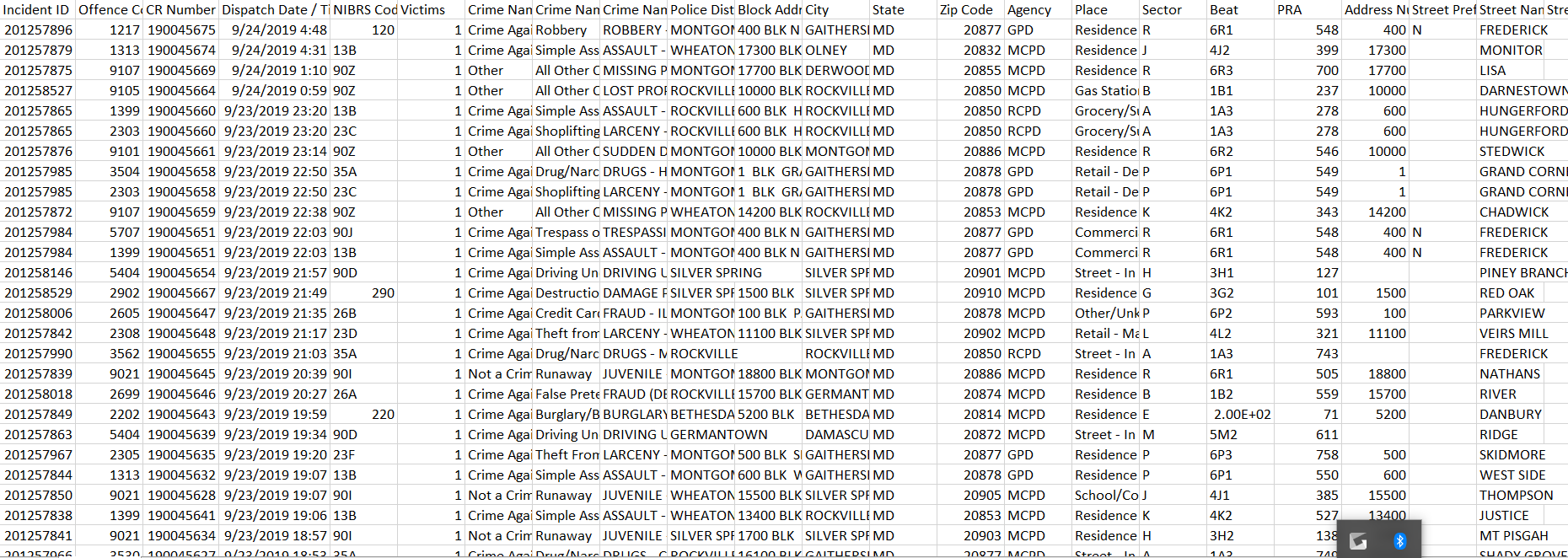
Jeffrey Hou

12/16/19

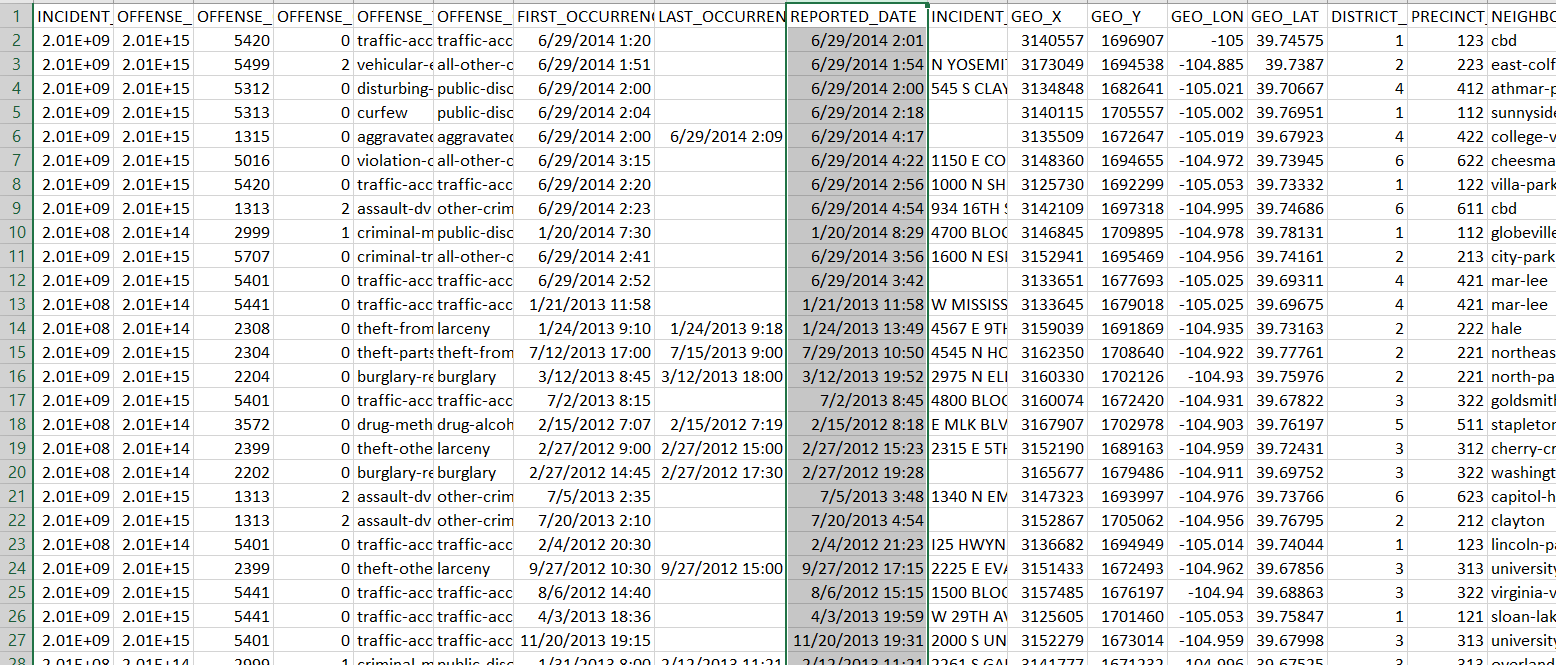
Data 205

The War on Crime

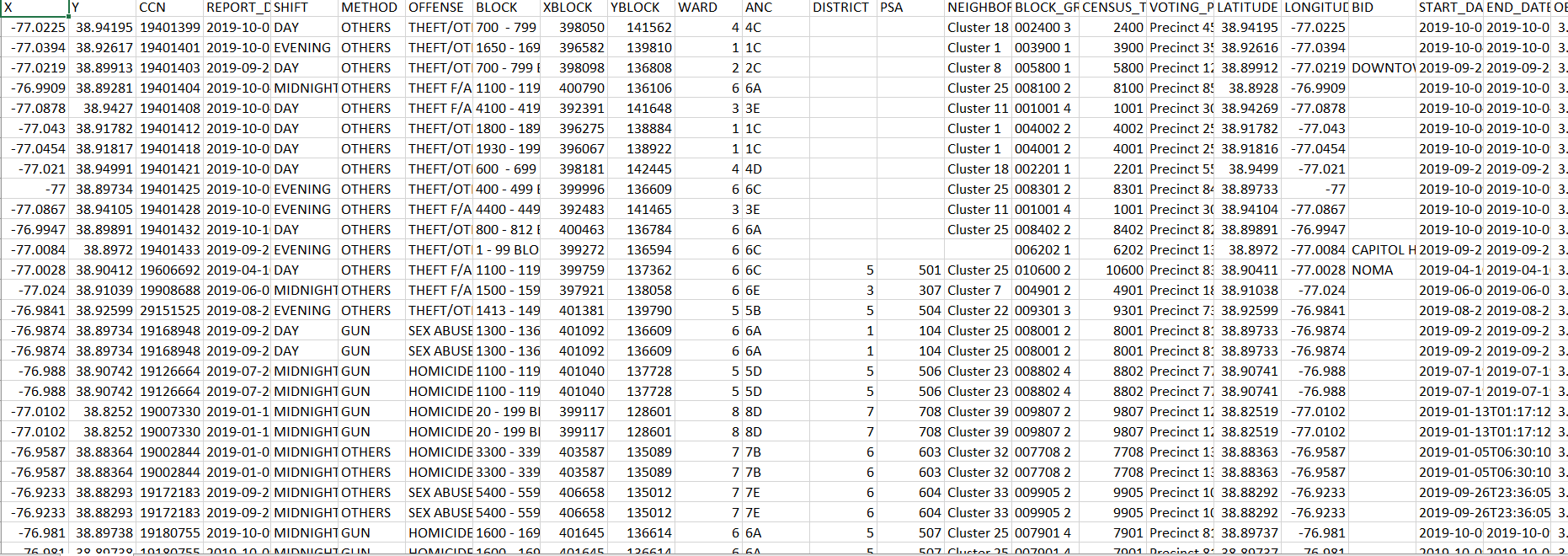
For this capstone project, I wanted to focus on Crime in general and Marijuana and Burglary crimes. I used 3 datasets other than the Data Montgomery dataset. My first dataset was the Crime dataset that was available on Data Montgomery.

This dataset contains 186680 rows and 30 columns. Some of the important columns in this particular dataset would be Crime Name 3 as well as Dispatch Date. These particular 2 columns are important for this project because they allowed me to be able to filter out specific drug names as well as make time-based graphs when needed.

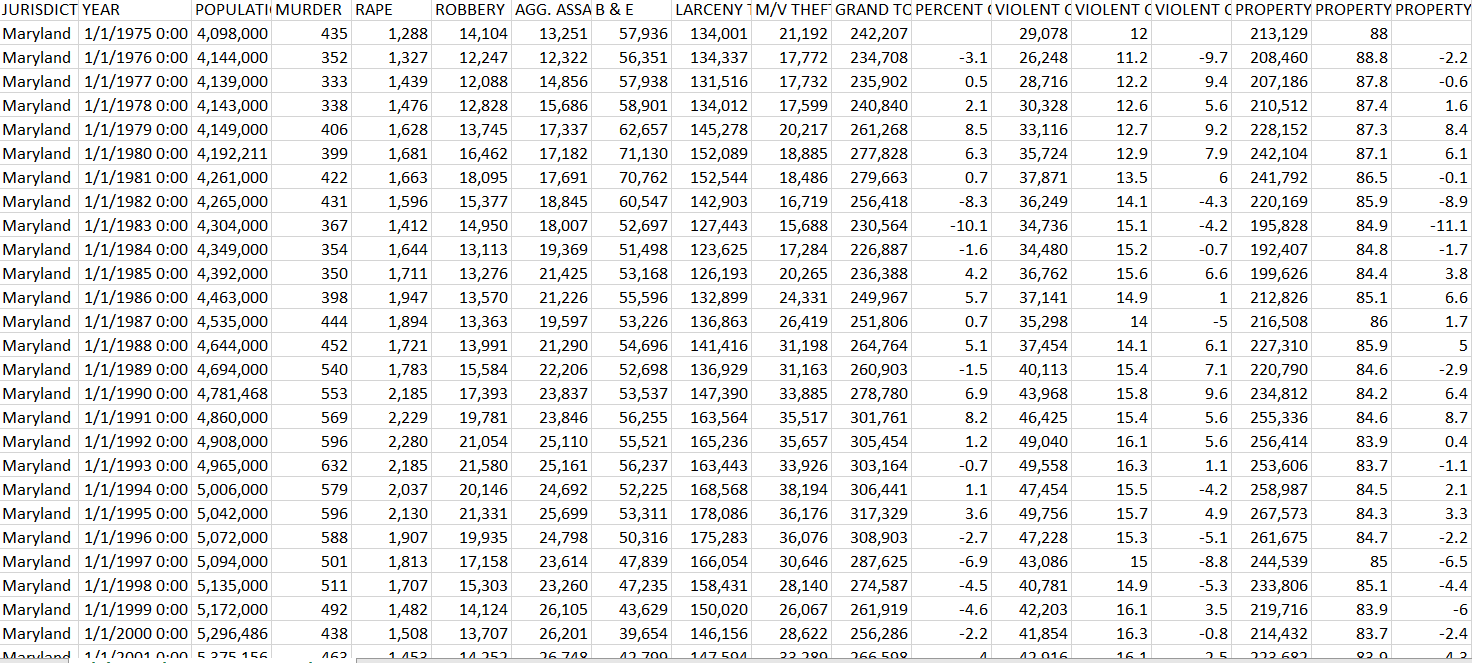
The second dataset I worked with was from Data.world and this was the raw crime data from Denver.



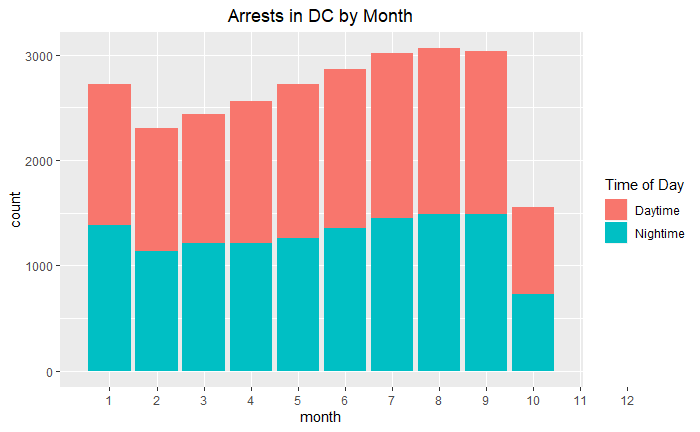
This dataset contained 460912 rows and 19 columns. This dataset represented the criminal offenses in the City and County of Denver for the previous five calendar years plus the current year to date. The data is based on the National Incident Based Reporting System (NIBRS) which includes all victims of person crimes and all crimes within an incident. The important variables to focus on for this dataset is Offense Type and Reported Date.

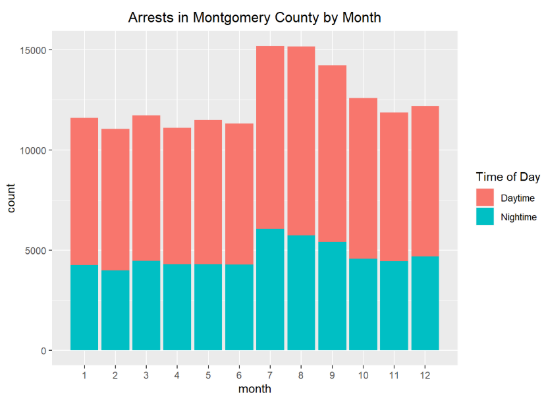
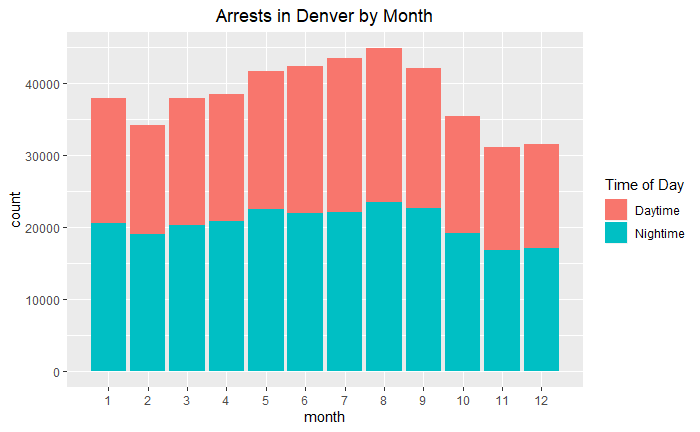
The third dataset I worked with was a DC. Gov dataset that contained crime info from Washington DC.

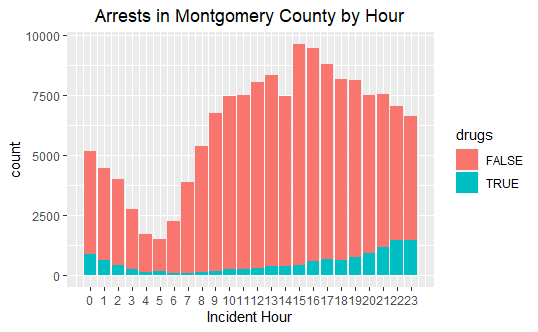
This dataset contains crime locations for the year 2019. This particular dataset has 32417 rows and 25 columns and was one of the smaller datasets that I used. There was a bit of trouble working with this dataset because originally, I wanted to focus on each neighborhood or cluster and see if there were any specific clusters that had increased rates of crime. The only problem was that even after doing research, I had no idea what each cluster in DC was. The important variables for this dataset are Offense and the Longitude and Latitude. Using the longitude and latitude, I was able to make geolocation maps which I was only able to filter using the Offense column.

The last dataset I used for this project was the Md.gov dataset that contained violent Crime & Property Crime Statewide Totals: 1975 to Present.

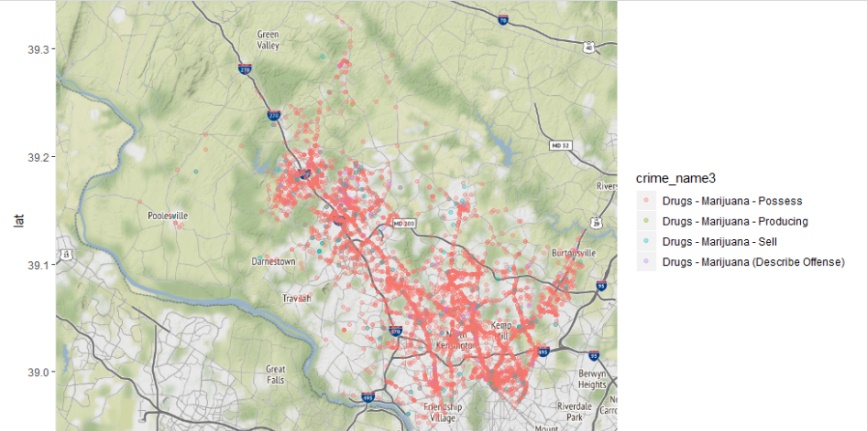
This dataset contained 42 rows and 38 columns. The unique thing about this dataset was the fact that each row represented a year, and the whole dataset was noncategorical variables. They were all statistical numbers and rates of changes.

For the Data Montgomery dataset, I found some key observations that I feel are quite interesting. The most popular crime type 1 was Crime against property. The most popular crime type 2 was Drug and Narcotic Violations. The most popular crime type 3 was Larceny from Auto. The last one was kind of surprising due to the fact I expected the type 3 to be some sort of Drug violation. The majority of the arrests are made by MCPD, which is the Montgomery County Police department. Another important observation I made is the fact that the month July was the most popular month for 3 different categories (General Arrests, Marijuana arrests, and Burglary arrests.

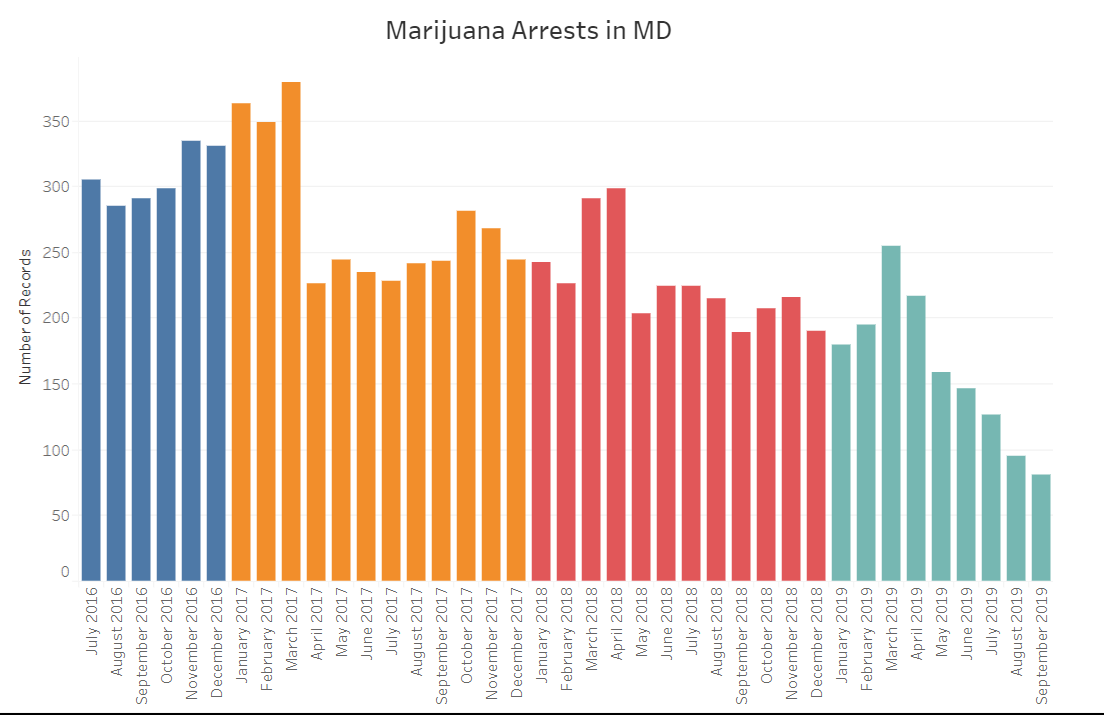


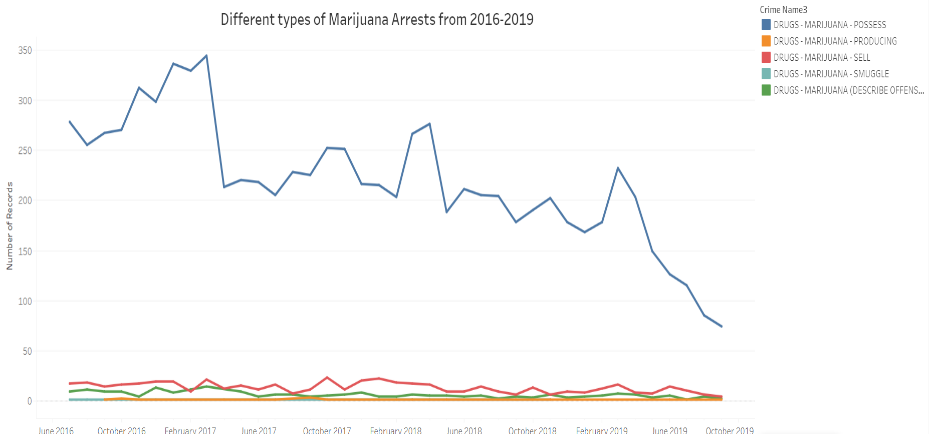
As you can see from the graphs above, the summer months tend to have the most number of crimes. The months July, August, and September are seen to have increased numbers of crimes for all 3 cities. Some reasons I believe this trend occurs is due to the fact that these 3 months tend to have increased temperatures as well as increased daylight hours. The longer daylight hours mean more time out in the public for criminals. This gives them more time to commit crime and therefore these 3 months always has heightened rates of crime.

A second question that arose while I was analyzing the datasets was the fact that drug arrests tend to be more popular at night. As you can see in the graph below, the most popular time for drugs arrests is from 9-11 PM. After thinking about this question, the main reason I can think why this happens is due to the fact that more people tend to use drugs or distribute drugs at night in order to avoid being seen by the police. It is far more likely for someone to be smoking and selling weed when it is dark outside, versus someone selling or smoking in broad daylight. Criminals tend to try to avoid attention to themselves, since they know what they are doing is already illegal. As you can see in graph below, the most times for Arrests in Montgomery County is 3pm-4pm, and this is right when people are getting off work. Very likely a lot of these arrests are on the road, as everyone is trying to get back home quickly after a long day of work.



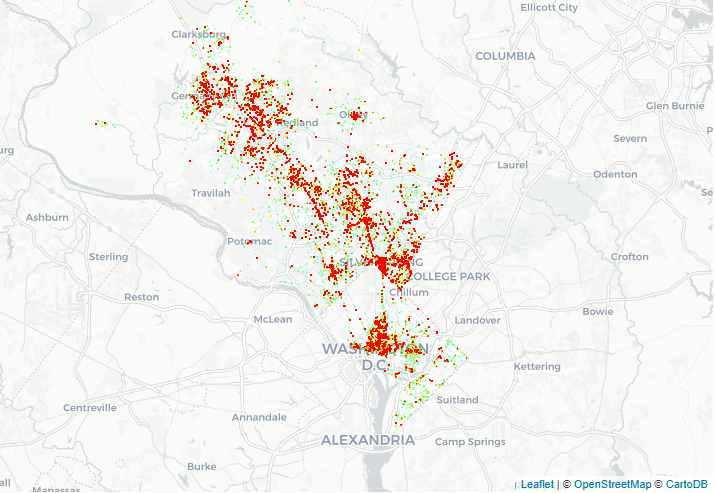
The majority of drug arrests in Montgomery County all tend to follow one path, and that is along I270. By looking at the graphic above, you can see the red dots all tend to follow the highway. Also, the majority of the marijuana arrests are possession. There are a couple reasons why I believe this is the case. Most marijuana possession charges are due to the fact that a driver is being pulled over. Very rarely does a cop pull a cop specifically for drugs. Most of the time the cop is just doing a routine traffic stop and find something suspicious and therefore begin to search the car for narcotics. Another important observation I made was that the majority of the Marijuana arrests happen near Silver Spring. This is not surprising since Silver Spring is also the city with the highest number of crimes. If MCPD could focus more officers near I270 alongside Silver Spring, this could prove to be beneficial to cleaning up the city and decreasing the number of narcotics that get around.

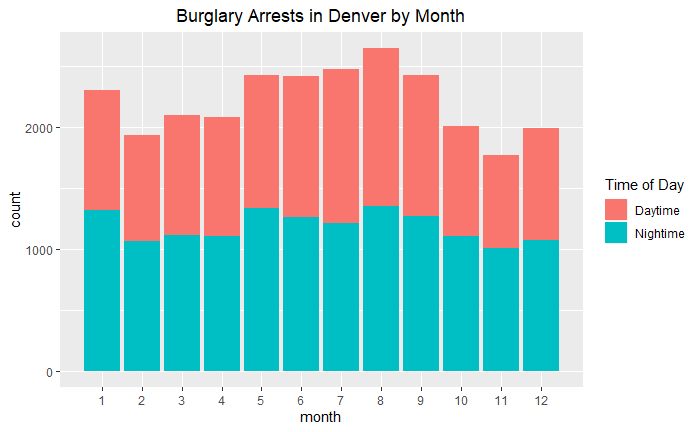


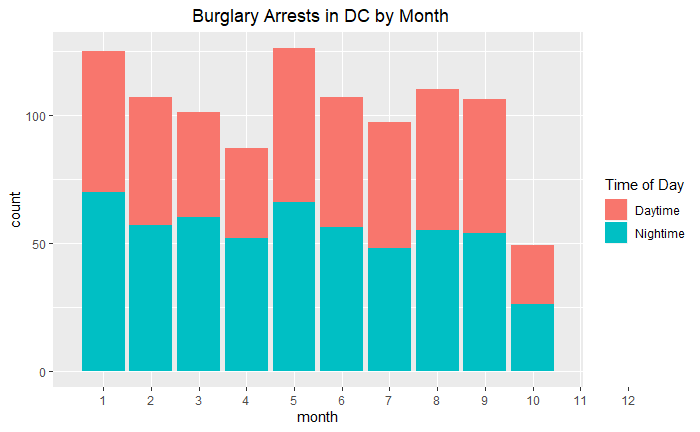
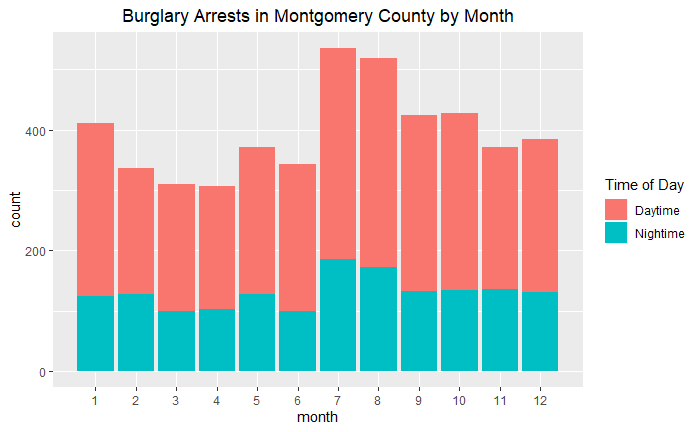
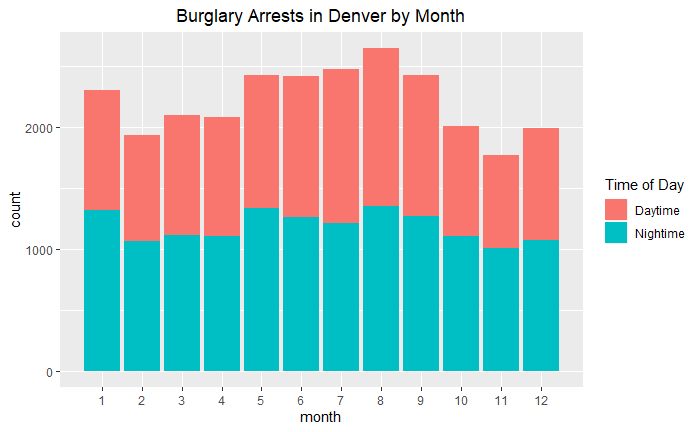


One positive thing about Marijuana in Montgomery County is the fact that it has been steadily decreasing since 2016. As you can see from the graph above to the left, from 2016 it was steadily rising till the beginning of the year of 2017. Ever since then, the amount of Marijuana arrests has been decreasing at a consistent rate. This is possibly due to the fact that Maryland has decriminalized Marijuana as a drug which means fewer illegal arrests are being made. You can also see on the graph to the right, the number of possession charges has dropped quite a bit since the beginning of the year this year to current. This shows that MCPD has done a great job cleaning our streets and getting rid of our drug problem in Montgomery County.

Looking at the DC dataset, there was some key takeaways that I think are quite interesting. The top offense for DC was thefts. For a city that lives so close to each other as apartment buildings are right next to each other, there is no surprise that theft is so high. The most popular month for crimes for DC was August, which coincides with our previous observations that the summer months tend to have the highest rates of crime. For DC however, it follows a pattern. There are steady increases in arrests from February till September, and the peak of the arrests is August.

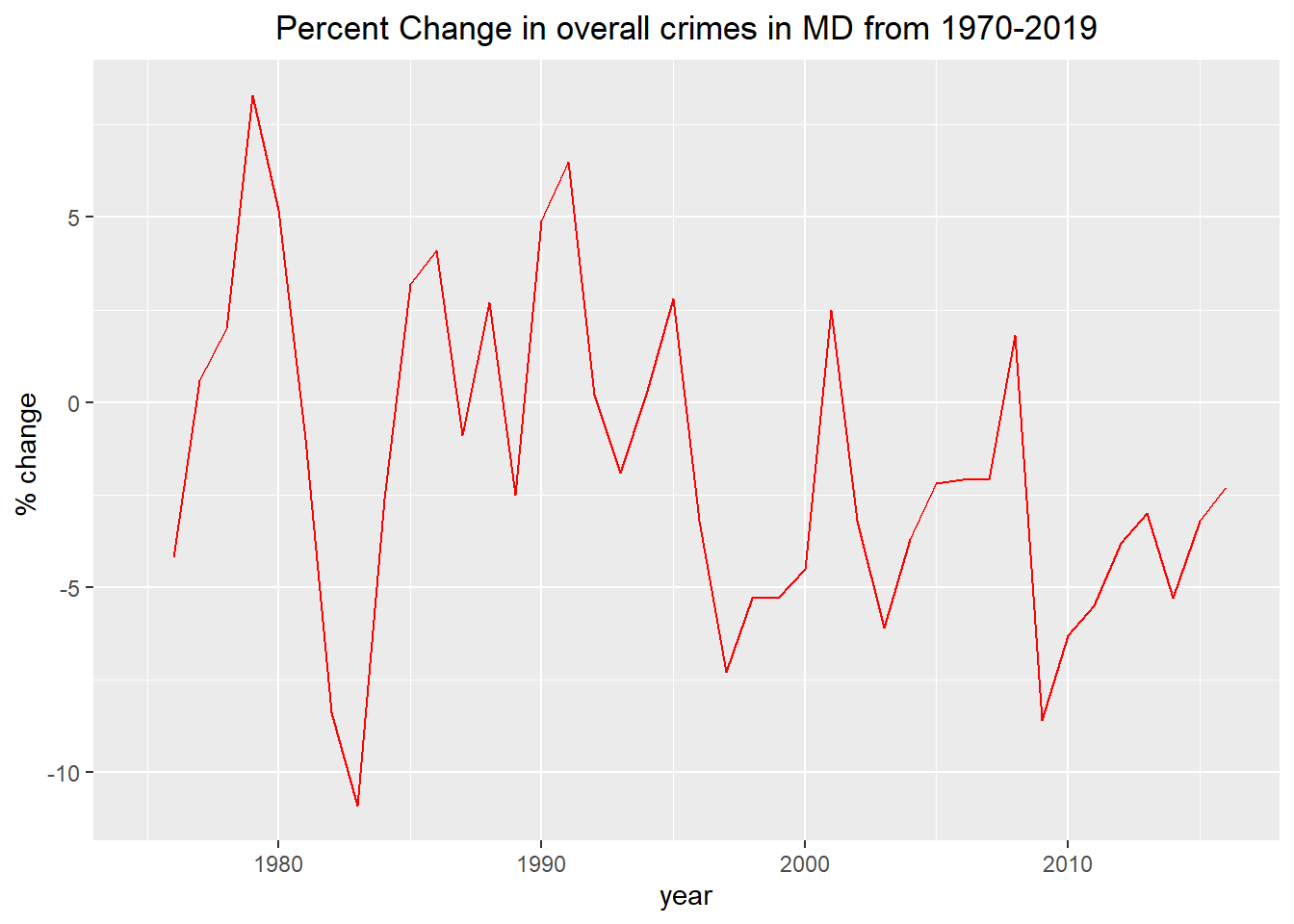
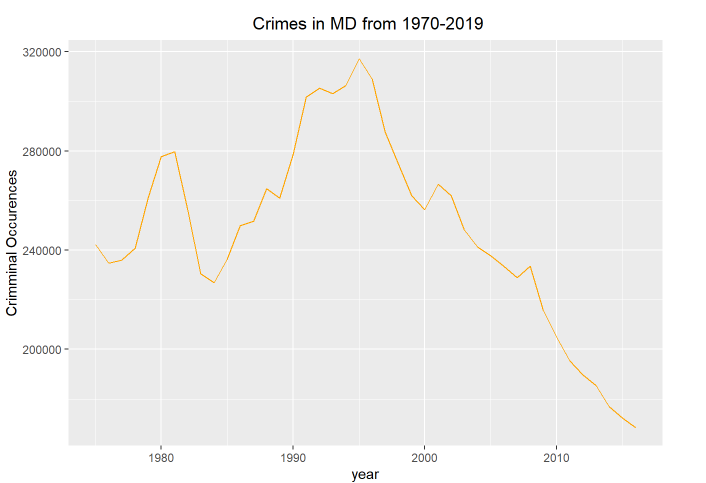


Comparing both DC and Maryland, you can see a couple patterns arise in the heat map I made below. First off, you can see the majority of the crimes that happen in MD tend to happen near the College Park area. This is to be expected as that area has one the highest population density in MD. That means more people in one area which increases the amount of crimes that can be committed. Having a bigger population also means more crimes being committed daily. The second pattern you can see, is the fact that between the border of MD and DC, there is a big chunk of white area that is shown. That means all the crime happens at the borders, but once people get into DC, the crimes tend to stop until they reach Northwest DC. As you can see, there is a big patch of red near Massachusetts Ave which shows that an abnormally high rate of crime being committed there. A suggestion I would give the MCPD would be to increase police activity near the border of DC and MD. The fact that marijuana is illegal in MD, but legal in DC can cause a lot of mistakes by criminals. The second suspected drug users are inside MD borders; they can be apprehended quickly.

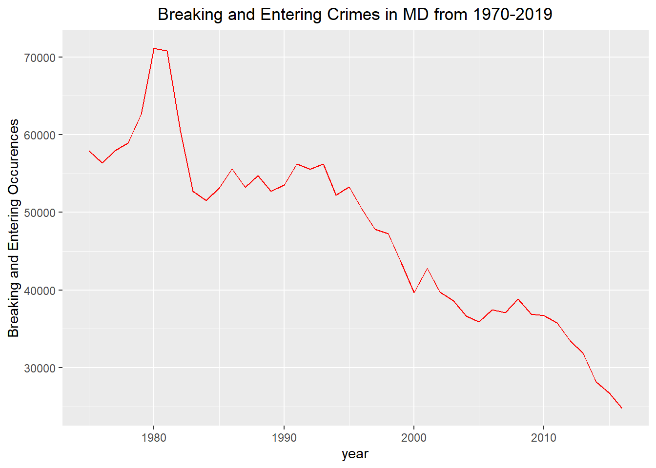


Taking a good look at Burglary arrests for the 3 different cities, you can see patterns emerge. For example, you can see that DC is the only city that follows 3 month intervals. Every 3 months, there tends to be a peak and then the next 2 months is slowly decreasing. For both Montgomery County and Denver, they tend to just have peaks at July and August but the rest of the year it is pretty steady.

Denver as a city is no stranger when it comes to crimes. The fact that they made Marijuana legal has done wonders for the city. Not only has the sales helped the state in general, but drug crimes as a whole has been decreasing. The top crime type for Denver is traffic accidents. Something interesting to note is that the fact many states with legalized marijuana are among the top in car accidents. This could be due to the fact that many drivers are driving high or under the influence when behind the wheel. One interesting thing about Denver is that there are more nighttime arrests compared to daytime. There are also more nighttime burglary arrests compared to daytime. When there are people at every corner where you walk, it is hard to burglarize a house in broad daylight compared to at night when you can barely see if anyone is walking in front of you.

The last dataset I worked on was the Maryland Violent Crime dataset. The main takeaway from this dataset was the fact that you can see the massive decrease in crimes since 1995. It has been decreasing at a consistent rate. It is currently at the lowest its been in 40 years, and this is all due to our wonderful police force.

As you can see from the graph above, there is a sharp decrease in crimes in MD from 1980-1983, almost a 20 percent difference in 5 years.

Breaking and entering crimes in MD are no exception to this trend as they also have been decreasing at an alarming rate since 1980. Some explanations for this can possibly be due to the fact of the Castle Doctrine, which allowed deadly force when in self-defense, which ultimately decreased the amount of people trying to break into other people’s homes. People aren’t willing to risk their lives to burglarize knowing that the homeowner is legal to take deadly force if they wanted to.

Using this data Montgomery dataset was an absolute joy, as it was also one of the cleaner datasets out of the 4 that I used. Some recommendations for this dataset I would say would be to include gender, race, age in the Arrest info. This would not only allow me to be able to make more detailed analyzations but also be able to filter based on race and age and see if that affected the person and the likelihood they got arrested. Another recommendation would be to include information such as if the person was jailed or not. With this information, I would be able to utilize machine learning and possibly predict if this specific crime would have jail time included or not. The last thing I would like to do with this dataset is to compare it to a city similar to Montgomery County, and see if the patterns I found are in line with smaller cities. Being able to predict pockets of crime in Montgomery County would be beneficial to the police force and help keep crime and narcotics of this beautiful country of ours.

Lastly, I would like to thank Dennis Linders for providing us with the direction and oversight of our projects! I would also like to thank Montgomery County Government for providing us these great resources to work on our projects. Last but not least, thank you to Victoria Lewis and data Montgomery for allowing us to use with their wonderful datasets and allowing us to analyze them!