

Assignment: Project 1

Class: DATA 601

Student Name: Nasim Aalemi

Data: 11/17/2023

**I. Introduction:** In this 6 years period crime dataset of Baltimore City I have analyzed, the overall crime rate for the entire City of Baltimore, the top 5 neighborhoods that have had the highest increase of violent crimes, and the top 5 neighborhoods that have seen the greatest reduction in overall violent crime rates per instructions provided for the project 1 assignment. Furthermore, in this analysis I have focused on three types of crimes which includes, Aggravated Assault, Homicide, and Shooting according to the assignment guidelines. The details of methodology, findings, and tools I have utilized to conduct this analysis are detailed below.

## II. Data Collection and Preprocessing Methods:

- a. **Data Source:** The data source for this project is the Open Baltimore portal. It is a dataset of crimes that have been compiled for the City of Baltimore.
- b. **Time Frame:** The time frame for this analysis is 6 years, 2017-2022.
- c. **Data Cleaning:** In order to analyze the data properly, I have taken the following steps to clean and prepare the data using Python and Microsoft Power Bi. First of all, I imported the necessary libraries including Panda, NumPy, and Matplotlib onto Google Colab platform. I then studied the data using the following Python syntax outlined in green color font.

```
crimes= pd.read_csv('https://raw.githubusercontent.com/naalemi/MyFirstRepo/')
crimes.head(5)
crimes.tail(2)
crimes.info()
crimes.describe()
crimes.isnull()
```

I then created or extracted an additional column from CrimeDateTime and named it 'Year' by using the following syntax.

```
crimes['Year'] = pd.to_datetime(crimes['CrimeDateTime']).dt.year
```

Given that I had downloaded additional columns with the raw data from the Open Baltimore portal, I used the following syntax to drop those unnecessary columns.

```
crimes.drop('New_District', axis=1, inplace=True)
crimes.drop('Ethnicity', axis=1, inplace=True)
crimes.drop('Age', axis=1, inplace=True)
crimes.drop('Gender', axis=1, inplace=True)
crimes.drop('PremiseType', axis=1, inplace=True)
crimes.drop('Inside_Outside', axis=1, inplace=True)
```

Next, I performed the following to drop all empty rows from the dataset.

```
crimes.dropna(inplace=True)
```

Lastly, I saved the result of the Crime dataset into an excel file and downloaded it to conduct further analysis using Microsoft Power Bi.

```
crimes.to_excel('data.xlsx', index=False)
from google.colab import files
files.download('data.xlsx')
```

**Overall Change in Crime Rates for Baltimore City:** I used Power Bi to calculate the overall change in crime rates for the entire City of Baltimore. For this calculation, I used DAX syntax as following:

The Overall Change in Crime Rates =

```
VAR __BASELINE_VALUE =
```

```

    CALCULATE (
        COUNTA('BaltCityCrimeData'[Total_Incidents]),
        'BaltCityCrimeData'[Year] IN { 2017 }
    )
VAR __MEASURE_VALUE =
    CALCULATE (
        COUNTA('BaltCityCrimeData'[Total_Incidents]),
        'BaltCityCrimeData'[Year] IN { 2022 }
    )
RETURN
    IF (
        NOT ISBLANK (__MEASURE_VALUE),
        DIVIDE (__MEASURE_VALUE - __BASELINE_VALUE, __BASELINE_VALUE)
    )

```

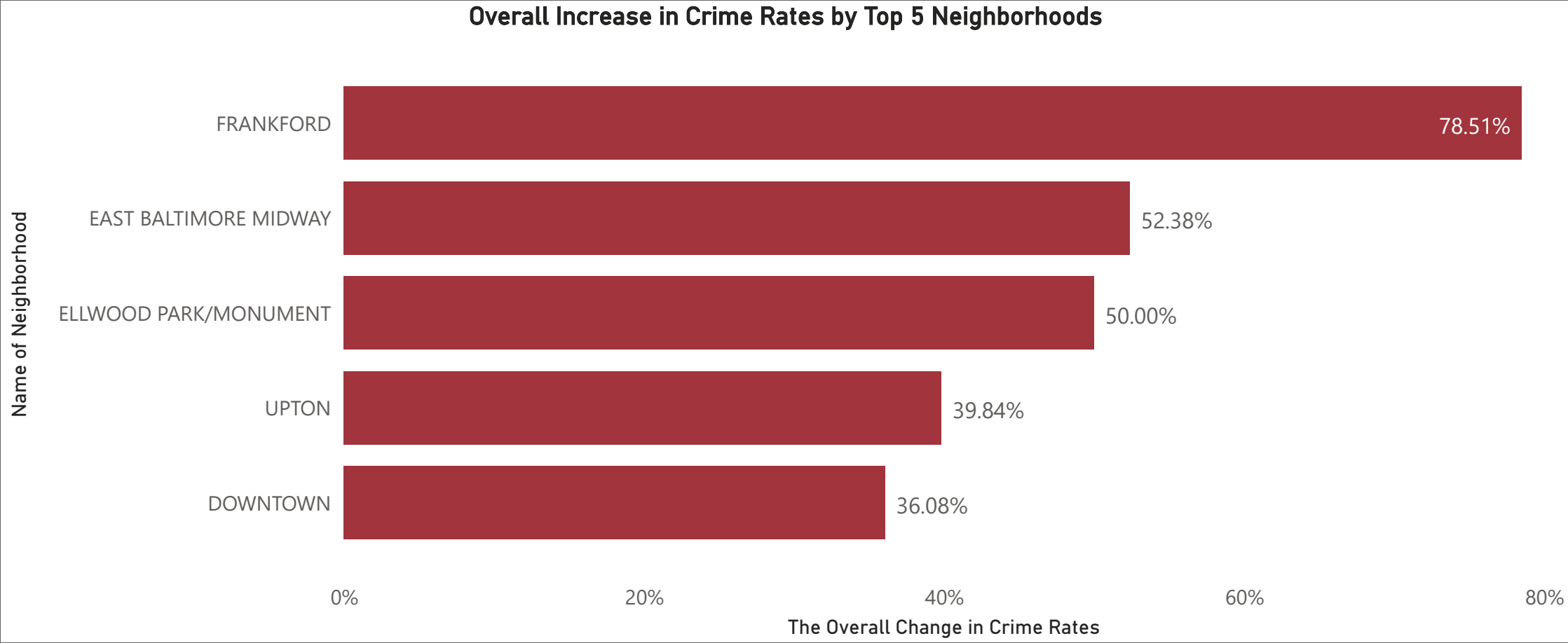
Overall Crime Rates for the Entire City of Baltimore

# 3.25%

The Overall Change in Crime Rates for Baltimore City shows a 3.25% percent increase in crimes.

**III. Top 5 Neighborhoods with the Highest Increase in Violent Crimes:** To get the top 5 neighborhoods that had the highest crime rates, I first sorted the neighborhood column by year, then took the difference of total incidents for the years 2017 and 2022 (by using DAX code: `Difference = [Total_Incidents for 2022]-[Count of CrimeDateTime for 2017]`), and then used filter functionalities (Top N) by applying `Difference` as filter criteria in Power Bi to get top 5 neighborhoods that had highest crimes. Additionally, to display Overall Change I used `The Overall Change in Crime Rates`, DAX codes on the first page of this report, in the column.

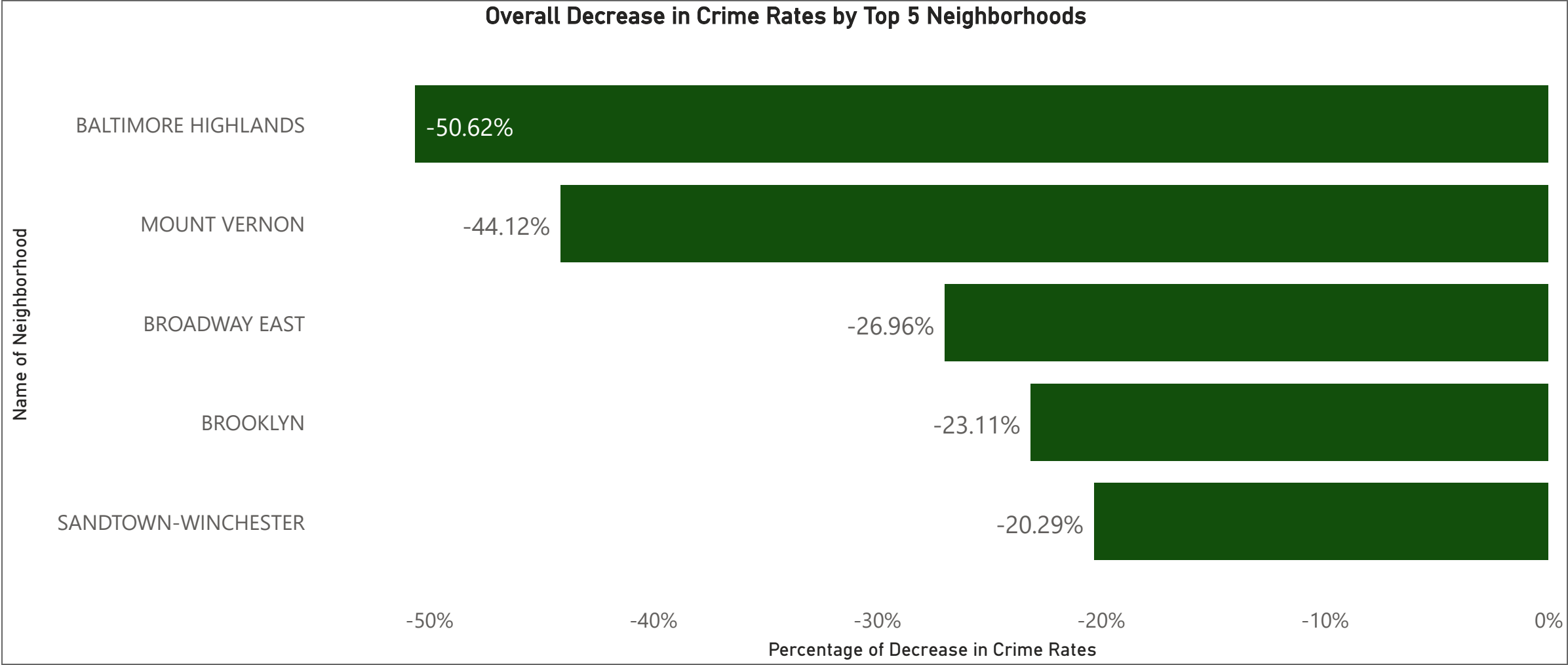
The top 5 violent neighborhoods are **Frankford**, **Downtown**, **Upton**, **East Baltimore Midway**, and **Ellwood Park/Monument**. Lastly, I used the result of analysis in Power Bi to create visualizations for the Top 5 Neighborhoods that had the Highest Increase in Violent Crimes. The bar chart and table below display the result of the analysis in more detail.



Top5 Neighborhoods with Increased Crime Rates	2017 Crimes	2022 Crimes	Overall Change
FRANKFORD	121	216	78.51%
EAST BALTIMORE MIDWAY	63	96	52.38%
ELLWOOD PARK/MONUMENT	58	87	50.00%
UPTON	123	172	39.84%
DOWNTOWN	158	215	36.08%

**IV. Top 5 Neighborhoods with Greatest Reduction in Overall Violate Crime Rates:** To get the top 5 neighborhoods that had the lowest crime rates, I applied a similar approach as with the ones that had the highest rates. I first sorted the neighborhood column by year, then took the difference of total incidents for the years 2017 and 2022 (by using DAX code: `Difference = [Total_Incidents for 2022]-[Count of CrimeDateTime for 2017]`), and then used filter functionalities (Bottom N) by applying `Difference` as filter criteria in Power Bi to get Top 5 neighborhoods that had the greatest reduction in crimes. Additionally, to display Overall Change I used `The Overall Change in Crime Rates`, DAX codes on the first page of this report, in the column.

The top 5 neighborhoods that experienced reduction in crime rates are **Baltimore Highlands**, **Mount Vernon**, **Broadway East**, **Brooklyn**, and **Sandtown-Winchester**. Lastly, I used the result of analysis in Power Bi to create visualizations. The bar graph and table below display the result of the analysis in more detail.



Top5 Neighborhoods with Crime Rate Reduction	2017 Crimes	2022 Crimes	Overall Change
BALTIMORE HIGHLANDS	81	40	-50.62%
MOUNT VERNON	68	38	-44.12%
BROADWAY EAST	115	84	-26.96%
BROOKLYN	225	173	-23.11%
SANDTOWN-WINCHESTER	207	165	-20.29%

**V. Crime Over Time, Summary Table, and Final Overall Change Percentage:** Following the provided instructions for this project, I created two separate line charts displaying the results of crimes for the top 5 neighborhoods that experienced increase and reduction in violent crime rates over six years. Needless to mention that the line graphs only include three specified crime types which are Aggravated Assault, Homicide, and Shooting.

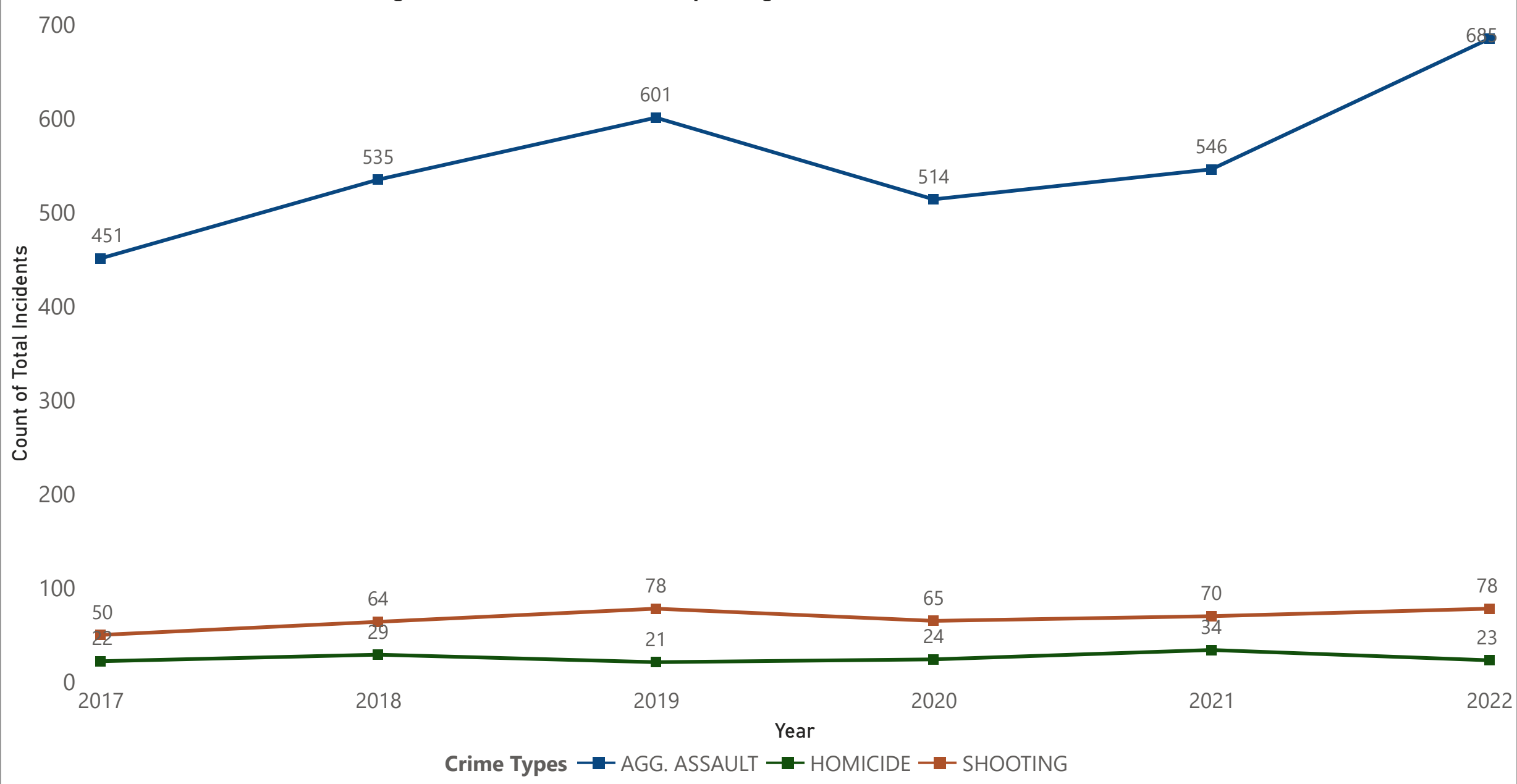
To create the line graphs, I used 'Year' column in X-axis, Count of 'Total\_Incidents' in Y-axis, 'Description' in Legend, and finally filtered by the neighborhoods of interests (those with high crime rates and low crime rates).

Additionally, per project instructions, I created a summary tables in the form of Pivot Tables (Matrix Table in Power Bi) that show the counts by year (all six years) for the 5 high and low crime rate neighborhoods. I used 'Description' and 'Neighborhood' as Rows, 'Year' as Column, and 'Total\_Incident' as values.

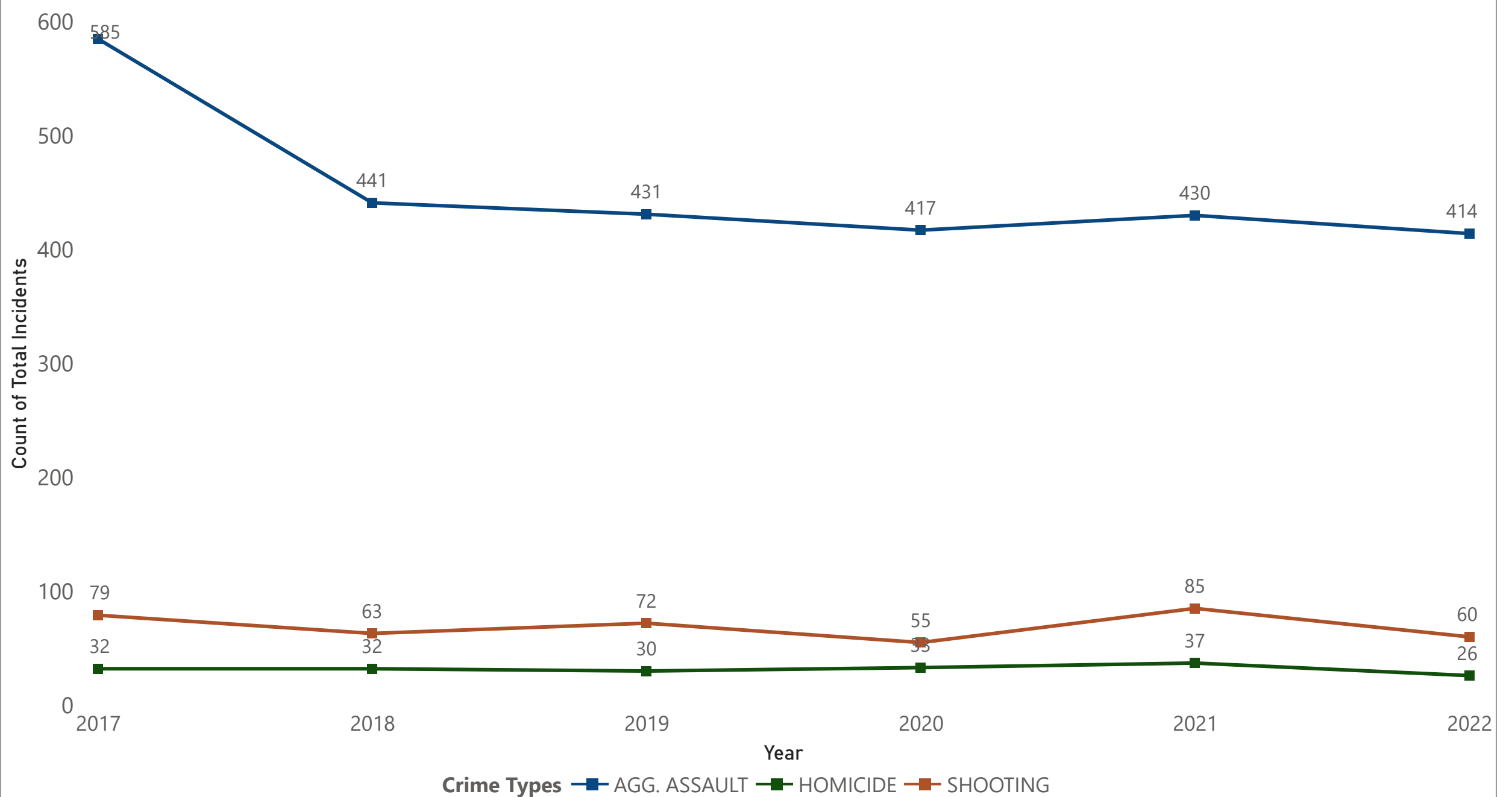
Lastly, I used [The Overall Change in Crime Rates](#), to which I have provided the DAX codes on the first page of this report, to calculate the `Final Overall Change Percentage`. The final overall change percentage is in the form of Card using Power Bi visualization capabilities. They have been filtered by the neighborhoods of interest, meaning those with high crime rate change during the six years 2017-2022 and those 5 neighborhoods with low crime rates for the same time frame.

The crime over time Line Graphs, Summary Tables, and Final Overall Change Percentage displayed in the following pages of this report.

Change of Crime Over Time for Top5 Neighborhoods that Crime Rates Increased



Change of Crime Over Time for Top5 Neighborhoods that Crimes Rates Decreased





**Summary Table: Counts by Years for Top5 High Crime Neighborhoods**

Crime Types & Neighborhoods	2017	2018	2019	2020	2021	2022
▲						
<input type="checkbox"/> <b>AGG. ASSAULT</b>						
DOWNTOWN	147	191	192	170	173	199
EAST BALTIMORE MIDWAY	48	51	97	59	69	83
ELLWOOD PARK/MONUMENT	49	81	54	62	62	70
FRANKFORD	109	120	141	146	133	186
UPTON	98	92	117	77	109	147
<input type="checkbox"/> <b>HOMICIDE</b>						
DOWNTOWN	4	2	3	4	7	5
EAST BALTIMORE MIDWAY	3	7	4	6	7	2
ELLWOOD PARK/MONUMENT	2	1	5	3	5	4
FRANKFORD	5	9	7	7	5	10
UPTON	8	10	2	4	10	2
<input type="checkbox"/> <b>SHOOTING</b>						
DOWNTOWN	7	14	15	14	20	11
EAST BALTIMORE MIDWAY	12	10	9	12	9	11
ELLWOOD PARK/MONUMENT	7	8	8	12	8	13
FRANKFORD	7	18	18	14	10	20
UPTON	17	14	28	13	23	23

**Summary Table: Counts by Years for Top5 Low Crime Neighborhoods**

Crime Types & Neighborhoods	2017	2018	2019	2020	2021	2022
▲						
<input type="checkbox"/> <b>AGG. ASSAULT</b>						
BALTIMORE HIGHLANDS	74	55	48	41	60	34
BROADWAY EAST	87	91	70	68	97	69
BROOKLYN	184	112	139	108	100	135
MOUNT VERNON	68	61	55	58	47	35
SANDTOWN-WINCHESTER	172	122	119	142	126	141
<input type="checkbox"/> <b>HOMICIDE</b>						
BALTIMORE HIGHLANDS	5	2	1		6	2
BROADWAY EAST	10	5	13	9	10	7
BROOKLYN	8	11	5	11	9	11
MOUNT VERNON		2		2		
SANDTOWN-WINCHESTER	9	12	11	11	12	6
<input type="checkbox"/> <b>SHOOTING</b>						
BALTIMORE HIGHLANDS	2	6	7	2	7	4
BROADWAY EAST	18	21	27	14	43	8
BROOKLYN	33	13	18	20	18	27
MOUNT VERNON		1	3	1	1	3
SANDTOWN-WINCHESTER	26	22	17	18	16	18

Final Overall Change Percentage for Neighborhoods with Increased Crime Rates

54.70%

Final Overall Change Percentage for Neighborhoods with Decreased Crime Rates

-28.16%

**VI. Conclusion:** As a result of the analysis conducted for this project 1 assignment, I have discovered that the entire City of Baltimore experienced an increase of violence by 3.25% between the years 2017 and 2022. The top five neighborhoods that have had the highest crime rates were Frankford, Downtown, Upton, East Baltimore Midway, and Ellwood Park/Monument. And the the top five neighborhoods that have had the lowest crime rates were Brooklyn, Sandtown-Winchester, Broadway East, Mount Vernon, and Baltimore Highlands. Note that the neighborhood analysis has been done under the assumption that all neighborhoods contained within this dataset follow a uniform naming convention and do not have any spelling issues, extra spaces, punctuation errors, etc. that would distort the aggregation by neighborhood. The analysis also revealed that the Aggravated Assault has been the leading type of crime during 2017-2022 for both neighborhoods with highest and lowest crime rates. Finally, the Overall Change Percentage for neighborhoods with increased crimes rates between 2017 and 2022 was 54.70% and for the neighborhoods with low crime rates was -28.16%.