NYC Restaurant Health Check:
Exploring DOHMH Inspection Results Dashboard

By

Tipu Sultan

Fordham University

Data Visualization Capstone Project

Spring 2023

Abstract:

Though it is renowned for its glamor and celebrity draw, the New York City restaurant industry faces significant difficulties maintaining a well functioning and hygienic backdrop. This project explores the myriad of hygiene challenges faced by the restaurants in New York City, and how they have changed over time, whether for better or for worse. By analyzing violation citations from restaurant and college cafeteria inspections over the years, we will identify key factors that contribute to a successful hygienic restaurant, why certain restaurants can't seem to stay open, and provide our suggestions for the successful operation of a Grade A restaurant. Our study will analyze the letter grading system (A-P), what these values mean and and how the inspection process and outcomes are analyzed and rated. The goal of our project is to provide valuable insights for restauranteurs, in addition to providing valuable information about where to eat for ordinary New Yorkers where is it safe? By utilizing both exploratory and explanatory charts, we gained a deeper understanding of the restaurant landscape in New York City and were able to answer many of our key questions we asked at the beginning of the project. We have been able to identify the most prevalent cuisine type, common violations, trends in inspection scores, and factors that correlate with inspection scores. Our visualizations explore the intricacies of the preceding factors in the hopes of identifying nuances that we believe will be invaluable for current and future restaurateurs and foodies alike.

Introduction:

In a city that has recently declared a war on rats, the hygiene problems facing our city have become ever so startling to local New Yorkers[1]. We all have, from one time to another, walked past our favorite local coffee shop, only to notice the startling gigantic yellow "C" displayed promptly in the front window, visible to passerbyers. This grade may perhaps catch your eye as a 70% isn't typically promising, but you still go into the restaurant, order your favorite coffee, as you do every day, and carry on. But what really does this grade mean? What are the contributing factors to the glaring "C" in the coffee shop window? In this project, we provide a comprehensive analysis of New York City's restaurant industry, focusing in particular on investigating the various hygiene facets of the industry with a strong emphasis on the food inspection process and the criteria for passing said inspections. Through exploring the data, we identify trends in the types of restaurants that consistently pass inspections in addition to those who struggle to meet the requirements. Simply stated, our overarching aim is to determine the factors contributing to the success or failure of a food operating establishment.

Background:

In order to have a clear understanding of the key indicators in this project, there is some domain information that needs to be considered. This dataset Data source provides a complete record of violation citations for active eateries and college cafeterias from the past three years in the 5 boroughs of New York City. This also includes new and violation-free establishments. It features a two-step Letter Grading Program, with grades assigned based on points and re-inspections possible after a week. Grade changes may occur after hearings or settlements; future inspections depend on prior scores. The grades are scaled based on the following measures: 0 to 13 is an A, 14 to 27 points is a B, and 28 or more points is a C. Each violation is associated with a number of points; the number of points is dependent on the type, extent, and risk posed to the public. The points are added together for an inspection score. Lower inspection scores indicate better compliance with the Health Code. A restaurant is typically closed should it receive 28 or more points three consecutive times[2]. There are also several violation codes that correspond to different violations. These violations are subsequently categorized into critical and non-critical; critical

violations being those factors which have been identified as violations that could potentially cause food-borne illness and non-critical violations are typically maintenance or non-food surface sanitation issues that are not likely to be the cause of a foodborne illness[3].

During the Visualization:

Discusses the visualizations designed, and their values. For example, they may help you validate certain hypotheses. If you use some data sets, you should also describe the data sets, and any other details related to the visualizations and any hypotheses validated by the visualizations. A discussion of the results may be included, or they could be included in a separate discussion section. Developed visuals: Description of them, relevant insights, or/and discussions. Our visualization process began with an extensive data cleaning initiative led by Tipu Sultan. There was a necessity for this as there typically is with real world data. This turned out to be the most grueling aspect of the visualization process. This involved:

1. The number of restaurants was specified for each of the five boroughs: Manhattan, Brooklyn, Bronx, Queens, and Staten Island.
2. The distribution of grades A, B, and C was indicated across these boroughs from 2018 to 2023.
3. Mapbox was utilized to display the evolving distribution of restaurants, grouped by grades A, B, C, D, P, and Z, over the past year.
4. Line charts were employed to showcase cleanliness trends throughout the years.
5. The top five cuisines in New York, ranked by violation scores, were identified: American, Italian, Chinese, Pizza, and Coffee/Tea. American cuisine had the highest violation score, while Coffee/Tea had the lowest.
6. The top 10 violations were illustrated based on severity.
7. An overview of critical and non-critical violations in New York restaurants was provided.
8. Notable increases in critical flags were displayed, along with forecasting for future potential spikes in red flag increments.
9. A trend line demonstrated the exponential growth of violations month by month from 2018 to 2023.

10. The correlation between violation scores and inspections for the top 10 cuisines was examined. A linear relationship was observed between the count of inspections and the sum of violation scores.
11. The top three popular restaurants—Dunkin, Starbucks, and Subway—were compared in terms of their violation scores relative to the average number of annual violations.
12. The months during which the Department of Health and Mental Hygiene typically conducts inspections were identified, providing valuable information for restaurants aiming to prevent low grades.
13. The relationship between inspection and re-inspection cycles from 2018 to 2023 was explored, revealing the strictness of DOHMH's re-inspection process.
14. Instances of restaurant closures by DOHMH due to high violation counts were documented.
15. The reasons for Starbucks' leadership in the coffee industry were analyzed.
16. The performance of restaurants at Fordham University was examined, highlighting their grades and associated violation scores.
17. Mapbox was used to create a geographic view, allowing users to search for nearby restaurants categorized by grade and violation score.

Closer Look into the Visualizations

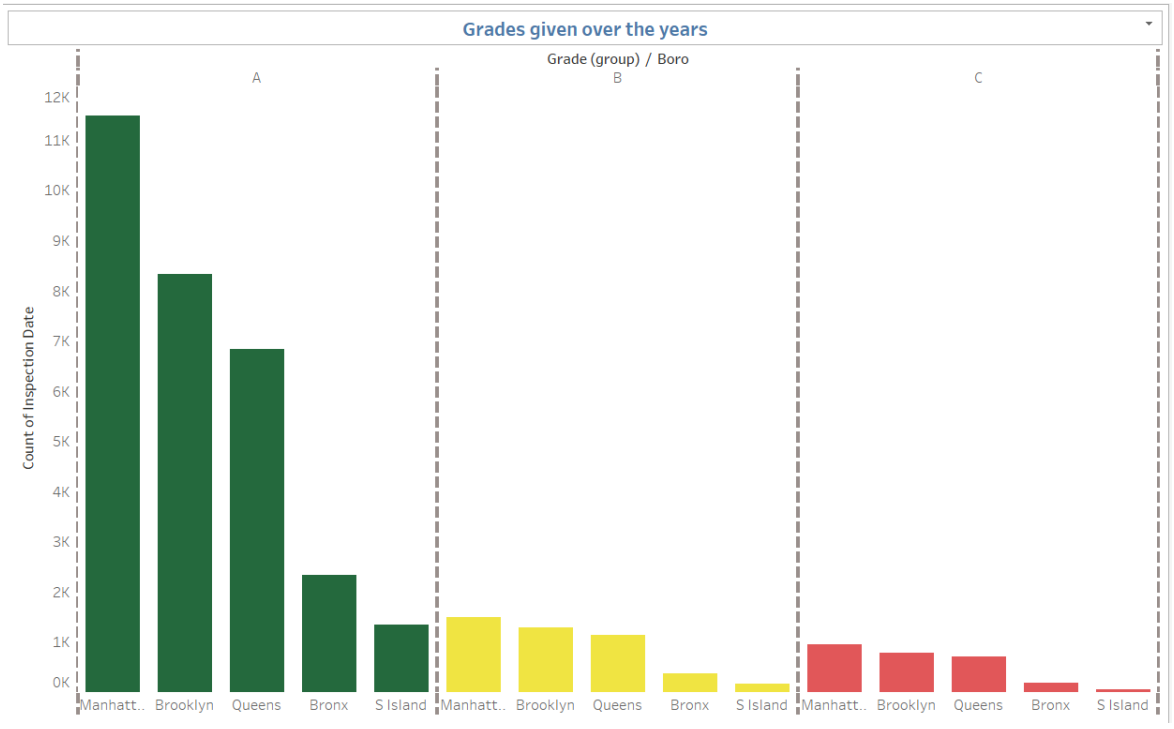
In this section we provide a closer look into all the visualizations that we provided in this project and we deliberately explain the required calculations for each visualization.

Visualization1: Number of restaurants based on each Area

Boro	
Manhattan	68.5K
Brooklyn	50.4K
Queens	43.6K
Bronx	17.3K
S Island	6.0K

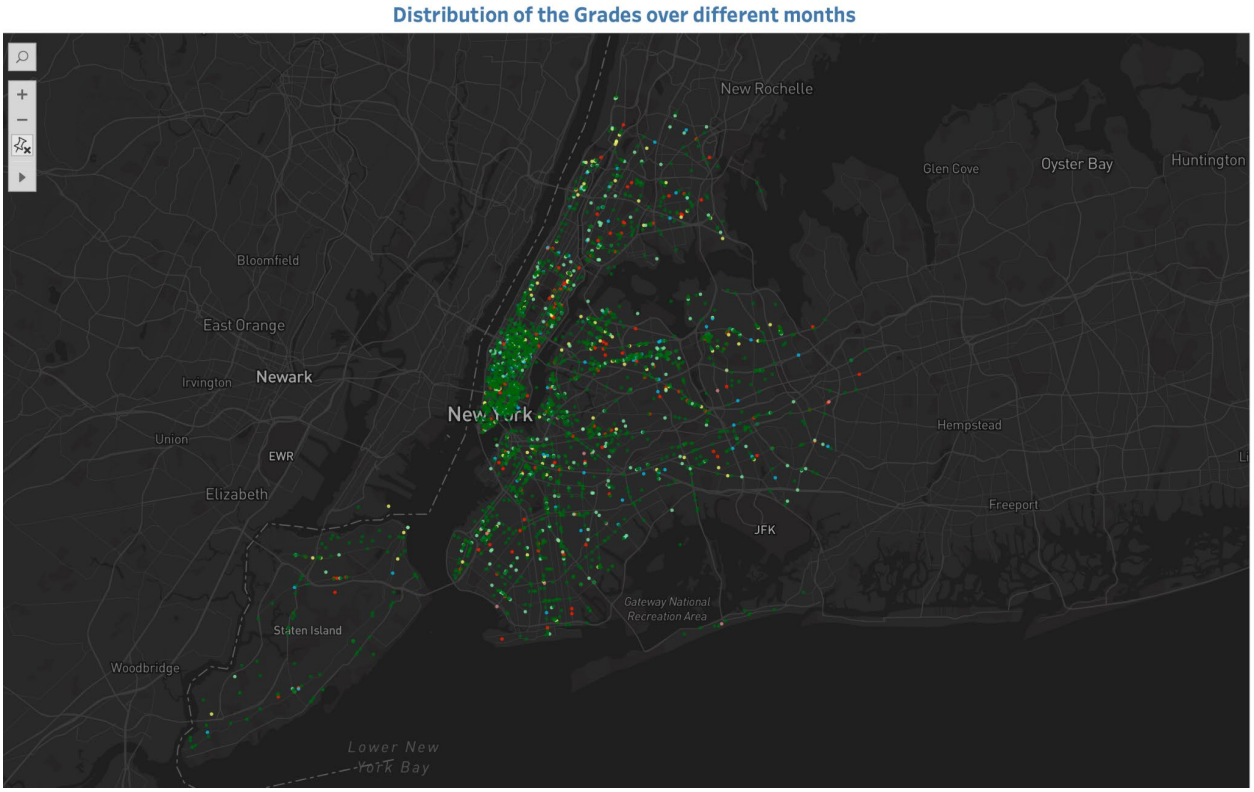
According to the table, Manhattan has the highest number of residents, with a population of 68.5k, followed by Brooklyn in second place with 50.4k residents. The remaining cities in the dataset are Queens with a population of 43.6k, the Bronx with 17.3k residents, and Stone Island with 6k residents.

Visualization 2: Grades Given Over the Year



This animated graph shows an overview of the total amount of “A”, “B”, and “C” grades handed out throughout the city over the years. This started as an exploratory visual in order to give us a quick insight into the data set and the city landscape. We were happy to see that there were more “A” grades handed out. It also allowed us to begin asking questions: what are the factors that go into each of these categories.

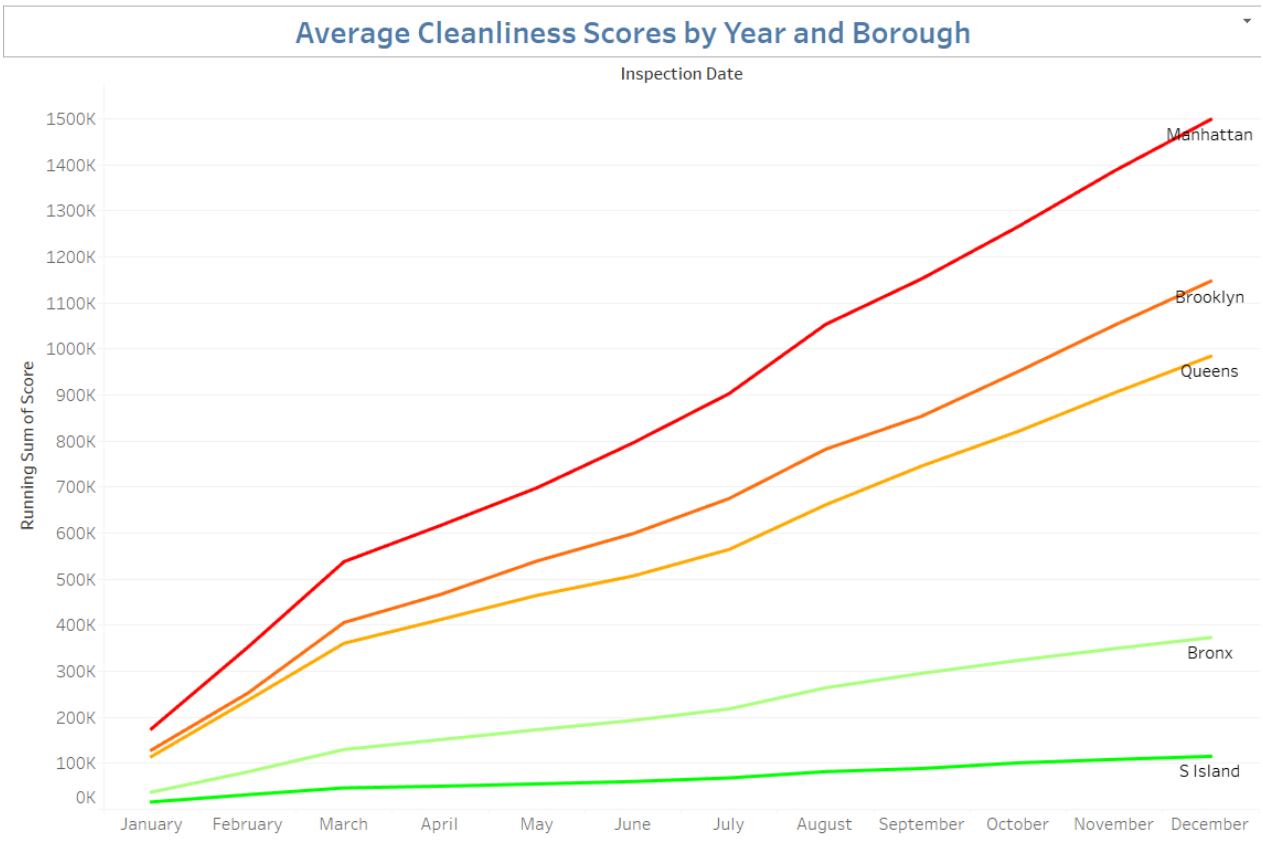
Visualization 3: Distribution of the Grades over different months



This animated map presents a vivid depiction of restaurant distribution in New York City based on their respective letter grades. In this compelling visualization, restaurants are organized into five distinct categories according to their grades. Dark green highlights the prevalence of grade A establishments across the city's five boroughs, while lighter green, blue, yellow, pink, and red correspond to grades B, C, D, P, and Z, respectively. As the animation progresses, viewers can effortlessly discern areas with grade A or grade C restaurants and also observe their improvement or decline in grades over the time in Manhattan, Brooklyn, Bronx, Queens and Staten Island. As

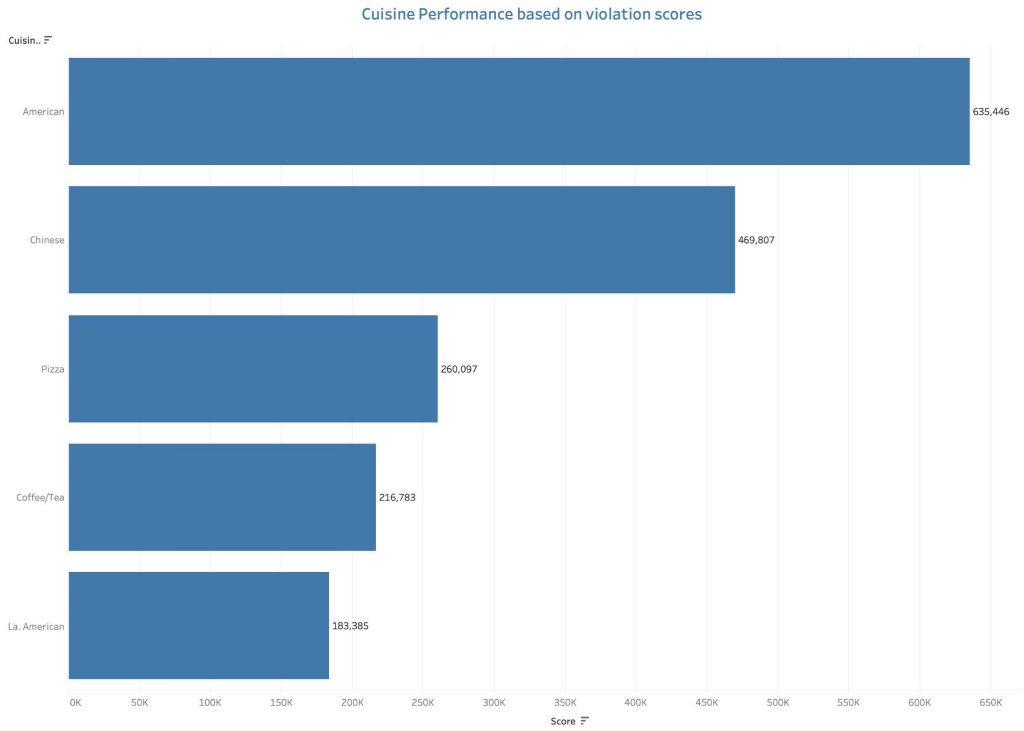
depicted in the map, the prevalence of grade A restaurants has experienced a steady rise from 2018 to 2023, signifying a continuous improvement in the quality of New York's restaurant industry.

Visualization 4: Average Cleanliness Scores by Year and Borough



This animation line graphs show the cleanliness score of the restaurants in five different boroughs based on the sum of their score. The visualization clearly reveals that among the five cities analyzed, Stone Island stands out as the cleanest. The Bronx closely follows as the second cleanest city. In contrast, Queens, Brooklyn, and Manhattan demonstrate comparatively lower levels of cleanliness when measured against Stone Island and the Bronx. The visualization highlights the inherent challenges faced by restaurants in maintaining cleanliness within densely populated areas.

Visualization 5 : Cuisine Performance over the years based on Violation Score

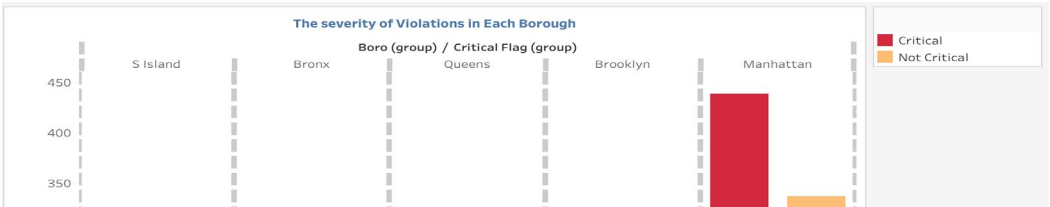


Based on the Violation Score, it is evident that American restaurants in New York City have a significantly higher violation score compared to Chinese restaurants. The American restaurants have the highest violation score, reaching 635,446. Chinese restaurants occupy the second position in terms of violations. On the other hand, Latin American, and pizza and coffee/Tea restaurants demonstrate much lower violation scores compared to Chinese and American restaurants.

Visualization 6: Most Common Violations in New York

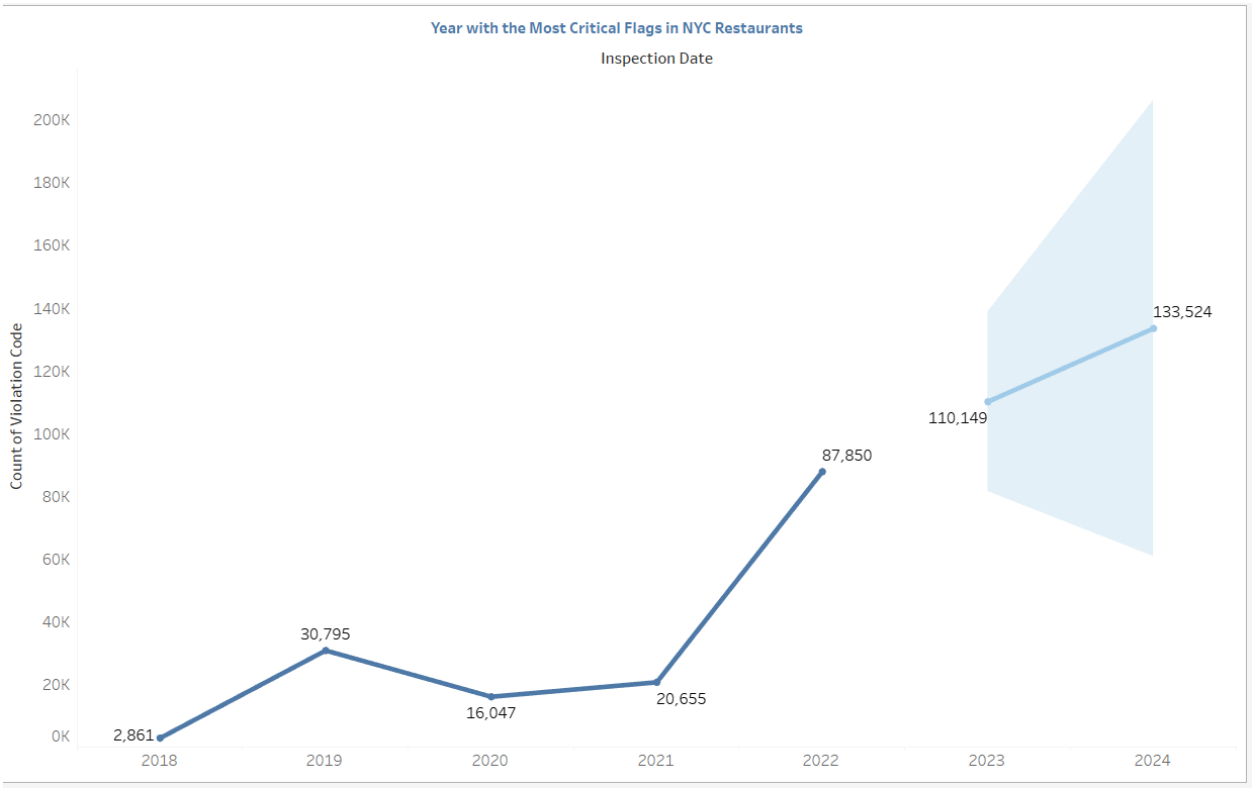
The analysis of restaurant violations in New York City reveals the top 10 most common violations. Vermin, such as insects or rodents, ranks as the most prevalent violation. Following vermin, contaminated food surfaces emerge as the second highest violation. However, the lowest occurrence is observed with mice-related violations.

Visualization 7 : severity of violations in each boroughs



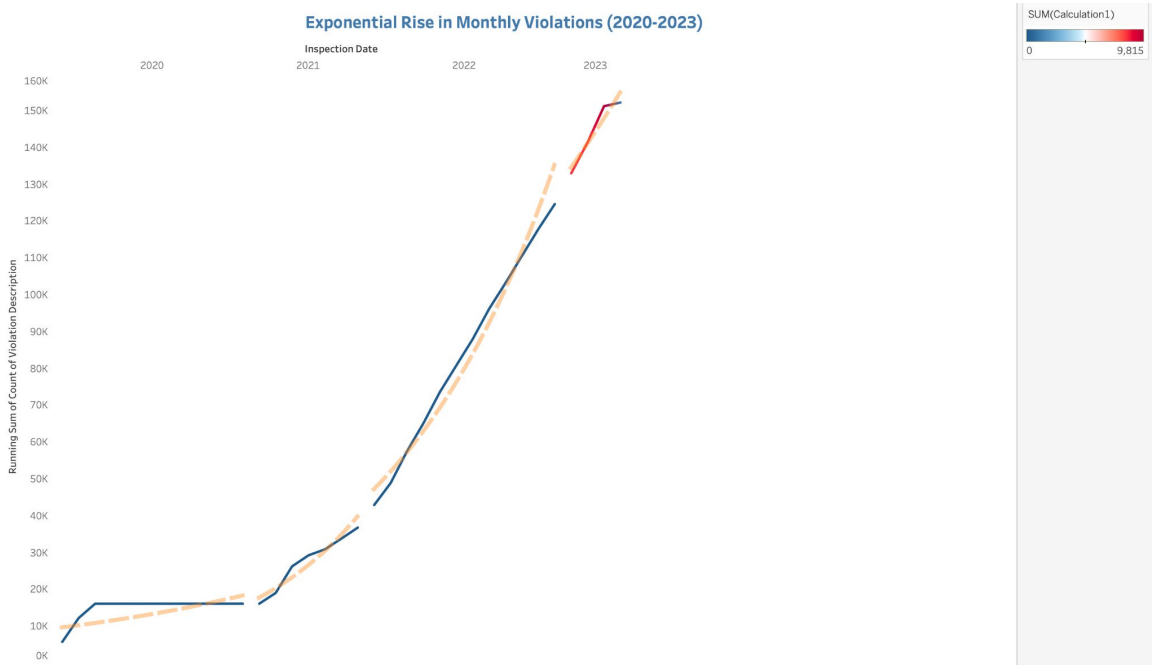
This visualization depicts the juxtaposition of Critical and Non-critical violations, strictly in terms of count. It is interesting to note the frequency at which these separate kinds of violations are handed out. It seems, for the most part, that priority is given to handing out critical violations, as those are significantly more severe violations that are attached to higher risk to public safety. So, while inspectors are on the lookout for violations, they are perhaps paying closer attention to the violations that most directly affect the public.

Visualization 8: The year with the Most Critical Flags in NYC Restaurants



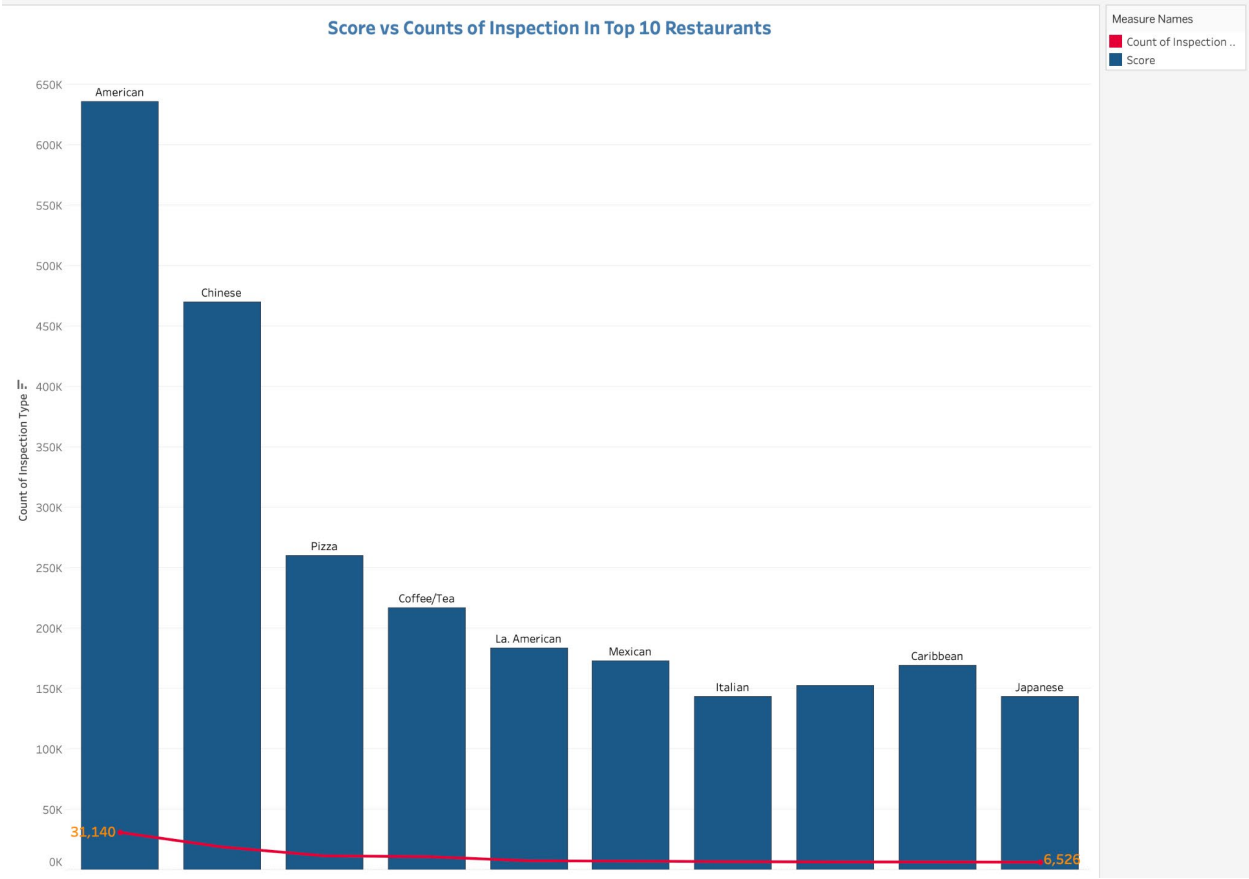
The analysis of the line graph depicting inspection rates in NYC restaurants reveals several trends. From 2018 to 2019, there was a gradual but noticeable increase in the inspection rate. However, the onset of the COVID-19 pandemic resulted in a significant drop in inspections until 2021. Following 2021, there was a sharp and rapid increase in the inspection rate. Based on these observed patterns over the past five years, it is highly likely that the future inspection rates will continue to rise. This prediction aligns with the trend of increasing focus on food safety and regulations in the restaurant industry.

Visualization 9: Exponential Rise in Monthly Violations (2020-2023)



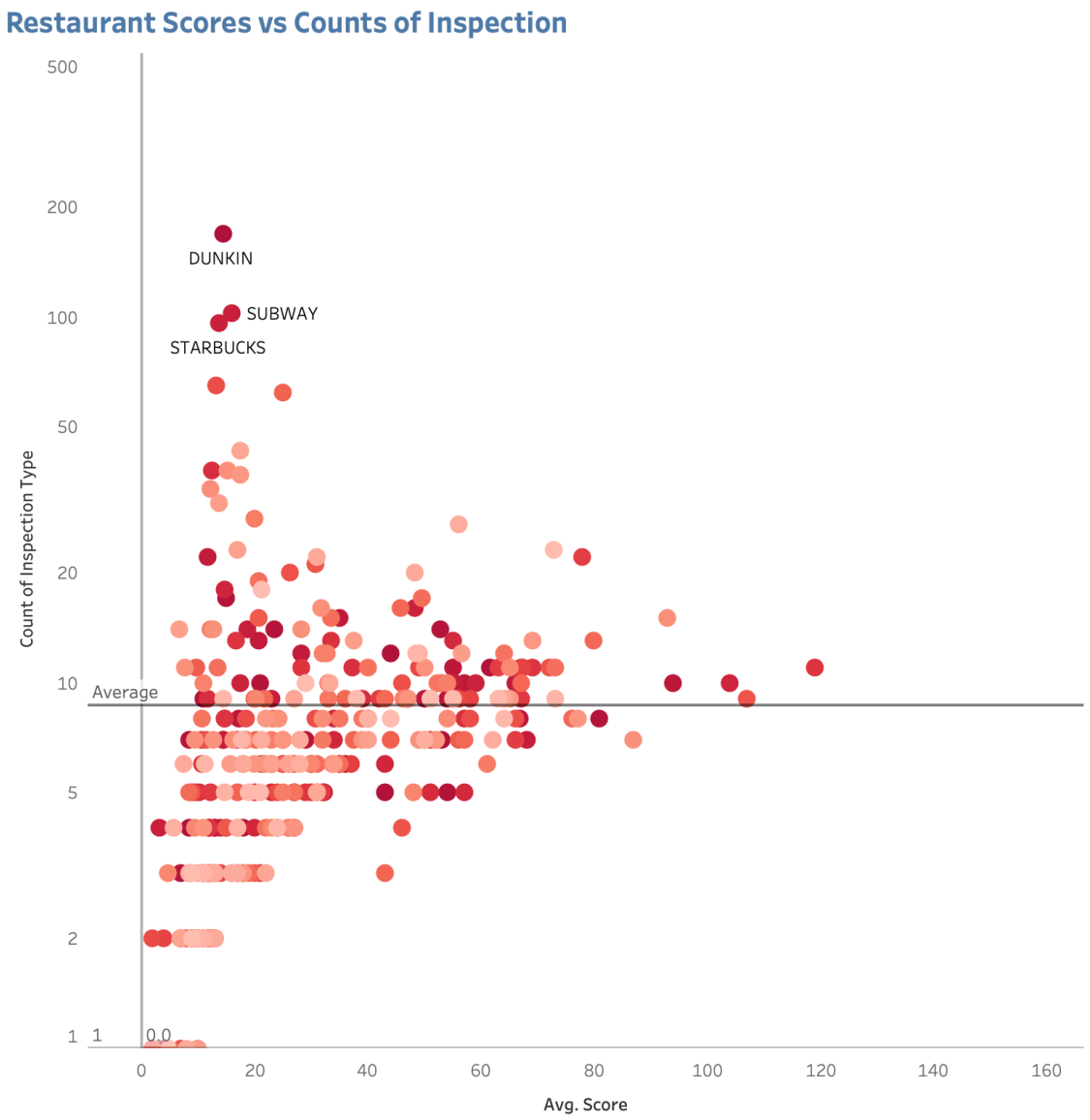
This visualization illustrates the exponential increase in violations among New York City restaurants over the past two years. The data reveals a relatively stable trend in violation growth from March 2020 to June 2021, coinciding with the COVID-19 pandemic. However, a significant surge in violations is evident following the reopening of restaurants post-pandemic. The well-fitted trend line effectively supports the observed exponential growth.

Visualization 10 : Score vs Counts of inspection in Top 10 Restaurants



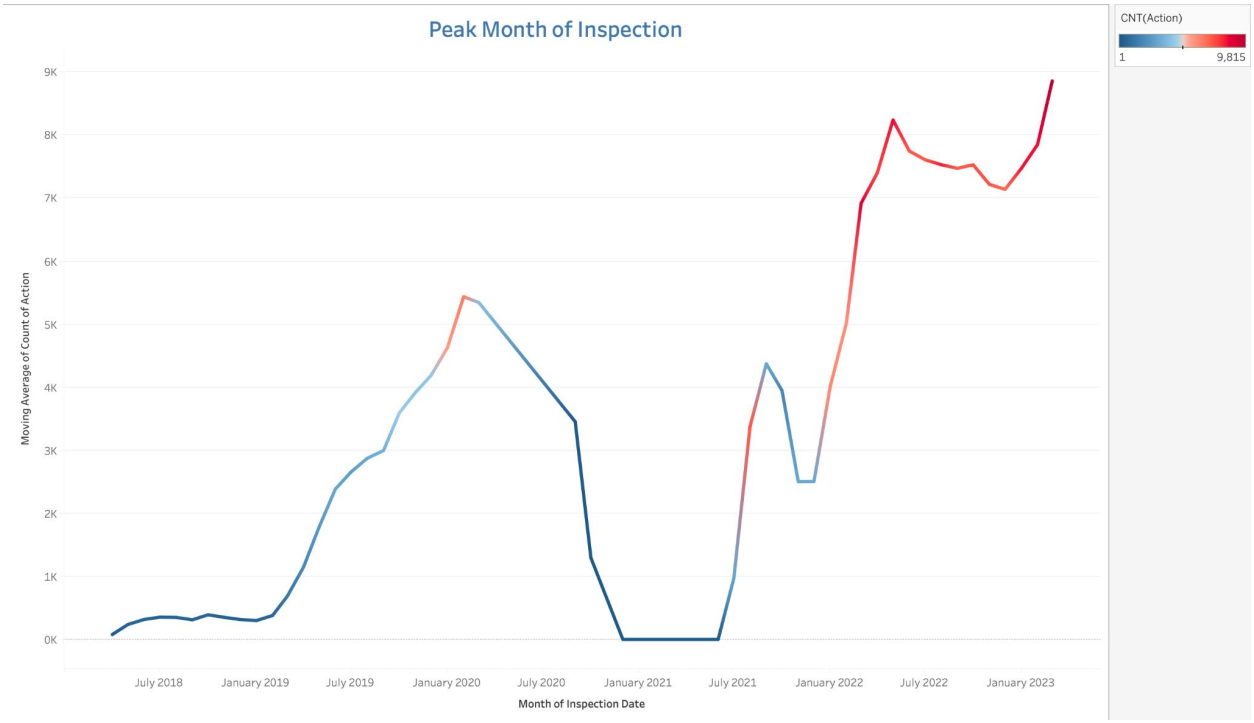
In this visualization, a combination of bar and line charts is used to display the relationship between violation scores and the count of inspection types for each cuisine among the top 10 cuisines. This dual-axis chart effectively demonstrates the correlation between increased violation scores and the number of inspection types. American cuisine, represented by a red line, has the highest violation score of 31,140 and the greatest number of inspection types. In contrast, the line chart reveals that Japanese cuisine has a smaller number of inspection types, totaling 6,256.

Visualization 11 : Restaurants Violation scores based on their count of inspection



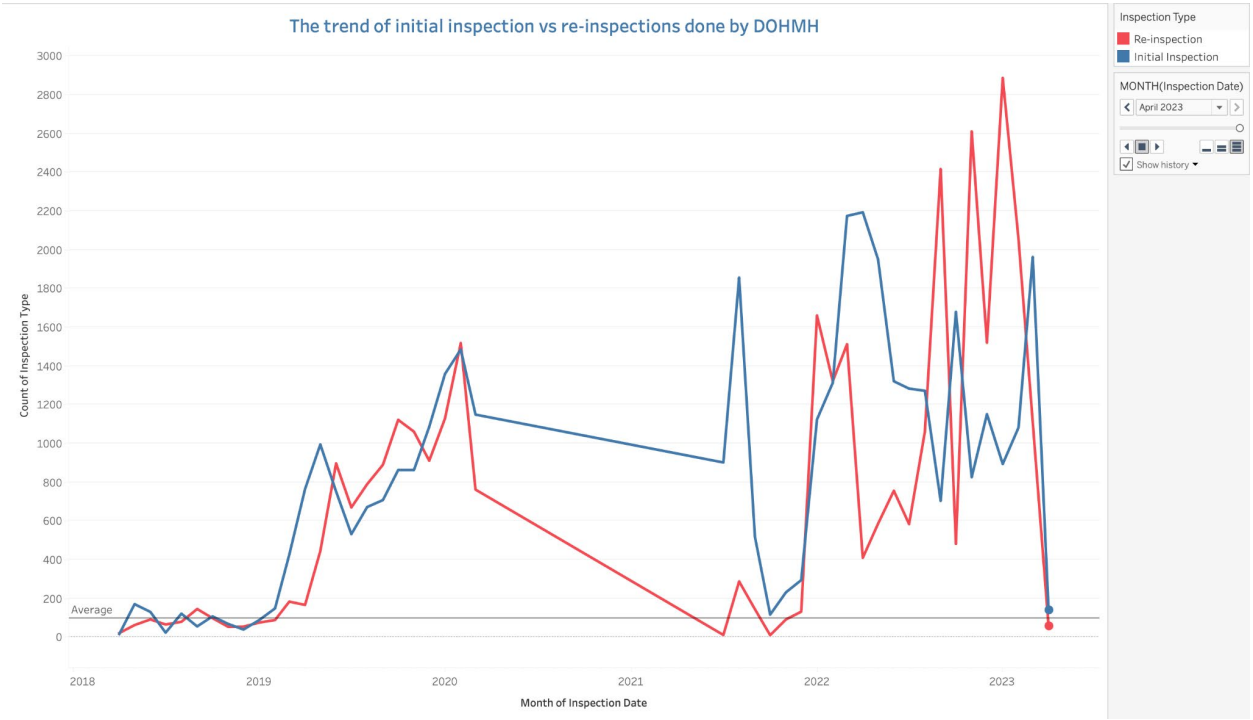
This animated scatter plot presents the monthly changes in inspection counts for three popular restaurants: Dunkin, Subway, and Starbucks. As the animation progresses, the movement of the average line indicates that these establishments consistently exceed the average number of inspections in most months. Consequently, these three restaurants can also be considered outliers

Visualization 12 : Peak Month of Inspection



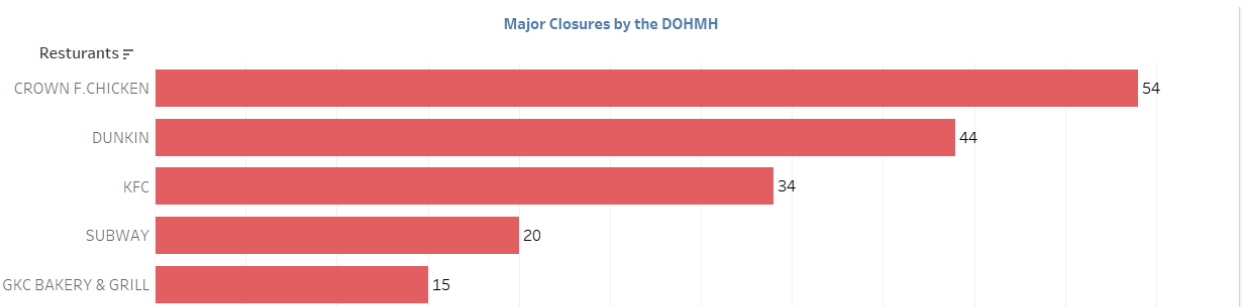
This line graph displays the peak months of inspection from 2018 to 2023. The data reveals that the number of inspections conducted by DOHMH tends to increase annually in January, September, and May. A notable decrease in inspections occurred from March 2020 to December 2020, likely attributable to the COVID-19 pandemic. However, a significant rise in inspections from December 2022 to the present indicates that DOHMH enforces stringent hygiene regulations for restaurants.

Visualization 13 : The trend of initial inspection vs re-inspections done by DOHMH



This animated line graph compares the number of inspections and re-inspections over a five-year period. Between 2018 and 2020, both inspection and re-inspection trends followed a similar pattern. However, there was a significant increase in initial inspections compared to re-inspections from February to May 2019, and again from March 2020 to August 2021. Starting in October 2023, the number of re-inspections surpassed that of initial inspections, suggesting that DOHMH has adopted a more stringent approach towards restaurants during the inspection cycle. In other words, DOHMH imposes additional requirements for restaurants to improve their grades.

Visualization 14: Major Closures by the DOHMH

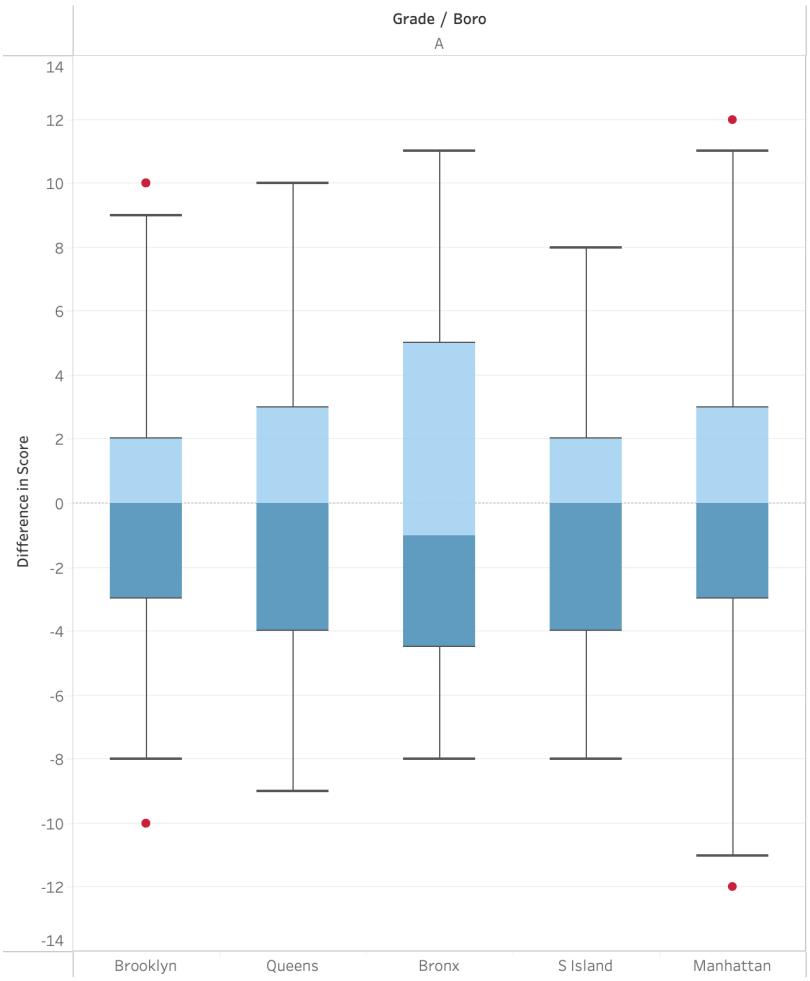


According to data from the Department of Health and Mental Hygiene (DOHMH), Crown F.Chicken restaurants have experienced the highest number of closures, with 54 instances since 2018. Following closely behind, Dunkin has been closed 44 times during the same period. On the other hand, Bluestone Lane and McDonald's have had significantly fewer closures, with only 3 and 2 instances respectively. These figures indicate that Crown F.Chicken and Dunkin have had a higher frequency of closures compared to Bluestone Lane and McDonald's.

Visualization 15 : Why Stabucks is Pioner in Coffee Indistry?

	Grade / Inspection Type	
	A	
Boro	Re-inspection	Initial Inspection
Bronx	39	88
Brooklyn	131	671
Manhattan	478	3,618
Queens	124	714
S Island	83	125

Why Starbucks is Pioner in Coffee industry?

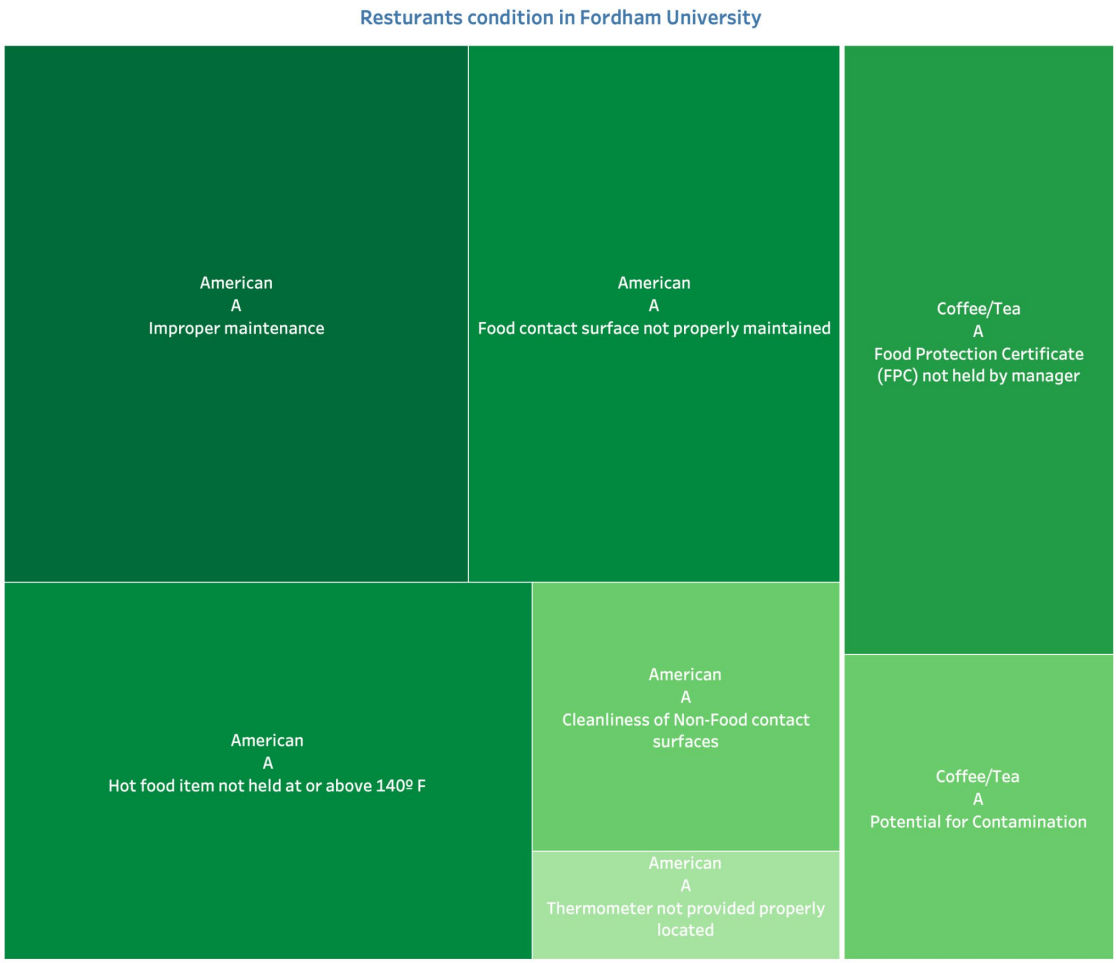


According to the Table, scores for the initial inspection and re-inspection has specified for Starbucks branches in five different boroughs : Manhattan, Brooklyn, Bronx, Queens and Staten Island. The performance of the Starbucks can be calculated by finding the difference between the

initial inspection and re-inspection scores. For instance, If the difference between the initial inspection score and the re-inspection score is negative, it means that the restaurant's re-inspection score is higher than its initial inspection score, indicating a decline in its food safety and sanitary conditions. Recall that a lower score is better, as it represents fewer violation points. A negative difference implies that the restaurant had more violation points during the re-inspection compared to the initial inspection, which may result in a lower grade (i.e., moving from Grade A to Grade B or from Grade B to Grade C).

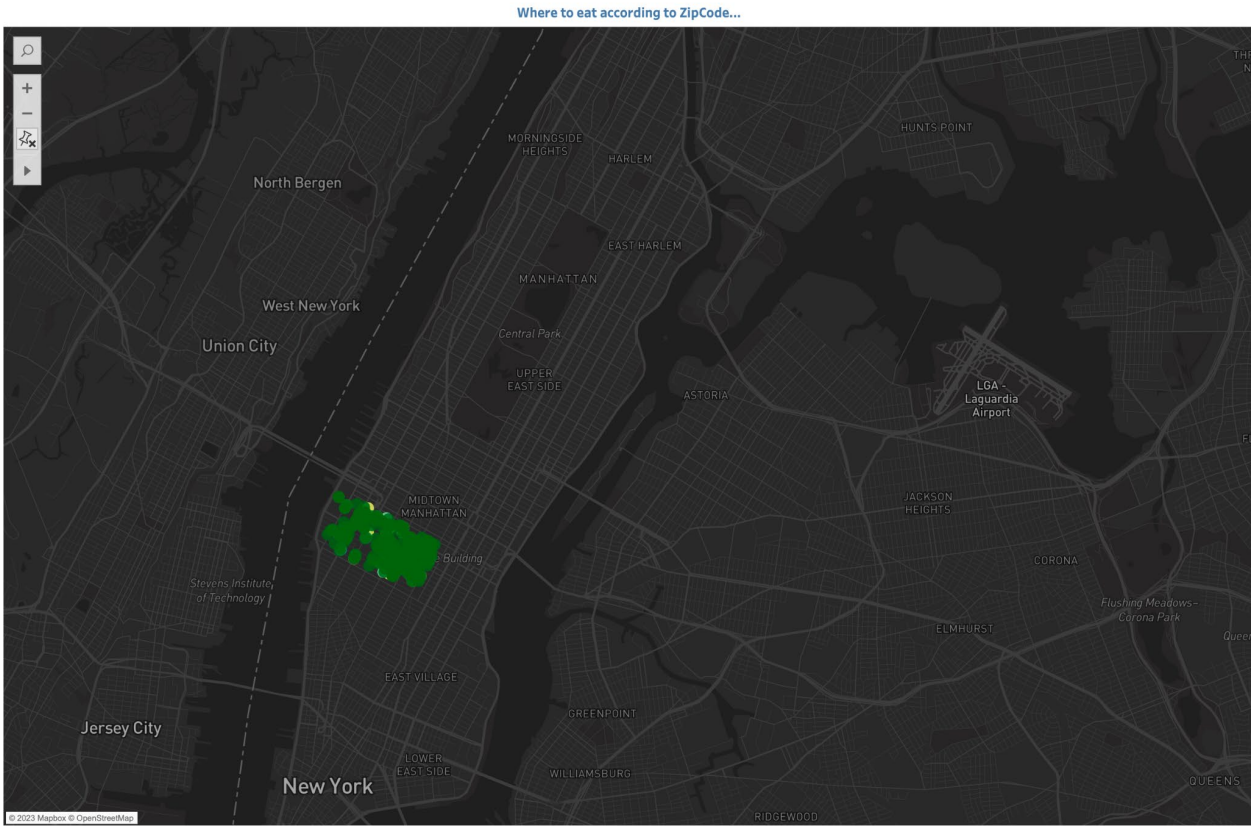
According to the box pots, these statistics suggest that all five boroughs have experienced notable improvements in food safety and sanitary conditions between the initial inspection and re-inspection. The largest improvement was observed in Manhattan, followed by Queens, Brooklyn, Staten Island, and the Bronx. This could be attributed to effective food safety regulations and enforcement, as well as increased awareness and compliance among restaurant owners and staff.

Visualization 16: Restaurants condition in Fordham University



This TreeMap summarized the situation of the restaurants at Fordham University. According to the map, All the restaurants preserve the A grade over the last five years and they got less than 14 points in their initial inspections. However, it doesn't mean that there is violations happening in these restaurants. For instance, the most common violations among these restaurants can be categorized into “Improper maintenance”, “Cleanliness of non-food contact” and “hot food item not help at 14 F”.

Visualization 17: Where to eat according to the Zipcode



This animated map depicts restaurants by zip code based on their respective letter grades. Here, restaurants are organized into five distinct categories according to their grades. Dark green highlights grade A establishments across the city's five boroughs, while lighter green, blue, yellow, pink, and red correspond to the remaining letter grades, in order of decreasing grade light green to red. The audience is able to look at the map, decide on a zip code to zoom in on and get a closer look at the area.

Conclusion:

This project is a comprehensive overview of the NYC Restaurant and food operating establishments across the city. We were curious to see how the inspection process and the results of inspections affect establishments that many of us frequent on a daily basis. For the next steps of the project, we would want to take a closer look at the last map. A really cool continuation of this research using this information is creating an application for foodies to be able to locate restaurants near them that tells them where the clean places are to eat. The visualization that allows the user to check the restaurants in their zip code could easily be integrated into this application. Additionally a feature that lets you check the stats for each restaurant would be a really useful tool. This would require updating the dataset yearly, as we have seen things can change drastically year to year. Up to date information is crucial for the reliability of this information. This visualization could be edited with additional detail and become an entire project itself. Some additions that would be useful in the visualization is a zoom in feature. For example, when someone chooses a zip code, if the animation could automatically zoom into that zip code. That would be an incredible tool for a user. This is a feature that could perhaps be integrated with companies such as YELP, or OpenTable, for example. There are many avenues that can be taken with the information that we have presented in this project.

Work Cited:

1. “Eric Adams, New York's Pied Piper, declares war on rats.” *The Economist*, 10 November 2022, https://www.economist.com/united-states/2022/11/10/eric-adams-new-yorks-pied-piper-declares-war-on-rats. Accessed 3 May 2023.

2.“How We Score and Grade_Layout 1.” *NYC.gov*, https://www.nyc.gov/assets/doh/downloads/pdf/rii/how-we-score-grade.pdf. Accessed 3 May 2023.

3. “How We Score and Grade_Layout 1.” *NYC.gov*, Accessed 3 May 2023.

4. “What to Expect When You’re Inspected: A Guide for Food Service Operators” June 2016 Michael R. BloombergMayor Thomas Farley, M.Commissioner