Samuel Guerrier, Ousmane Diallo 12/21/2022

Data Analytics – Choosing the Best Restaurant

Objective

Can we find or open a restaurant that has longevity?

Questions That Can Help Define Longevity

- Is the restaurant open?
- What kind of data is a good indicator of the current standing of the restaurant (is it open or closed)?

Description of the SQL Databases Before The Data Analysis

The data is presented in two main database tables. It is presented in the Microsoft Azure SQL Server. The data is raw. It is categorized as restaurant_data and violation_data. For this analysis, the following tools that were used include:

- SQL
- Excel Spreadsheets
- Microsoft Azure SQL Analytics

Info on the data

-restaurant has 27,960 rows

-violation has 248,020 rows

--average score by camis

Metrics In The Restaurant SQL Databases That Characterizes Longevity

- The grade of a restaurant.
- The score of a restaurant.
- Actions that was taken by the city government (closing a restaurant)
- How was the restaurant inspected? Some inspection techniques may be more effective than others.
- The date in which the restaurant was inspected and graded. If the restaurant is open and has other positive attributes that may be indicative of its longevity, was that data taken at a recent date?

- How safe is the neighborhood in which the restaurant operates?

Procedure for Data Cleansing

For the analysis the data is first analyzed using SQL before being exported to CSV files or charts. This was done in SQL as the following results show. The CSV files generated are based on the following analysis:

- Get all restaurant data.
- Get all Violation data.
- Count the number of restaurants by Grade.
- Count the number of restaurants by score.
- Count the number of restaurants by Critical Flag.
- Count the number of restaurants by Violation Code.
- Count the number of restaurants by Location.

Data Analysis Approach

For this analysis the following were done:

- Restaurants Graded 'A' and critical flag 'Not Critical'.
- Restaurants Graded 'A', 'Not Critical', Score less than 15.
- Restaurant location Graded 'A', 'Not Critical', Score less than 15.

Graphical Results

Here, we have some graphical results to help us choose the restaurant that has the best longevity.

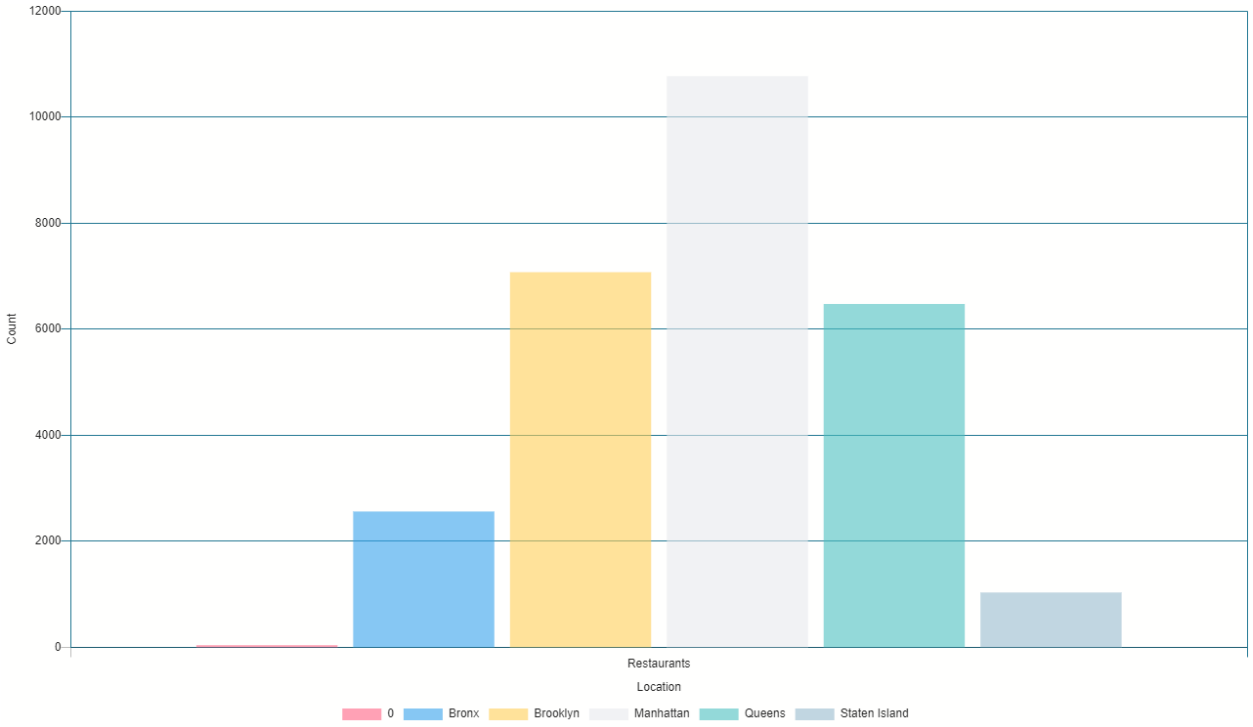

Figure 1. Number of Restaurants based on location

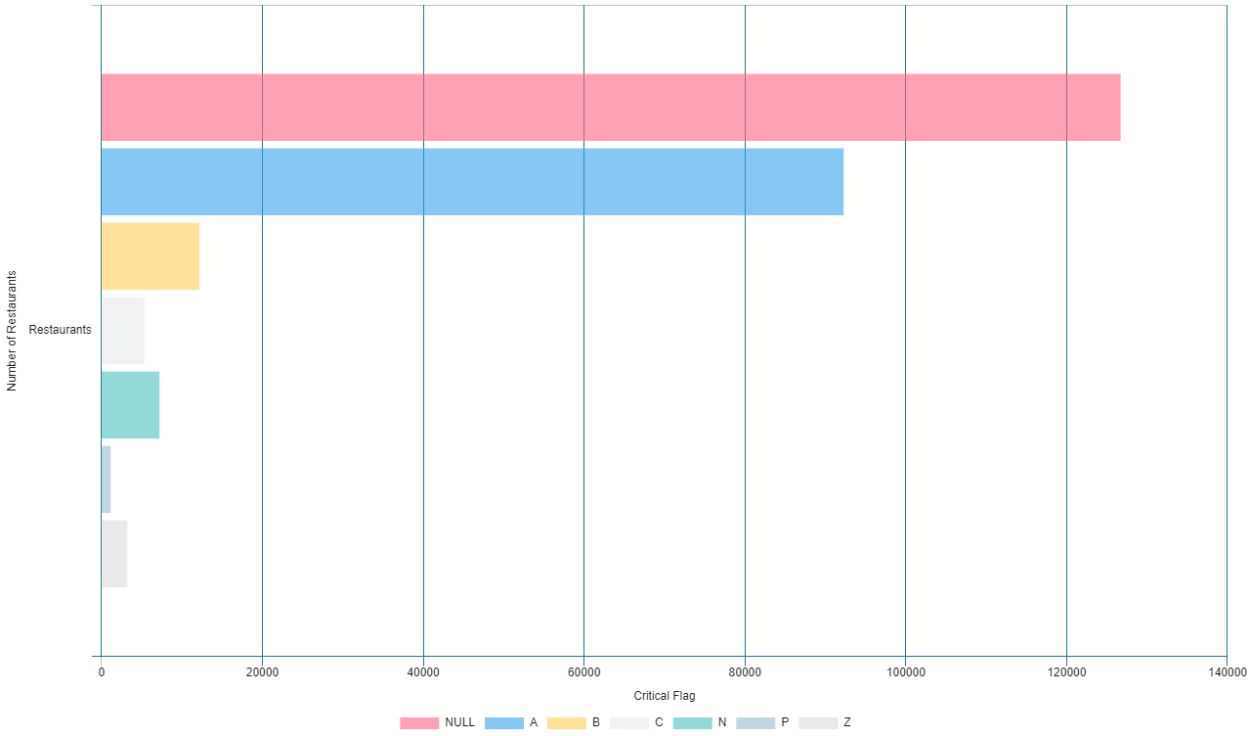

Figure 2. Number of Restaurant Based on Grade

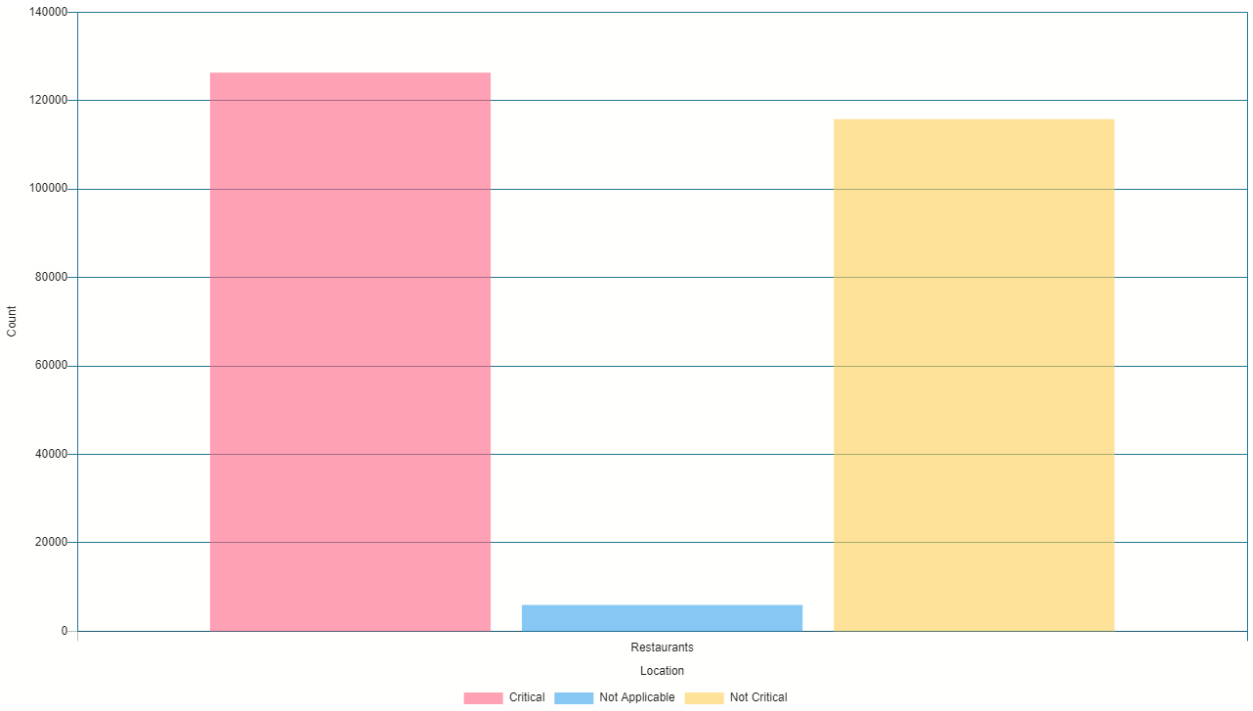

Figure 3. Number of restaurants based on critical flag assigned

Insights From Our SQL Queries

- We were able to filter out restaurants that were recently closed down by the New York City Department of Health and Mental Hygiene. One relationship among these restaurants that were recently closed down is that they were all recorded to be closed down on August 27th, 2022.
- We were able to filter out the restaurant with the lowest score, highest grade, and no violations were recorded for the restaurant on the most recent day the restaurant was inspected. This restaurant is called Maman, which is located in Manhattan.
- We were able to filter out the restaurant with the lowest score, highest grade, and no violations were recorded for the restaurant on the most recent day that the data was pulled. This restaurant is called PPL Coffee and it is located in Brooklyn.
- We were able to filter out the restaurant with the lowest score, highest grade, and no violations were recorded for the restaurant on the most recent day that the restaurant was graded. Maman, the restaurant that was previously mentioned in one of our previous insights also fulfills this filter result.

- By running another SQL query, we were able to filter out the inspection method with the average lowest score given to restaurants. This inspection method is called the Pre-permit (Non-operational) / Second Compliance Inspection method.

Conclusion

Ultimately, we have decided to choose Maman as our restaurant to do business in because it was given a grade of A and a score of 0 on the most recent day that any restaurant was inspected and on the most recent day that data for any restaurant was recorded into the SQL database.