

Jay Mudgal

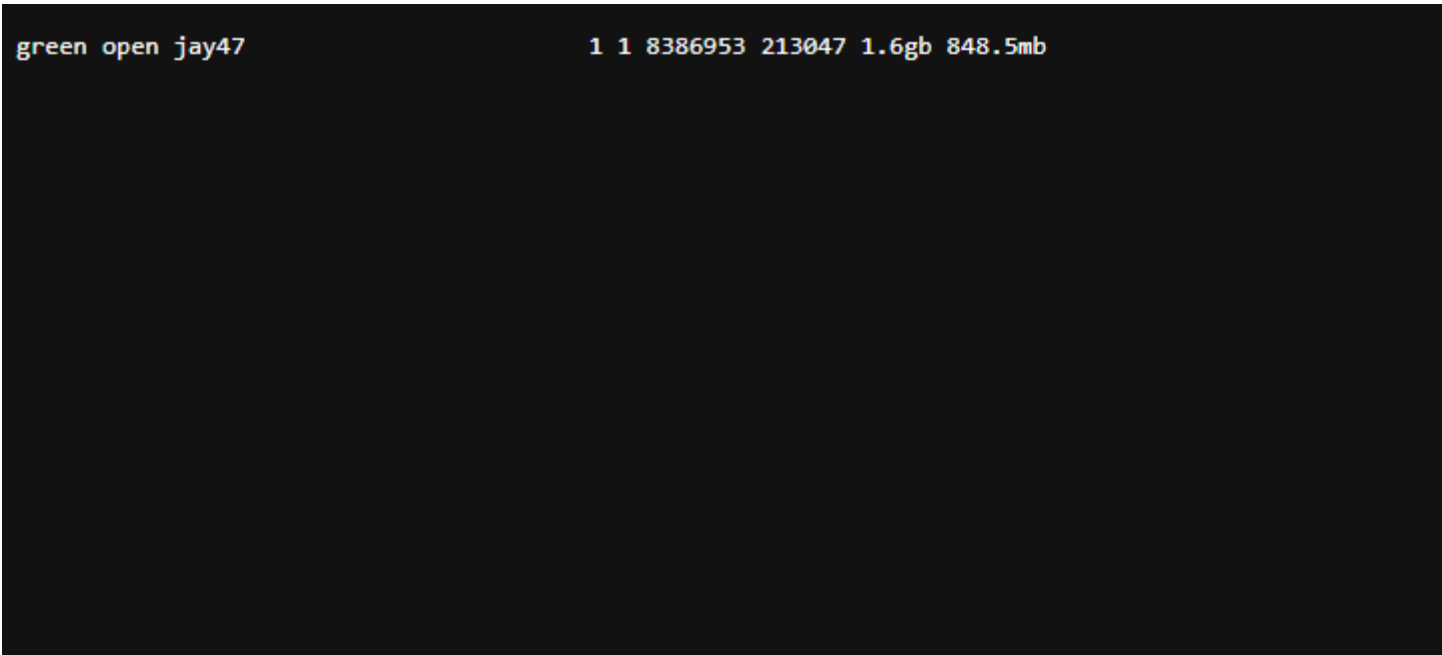
Business Analyst || Data Analyst || Project
Manager.

OpenSearch and Kibana Visualization

We have utilized AWS Kibana to visualize data on Fire incidents sourced from NYT APT. To extract the data using the API, they created a Python script named "Dockerfield, Main.py" **(the code for which is available upon request via email)**.

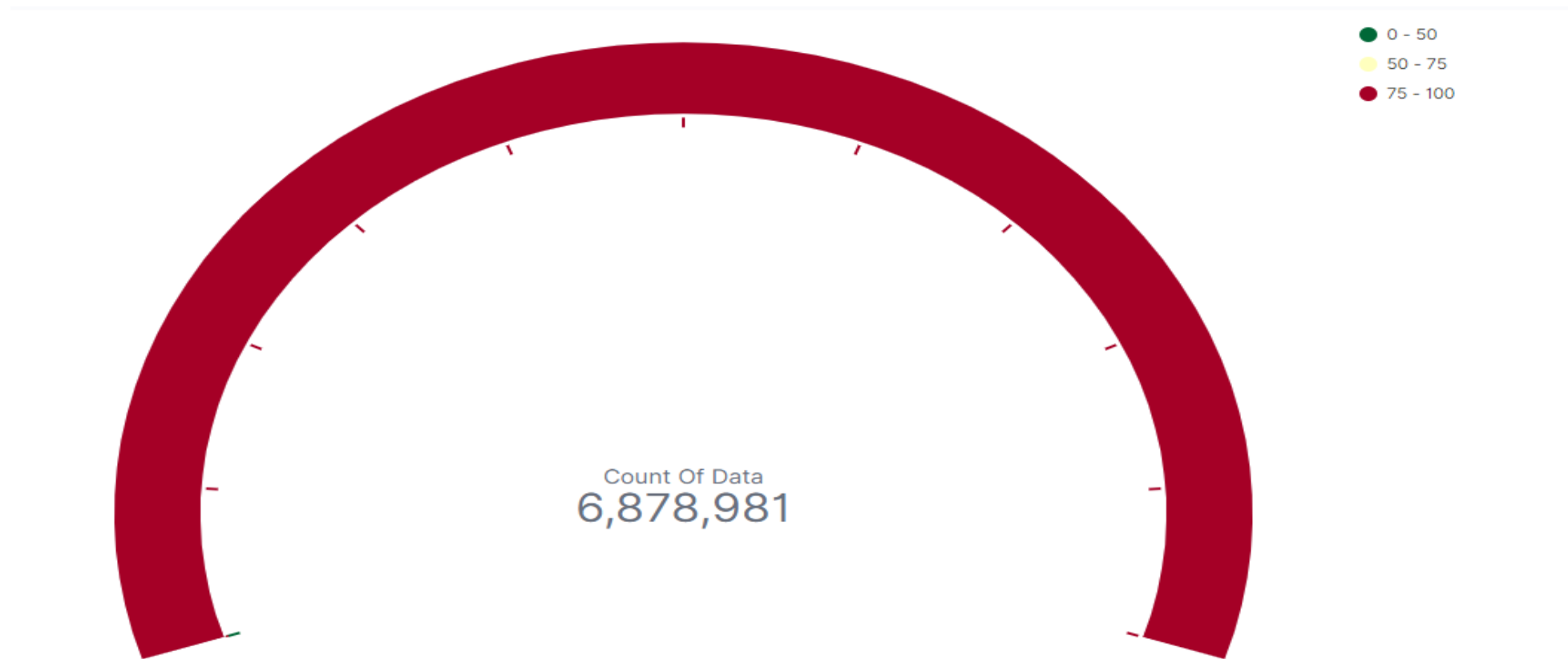
Subsequently, the extracted data was uploaded onto OpenSearch, enabling the team to leverage AWS Kibana for visualization purposes.

We have extracted 8M data points with the API but for the visualization, we have used the Index with the 6.8M data points.

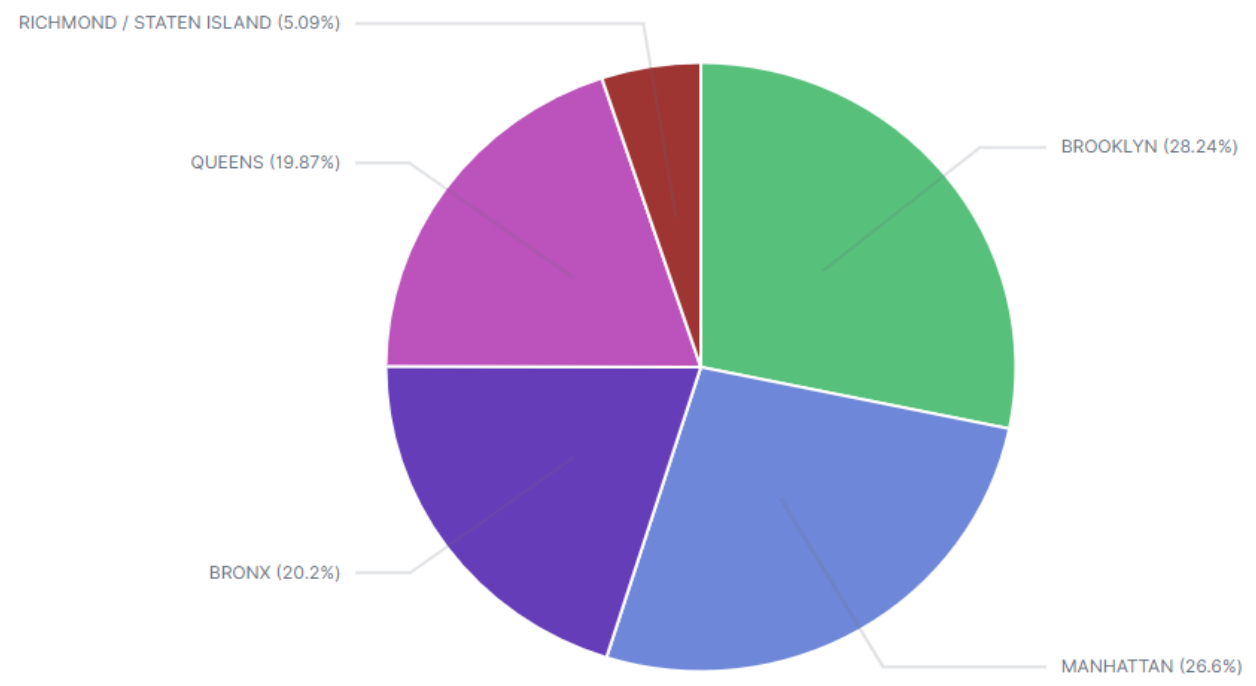


The entire dataset was migrated to OpenSearch for further analysis. Utilizing AWS Kibana, various visualizations were created to gain insights and deeper understanding of the data.

We fetched 6878981 Data points

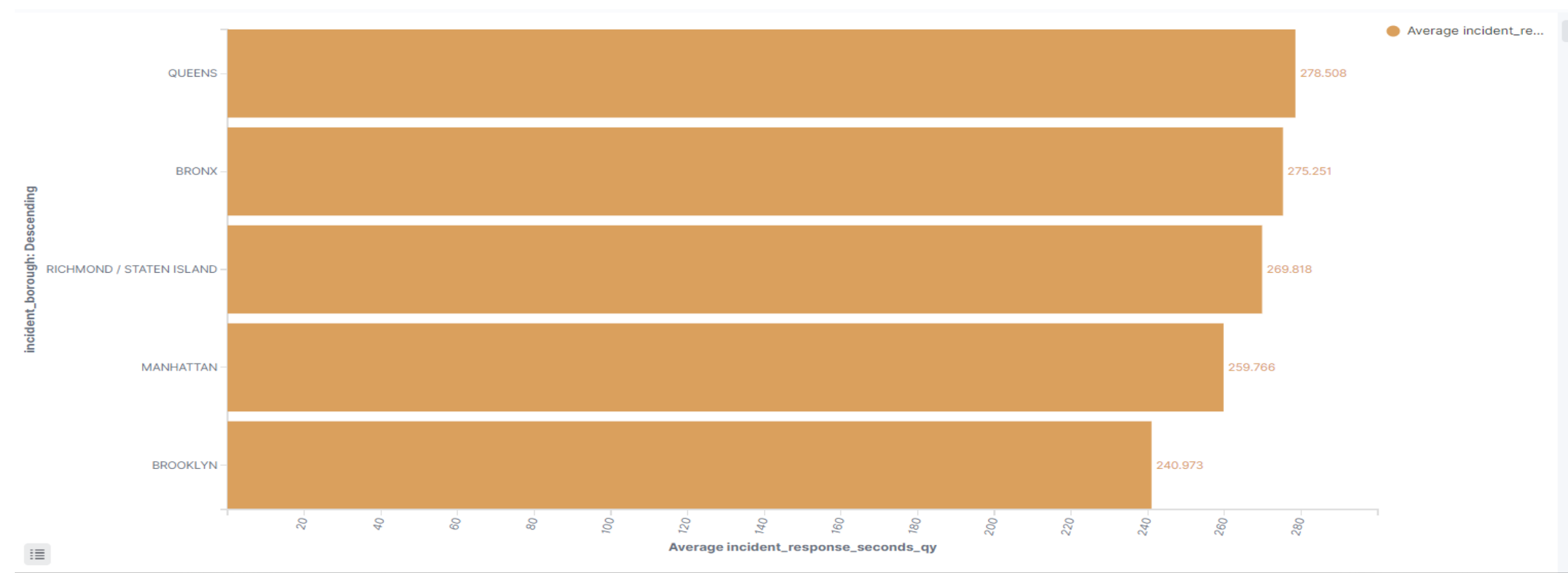


Based on the analyzed data, which area of New York City should the fire department prioritize for improving their infrastructure?



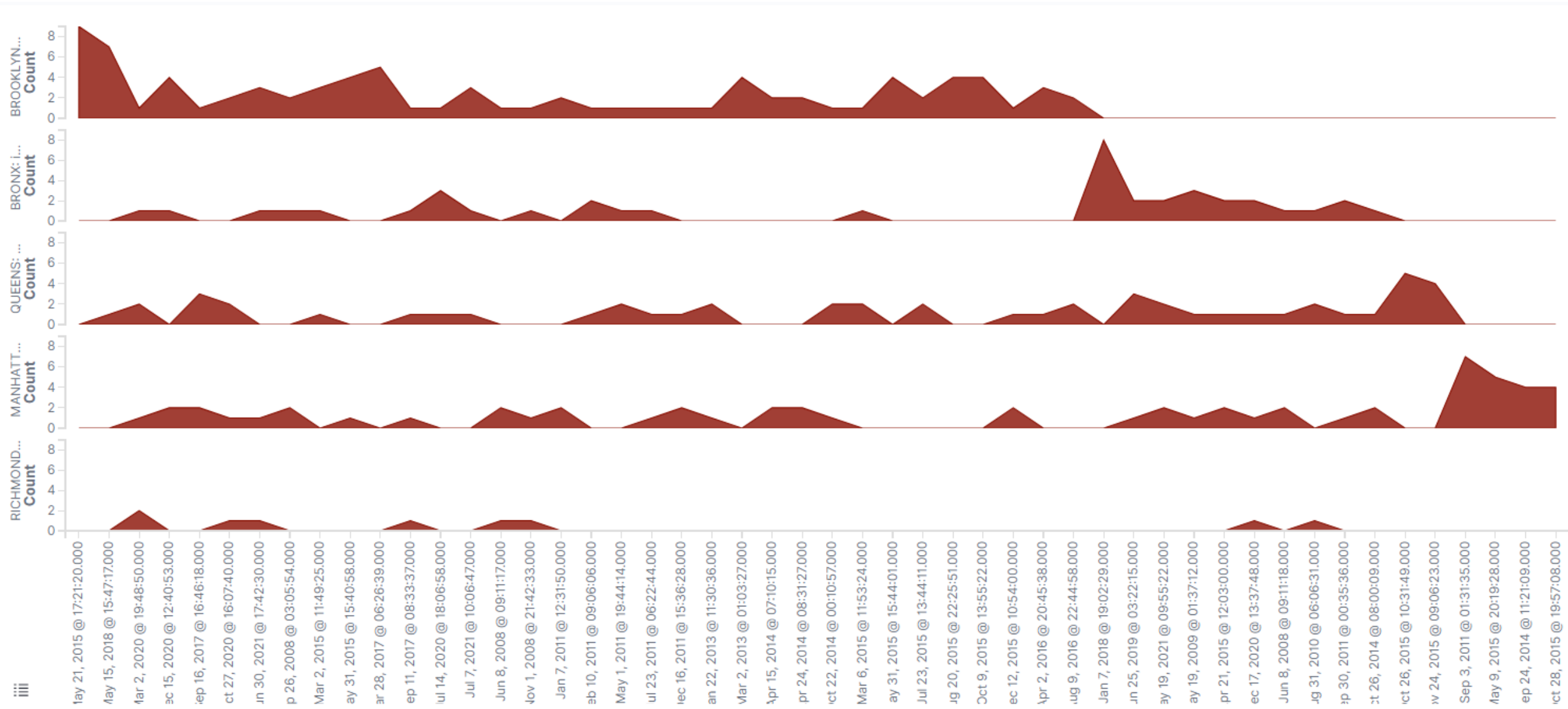
The pie chart clearly illustrates that over **54%** of all recorded fire incidents took place in **Brooklyn** and **Manhattan**. This highlights the need for increased forestation and infrastructure development in these areas by the fire department.

Considering response time, in which area of New York City should the fire department focus on improving?



Upon analysis of the simple vertical bar chart, it is evident that the fire departments in Brooklyn and Manhattan have the fastest average response time. However, there is a need to improve the average response time in **Queens** and the **Bronx**, as it is significantly higher in comparison to the other areas.

Perform a time series analysis to identify patterns of frequent fire incidents based on the areas in New York City.

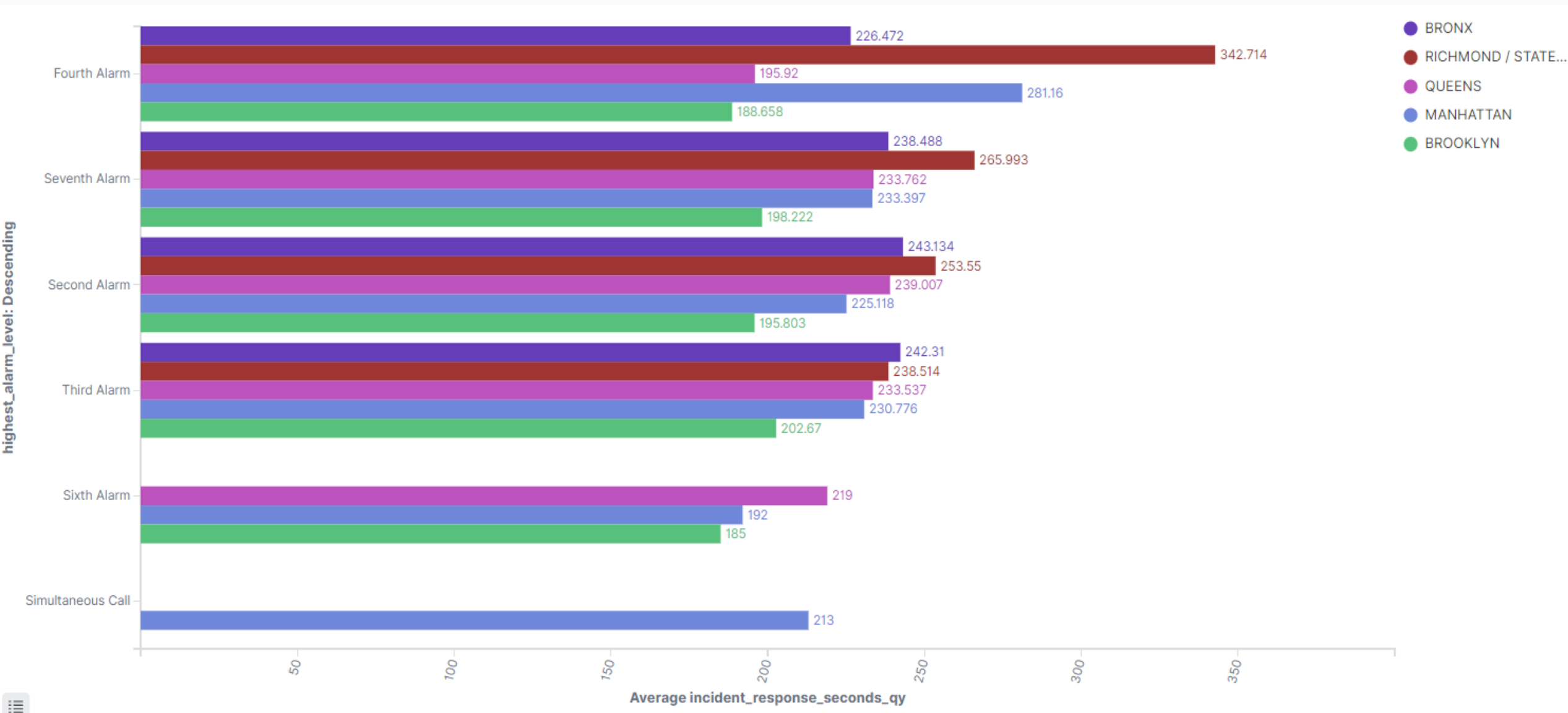


Perform a time series analysis to identify patterns of frequent fire incidents based on the areas in New York City.

Based on the above graph, it can be observed that over time, the number of fire incidents in Manhattan has decreased. This suggests that the fire department may have implemented new policies and regulations to prevent such incidents.

However, the situation in Brooklyn is different as the number of fire incidents has increased compared to other areas over time. As a result, the fire department should consider issuing new policies and guidelines to address this issue.

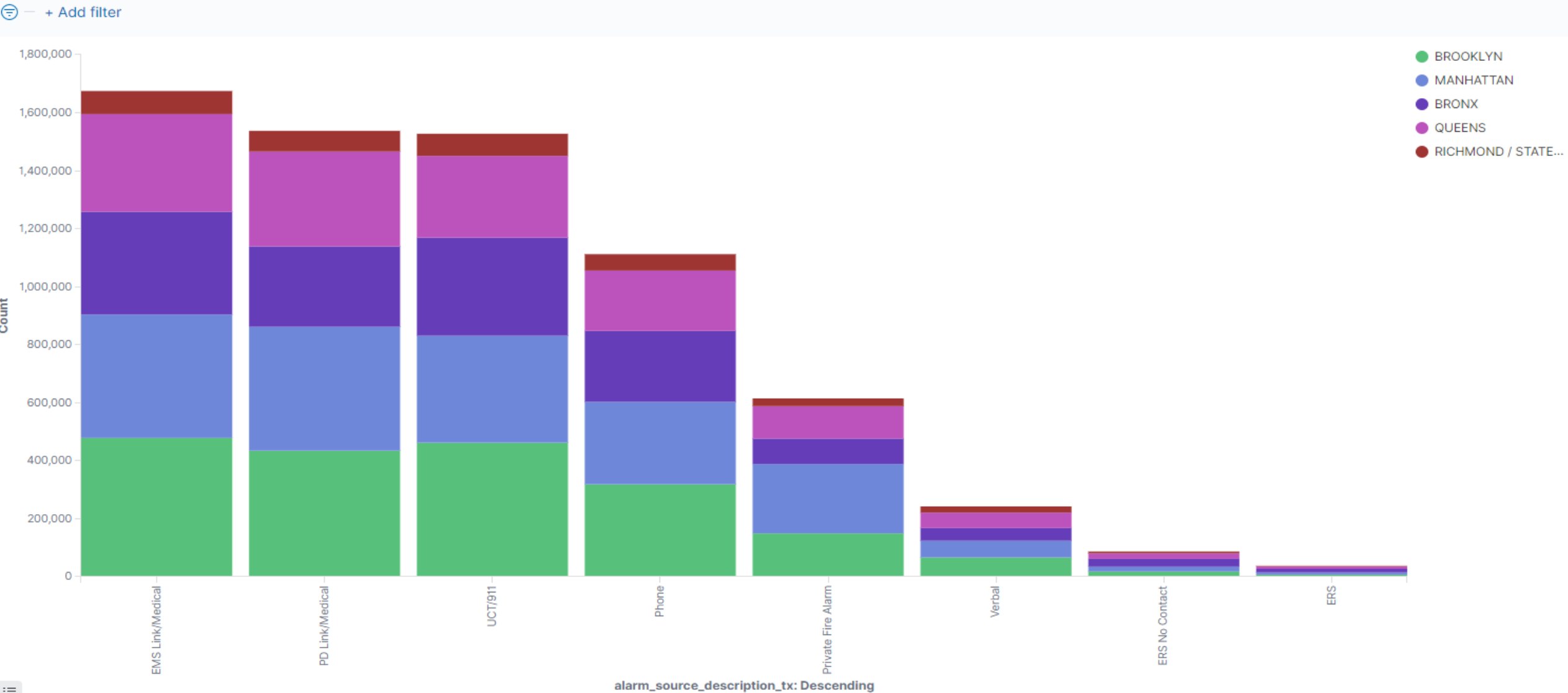
Create a visualization to showcase the distribution of fire alarm levels across different areas in New York City.



Create a visualization to showcase the distribution of fire alarm levels across different areas in New York City.

- Based on the above visualization, it can be inferred that the highest number of fire incidents reported in **Richmond** is at the Fourth Alarm level.
- The data reveals that the Bronx has the highest number of fire incidents reported at the Third Alarm level.
- Simultaneous call alerts are from only Manhattan

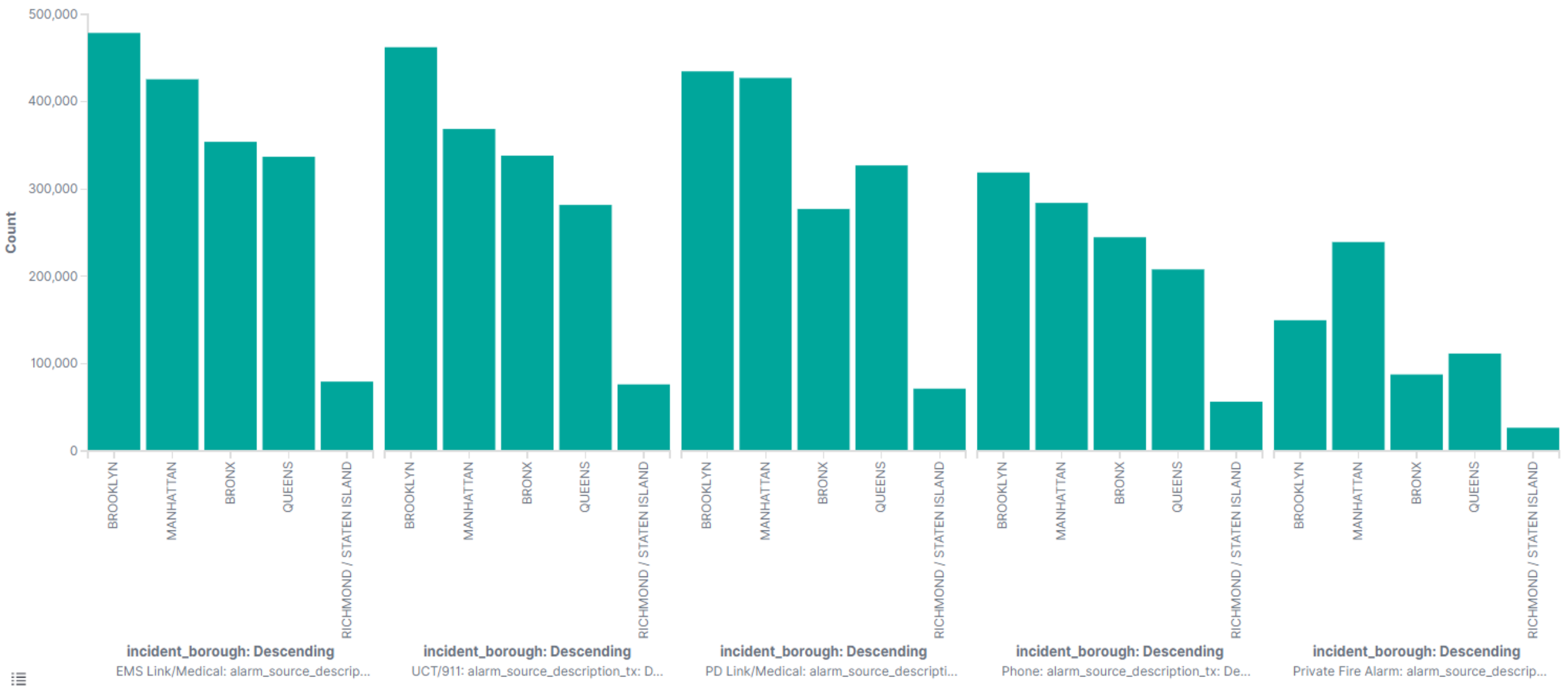
Create a visual representation depicting the number of fire incidents in New York City categorized by area, with further subcategorization based on the source of description.



Create a visual representation depicting the number of fire incidents in New York City categorized by area, with further subcategorization based on the source of description.

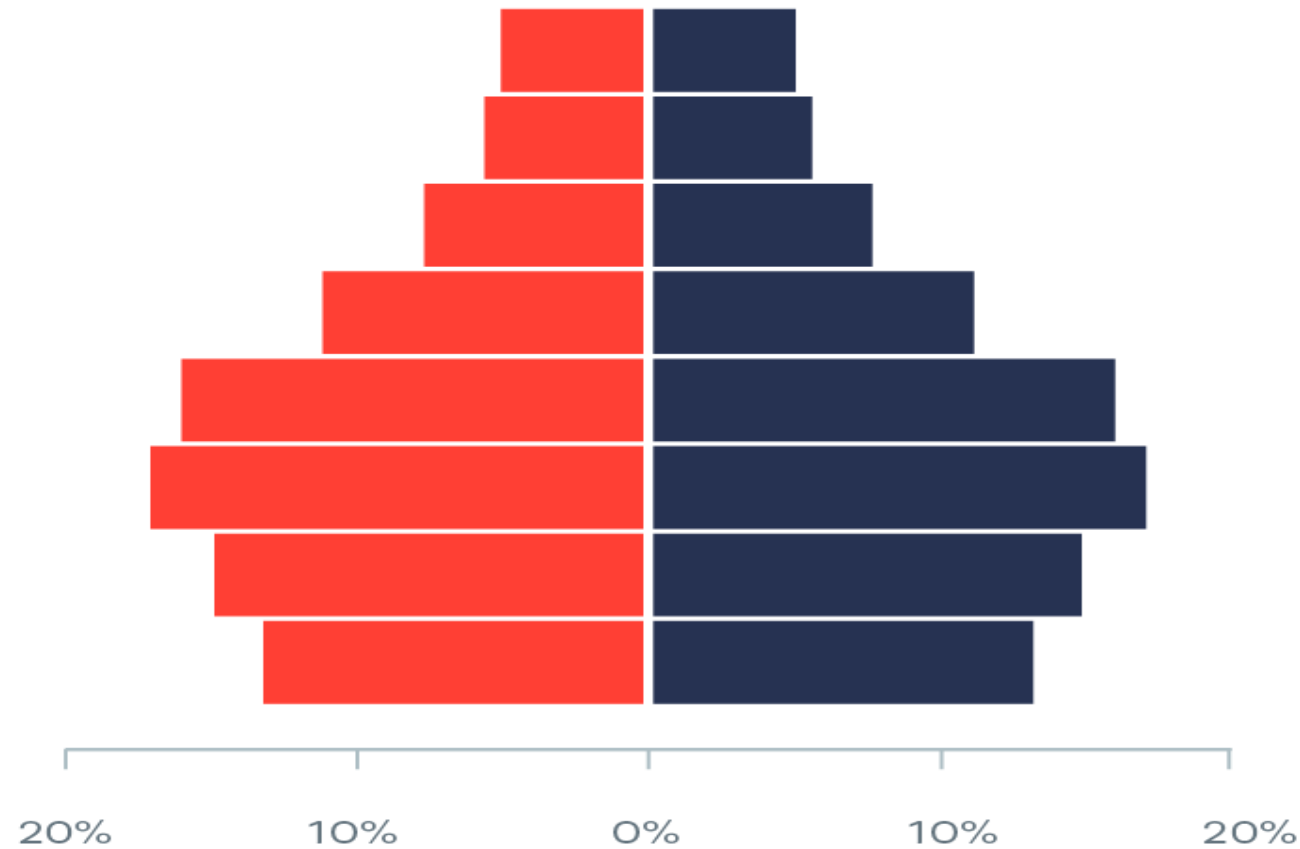
- We can see that the number of sources of description is all most the same in all the areas.

Create a visual representation illustrating the count of fire incidents in New York City and categorize it further based on the source of description.



- Staten Island has the least incidents across all boroughs for all sources of reporting of incident

To perform a comparative analysis between two distinct areas, the data will be subdivided based on a predetermined classification system



The above visualization tells the count percentage of the fire incident in Manhattan and Brooklyn and divides it on the bases of classification

The chart displays the daily number of new COVID-19 cases in the United States from October 17 to November 11, 2019. The y-axis represents the number of cases, ranging from 0 to 100,000. The x-axis shows dates from 10-17 to 10-19, 2019, and 10-21 to 10-23, 2019, and 10-25 to 10-27, 2019, and 10-29 to 10-31, 2019, and 11-03 to 11-05, 2019, and 11-07 to 11-09, 2019, and 11-11, 2019. The chart shows a significant peak in cases around November 3, 2019, reaching over 100,000 cases per day. The data is broken down by state, with the top states being California, New York, and Texas.

