

Final Assignment Paper Analysis of NYPD Arrests

Group: 04

Sharvil Kishor Wadekar
Mohit Ravindra Kamble
Soham Santosh Mane
Jiangyu Qin

College of Professional Studies – Northeastern University
ALY6070: Communication and Visualization for Data Analytics

Prof. Jack Bergersen
February 16, 2024



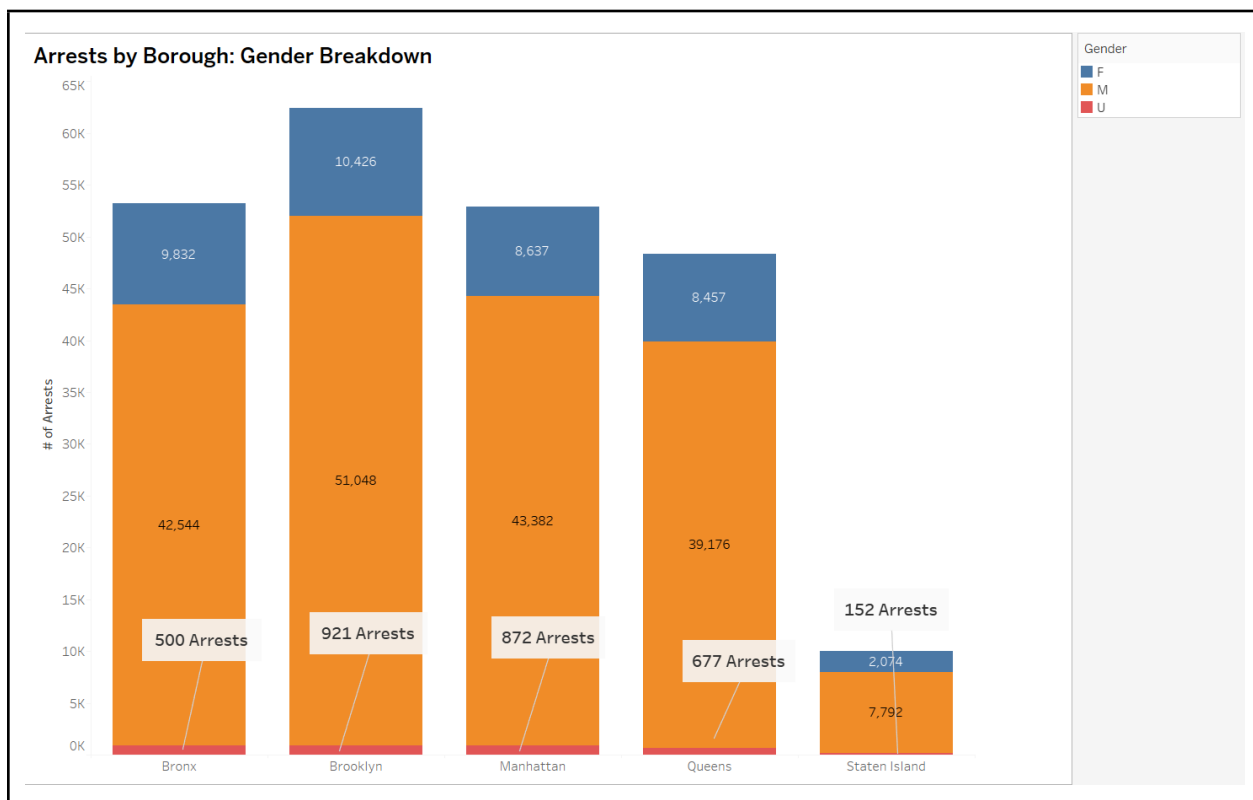
Overview:

This report provides a comprehensive analysis of the NYPD Arrests Dataset, examining various aspects such as gender breakdown, primary offense categories, spatial distribution across boroughs, and age demographics over time. Through the utilization of Tableau visualizations, insights into policing practices, resource allocation, and societal dynamics in New York City are uncovered.

Analysis:

Research Question 1:

How do arrest rates differ across New York City boroughs, considering the gender of the individuals involved?



Questions we asked about the data:

- What is the frequency of arrests for men and women in each of the five boroughs of New York City in 2023?
- Which borough had the highest number of arrests in 2023?
- How does the gender disparity in arrests vary across the five boroughs in 2023?

Variables we used on analysis:

- PERP_SEX
- ARREST_BORO
- COUNT(NYPD_ARREST_DATA)

Methods we used to analyze the data:

- Conducted a gender-based analysis of arrest frequencies in each borough for the year 2023.
- Determined the overall arrest rates for each borough to identify the one with the highest number of arrests.
- Calculated the gender disparity ratio by comparing the arrest rates for men and women in each borough.

Possible Findings:

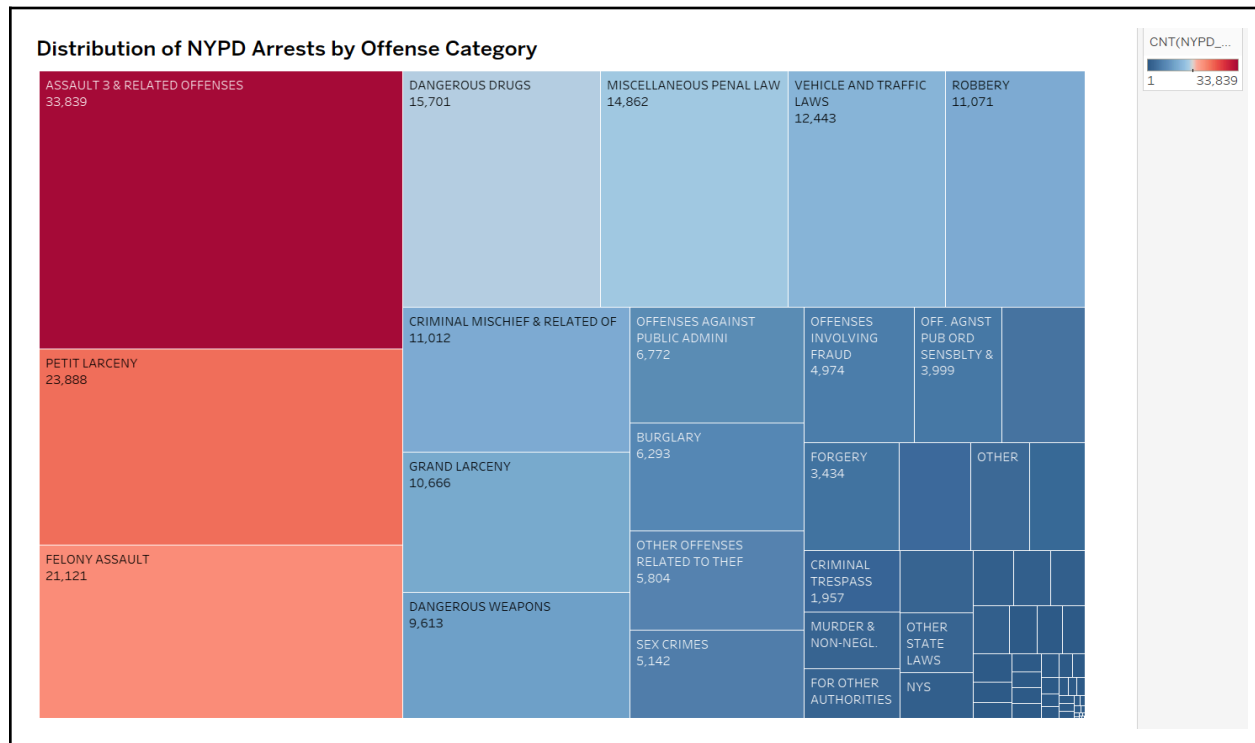
- Analyzed arrest data for New York City in 2023, revealing that men were arrested more frequently than women in all boroughs, with the highest gender disparity observed in Staten Island.
- Utilized variables such as gender, boroughs, and arrest frequency to conduct the analysis, employing bar charts, pie charts, and line charts to present the findings visually.
- Emphasized key design decisions, including color coding and visual selection, to effectively communicate the disparities, highlighting the need for discussions on policing strategies and interventions.

Design decisions we made:

- **Visual Selection:** Chose bar charts for gender-based analysis to effectively compare arrest frequencies. Utilized pie charts for total arrests to emphasize proportions. Employed line charts or radar charts for displaying gender disparity ratios across boroughs for clarity.
- **Color Coding:** Used distinct colors for men and women in bar charts to enhance visual differentiation. Employed different shades to represent each borough consistently across visuals for easier interpretation.
- **Emphasis on Disparities:** Focused on highlighting the gender disparities in arrests across boroughs in the summary to draw attention to potential areas for discussions on policing strategies and interventions.

Research Question 2:

What are the primary categories of offenses in the NYPD arrest dataset, and how do their frequencies compare?



Questions we asked about the data:

- What are the primary offense categories and their frequencies?
- What are the trends in arrest rates over time?
- How do arrest rates vary by borough or neighborhood?
- What are the demographics of the people who are arrested?
- What are the outcomes of the arrests? For example, how many people are convicted of crimes, and what are their sentences?

Variables we used on analysis:

- OFNS_DESC
- COUNT(NYPD_ARREST_DATA)

Methods we used to analyze the data:

- Treemap chart visualization to depict the distribution of arrests across offense categories.
- Extraction of specific numbers for each offense category to identify the highest number of arrests.
- Comparison of arrest frequencies for property crime and drug-related offenses.

Possible Findings:

- Treemap chart illustrating the distribution of arrests across different offense categories.
- Assault is the most common offense category with over 33k arrests, followed by petty larceny with over 23k arrests, and miscellaneous arrests with over 14k arrests. Property crime and drug-related offenses are prevalent, with notable figures for robbery, sex crimes, and various traffic violations.

Design decisions we made:

- Selection of Treemap chart: Chose this visualization to effectively showcase the distribution of arrests and highlight the most common offense categories.
- Modification of visuals: Emphasized specific offense categories in the Treemap chart to draw attention to key findings, such as assault, petty larceny, and miscellaneous arrests.
- Consideration of limitations: Acknowledged that the figures represent arrests, not convictions, and emphasized the importance of interpreting the data in the context of overall crime trends in New York City.

Research Question 3:

How do arrest counts vary spatially across the boroughs of New York City?



Questions we asked about the data:

- What is the overall distribution of NYPD arrests across the boroughs of New York City?
- Which boroughs have the highest and lowest arrest counts?
- How do arrest rates vary spatially across the city?
- Are there any discernible patterns in the distribution of arrests based on the provided map?

Variables we used on analysis:

- Latitude
- Longitude
- ARREST_BORO
- COUNT(NYPD_ARREST_DATA)

Methods we used to analyze the data:

- Visual inspection of the map to identify patterns and variations.
- Utilization of color intensity to highlight differences in arrest counts.
- Comparison of boroughs to determine relative arrest rates.
- Incorporation of background map layers for geographical context.

Possible Findings:

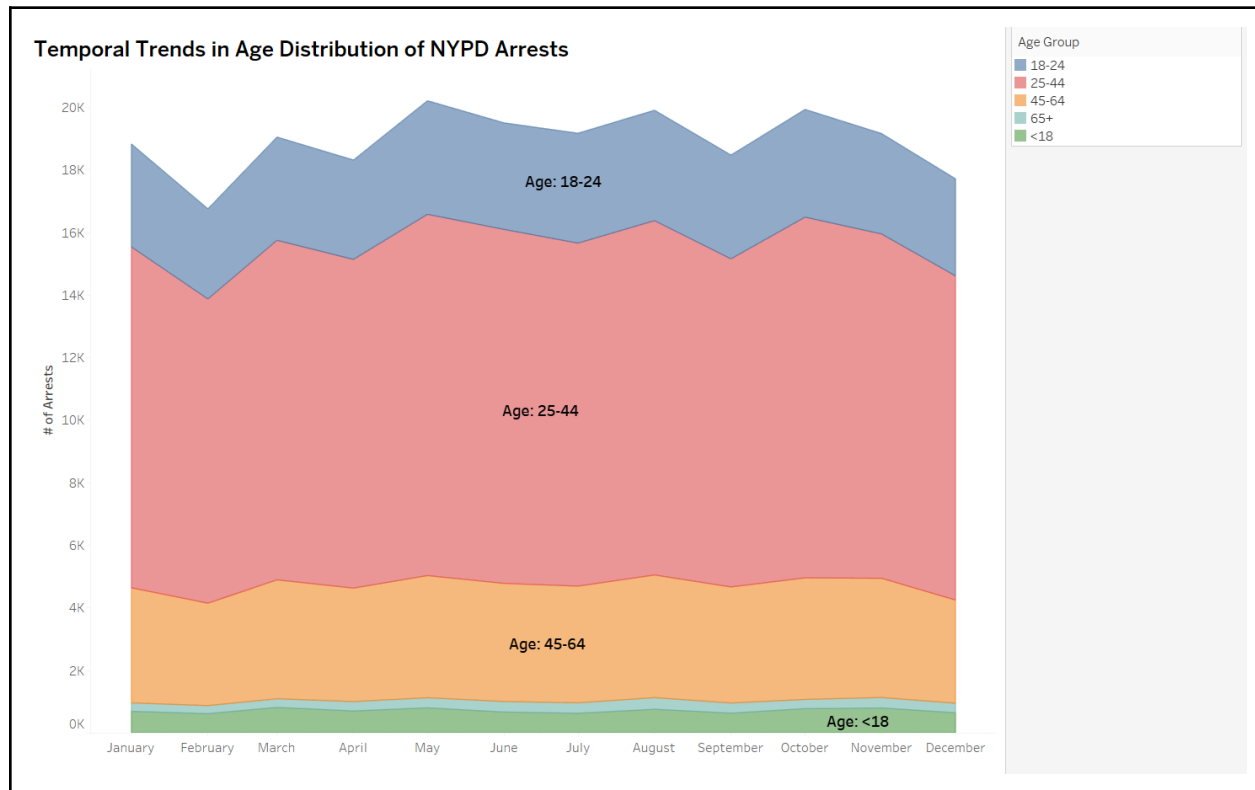
- Brooklyn and Manhattan exhibit the highest number of arrests, with distinct dark shades on the map.
- The Bronx and Queens closely follow with relatively high arrest counts, depicted by darker shades as well.
- Staten Island has the lowest arrest count, represented by a lighter shade on the map.
- Spatial analysis reveals concentrated areas of high and low arrest rates, providing insights into law enforcement dynamics across different boroughs.

Design decisions we made:

- **Color Coding:** The use of color, with darker shades indicating higher arrest counts, facilitates quick interpretation of the data and highlights variations among boroughs.
- **Background Map Layers:** Background map layers in a normal style were incorporated to provide geographical context, enhancing clarity and readability.
- **City Names:** Inclusion of city names on the map aids easy identification of boroughs, contributing to a user-friendly visualization.
- **Overall Map Design:** The map design prioritizes clarity and simplicity to effectively communicate the spatial distribution of NYPD arrests, ensuring that patterns are easily discernible for comprehensive analysis.

Research Question 4:

How does the distribution of arrests vary by age group over time, as depicted by the area chart visualization?



Questions we asked about the data:

- How does the age distribution of individuals arrested by the NYPD vary over time?
- Are there any seasonal variations in the age distribution of arrests?
- Are there any differences in the age distribution of arrests across different age groups?

Variables we used on analysis:

- ARREST_DATE
- AGE_GROUP
- COUNT(NYPD_ARREST_DATA)

Methods we used to analyze the data:

- We created an area chart with months on the x-axis and counts of arrests on the y-axis, with different age groups represented by colors.
- We examined the chart to identify any patterns or trends in the data.

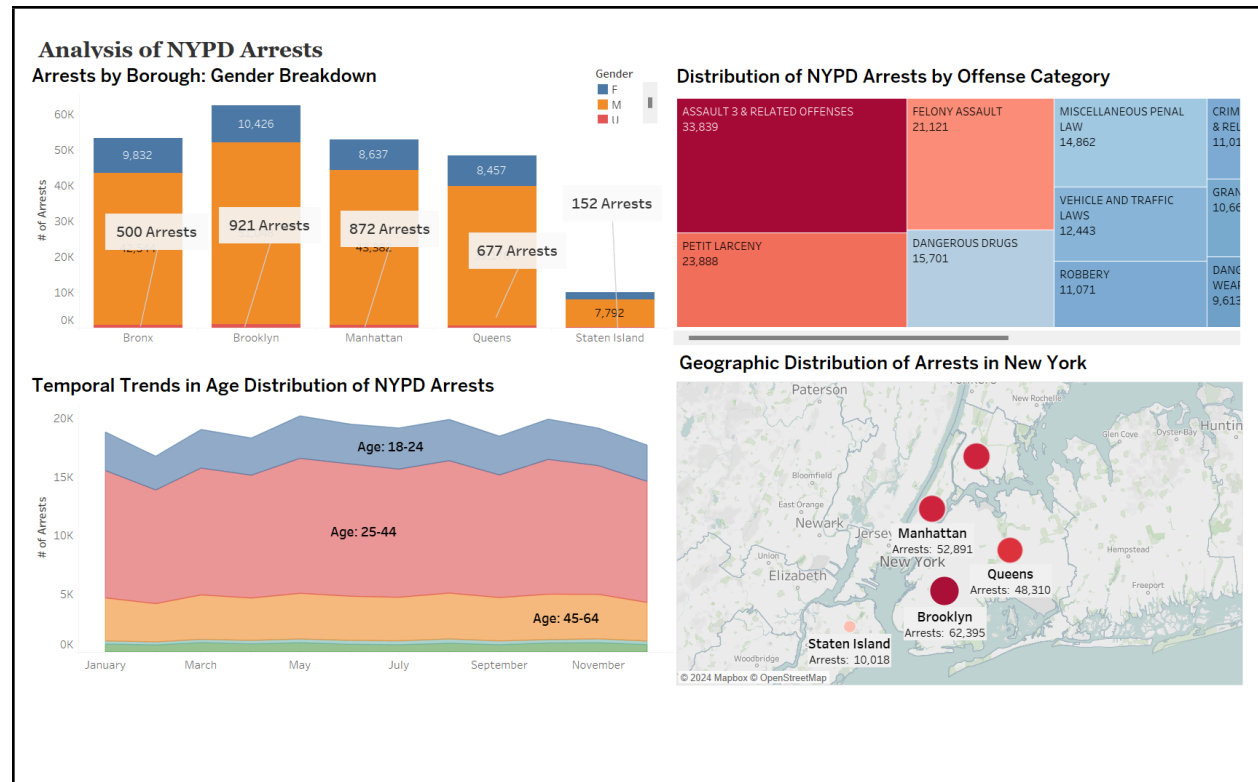
Possible Findings:

- The 25-44 age group had the highest number of arrests throughout the year, followed by the 18-24 age group.
- Arrests for the 45-64 age group were considerably lower, while arrests for those under 18 and over 65 were the lowest.
- There were some variations across the months. For example, in January and February, the gap between the 25-44 and 18-24 age groups was smaller than in other months. Additionally, arrests for the under-18 age group peaked in April and May, before dropping significantly in the summer months.

Design decisions we made:

- **Choice of Visuals:** Selected an area chart for its effectiveness in displaying the distribution of arrests over time, allowing for easy identification of patterns.
- **Color Coding:** Assigned different colors to age groups for clarity, enabling viewers to distinguish between segments in the chart easily.
- **Monthly Aggregation:** Opted for monthly data aggregation to capture nuances and patterns at a more granular level.
- **Focus on Age Groups:** Prioritized age groups as a key variable, highlighting potential insights into demographics related to arrests.

Dashboard:



Findings:

- Analyzing NYC arrest data revealed that men were arrested more across all boroughs, notably in Staten Island.
- We utilized gender, borough, and arrest frequency, presenting findings visually with bar, pie, and line charts.
- Assault topped offenses with 33k+ arrests, followed by petty larceny and miscellaneous offenses.
- Brooklyn and Manhattan had the most arrests, while Staten Island had the least. Spatially, arrests clustered differently across boroughs.
- Arrest trends fluctuated over time, and age groups showed consistent arrest patterns, suggesting focus areas for law enforcement.
- The data underscores the need for discussions on policing strategies and interventions.

Conclusion:

In conclusion, our analysis of NYPD arrest data using Tableau uncovered notable gender disparities, revealing higher arrest rates for men across all boroughs, particularly in Staten Island. We delved into offense categories, highlighting assault as the most prevalent. Brooklyn and Manhattan exhibited the highest arrest rates, showcasing spatial variations. Arrest trends fluctuated over time, emphasizing potential focus areas for law enforcement. Our findings underscore the importance of discussions on policing strategies and interventions. The visualizations, including bar, pie, and line charts, provided a clear understanding of the complex dynamics, fostering awareness for informed decision-making and societal dialogue.

Citations:

- **Dataset:** [source](#).
- **Bar Chart:** [source](#).
- **Map Visualization:** [source](#).
- **Tree Maps:** [source](#).
- **Area Chart:** [source](#).