

NYPD Arrest Data

Year to Date (2022)



Interact with the dashboard here: <https://www.novypro.com/project/nypd-arrest-data-2>

Introduction

This study aims to analyze and recognize patterns and trends in New York City police enforcement activity as represented in statistics on arrests made by the NYPD. The information comprises the type of crime, the location and date of enforcement, and perpetrator demographics. By examining this data, it is intended to obtain insight into the nature of police activity in the city and how it may change depending on criteria.

Data sources

The data for this project was collected from the New York City Police Department's NYPD Arrest Data (Year to Date) dataset. This dataset comprises records of every arrest made in New York City by the NYPD during the current fiscal year and is updated quarterly.

- *Why is this data collected?* This is a breakdown of every arrest effected in NYC by the NYPD during the current year.
- *How is this data collected?* This data is manually extracted every quarter and reviewed by the Office of Management Analysis and Planning.

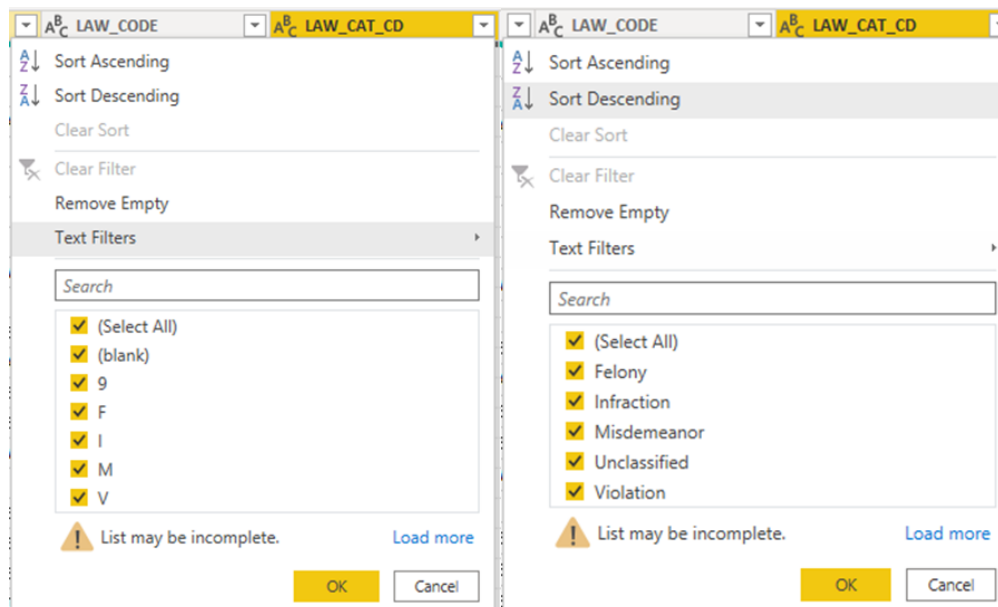
- *What does each record represent?* Each record represents an arrest effected in NYC by the NYPD and includes information about the type of crime, the location, and the time of enforcement.
- *How can this data be used?* In addition, information related to suspect demographics is also included.
- *What are the idiosyncrasies or limitations of the data to be aware of?* This data can be used by the public to explore the nature of police enforcement activity.

Please refer to the attached data [footnotes](#) for additional information about this dataset.

Data cleaning and preparation

The main tool that is used for data analysis is Power BI. The NYPD Arrest Data (Year to Date) dataset needed to be cleaned and processed before analysis could begin. This involved several processes, including:

1. Removed Errors
2. In the LAW_CAT_CD (level of offense) column, the values only include a single letter, with “F” representing a felony, “M” representing a misdemeanor, “I” representing an infraction, and “V” representing a violation. *Replace values* were used to replace those values to show their meaning. The blank values were also replaced with “Unclassified”. However, there is a value of “9” that does not correspond to any of these categories and should be removed.



3. Replaced values of ARREST_BORO (Borough of arrest) column. B(Bronx), S(Staten Island), K(Brooklyn), M(Manhattan), Q(Queens).

The screenshot shows a data tool interface with two columns: LAW_CAT_CD and ARREST_BORO. The ARREST_BORO column is selected. The filter settings for ARREST_BORO are displayed, showing a list of boroughs: (Select All), Bronx, Brooklyn, Manhattan, Queens, and Staten Island. All are checked. The interface includes buttons for Sort Ascending, Sort Descending, Clear Sort, Clear Filter, Remove Empty, Text Filters, and a Search box. OK and Cancel buttons are at the bottom.

4. Added a conditional column to show the jurisdiction responsible for the arrest. Jurisdiction codes: 0(Patrol), 1(Transit) and 2(Housing) represent NYPD whilst codes 3 and more represent non-NYPD jurisdictions

Add Conditional Column

Add a conditional column that is computed from the other columns or values.

New column name

JURISDICTION_ARREST

	Column Name	Operator	Value ①		Output ①
If	JURISDICTION_C...	equals	ABC 123 0	Then	ABC 123 Patrol
Else If	JURISDICTION_C...	equals	ABC 123 1	Then	ABC 123 Transit
Else If	JURISDICTION_C...	equals	ABC 123 2	Then	ABC 123 Housing ...

Add Clause

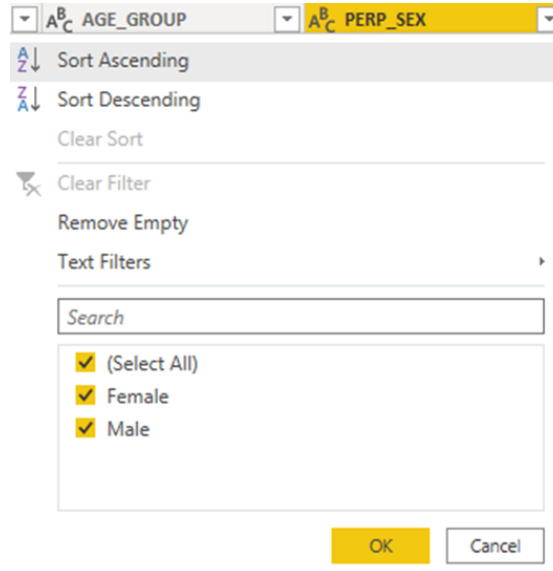
Else ①

ABC 123 Non-NYPD

OK

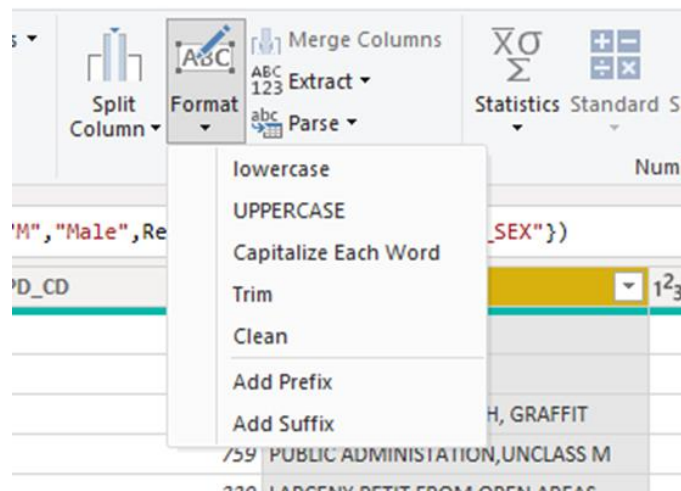
Cancel

5. Replaced Values of PERP_SEX column, F to Female and M to Male.

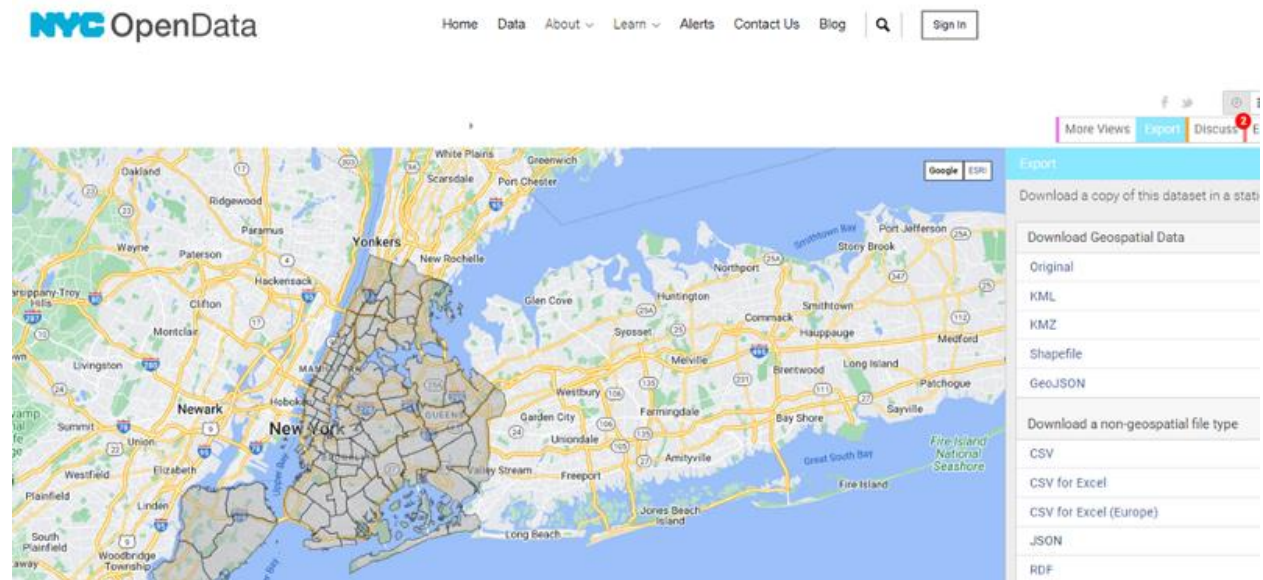


6. Remove columns: X_COORD_CD, and Y_COORD_CD since there's already Latitude and Longitude for Global Coordinate System.

7. To make values uniform, change the text format to *Capitalize Each Word* for all text columns.

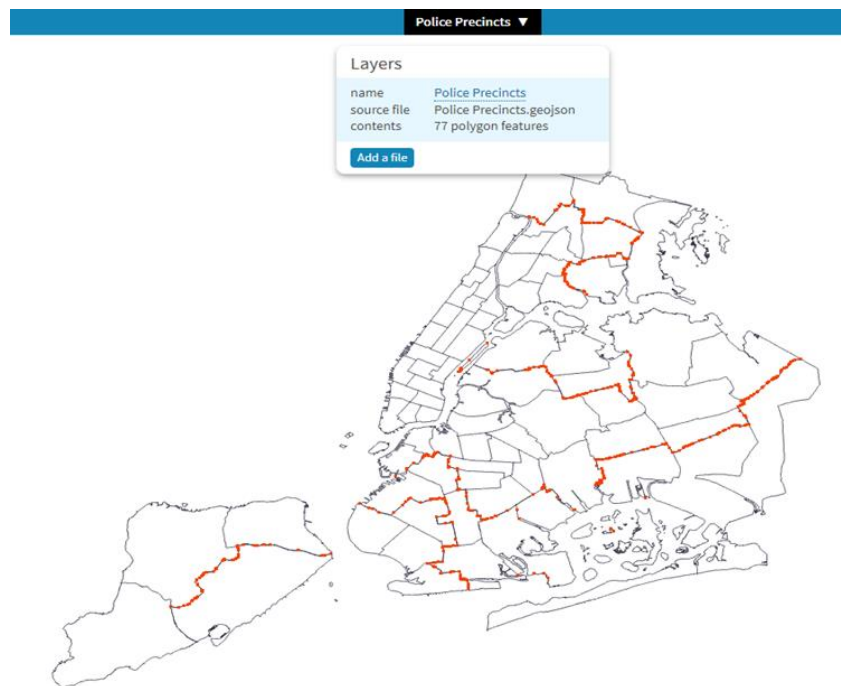


8. The shape map that will be created later on for the dashboard needs to show only New York. NYC Open Data provides a JSON file providing a map of New York. Click on GeoJSON to download the file.

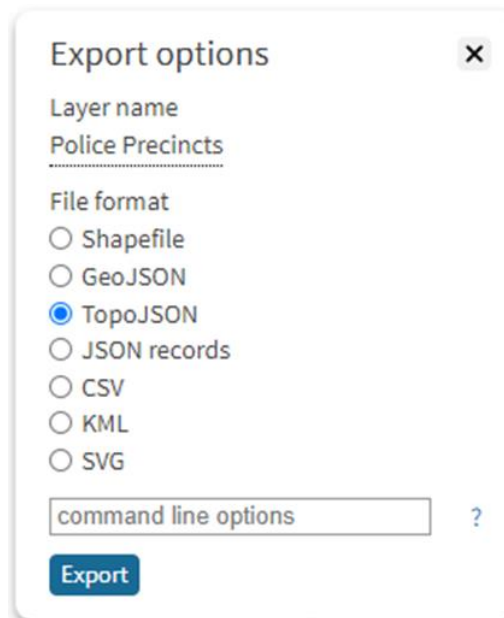


NYC Open Data website

9. After downloading the file, it should be imported into [mapshaper.org](https://www.mapshaper.org/) and converted to TopoJSON format. This is necessary since Power BI only supports TopoJSON for shape maps.



10. Click Export on the rightmost part of the website. A message box will appear, select TopoJSON, and then click Export.

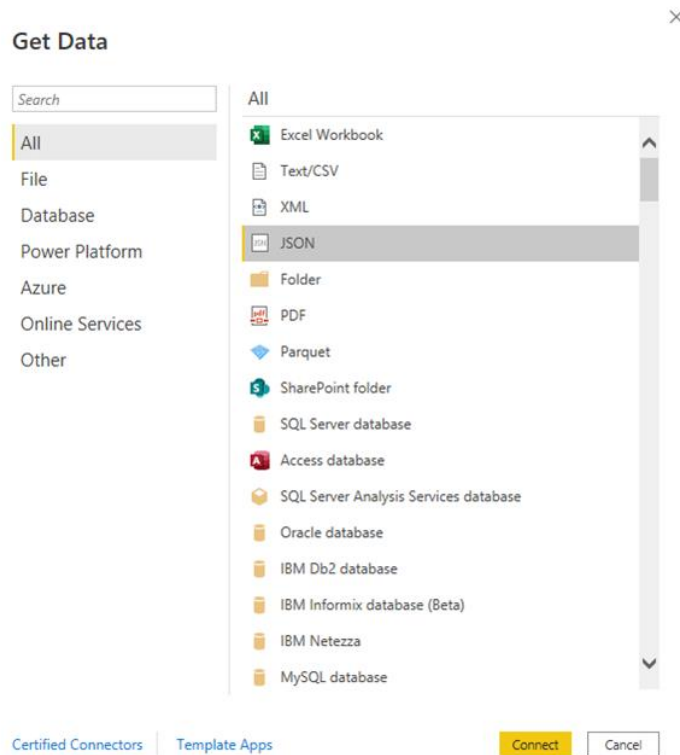


The image shows a dialog box titled "Export options" with a close button (X) in the top right corner. Inside the dialog, there are two sections. The first section is labeled "Layer name" and contains the text "Police Precincts" followed by a dotted line. The second section is labeled "File format" and contains a list of radio button options: Shapefile, GeoJSON, TopoJSON (which is selected with a blue dot), JSON records, CSV, KML, and SVG. Below the radio buttons is a text input field labeled "command line options" with a question mark icon to its right. At the bottom of the dialog is a blue button labeled "Export".

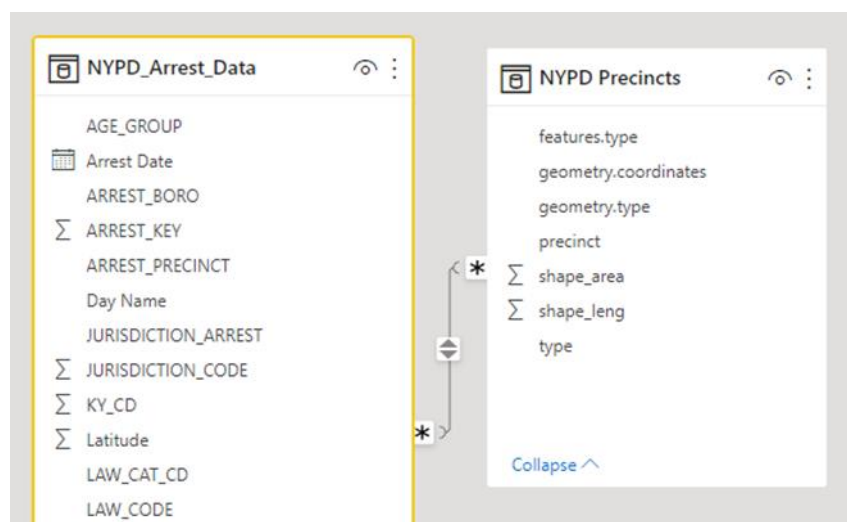
11. To show which day of the week has the most number of arrests, go back to the Power Query editor and duplicate the table. Then, use the Group By function. Select Day Name and then Count for the aggregate function.

Data exploration and visualization

1. Starting with the shape map, import the JSON file that was downloaded by clicking on *Get Data* and then selecting JSON.



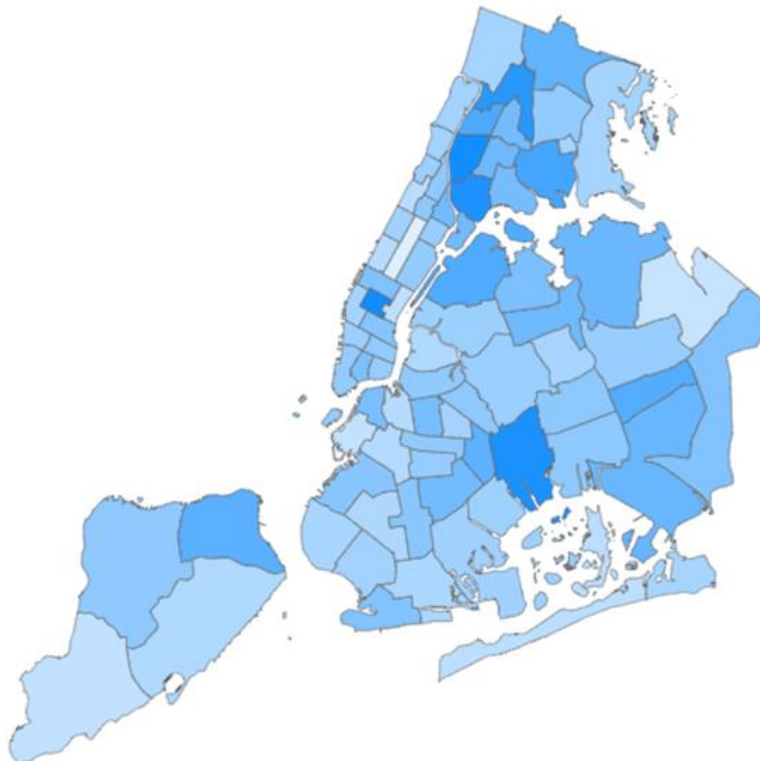
2. Go to the *Model* tab and then drag the *precinct* column from the NYPD to the main table to have a relationship.



3. Next, add a shape map to the dashboard. Shape maps need to be turned on from the settings (File>Options and Settings>Options>Global>Preview Features). After adding the shape map, add *Precinct* to the Location, and then in Color saturation add the *number of arrests*. Go to the Visual tab, change Map type to Custom map and then upload the converted TopoJSON file.

The screenshot displays the dashboard configuration interface. On the left, the 'Location' field is set to 'Precinct', the 'Color saturation' field is set to 'Number of Arrests', and the 'Tooltips' field is set to 'Borough'. The 'Legend' field is empty with the text 'Add data fields here'. On the right, the 'Visual' tab is selected, and the 'Map settings' section shows 'Map type' set to 'Custom map', 'View map type key' link, 'Add a map type' button, and 'Projection' set to 'Mercator'. The 'Police Precincts1.j...' file is listed under 'Add a map type'.

The shape map should look like this:

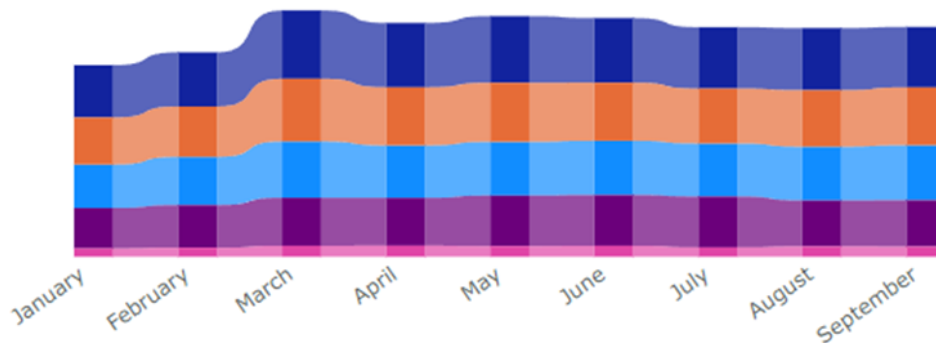


4. The next chart to add is a ribbon chart, which is a type of chart that shows trends over time and can be used to compare multiple series of data. To create the ribbon chart, add the *Arrest Date* column to the X-axis, *Arrests* in the Y-axis, and *Borough* to the Legend. This will allow us to see how the number of arrests varies over time for each borough.

X-axis	
Arrest Date	▼ ×
Year	×
Quarter	×
Month	×
Day	×
Y-axis	
Arrests	▼ ×
Legend	
Borough	▼ ×

Arrests by Month and Borough

Borough ● Bronx ● Brooklyn ● Manhattan ● Queens ● Staten Island

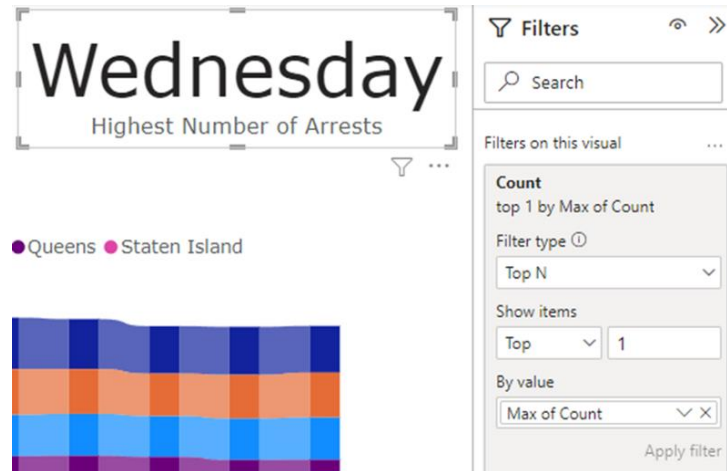


5. There are different kinds of levels of offenses, which can be shown by using a funnel chart.

6. Next, three filters are added: Borough, Precinct, and Jurisdiction.

Borough	Precinct	Jurisdiction ▼
All ▼	All ▼	All ▼

7. Two cards have been added to display the borough with the highest number of arrests at the current time and the day of the week with the most arrests. For the card showing the highest number of arrests, add a filter on the visual and choose filter type to *Top N*, input Top 1 and then choose *Max of Count* under *By value*.

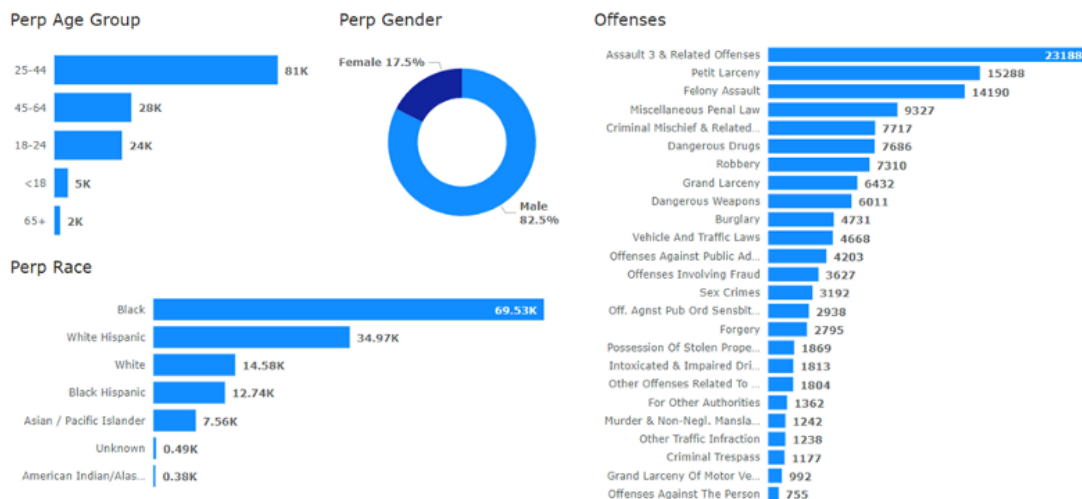


Interactions for these cards have been disabled, so they will not be affected by any changes made to filters.

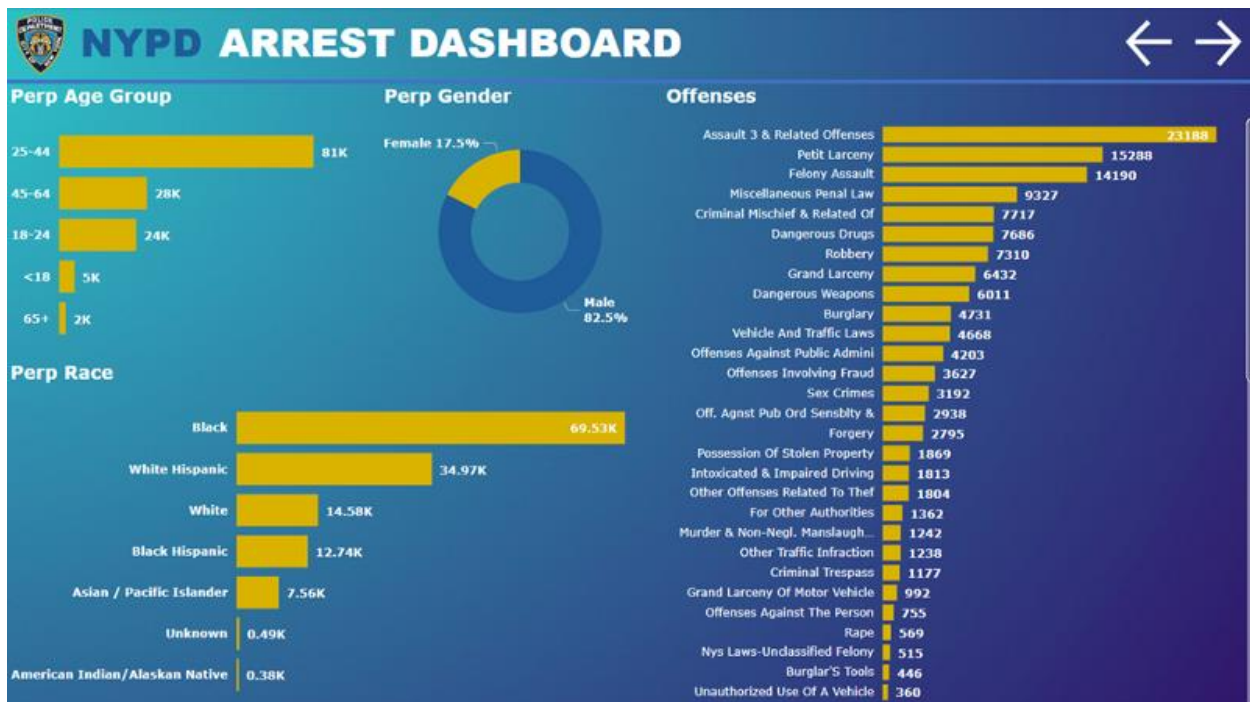
Brooklyn Wednesday

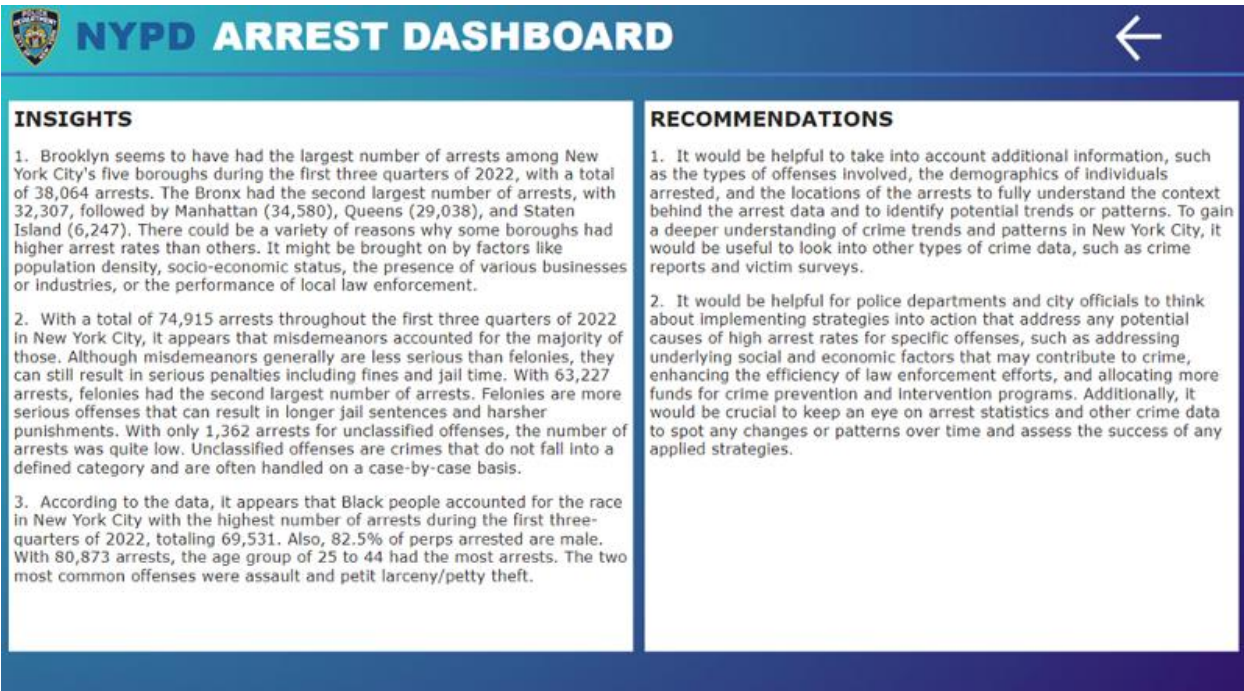
Borough with Most Arrests Highest Number of Arrests

8. The next dashboard, called the “Perp Dashboard,” displays information about the demographics of the perpetrators. It includes two stacked bar charts for the Perp Age Group and Perp Race, a pie chart for the Perp gender, and a bar chart that shows all offenses. All of these charts are used to visualize the data about the perpetrators.



9. All charts are now placed in the dashboard. After editing the theme and the background, here are the final dashboards:





Insights

1. Brooklyn has the largest number of arrests among New York City's five boroughs during the first three quarters of 2022, with a total of 38,064 arrests. The Bronx had the second largest number of arrests, with 32,307, followed by Manhattan (34,580), Queens (29,038), and Staten Island (6,247). There could be a variety of reasons why some boroughs had higher arrest rates than others. It might be brought on by factors like population density, socio-economic status, the presence of various businesses or industries, or the performance of local law enforcement.
2. With a total of 74,915 arrests throughout the first three quarters of 2022 in New York City, it appears that misdemeanors accounted for the majority of those. Although misdemeanors generally are less serious than felonies, they can still result in serious penalties including fines and jail time. With 63,227 arrests, felonies had the second largest number of arrests. Felonies are more serious offenses that can result in longer jail sentences and harsher punishments. With only 1,362 arrests for unclassified offenses, the number of arrests was quite low. Unclassified offenses are crimes that do not fall into a defined category and are often handled on a case-by-case basis.
3. According to the data, it appears that Black people accounted for the race in New York City with the highest number of arrests during the first three quarters of 2022, totaling 69,531. Also, 82.5% of perps arrested are male. With 80,873 arrests, the age group of 25 to 44 had the most arrests. The two most common offenses were assault and petit larceny/petty theft.

Recommendation

It would be helpful for police departments and city officials to think about implementing strategies into action that address any potential causes of high arrest rates for specific offenses, such as addressing underlying social and economic factors that may contribute to crime, enhancing the efficiency of law enforcement efforts, and allocating more funds for crime prevention and intervention programs. Additionally, it would be crucial to keep an eye on arrest statistics and other crime data to spot any changes or patterns over time and assess the success of any applied strategies.

References

NYC Open Data (2022). GIS data: Boundaries of Police Precincts. Police precincts.
<https://data.cityofnewyork.us/Public-Safety/Police-Precincts/78dh-3ptz>

Police Department (NYPD) (2022). NYPD Arrest Data (Year to Date).
<https://data.cityofnewyork.us/Public-Safety/NYPD-Arrest-Data-Year-to-Date-/uip8-fykc>

Visit my website: <https://bit.ly/neilmartinez>

BI Portfolio: <https://bit.ly/neilmartinez-novypro>

Check out my other blogs: <https://bit.ly/neilmartinez-medium>

LinkedIn: <https://bit.ly/neilmartinez-linkedin>