**Caribbean Restaurant Chain Looking to Expand in New York City**

**Introduction**

A Texas based chain of Carribean restaurants is planning to expand in Brooklyn and Queens boroughs of New York City. They have approached a data scientist to analyze which neighborhood(s) of these boroughs have heavier footfall as far as Carribean food is concerned.

Data analysis will be performed to determine clusters of neighborhoods where Carribean food is popular, and hence will be a a good idea for the restaurant chain to first explore these neighborhoods for expansion of business.

**Data Sources**

Data for neighborhoods in Brooklyn and Queens will come from

https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN-SkillsNetwork/labs/newyork\_data.json

Data for venues will come from FourSquare.

**Methodology**

Project methodology involved creating clusters of neighborhoods in Brooklyn and Queens boroughs and find the cluster(s) of neighborhoods where Caribbean restaurants were a preferred venue.

Project was completed using Python and involved the following steps:­­

* Imported needed Python libraries
* Downloaded dataset for New York City containing information about boroughs, neighborhoods, and their longitudes and latitudes.
* Two separate dataframes were created, one each for Brooklyn and Queens boroughs.
* For each borough, information form FourSquare was used to obtain the list of venues and venue categories within a 5000 meters radius of different neighborhoods. This exercise yielded a list of 7000 and 8006 venues for Brooklyn and Queens respectively. There were 230 unique venue categories for Brooklyn while Queens had 259 unique categories.

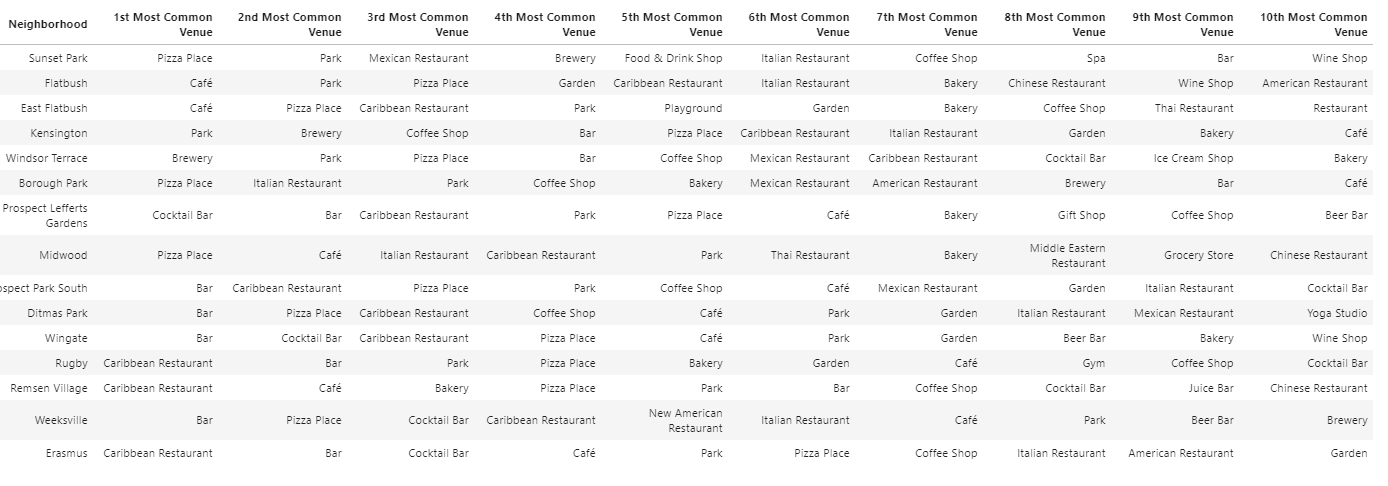
**Results and Observations**

(Note: For detailed results see,

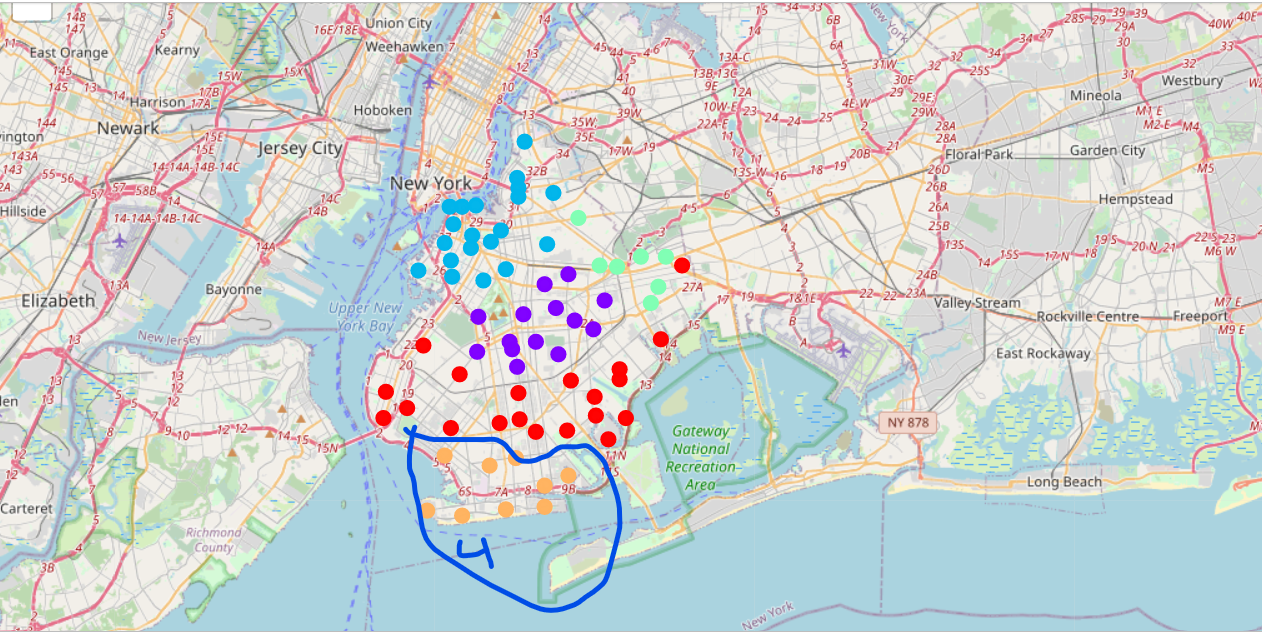
<https://github.com/vinaykansal/testing/blob/main/Brooklyn_Vs_QueensCleaned.ipynb>)

Clustering exercise identified that Cluster 4 in each of the two neighborhoods had a preponderance of Caribbean restaurants as first/second/third venue. Other neighborhood clusters did not have this marked preference for the Caribbean food. These neighborhood clusters for Brooklyn (Artifacts #1 and #2) and Queens (Artifacts #3 and #4) are shown below.

**Artifact #1: Neighborhoods in Cluster 4 for Brooklyn**



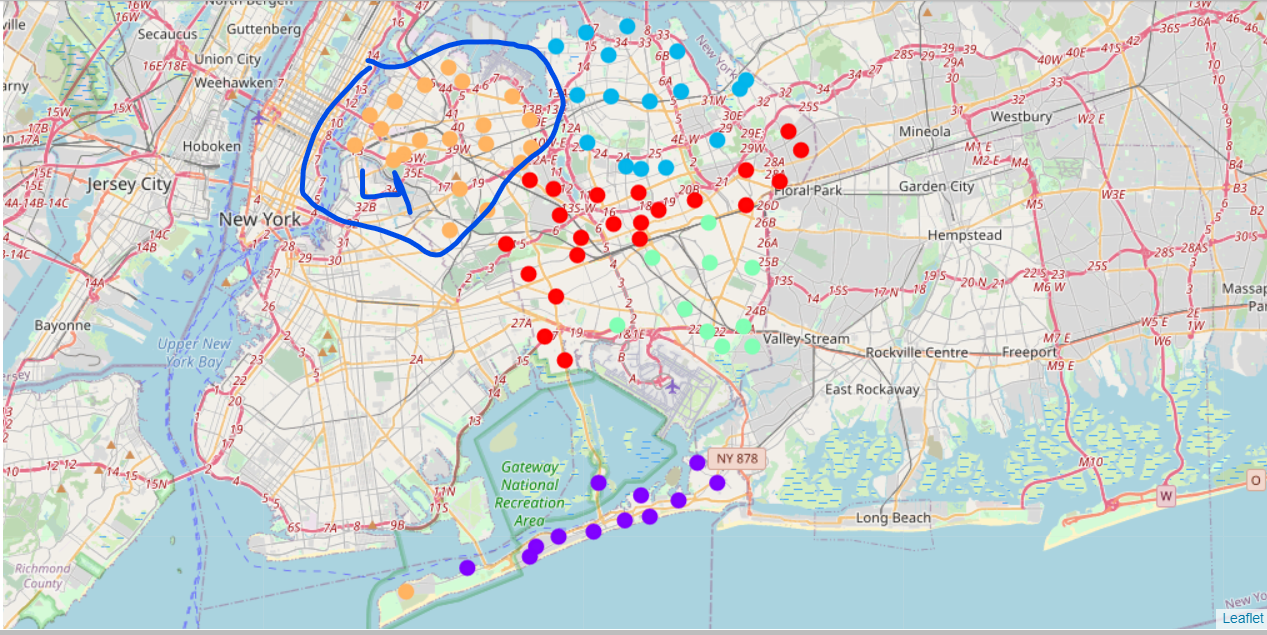
**Artifact #2: Mapping Neighborhoods in Cluster 4 for Brooklyn**



**Artifact #3: Neighborhoods in Cluster 4 for Queens**



**Artifact #4: Mapping Neighborhoods in Cluster 4 for Queens**



**Conclusion**

Based on the results of this thorough data analysis, the data analyst was able to identify the clusters in Brooklyn and Queens boroughs that had a preference for the Caribbean food. The Caribbean food chain from Texas will be better off focusing its expansion efforts in these neighborhoods.