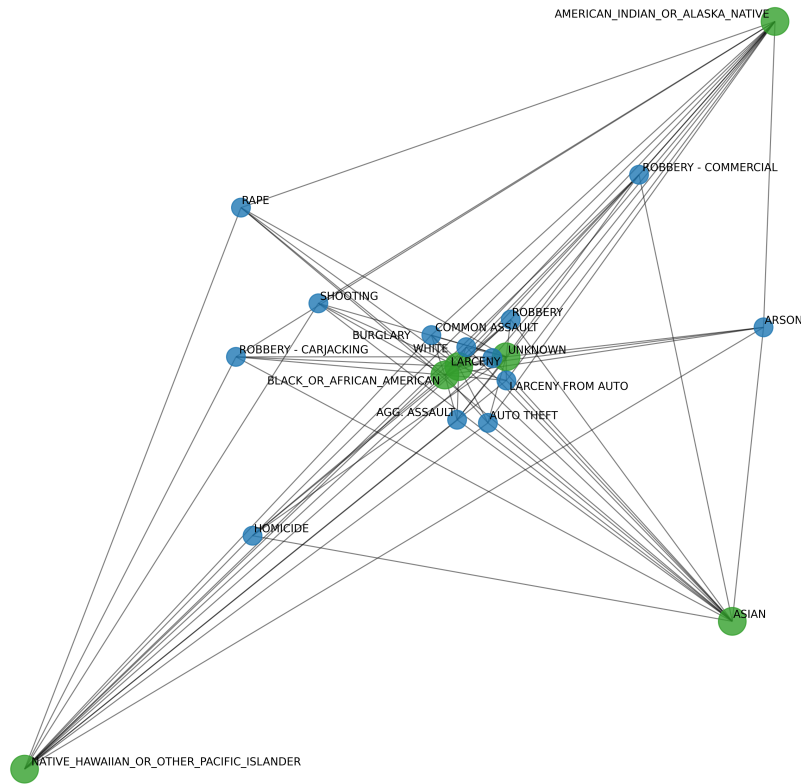


# INFSCI 2415 Final Report

## "Criminal Mosaic: Unraveling Interconnections in Crime Attributes"

Figure 1: Node network

Relationship between Crime Type and Race



### legend explained

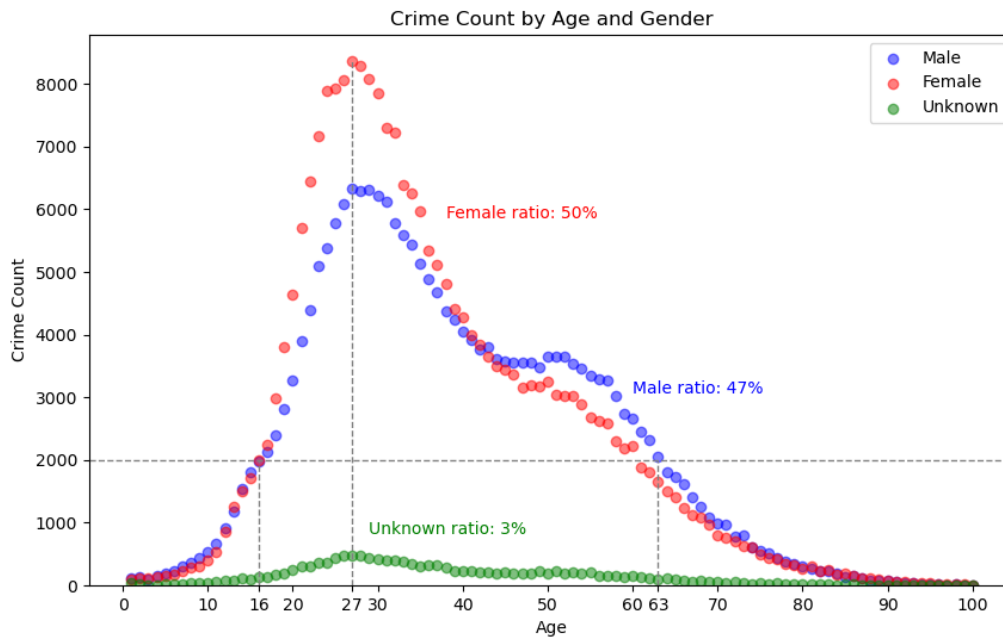
- Blue nodes: represent different types of crime, e.g. "ROBBERY - CARJACKING", "LARCENY", "ROBBERY", etc.
- Green nodes: represent different racial categories, e.g. "BLACK\_OR\_AFRICAN\_AMERICAN", "ASIAN", "WHITE", etc.
- Lines: The line represents the relationship between two nodes. In this case, it represents the association between a race and a specific crime type. The thickness of the line can indicate the strength or frequency of the relationship, with thicker lines representing more frequent associations.

### Findings text introducing highlights of the produced figure in bullet points

- Crime is a phenomenon that spans across all racial categories, underscoring its cross-racial nature.
- There is a notable correlation between the "BLACK\_OR\_AFRICAN\_AMERICAN" group and various crime types within the dataset.
- Crimes like "LARCENY", "AGG ASSAULT", and "LARCENY FROM AUTO" are prevalent across all racial groups, suggesting these are common offenses.
- Specific crime types exhibit a higher frequency within certain racial groups, with "ASIAN" more associated with "ROBBERY - COMMERCIAL" and "AMERICAN\_INDIAN\_OR\_ALASKA\_NATIVE" with "ROBBERY - CARJACKING".

*It's important to note that these are only associations shown in the data and do not imply that a particular race is more likely to commit a specific crime in actual world.*

**Figure 2: Scatter Plot**



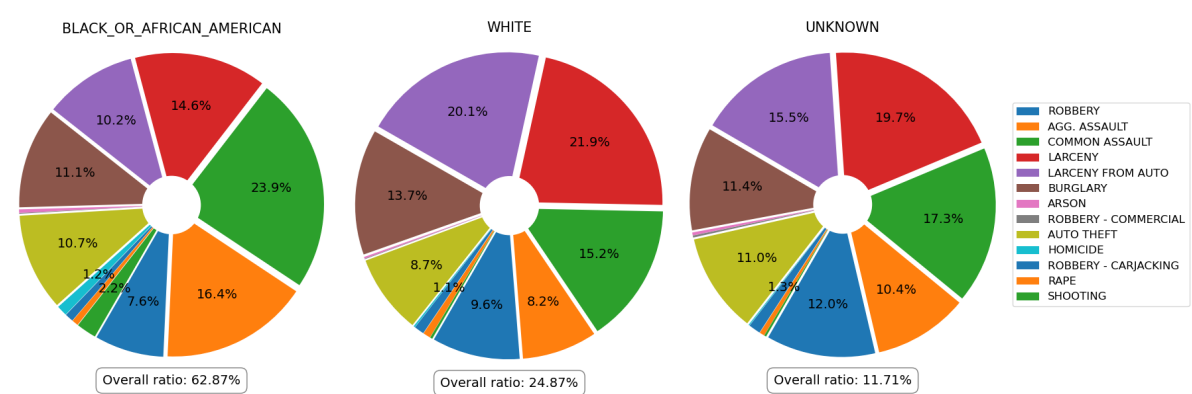
**legend explained**

- Blue scatter plots: Represent the number of crimes committed by males.
- Red scatter plots: Indicate the number of crimes committed by females.
- Green scatter plots: Show the crime numbers where the gender is unknown.
- Horizontal axis: Represents the age range of the crime demographics, ranging from 0 to 100 years.
- Vertical axis: Indicates the total number of crimes committed by each gender and age group.

**Findings text introducing highlights of the produced figure in bullet points**

- The chart illustrates the variation in crime numbers among males and females across different ages in Baltimore city, encompassing data from 2013 to 2023. It covers offenses ranging from minor crimes like "Larceny" to major ones like "Homicide."
- An interesting observation is the near equality in the total number of crimes committed by males and females across all crime types.
- Both male and female crime numbers peak at age 27, indicating a significant period for crime occurrence in both genders.
- The data reveals distinct crime trends across various age groups. For ages 16 to 63, males and females constitute 88.22% and 90.34% of the total crimes for their respective genders. In the 22 to 40 age group, these figures are 46.32% for males and 53.43% for females. Furthermore, crimes committed by young adults (ages 18 to 29) account for 26.32% and 33.34% of the total crimes for males and females, respectively.
- Subsequent in-depth analyses focus primarily on the crime statistics of individuals aged between 16 to 63 years.

Figure 3: Pie Chart



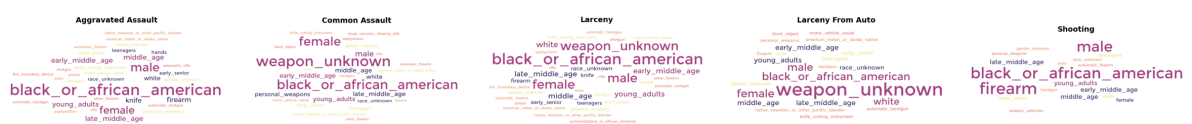
legend explained

- The chart represents the proportion of various types of crimes committed by different racial groups in Baltimore city from 2013 to 2023.
- The "overall ratio" at the bottom of the chart indicates the proportion of each racial group in the total crime statistics.
- The chart's title at the top specifies the racial group being analyzed.
- The pie chart's different colored sectors represent various crime types, including 13 varying levels of seriousness, such as "Robbery," "Larceny," and "Burglary."

Findings text introducing highlights of the produced figure in bullet points

- The chart indicates that the BLACK\_OR\_AFRICAN\_AMERICAN group has a higher representation in crime statistics compared to other groups in Baltimore city.
- Combined, the "BLACK\_OR\_AFRICAN\_AMERICAN" and "WHITE" groups account for over 87% of the total recorded crimes. Other racial groups like "ASIAN" and "AMERICAN\_INDIAN\_OR\_ALASKA\_NATIVE" have smaller representations in the data.
- The most frequently recorded crimes for the "BLACK\_OR\_AFRICAN\_AMERICAN" group are "COMMON ASSAULT," "AGG. ASSAULT," and "LARCENY."
- For the "WHITE" and unknown race groups, the most common crimes are "LARCENY," "LARCENY FROM AUTO," and "COMMON ASSAULT."
- The chart provides a comparative perspective on the proportion of different crimes committed by various racial groups, highlighting the distribution of crime types among these groups.

Figure 4: Word Cloud



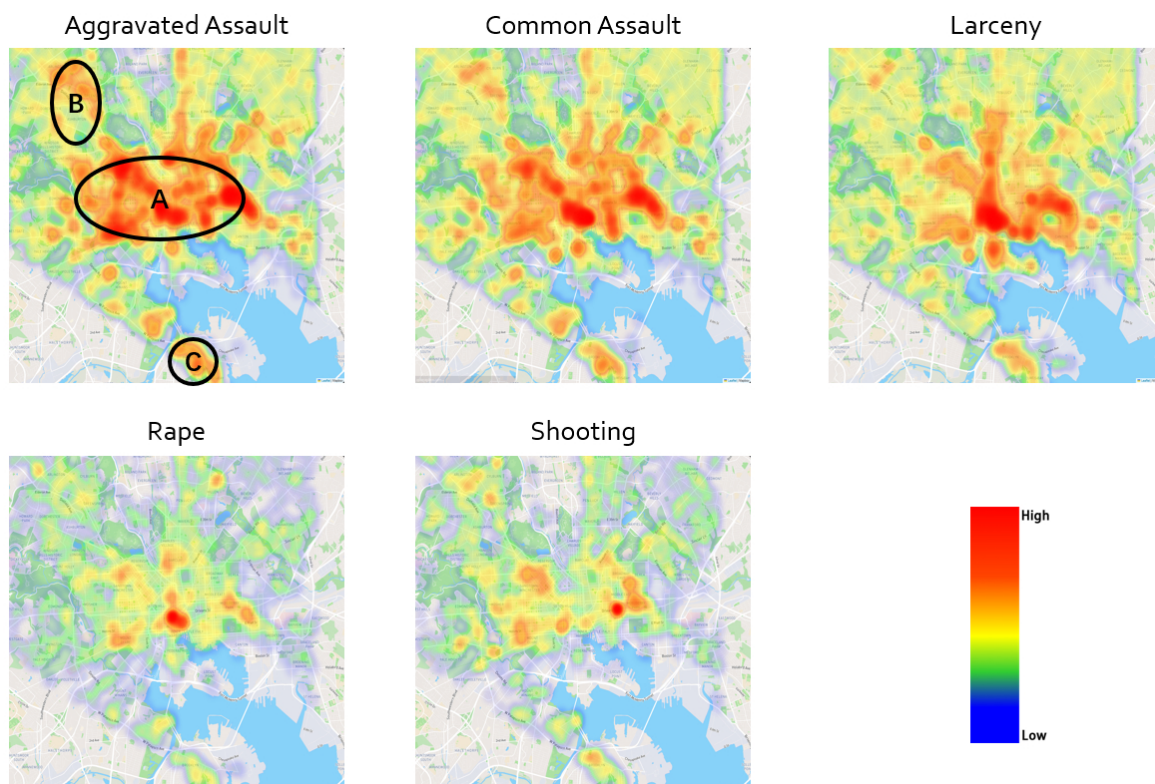
legend explained

- The chart displays word clouds representing key terms associated with different types of crimes in Baltimore city from 2013 to 2023.
- The title at the top of the chart identifies the specific crime type, with five different crime types of varying severity chosen for representation.
- The word clouds contain four categories of keywords: "Weapon," "Age," "Race," and "Gender," indicating their relation to different crime types.
- The size of the words in the word clouds signifies their relevance, with larger words indicating greater significance.

#### Findings text introducing highlights of the produced figure in bullet points

- The word clouds show the term BLACK\_OR\_AFRICAN\_AMERICAN appearing in association with all five types of crimes.
- In the word clouds for "Larceny" and "Larceny From Auto," there is an absence of weapons and no clear connection to specific age or gender groups, which aligns with typical characteristics of theft-related crimes.
- The word clouds for "Aggravated Assault" and "Assault" highlight a higher frequency of these crimes in the "Early Middle Age" and "Young Adult" categories, with "Assault" showing a more significant association with females.
- The "Shooting" word cloud shows a strong association of this crime with "BLACK\_OR\_AFRICAN\_AMERICAN," "Male," and "Firearm" as a weapon.
- These word clouds provide a multi-dimensional and rapid overview of the relationships between different attributes and the extent of these connections in the data.

**Figure 5: Geospatial Heatmap**



#### legend explained

- The chart maps the specific locations and distribution of various types of crimes in Baltimore city from 2013 to 2023.
- The title at the top of the chart identifies the different crime types, with five selected based on their varying levels of severity.
- The colors on the map indicate the density of crime occurrences, with shades closer to red denoting frequent crime activity, and hues closer to blue-green suggesting fewer incidents.

#### Findings text introducing highlights of the produced figure in bullet points

- The map reveals that the majority of crimes are concentrated in the central area of Baltimore city, with a notable prevalence of shootings around the outskirts of this central region.
- In the "Aggravated Assault" map, the city is mainly divided into three zones: A, B, and C. Zone A, possibly representing downtown Baltimore, experiences frequent crimes. Zones B and C are also crime-prone, with Zone C being further from the city center, potentially indicating it as a port or commercial area.
- For severe violent crimes like "Rape," a higher frequency is observed in the city center, with notable occurrences also in the far northern parts of Baltimore city.
- The "Shooting" incidents are distributed not in the city center but primarily around the outskirts. The pattern suggests a higher frequency of shootings in the southern and northwest areas of Baltimore.

- The utilization of a geospatial heatmap facilitates swift analysis and visualization of crime hotspots, offering valuable insights into the spatial distribution of various crimes throughout Baltimore city.

### Findings and Suggestions Based on the visualization

- Implementing gun control measures may be beneficial, particularly in areas with high firearm-related crime rates. Our visualization and analysis suggest a correlation between firearms and criminal activities, and regulating guns could potentially reduce these crimes.
- We advise considering an increase in police patrols in residential neighborhoods to deter property crimes, such as break-ins and thefts. An enhanced police presence might act as a deterrent and contribute to a community atmosphere less conducive to crime.
- Deploying additional police forces in strategic areas could improve quick response capabilities and intelligence gathering, which may help prevent future crimes in high-crime regions.

### Data and Method Description

- The dataset used encompasses information on major crimes against individuals, like shootings and robberies, within the City of Baltimore. It includes comprehensive details on the crimes, covering aspects like gender, age, race, locations, and more. The dataset contains a total of 587,356 records, with crime data starting from 2013.
- To create the visualizations, I utilized Matplotlib, NumPy, json, folium, and wordcloud in Pycharm.
- Pycharm is a Python IDE (Integrated Development Environment) with a full set of tools to help users improve their efficiency when developing in the Python language, such as debugging, project management, and more.
- I employed Matplotlib's scatter() for scatter plots, pie() for pie charts, and Circle() to create central circle charts. The wordcloud.WordCloud() function was used to generate word clouds from processed feature text, while folium.Map() was employed for creating geospatial heatmaps.
- Each visualization's title, x-axis, y-axis, and legend were individually adjusted for better aesthetics. Matplotlib's text() was used to add relevant annotations and useful information to the charts. Functions like plt.axhline() and plt.vlines() were utilized to add helpful dashed lines for clarity in scatter plots. Lastly, Matplotlib's tight\_layout() function helped adjust the subplot layouts to accommodate legends appropriately.

### Significance statement on why the presented figure is important

1. The visualization of crime data in Baltimore is essential for understanding the city's unique social dynamics and aids in formulating effective public policies and law enforcement strategies. This is particularly important given the city's rich history and diversity.
2. The figure helps in identifying patterns between criminal activities and socio-demographic factors such as race and age. This is crucial for developing targeted crime response strategies and proactive prevention measures, especially for vulnerable age groups.
3. By comparing crime data across different racial groups within specific age brackets, such as 20 to 29 years, the figure sheds light on underlying inequalities and structural social issues. It also highlights disparities in the distribution of educational and employment opportunities, as well as community resources, which are pivotal for shaping the future of the youth and the community at large.

*It's important to note that these are only associations shown in the data and do not imply that a particular race is more likely to commit a specific crime in actual world.*

### References

Data Source: [https://services1.arcgis.com/UWYHeuuJISiGmgXx/ArcGIS/rest/services/Part1\\_Crime\\_Beta/FeatureServer](https://services1.arcgis.com/UWYHeuuJISiGmgXx/ArcGIS/rest/services/Part1_Crime_Beta/FeatureServer)

Codes: [https://github.com/AmazingYanyan/Information\\_visualization\\_final\\_project](https://github.com/AmazingYanyan/Information_visualization_final_project)

Yanyan Li (YAL185)