



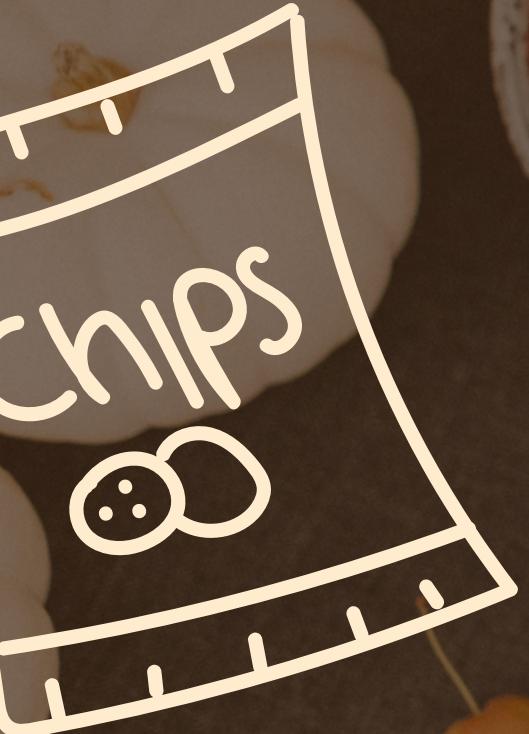
PAUCEK  
AND LAGE



# PROJECT 2

Mohammed sabt  
Ali Sameer

WWW.REALLYGREATSITE.COM



# OBJECTIVES

- 1 Data Cleaning & Preparation
- 2 Exploratory Data Analysis (EDA)
- 3 Visualization & Reporting

# PROBLEM STATEMENT

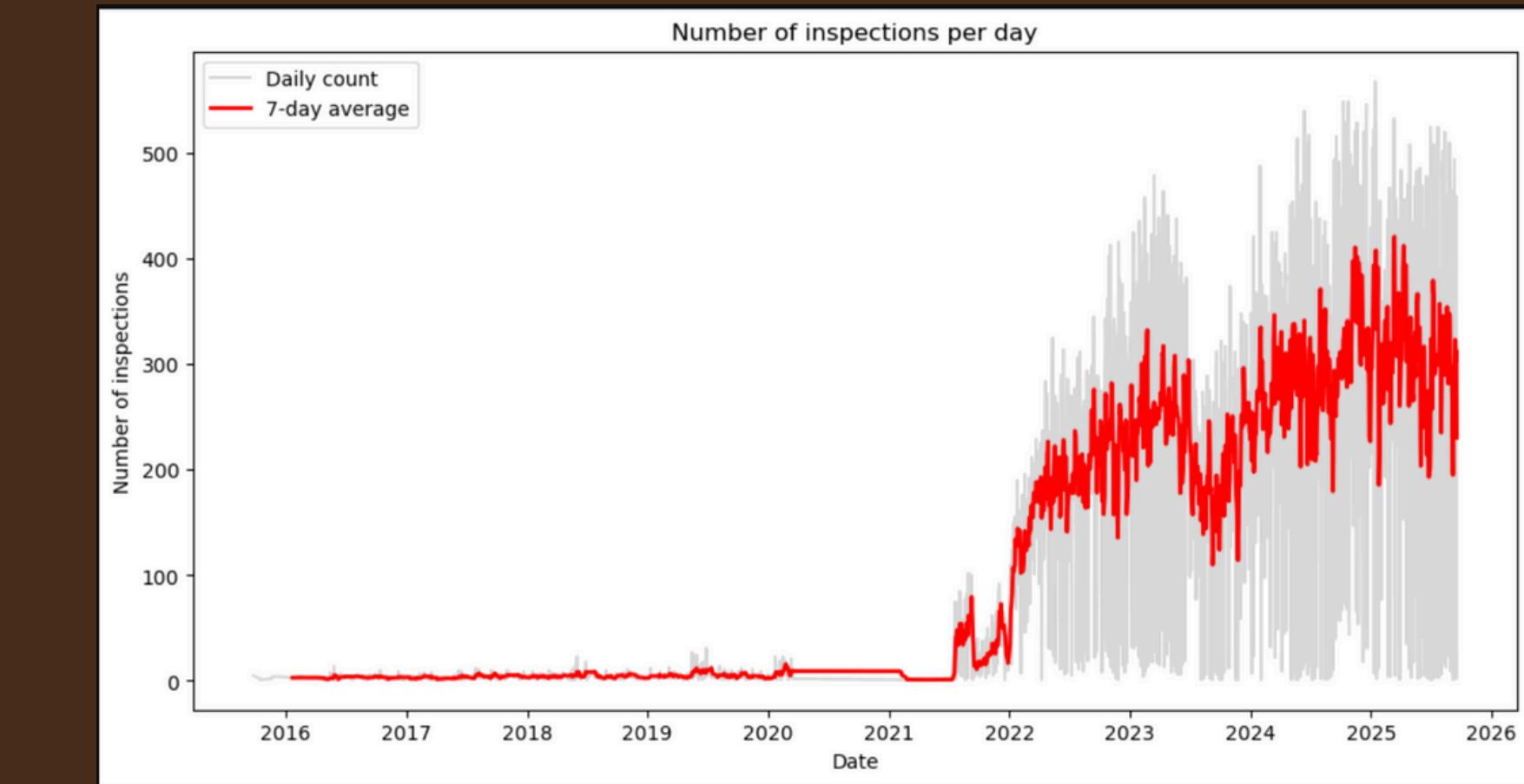
Food safety violations in NYC restaurants remain a public health risk, and inspection data is underutilized for identifying trends.



# FIRST STEP

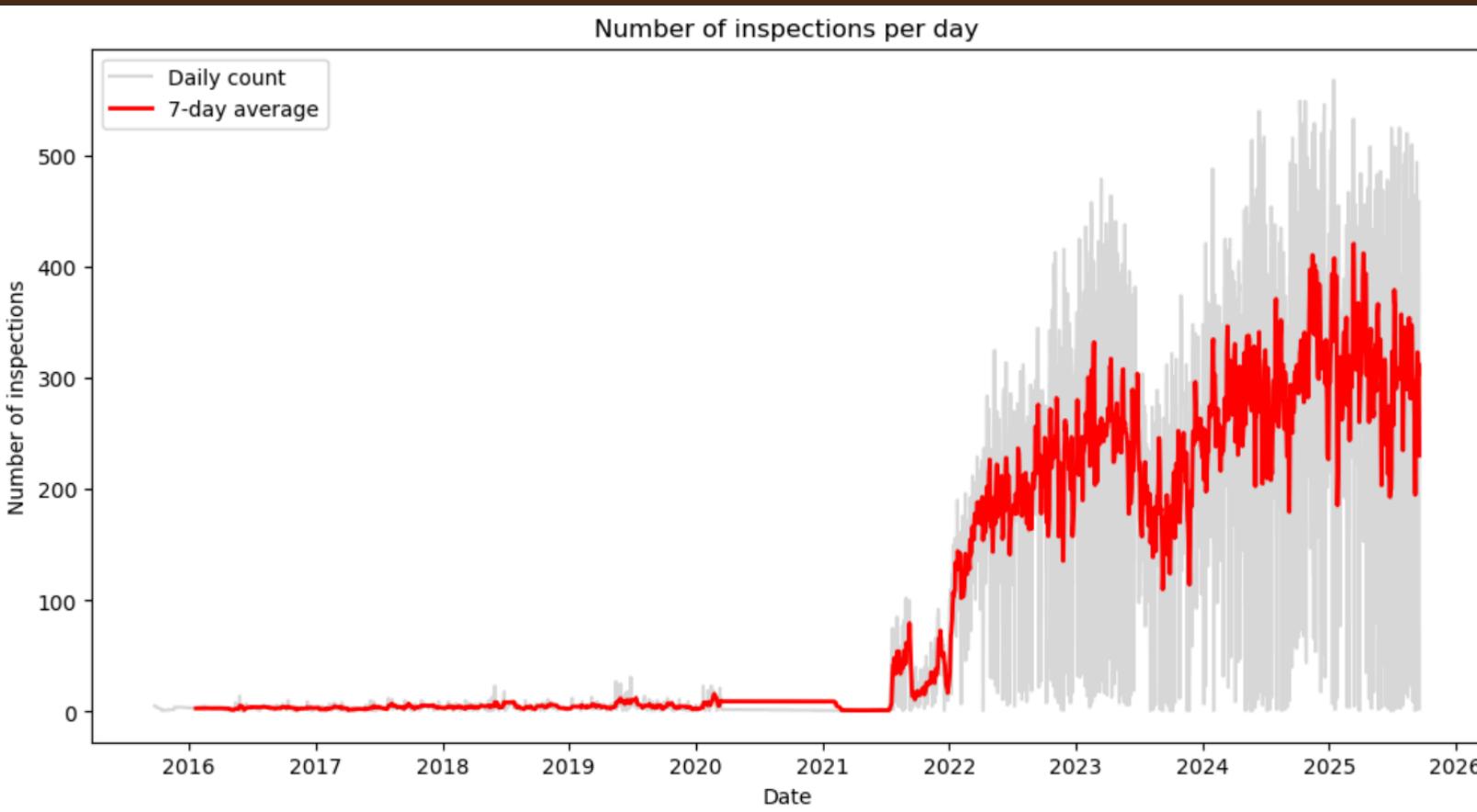


This graph shows the number of inspections over the years. Where the red line represents the rolling average while the grey line is the daily number of inspections.



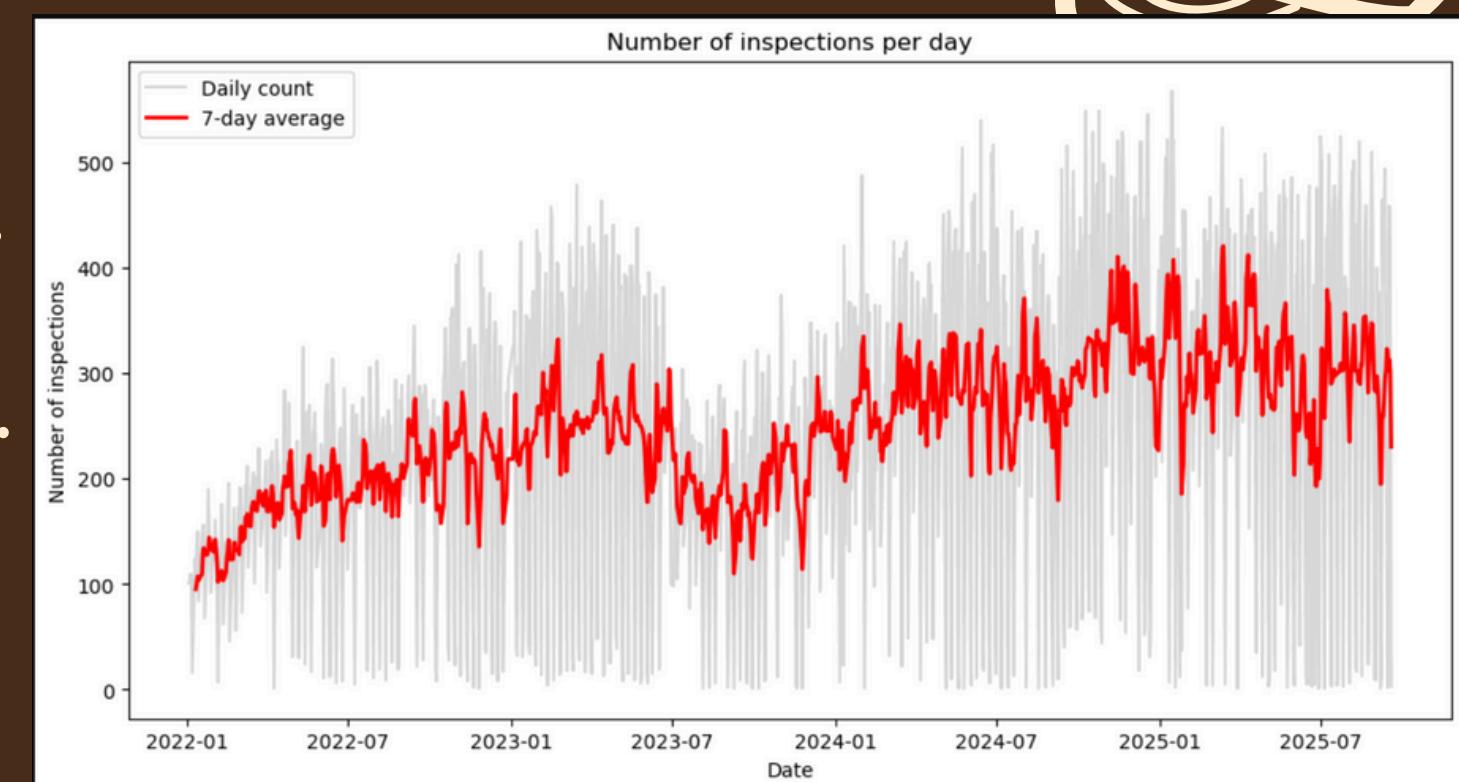
# CONTINUE

- 1 It can be seen that number of inspections has increased in the year 2022 which possibly means that there no proper records before that year.

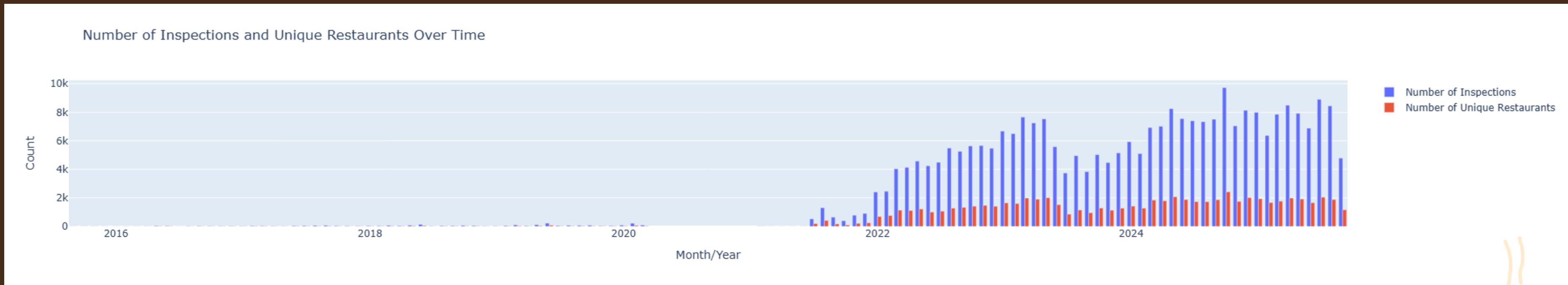




- 1 So, this graph shows the years after 2022.
- 2 The trend line is more obvious than before.
- 3 Other graphs are plotted using these dates.

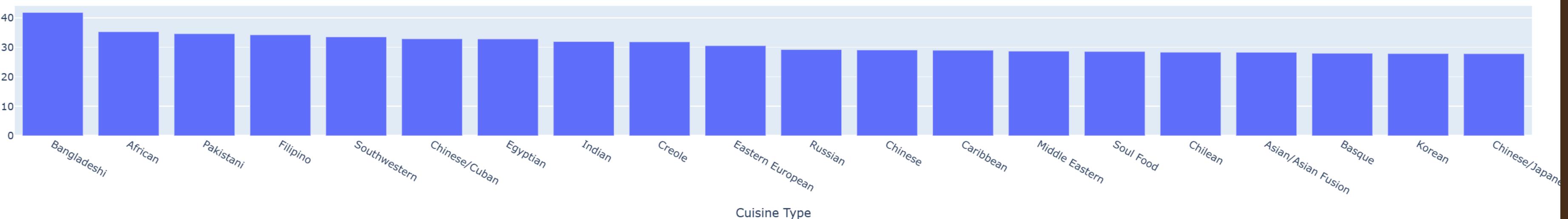


However, the data before the year 2022 do not show lots of records for restaurants which explains the low count of the inspection

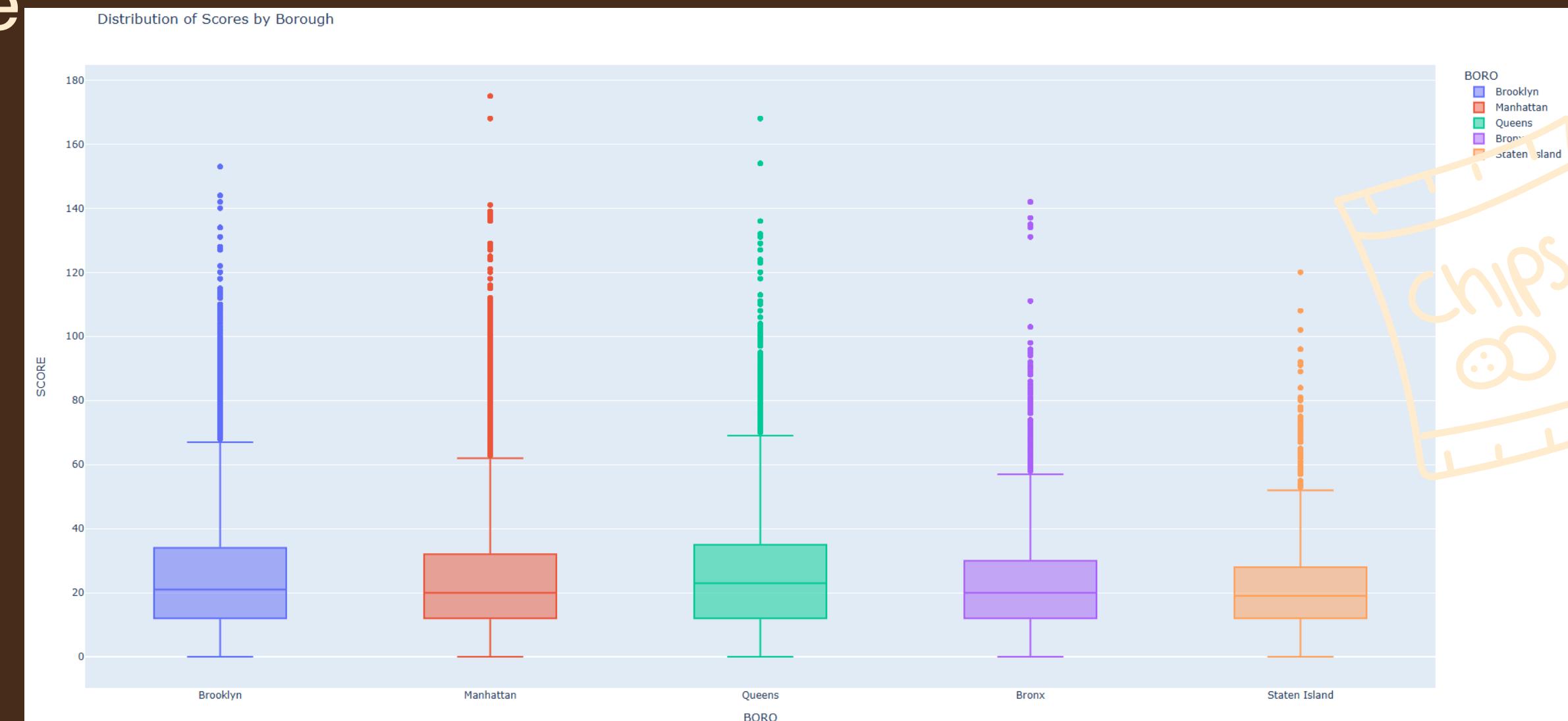


This graph shows the average score by cuisine. These are the top 20 cuisine with highest score average, it can be seen that Bangladeshi cuisine has the high average score,

Average Score by Cuisine Type (Top 20)



In this graph we can see  
a box plot of the  
distribution of the  
scores across the  
borough(BORO).





In conclusion, the data set might have some issues with the data before the year 2022, however we have seen that it could be just poor data collection in those years affected by the missing restaurant records. Also, the Bangladeshi cuisine seems to have the highest score.