Restaurants in NYC

Amoli Vani - anv282
Shradha Taneja - st3277

Project goal
The main goal of the project is to find restaurants in different boroughs depending on the different cuisines available. It also helps us filter out restaurants depending on their ratings.

Data Set
https://www.kaggle.com/new-york-city/nyc-inspections

Analytical Questions and Proxy Tasks

1. Do different regions/boroughs favour different cuisines?
   Proxy Task: Finding out the number of restaurants of each cuisine in different regions/boroughs ex: Chinese food, Indian food, etc.
   Attributes: boro, cuisine description

2. Which are the 10 most famous restaurants in different boroughs?
   Proxy Task: Find the 10 restaurants which have the best rating in respective borough.
   Attributes: boro, score, DBA

3. Which restaurants in all boroughs are the most hygienic?
   Proxy Task: Find the restaurants in all boroughs that have the highest grade in terms of cleanliness and hygiene.
   Attributes: boro, grade(given by health government), DBA

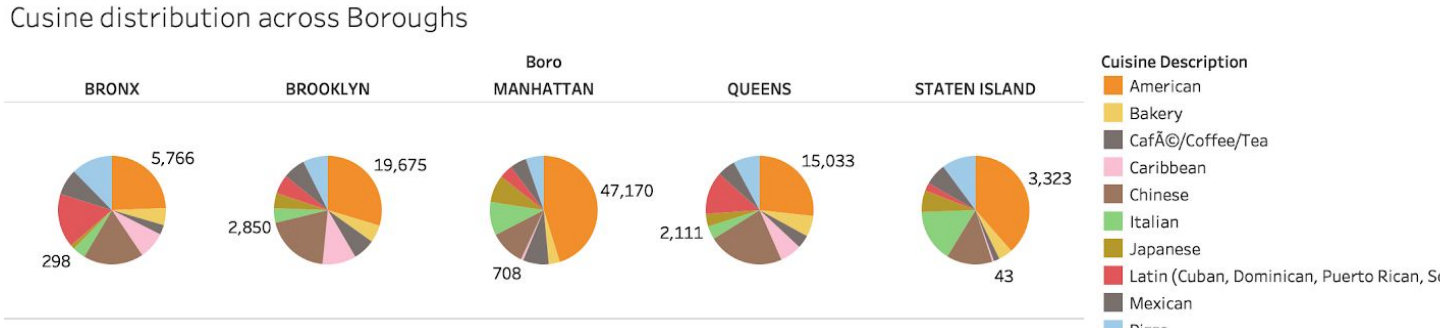
4. Which are the most famous restaurants for each cuisine in different boroughs.
   Proxy Task: For each cuisine in each borough, find the restaurant with highest score.
   Attributes: boro, score, cuisine description, DBA

5. Which restaurants in each borough can be shut down due to bad hygiene?
   Proxy Task: Find the restaurants in each borough due to not complying health standards.
   Attributes: boro, DBA, violations description(not comply health standards)

Story Design
Data Analysis

We have created the following visualizations after performing the data analysis for the questions above:

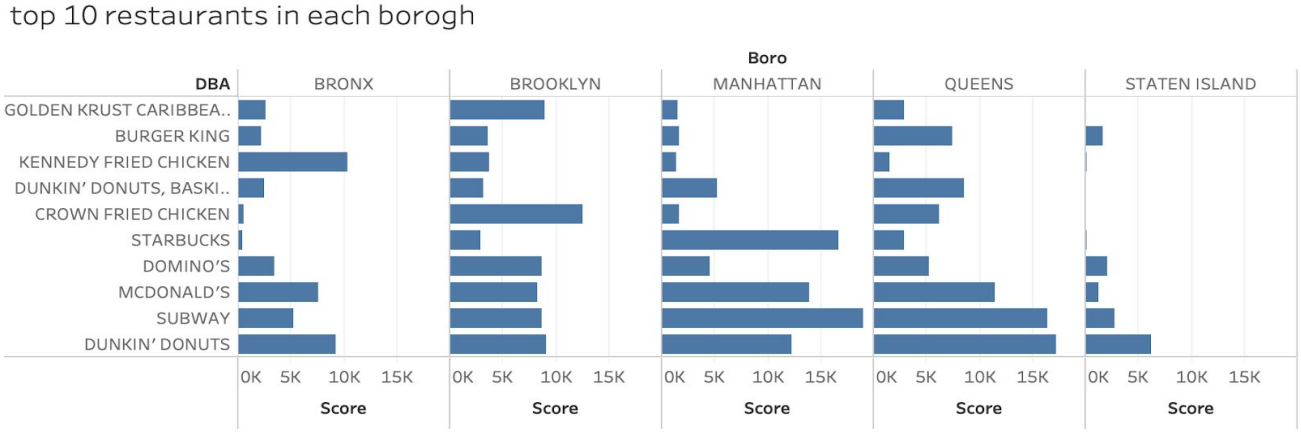
Visualization for Question 1:



Cuisine Description (color) broken down by Boro. The view is filtered on Boro and Cuisine Description. The Boro filter keeps BRONX, BROOKLYN, MANHATTAN, QUEENS and STATEN ISLAND. The Cuisine Description filter keeps 10 of 84 members.

In the first question we want to visualize which boroughs favour which cuisine. From the above graph we can visualize that all the boroughs favour American cuisine but in Queens Chinese cuisine is favoured more as compared to other boroughs. Similarly Italian is favoured more in Staten Island and Latin is favoured more in Bronx.

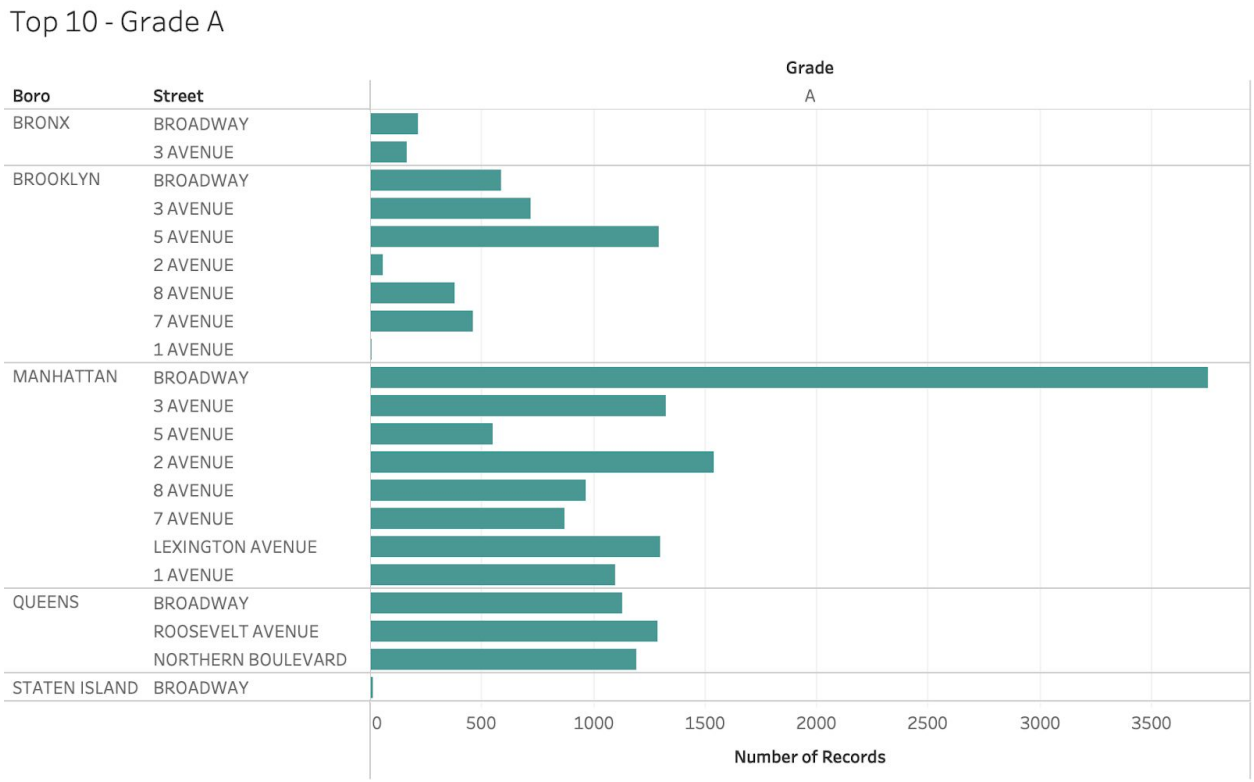
Visualization for Question 2:



Sum of Score for each DBA broken down by Boro. The view is filtered on DBA, which keeps 10 of 20,937 members.

In the graph above we can visualize the 10 most famous restaurants in each borough. For example, Subway is the most famous in Manhattan where as KFC is the most famous restaurant in Bronx.

Visualization for Question 3:



Sum of Number of Records for each Street broken down by Grade vs. Boro. The view is filtered on Grade and Street. The Grade filter keeps A. The Street filter keeps 10 of 3,329 members.

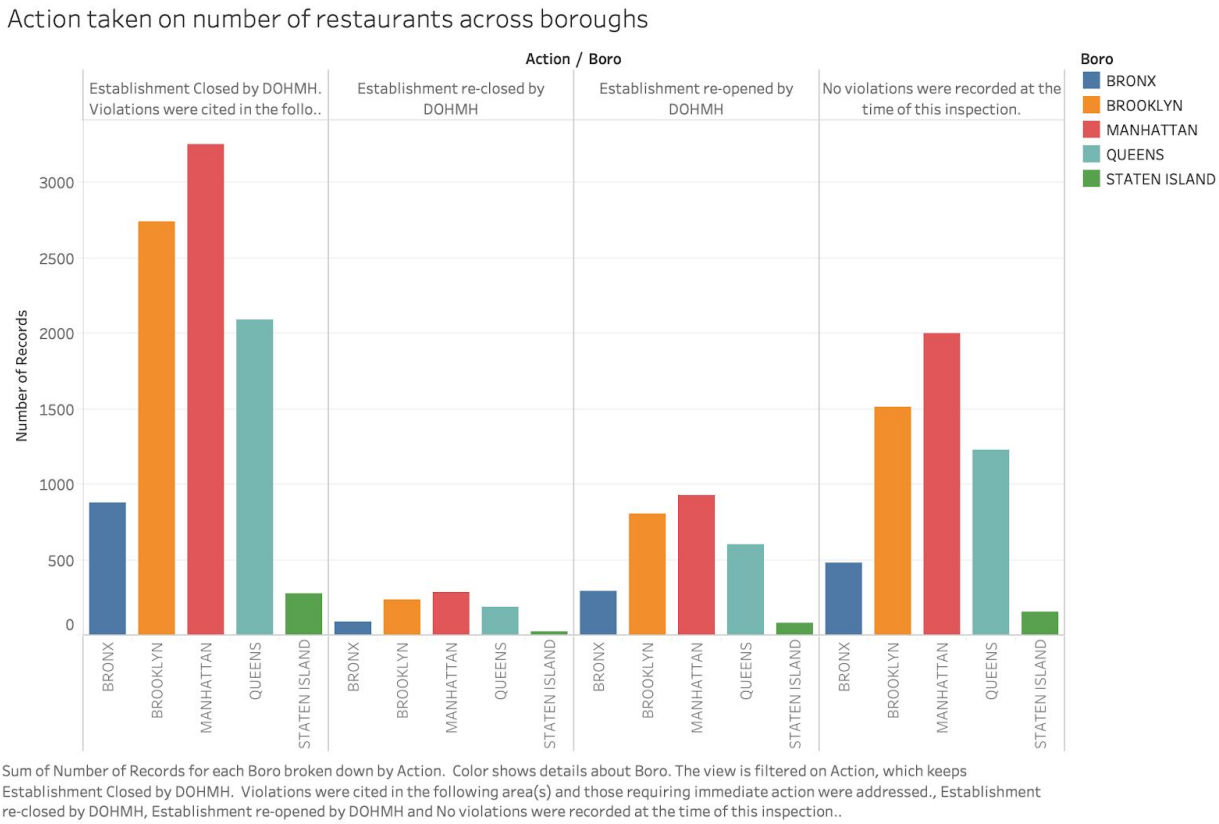
We have altered this question a little as per the feedback given to us (details mentioned in the changelog) to displaying number of restaurants that are most hygenic, i.e., have A grade across boroughs and streets. Thus we can see from the visualization that Broadway in Manhattan has the most hygienic restaurants followed by 2nd Avenue in Manhattan and 5th Avenue in Brooklyn has the most hygienic restaurants in Brooklyn. Thus this helps us identify the most hygienic restaurants not only in a particular borough but also across boroughs and in comparison of streets across all Boroughs.

Visualization for Question 4:

We have made alteration in the 4th question as while performing the data analysis we realised that the graph created for different restaurants with different cuisines is a very complex one without any relevant information extracted. Thus we have identified a new question that helped us get new insights. The question considered is:

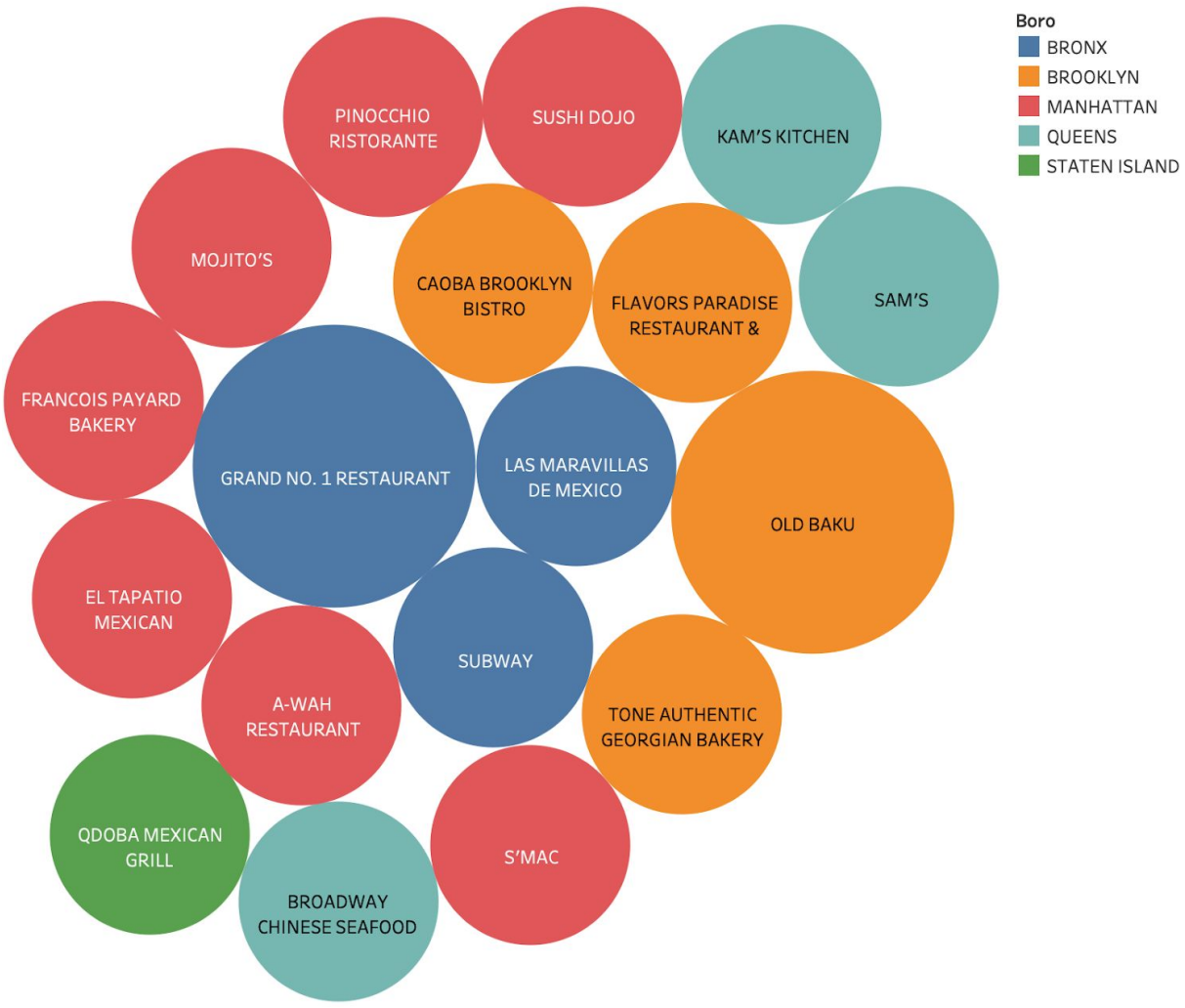
What were the results of the inspection, were restaurants closed in any of the areas? How many restaurants in each borough didn't have any violations?

Proxy Task: Find the number of restaurants in which different actions were taken.
Attributes Used: Action, Boro



In the above visualization, we can see that more than 3000 restaurants were closed in Manhattan as a result of the inspection. Where as Staten Island encountered the least number of closings. Also, as compared to other boroughs, Manhattan had the maximum restaurants where there was no violation. The reason of Manhattan being the top of these opposite questions can be explained by the total number of restaurants there being much more than total number of restaurants in any other borough.

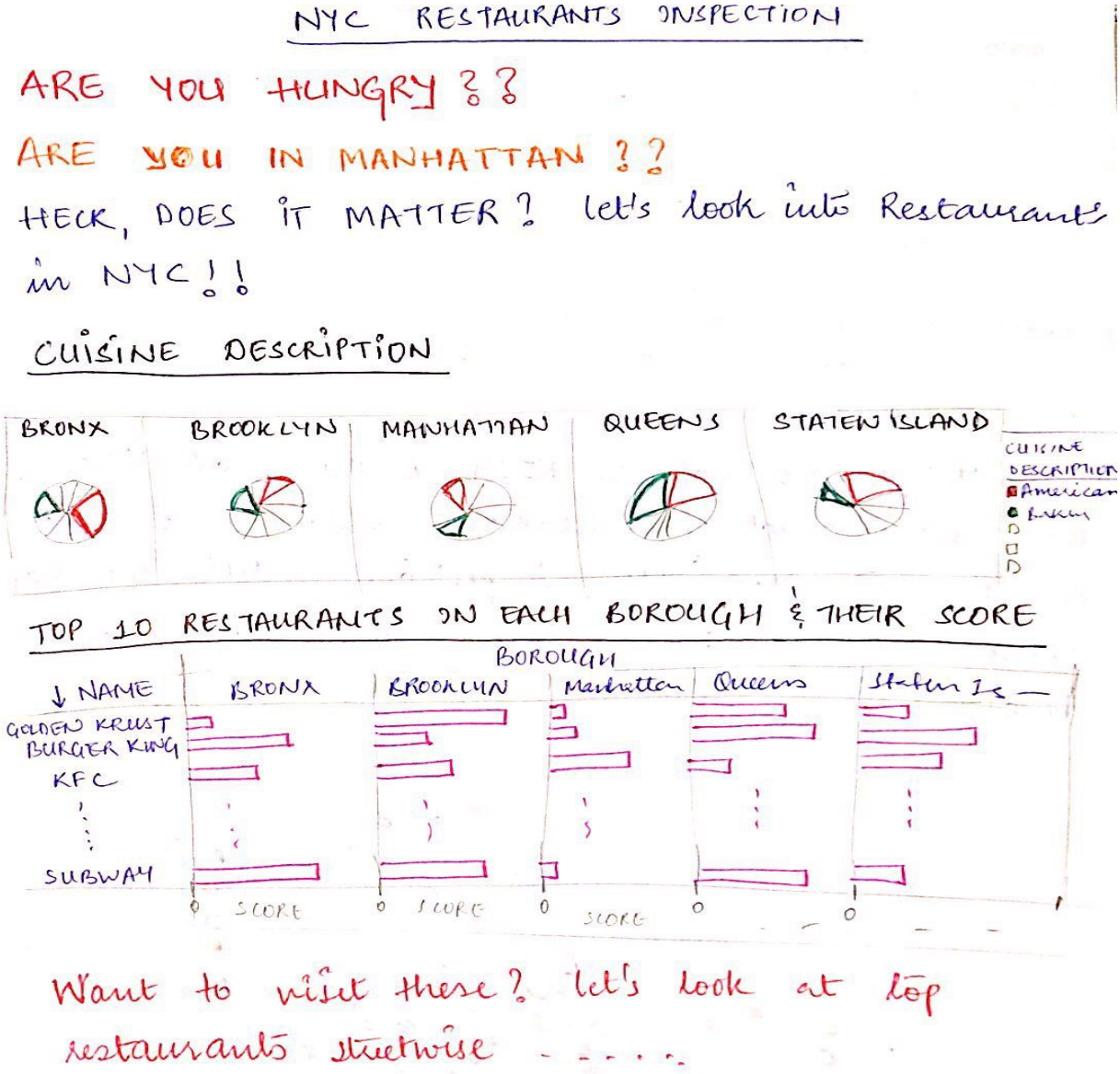
Visualization for Question 5:



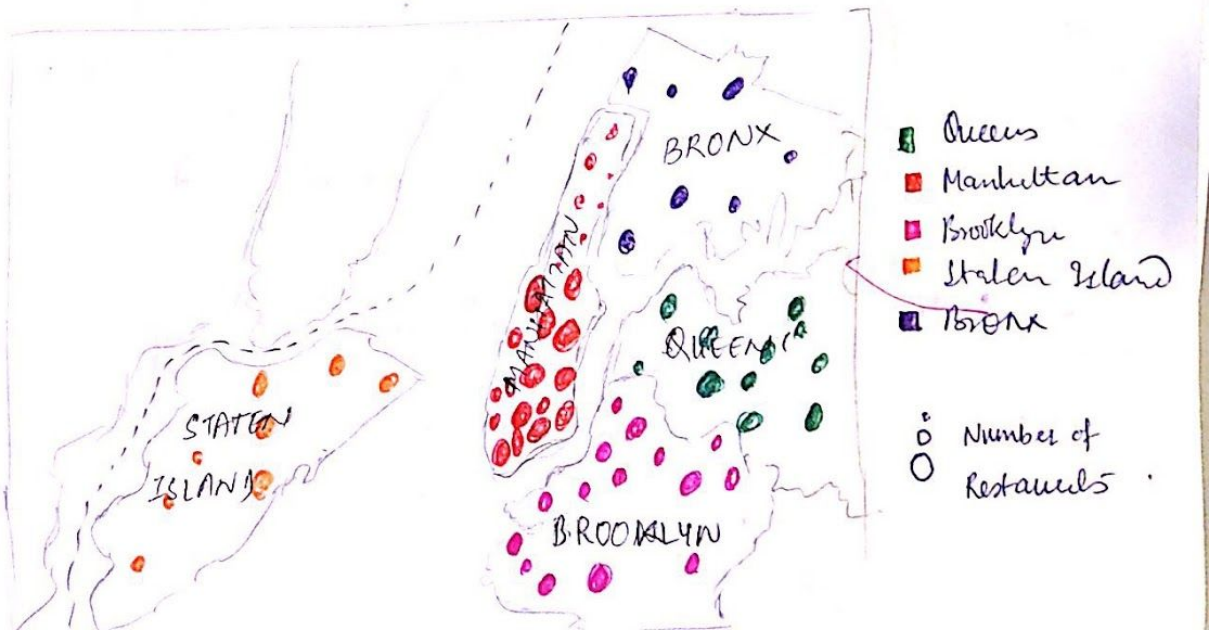
As we needed to visualize the restaurants with bad hygiene, we filtered the violation description to the one that does not comply to the health standards. In the graph above we can visualise the name of the restaurants across the 5 boroughs that did not comply to the health standards and that should be shut down.

Storyboard

The end result of our visualization is to display the inspection data in a way that we can perform various analysis. We have created a sketch of our visualization that we intend to create.



Well, things are not as ROSY as they look like. there were a lot of restaurants which were given a CRITICAL flag after inspection. let's see how critical restaurants distribute over different boroughs.

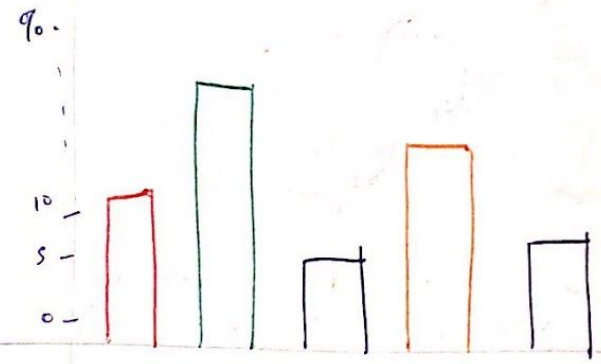


wondering why restaurants have critical flag?
let's look at different violations made by restaurants

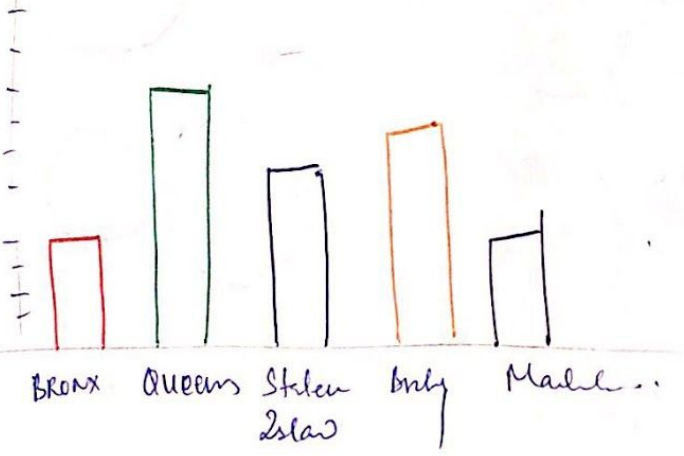
Toilet facility not available for employees

Manhattan
Bronx
Queen

Evidence of Mice in facility w food



Evidence of Rat in facility w food



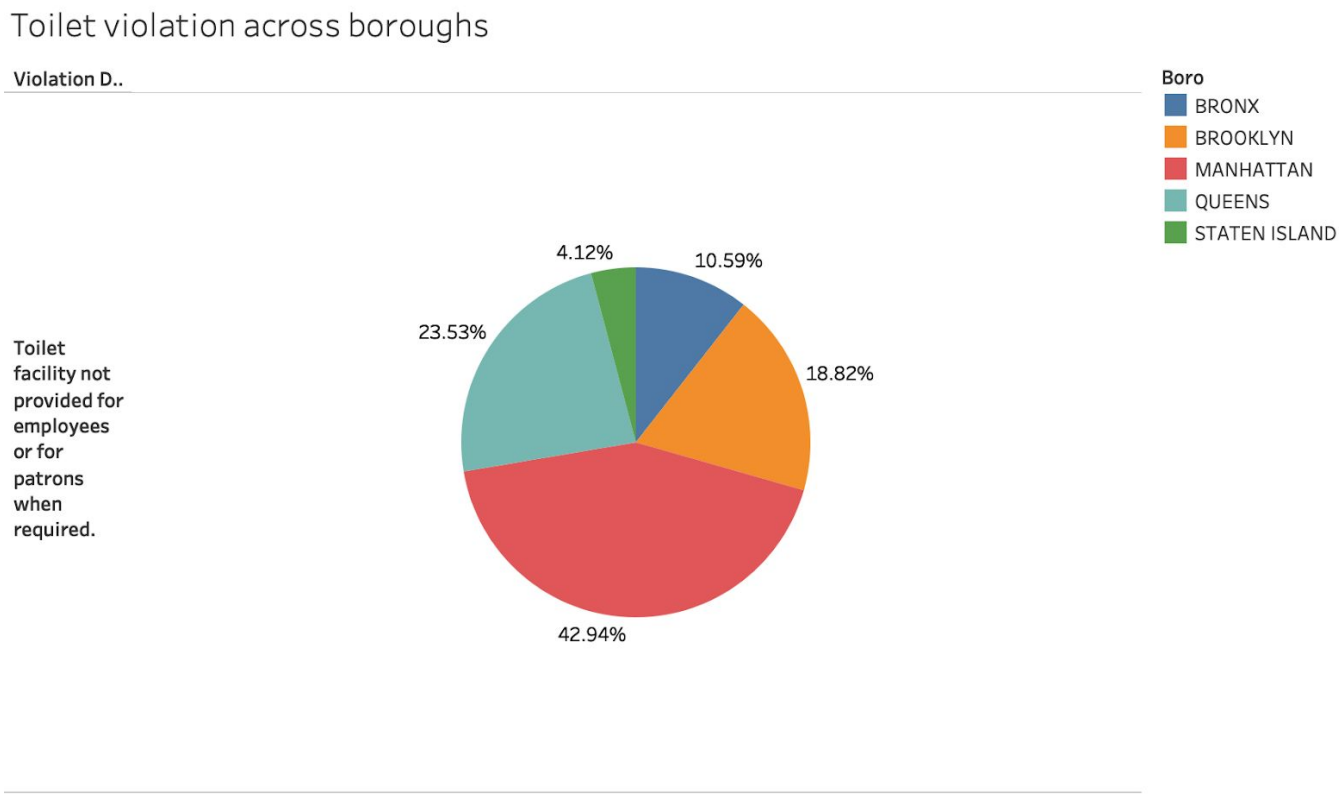
The above visualizations tell a story from start to the end. It begins with the user being hungry and looking for restaurants according to location. Initially the user is shown

different cuisines in different boroughs of New York, followed by top 10 famous restaurants in each borough and their score. It also provides the user with A grade restaurants as per the streets.

The next visualizations are related to give a perspective to the user of which restaurants to be avoided by showing them the restaurants with critical flags, toilet violations, rat and mice violations and even shows the user the actions taken by the Health Board like shutting down few restaurants.

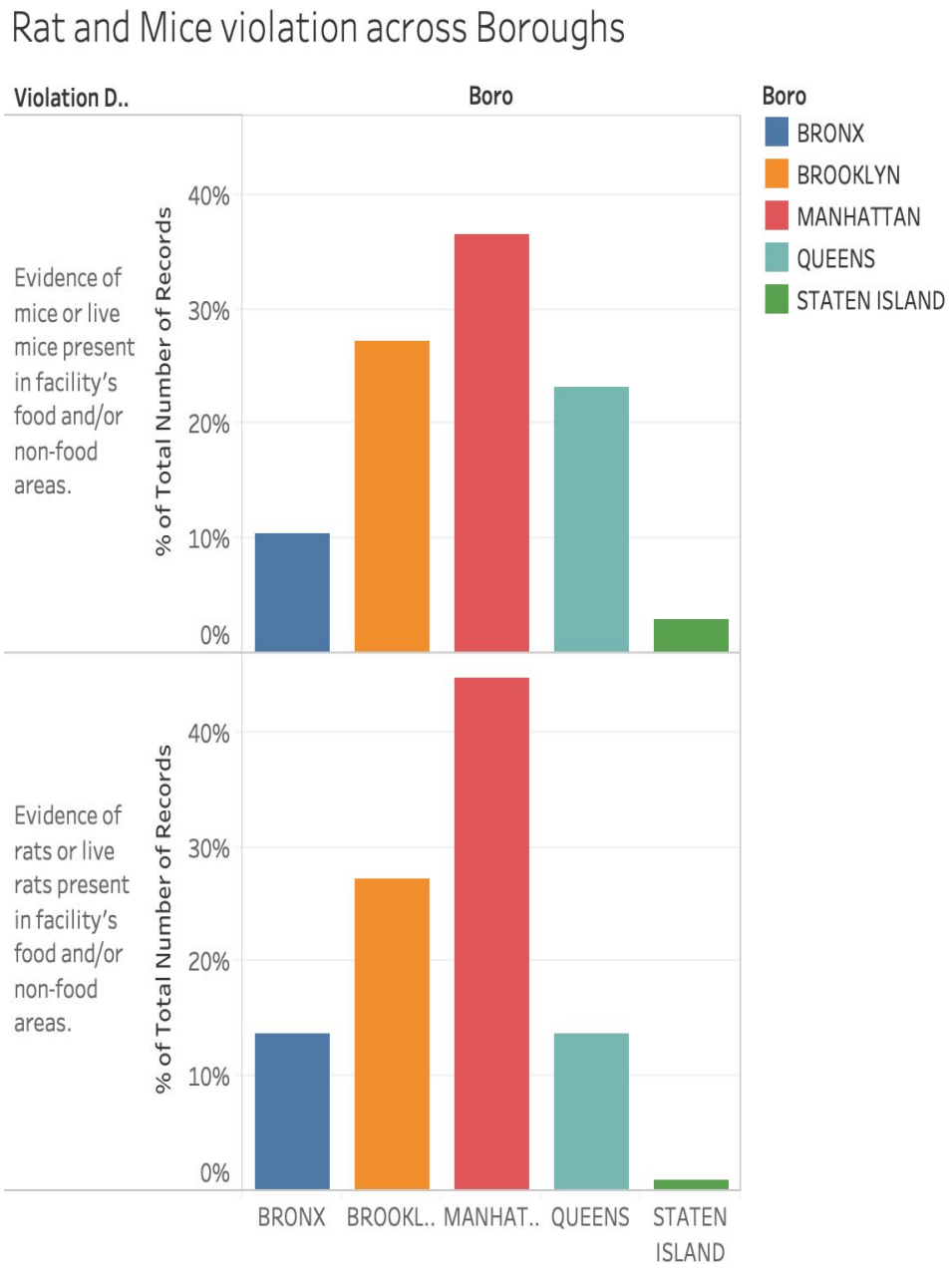
Specific Insights Observed:

- We observed that a lot of restaurants don't have toilets for employees over all the boroughs. In the visualization seen below, we can see that of the restaurants with a violation of toilet availability, more than 40% of the violations are in Manhattan.

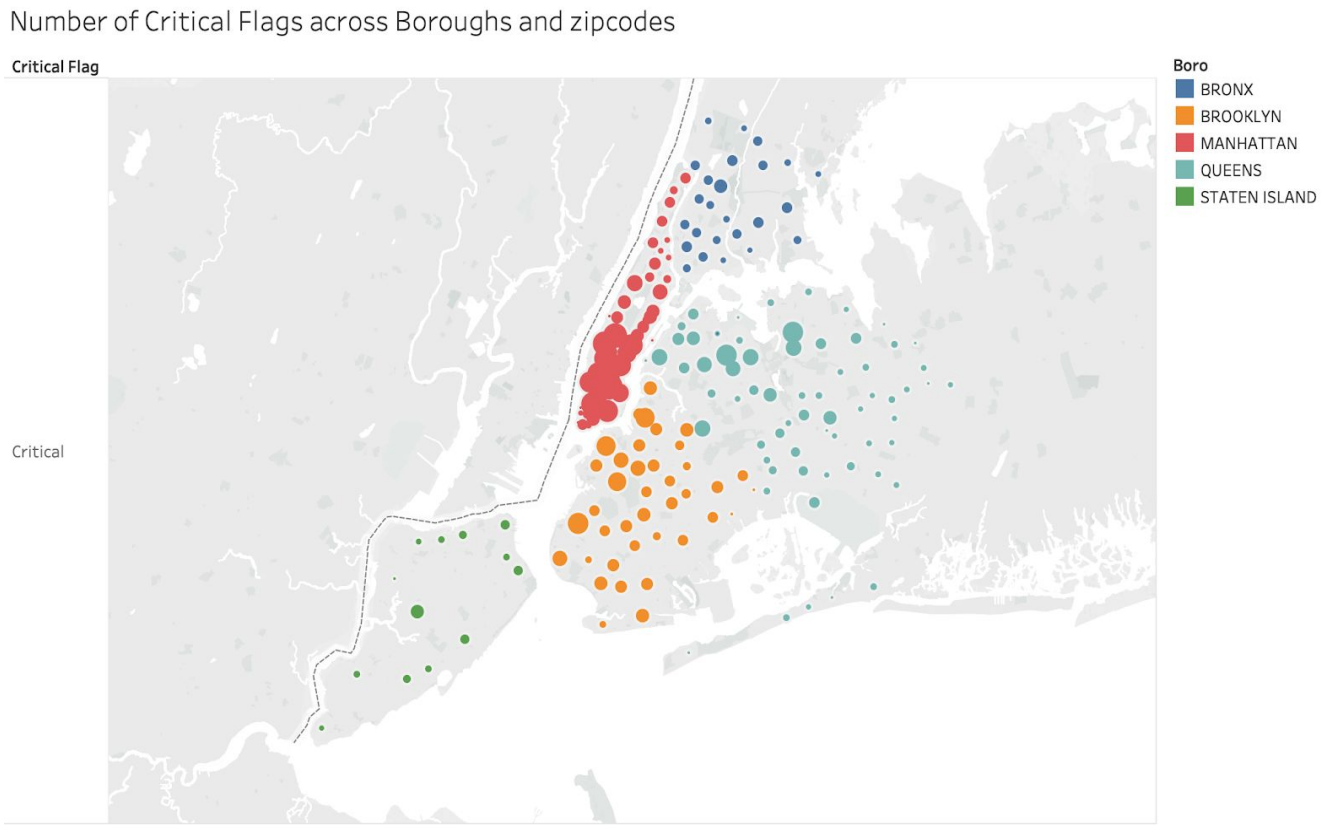


- Another observation is the evidence of mice and rats present in restaurants in all the boroughs. This was surprising because we didn't expect these violations to

occur at such a high rate. Hence this piece of information is very important and should be taken care of.



- The visualization below shows the distribution of critical violations across different boroughs. It can be seen that Manhattan has a lot of the critical violations as compared to Staten Island, whereas Brooklyn has quite some number of critical violations.



Changelog

Question 3:

Altered the question as per the feedback received of using neighbourhoods(here streets) for visualizing rather than just using boroughs. As the number of restaurants on streets are relatively very high and difficult to visualize hence we are displaying the number of most hygenic, i.e. A grade restaurants in different boroughs on different streets.

Question 4:

The information gathered from this will be a repetition from questions 1 and 2, hence we decided to remove that question. Also while performing the data analysis we realised

that the graph created for different restaurants with different cuisines is a very complex one without any relevant information extracted. Thus we have identified a new question that helped us get new insights. The question considered is:

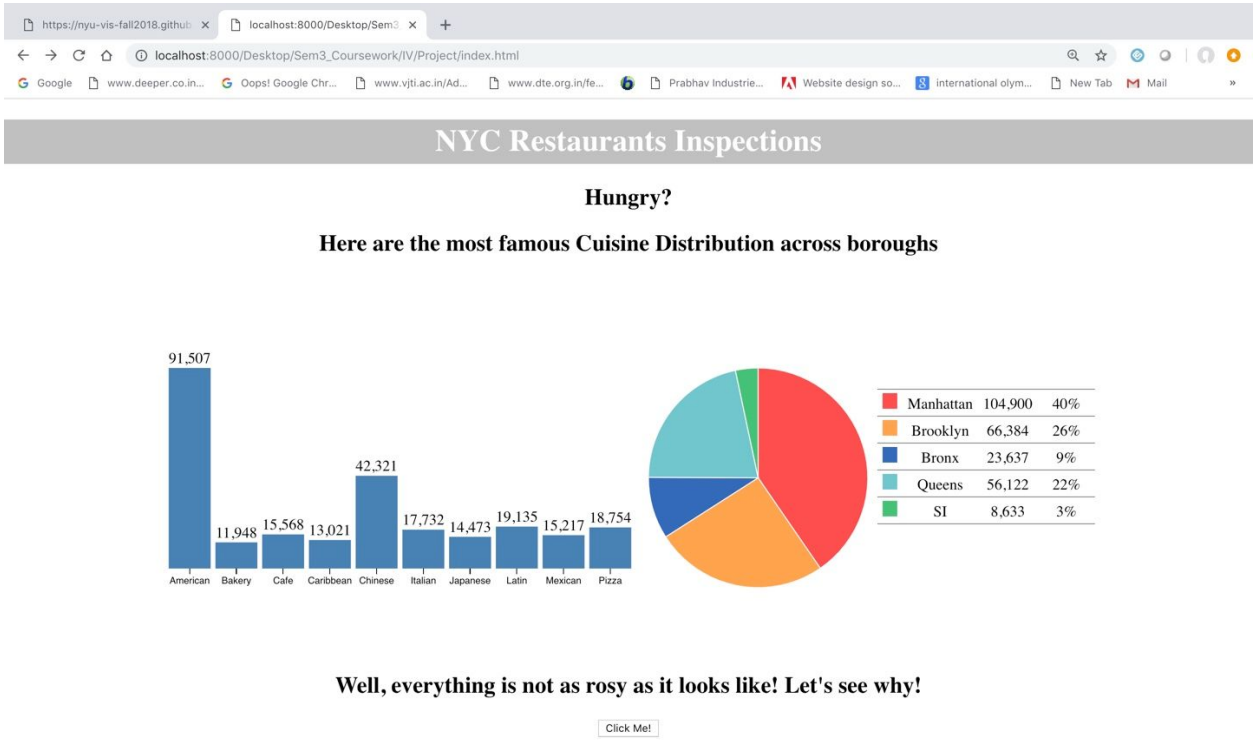
What were the results of the inspection, were restaurants closed in any of the areas? How many restaurants in each borough didn't have any violations?

Updates:

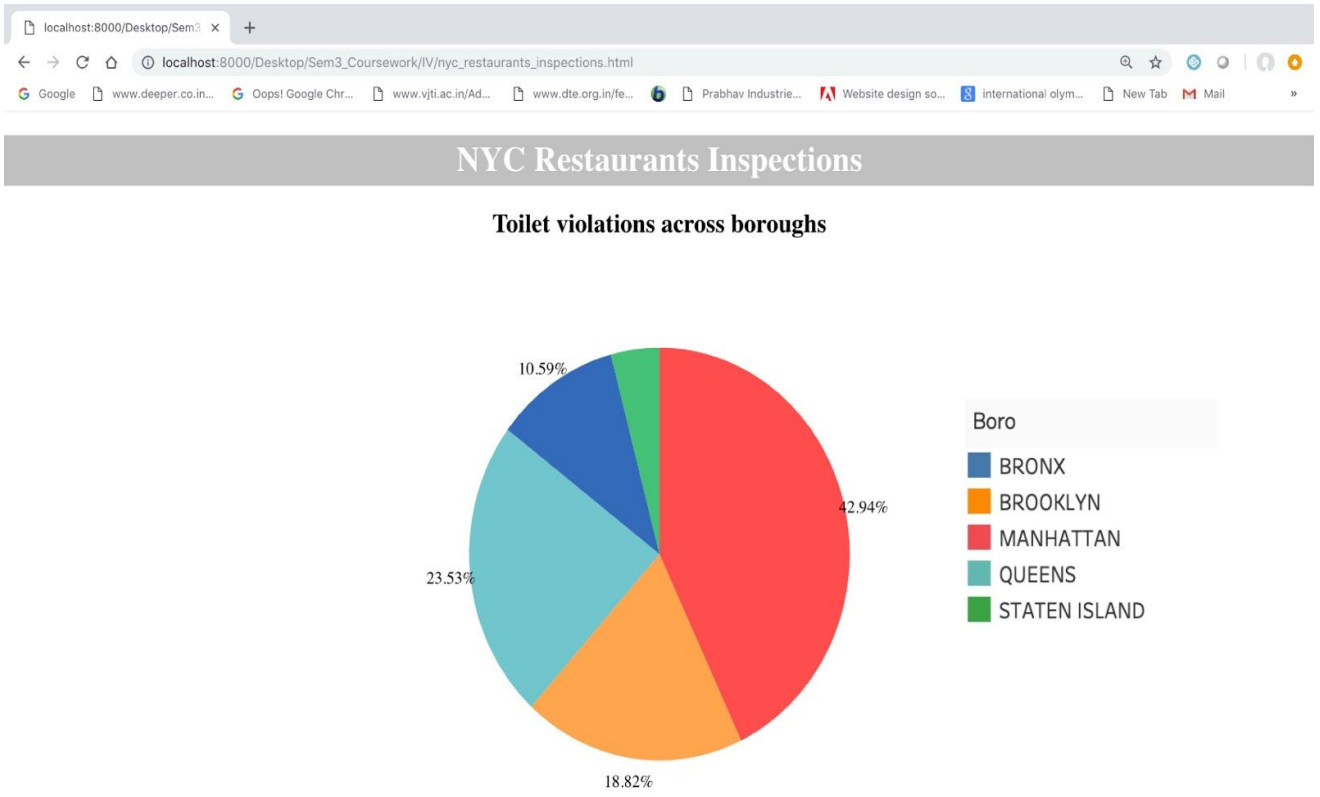
The sketch of our final project has been updated above as per the feedback received on phase two of the sketch being more about displaying the insights rather than letting the user use it as a tool.

Items already implemented:

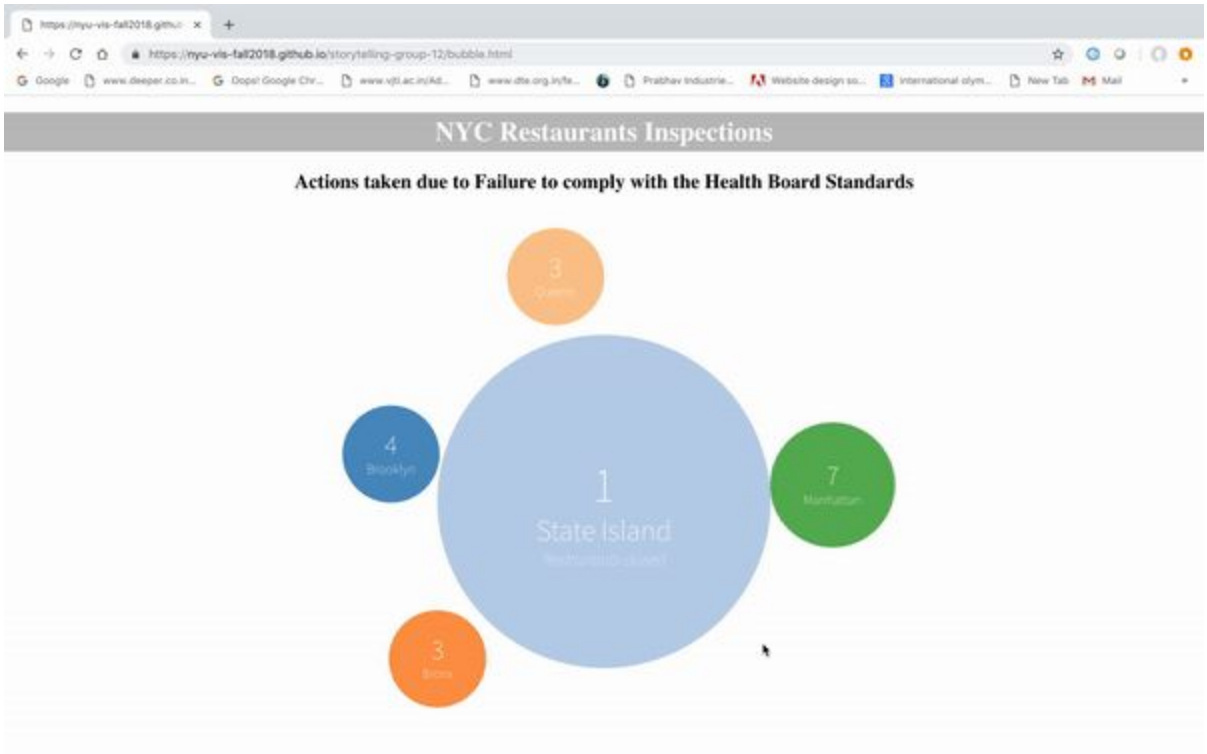
- Cuisine distribution across boroughs.



- Insight for Toilet violations across boroughs.



- Number of restaurants that have been closed in all the boroughs as they didn’t comply with the Health Board Standards.



Items yet to be implemented:

- 10 most famous restaurants in all boroughs.
- Restaurants with Grade A as per streets across boroughs.
- Number of restaurants with Actions taken across boroughs.
- Rats and mice violations across boroughs.
- Number of critical flag restaurants as per zip codes across boroughs.

Link to the current version of the page on GitHub:

https://nyu-vis-fall2018.github.io/storytelling-group-12/