**Final Project Report**

**DATA 205**

**Hye Young Park**

**May 13, 2020**

Crime Density by City in Montgomery County, MD

This journey started when I found the dataset that contained police dispatched incidents on dataMontgomery. The dataset was very thorough and contained a lot of information in it, so I knew it had to be one of the datasets in my project. Also, I always wanted to know if there was a (negative) correlation between crime rate in a city and its’ median house price value. There’s always a thought in people’s minds that neighborhoods with higher median house prices have less crime, so I wanted to test that theory. Like that, my first hypothesis for this project was born – is there a negative correlation between median house price in a city and the crime rate in that city? The second hypothesis was born in a brainstorming session later, when I noticed that the police dispatch dataset contained dispatch calls, time of arrival, time of close of the case, etc. I started wondering if there was a certain “peak” time for certain types of crime, and that is how I came to my second hypothesis – when do most crimes happen and are there any trends in that?

A close up of text on a white background

Description automatically generated When I started out, I began working with the two datasets I mentioned, police dispatched incidents(Table 1) and median house price by city (Table 2).

Table 2

Table 1

A screenshot of a cell phone

Description automatically generatedA close up of text on a white background

Description automatically generated After some trial and error and manipulation of the data, there were results, I could visualize the data and what I wanted to see, the amount of crimes by city(Figure 1). I then quickly realized that different cities will have vastly different crime amounts due to the size of that city. That is when I needed to find a dataset that contained the population(Table 3) of those cities so that we could get to the most important variable of this project, the rate of crime in that city. The rate of crime is the total amount of crimes committed divided by the population in that city(Figure 2). Throughout this phase of the project, I found that the rate of crime was vastly different between cities. I’m sure there are many factors that play a role, but I had a hunch that the median house price value was one of those factors. Now that we A screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generatedhad our most important variables, we could begin answering our hypotheses.

Table 3

Figure 2

Figure 1

A screenshot of a cell phone

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generated

Table 4

Figure 5

Figure 4

Figure 3

A close up of a map

Description automatically generatedThe first hypothesis – negative correlation – required us to look at this data and find a way to visualize the correlation between the different datasets. I began to merge the datasets and remove unnecessary data from them until we had the two most important columns, median house value in a city and the associated rate of crime in that city(Table 4). The correlation between these for 2017, 2018, and 2019 was surprisingly close, but surprisingly not very big, averaging -0.15 between the three years. This is far from proof of the hypothesis, but I feel like it’s leaning in the right direction and could perhaps be refined with further study/work on these datasets(Figure3, 4, 5).

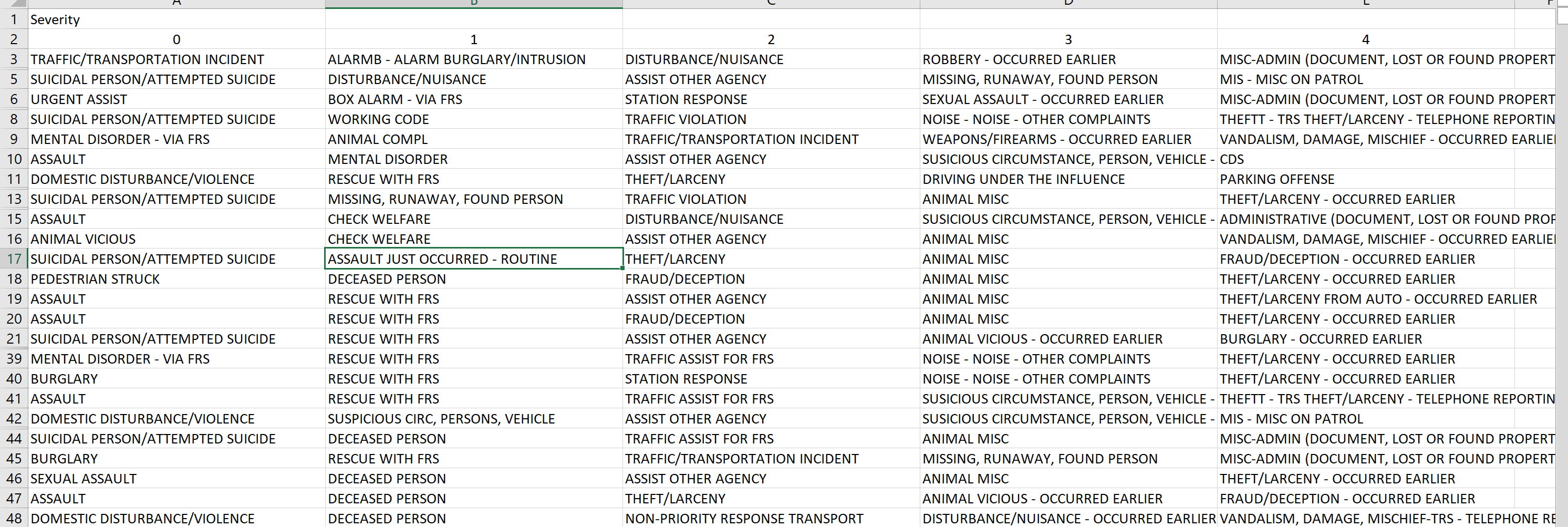
The second hypothesis is where a lot of the time was spent, working with the data and finding appropriate visualizations for it. There were several iterations of these visualizations(figures 6-13), but in the end, I felt that a heat map visualization showed best what I wanted to show(figure 6, 7, 8) - the rate of crime and the hour of the crime occurrence. As you can see, most of the crime does occur in the evenings to late night hours. Crimes of severity 1, 2, and 4 seem to occur for most of the day, peaking in the evening/nighttime. Crimes are in order of severity, 0 being most severe, and 4 being least severe(Table 5).

Table 5

A screenshot of a cell phone

Description automatically generatedA screenshot of text

Description automatically generated

Figure 7

Figure 6

A close up of text on a white background

Description automatically generatedA close up of a map

Description automatically generated

Figure 9

Figure 8

A close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generated

Figure 14

Figure 13

Figure 12

Figure 11

Figure 10

Further looks at this data could potentially reveal more insights. If I had more time, I would look at trends between years and maybe look deeper into the connection between median house price and different severities of crime and connect that with the hour of crime occurrence. There are potential insights there that could lead to increased patrolling of different areas at different times as well as being on the lookout for specific types of crime at specific times of day. One other area of potential information would be using 2020 census population data as that would provide more accurate information whenever it is available.

**Acknowlegement:**

Thanks to Dennis Linders, Montgomery County Government, and Victoria Lewis, dataMontgomery for the support and direction!!

**Data source information:**

1. Median House Price from Zillow

<https://www.zillow.com/research/data/>

housing data: home values – Data type (ZHVI All home (SFR, Condos/ co-op)Time series) / Geography (city): by city

1. Police Dispatched Incidents from dataMontgomery

<https://data.montgomerycountymd.gov/browse?q=Police%20Dispatched%20Incidents&sortBy=relevance>

1. Population from CensusViewer

Population in USA by Cities: Annual Estimates of the Resident Population for Incorporated Places: April 1, 2010 to July 1, 2018

<https://www.census.gov/data/tables/time-series/demo/popest/2010s-total-cities-and-towns.html>

Silver Spring – File name: Quickfacts April 13 20

<https://www.census.gov/quickfacts/fact/table/MD,silverspringcdpmaryland,US/PST045218>