5 and 7 tend to have fewer numbers of crimes reported at their precincts compared to the other districts.

## Bubble Map

Crime by Location in Denver

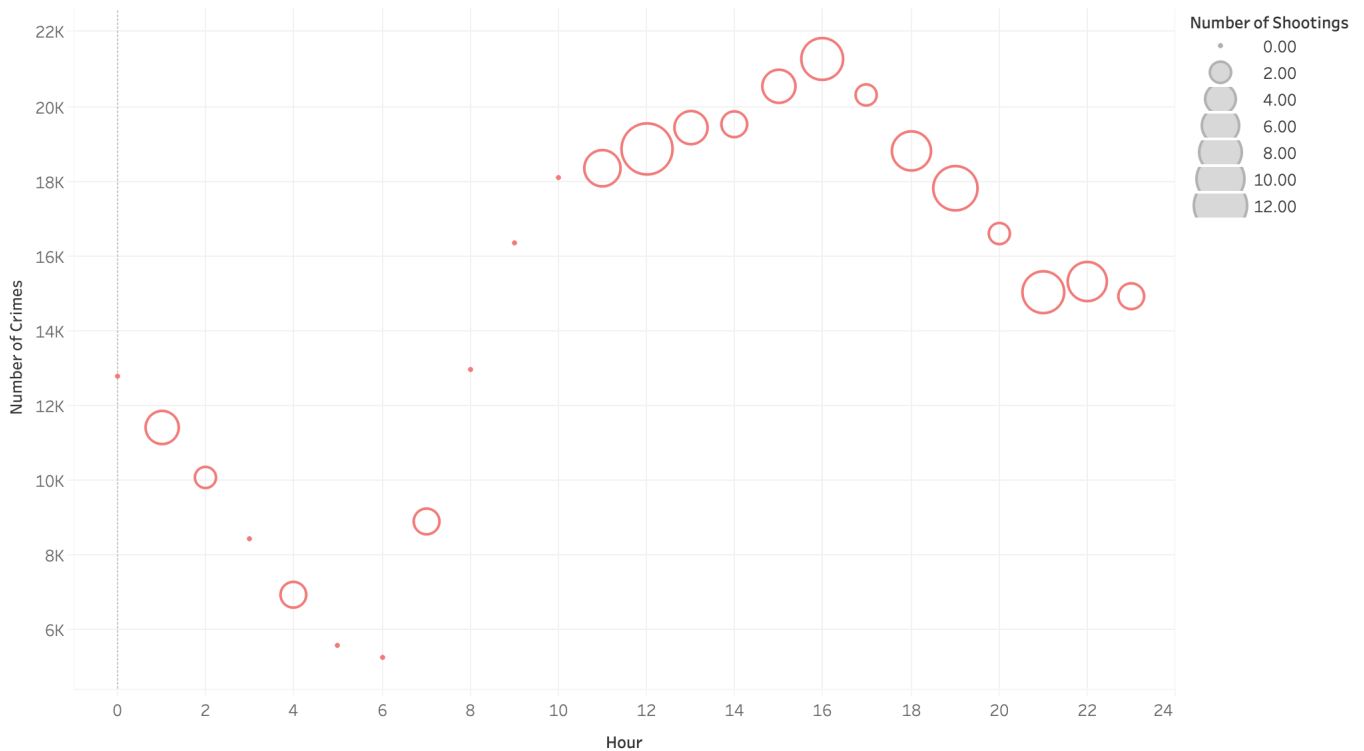


Map based on Geo Lon and Geo Lat. Color shows sum of Number of Records. Size shows sum of Number of Records. The view is filtered on Geo Lon, which keeps non-Null values only.

Since all the points are in the same city/county, I use Latitude and Longitude for my Bubble Map locations. Here we can see that the most crimes (not including traffic) in the city happen where the bubbles are biggest/darkest. This is true near downtown and near Aurora – two places where I would expect a higher number of crimes.

## Bubble Plot

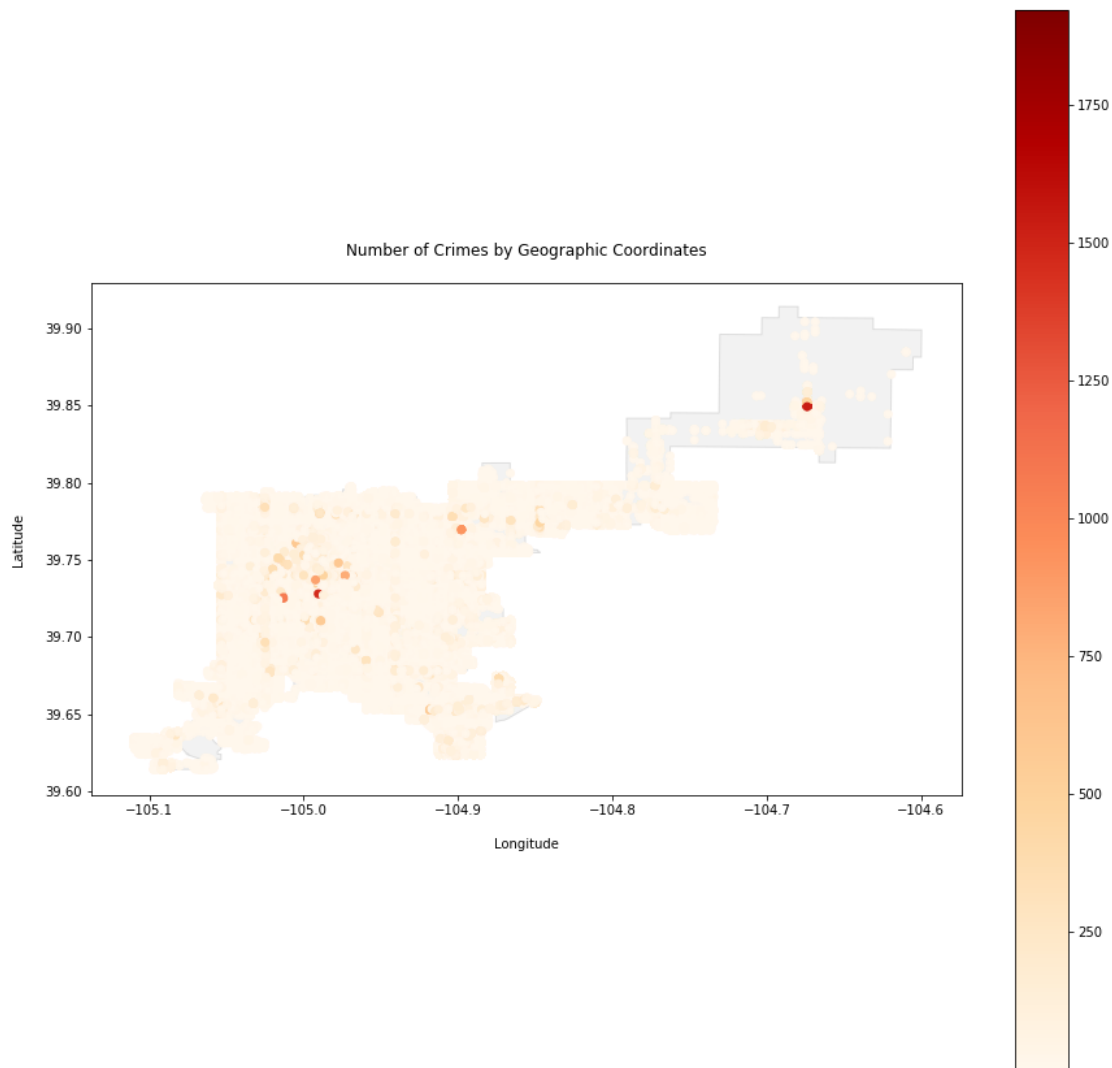
Number of Crimes and Shootings by Hour



Hour vs. Number of Crimes. Size shows details about Number of Shootings.

For the bubble plot, I looked at the number of crimes (not including traffic) by hours, and then made the size of the bubbles the number of shootings at those same hours. There is a spike in crimes in the afternoon and evenings and fewer crime in the early morning. The number of shootings is more evening distributed among the hours, however there are surprisingly no shootings between 8 AM and 10 AM, even though the number of crimes increases at a sharp rate during these hours.

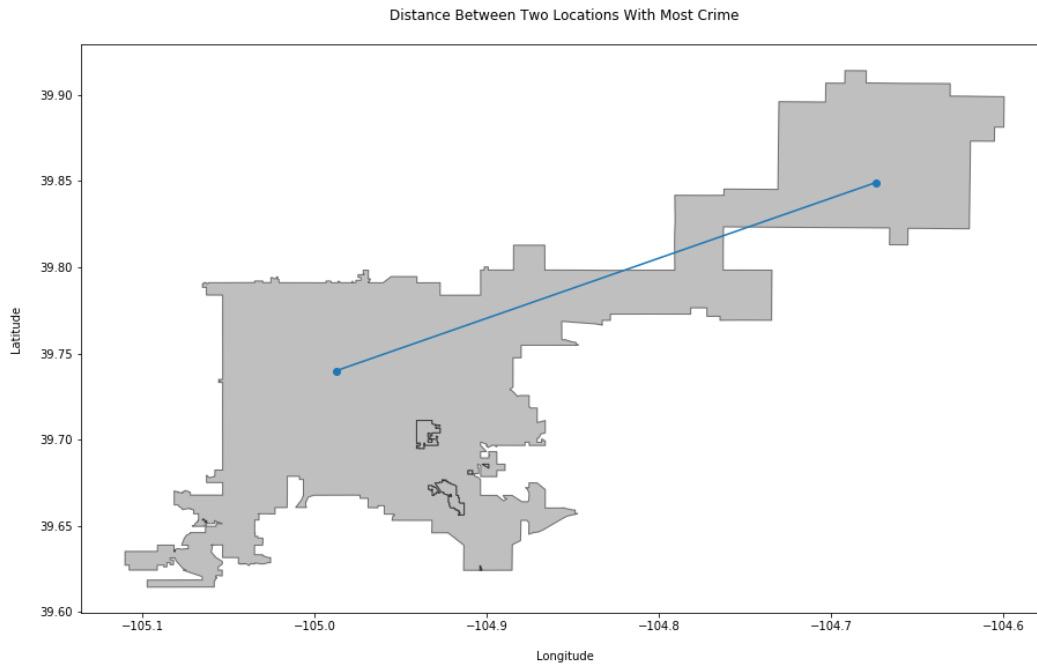
## Choropleth Map



Because the datasets are all within the same city, I decided to do a Choropleth based on the latitude and longitude coordinates. The grey background represents the Denver city limits – which I realized after doing this that Denver is a very weird shape. The darker red the point is, the greater the number of crimes for that location. There are a few hotspots where more than 1,500 crimes have been committed.

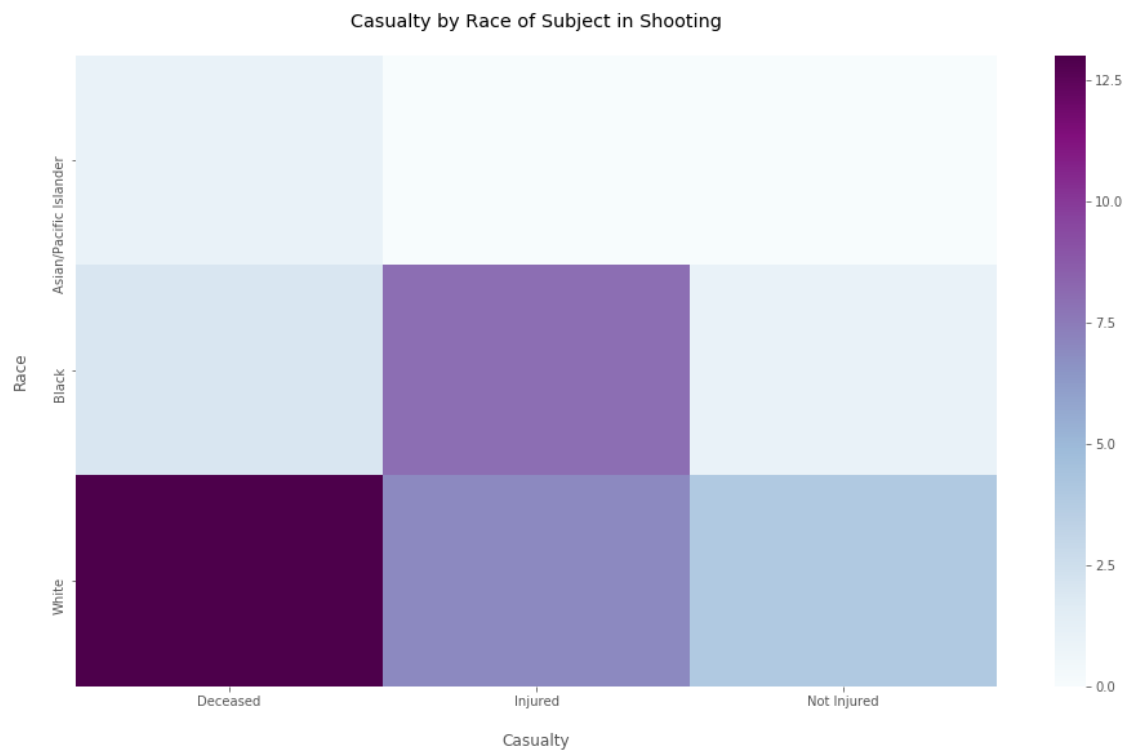


## Connection Map



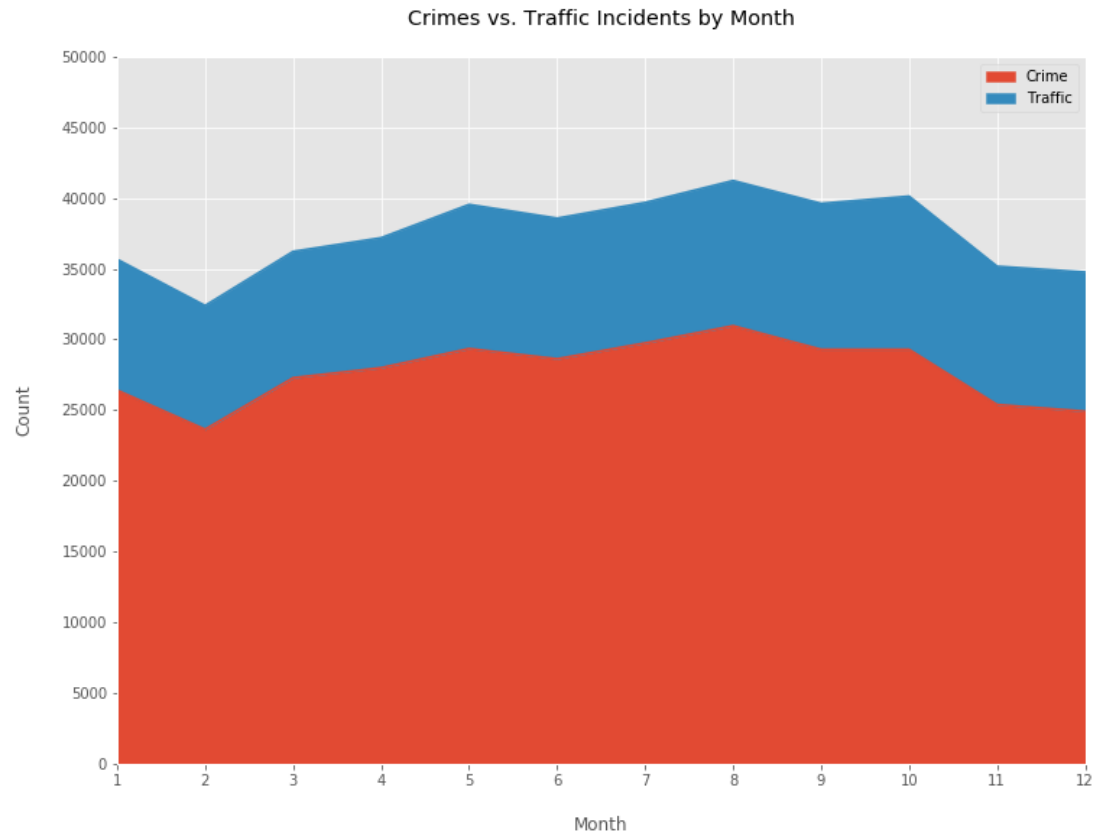
One thing that I thought was interesting is that the two locations (by Latitude/Longitude) with the most crimes are pretty far away from each other. Here I've showed that connection between these two locations with a line connecting the two points.

## Heat Map



In the above heat map, I looked at the relationship between race of the subjects involved in a shooting and the casualty that resulted from it. The majority of the subjects involved in the shootings are White. Out of the White subjects, most of the time the subject dies during the incident with the next highest category being injury.

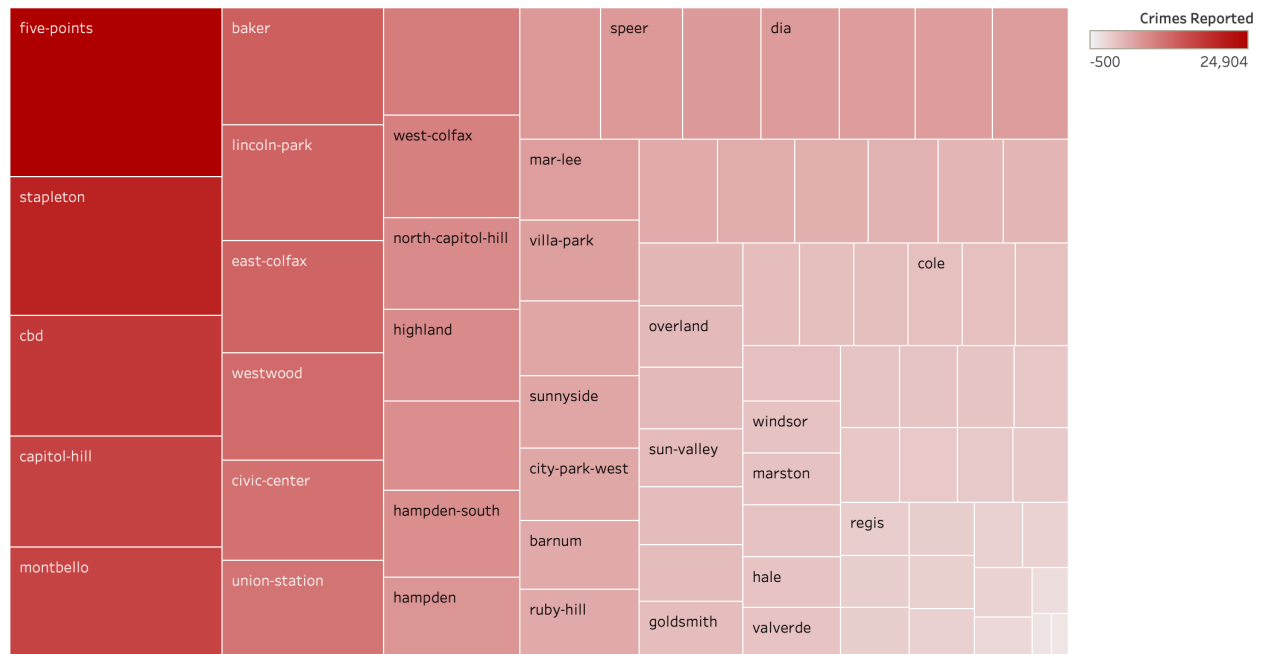
## Stacked area



In the crimes dataset there are columns to indicate whether the incident was a crime or simply traffic. It is important to note that an incident can have the `Offense_Category_ID` "traffic-accident" but be severe enough to where it is actually considered to be a crime. The stacked area above shows that there are almost triple as many crimes as traffic incidents in this dataset. Interestingly, both traffic and crime have the same trend across month; both are typically higher in the warmer months and lower in the snowier months.

## Treemapping

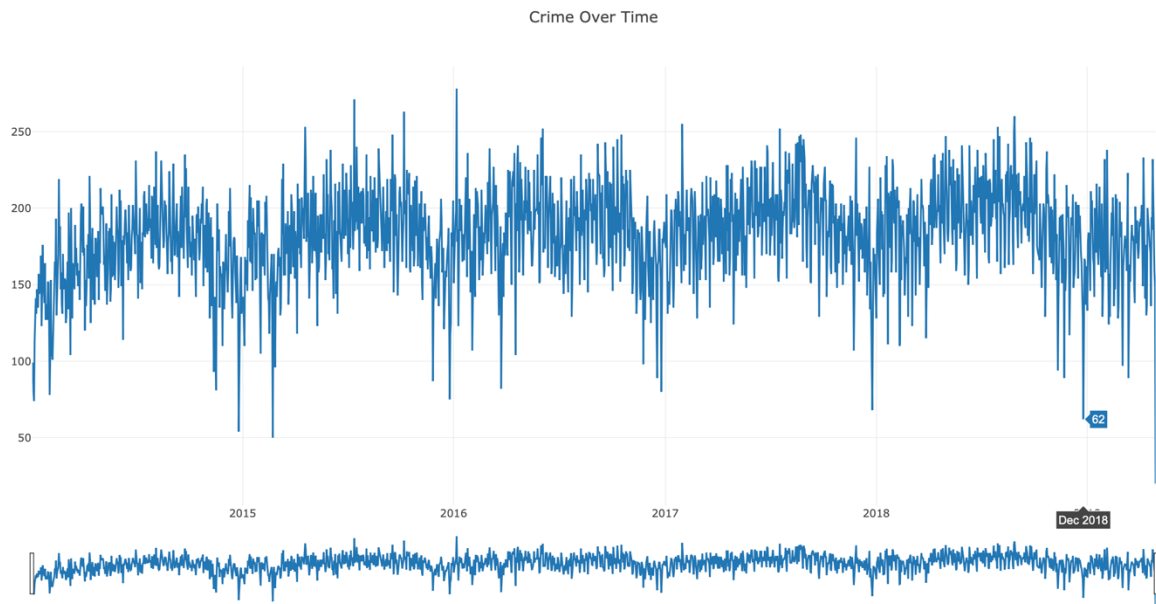
Denver Crimes Reported By Neighborhood



Neighborhood Id. Color shows sum of Number of Records. Size shows sum of Number of Records. The marks are labeled by Neighborhood Id.

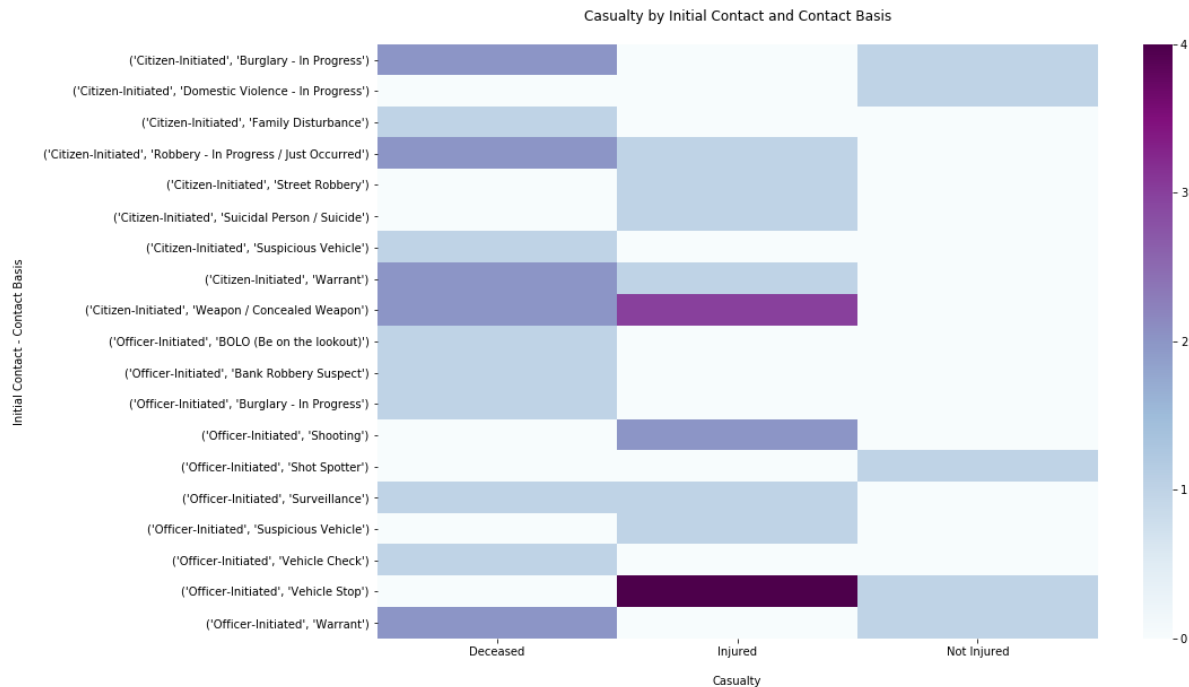
In the above plot, the bigger/darker the square, the more crimes reported in that neighborhood. Five Points is a very popular neighborhood to live in among my friends back home because it's close to the bars and the baseball stadium, so it's interesting that it is the neighborhood with the most crime.

## Interactive plot – Plotly Time Series



To play with the interactive graph, open the file “InteractiveTimeSeries.html”. I used Plotly to create an interactive time series of crimes (not including traffic) by day over time. There appears to be a slight trend through the years where crime goes up and then back down in a slight semi-circle shape. This follows the same general pattern as the stacked area plot above. After playing with the plot, I found that many of the days with very little crimes reported are (not surprisingly) Christmas Day.

## Storyline



One of the most interesting components to study in Police behavior is intent/motivation behind a Police officer's decision to initiate an incident. We hope that the officer did not make a decision due to racism nor prejudice. Earlier in this assignment I addressed this potential motivation when I looked at the relationship between a subject's race and the casualty that resulted from the incident.

In this plot, I dug deeper into the relationship between motivation to initiate contact and the resulting casualty. An incident's initial contact made by either a police officer or a citizen. That initial contact may be reported on one of sixteen different bases. The heatmap above looks at how the different combinations of initial contact and contact basis relate to the three levels of casualty.

There are two really interesting insights that came out of this heatmap. The first is that an Officer Initiated – Vehicle Stop resulted in the most injuries out of any other category while

also having zero deaths. This makes me curious as to why the officer stopped each of these vehicles in the first place. Were they driving suspiciously? Have expired plates? Also why were there so many injuries but no deaths? The second highest combination overall was for the injury casualty by Citizen Initiated – Weapon / Concealed Weapon. This one makes a little more sense to me. If a subject had a weapon and initiated a conflict with an officer, it seems within reason that they may need to injury said subject to protect themselves.

### **Results/Summary/Conclusion**

When I began my analysis on this data, I initially wanted to investigate whether there was a relationship between the type/location of crimes in Denver and police shootings. After working with the data for the past few weeks, I realized this is not the right question to be asking. Thankfully, in Denver there have only been 37 incidents involving the police in which a gun was fired. None of those 37 incidents have occurred in the same location as each other, making it nearly impossible to assess the rate of the shootings for a given area. A better question to ask instead is what caused these shooting incidents. Was it the subject's age/race? Was it the type of crime? Was it due to the time of day/year? These were the questions I ended up investigating with my visualizations. For future analysis, I would like to explore body camera data and the effect it's had on police shootings. Also, if data permits, I would love to expand my analysis to other issues such as police brutality.

## **Code Appendix/GitHub**

All code I used to generate the above plots can be found on my GitHub at the following link:

[https://github.com/nkacirek1/crime\\_visualizations](https://github.com/nkacirek1/crime_visualizations)

## **Sources**

- Crime Dataset - <http://data.opencolorado.org/dataset/city-and-county-of-denver-crime>
- Police Officer Involved Shootings Dataset - <http://data.opencolorado.org/dataset/city-and-county-of-denver-denver-police-officer-involved-shootings>
- <https://python-graph-gallery.com>
- <https://towardsdatascience.com/geopandas-101-plot-any-data-with-a-latitude-and-longitude-on-a-map-98e01944b972>