The Battle Of Neighborhoods

Introduction:

- New York is the most densely populated major city in the United States.
- It is a global hub of business and commerce.
- The NYC provides a lot of opportunities to the emerging businesses.
- At same time it also means that the market scenario in NYC is highly competitive, which would involve a lot of risk factor for the new businesses.
- This requires a thorough market analysis and background research.

Business Problem:

- Shifting demographics and changing lifestyles are driving the surge in food-service businesses.
- This has led to the flooding of markets making it highly competitive.
- So, for a new restaurant to venture in such a competitive market would involve a lot of risk.
- Therefore, a thorough market analysis is required.
- The main aim of the project is to recommend a good neighborhood alternative to the Restaurant company to start their business, where there are lesser number of similar food-service businesses..

Data Description:

- We have used data for this project from the given link!
  https://cocl.us/new_york_dataset
  It's a json file.
- NYC has 5 boroughs and 306 neighborhoods .
- We'll be focussing on the neighborhoods in Brooklyn and Mnahattan.
- **Foursquare API Data:** NYC geographical coordinates data will be given as an input to the Foursquare API, in order to get information(Data) about different venues around each neighborhood in Brooklyn and Manhattan.

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

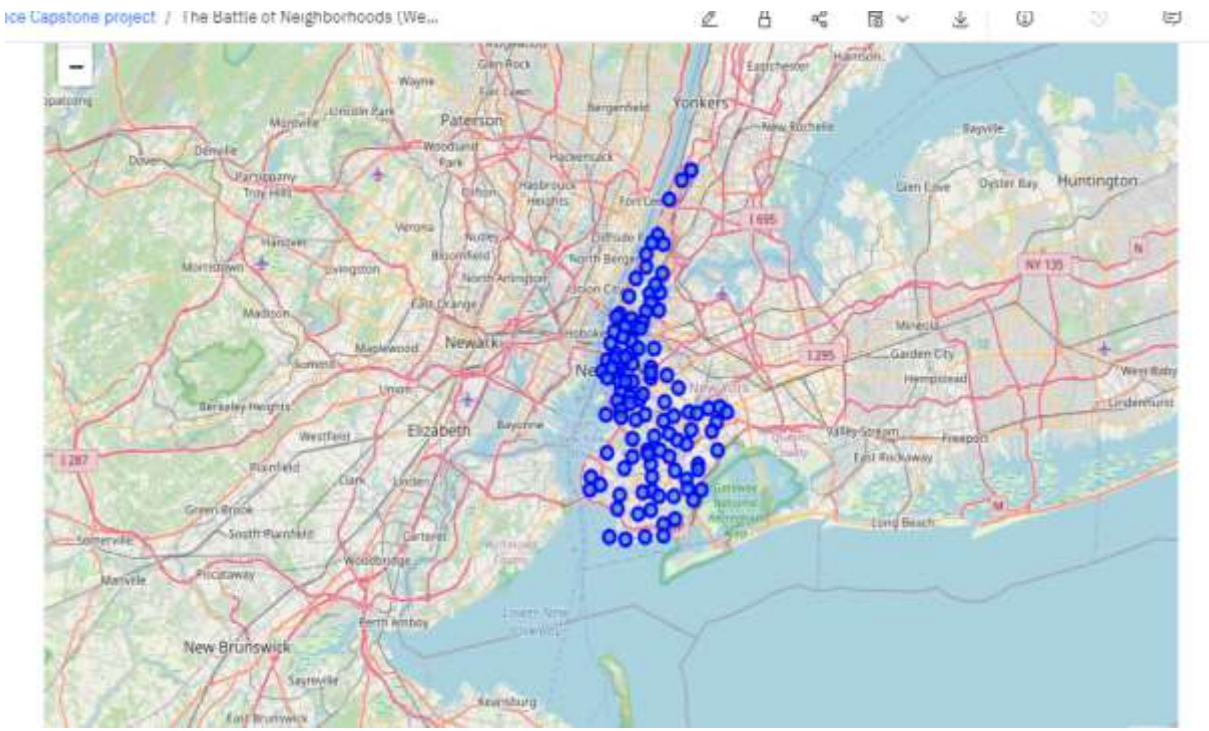
Analytic approach:

- NYC has 5 boroughs and 306 neighborhoods.
- Exploratory data analysis of the venues in Manhattan and Brooklyn.
- Only the Restaurant data is filtered from the foursquares venue data for the scope of this project.
- Then we have clustered different restaurant venues in these neighborhoods.

Methodology: slide1

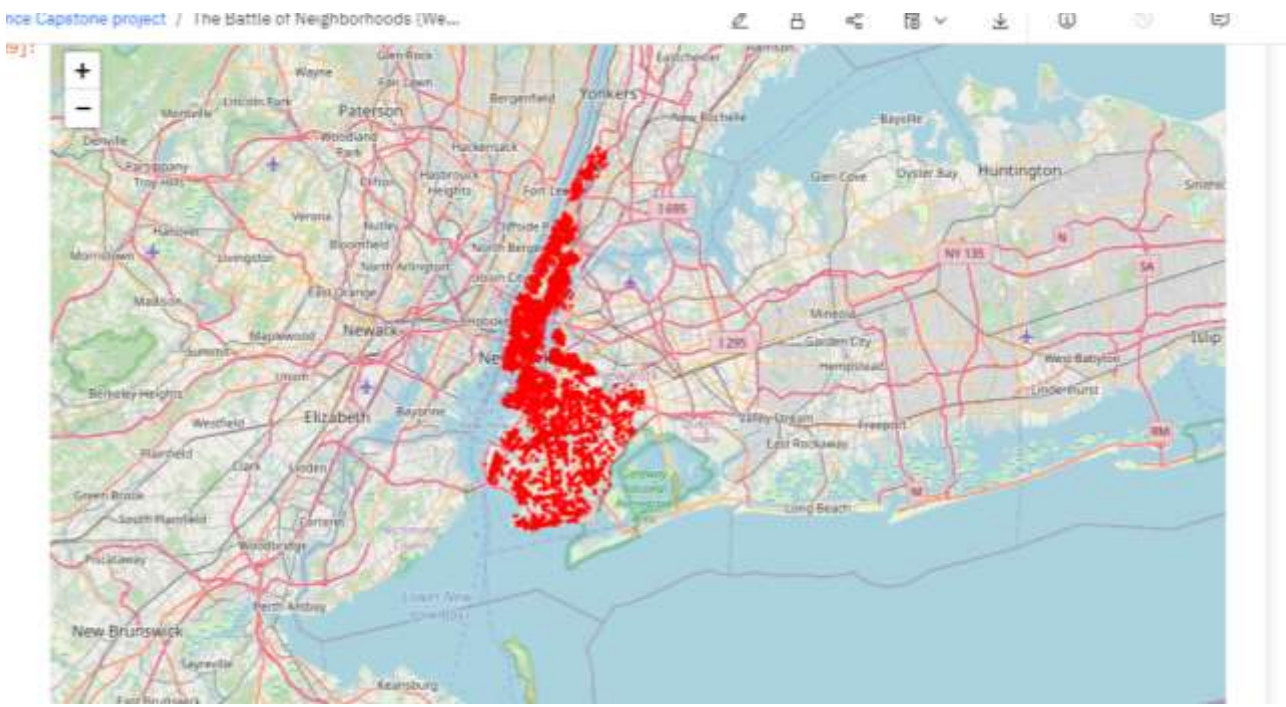
NYC's Geographical coordinates data:

- Load the data from the .json file.
- Transform the data into the pandas dataframe.
- Dataframe consists the geographical coordinates of NYC.
- We'll focus on Manhattan and Brooklyn data .
- Data will be used to get venues from foursquare api.
- Folium and geopy libraries are used to visualise the neighborhoods data .



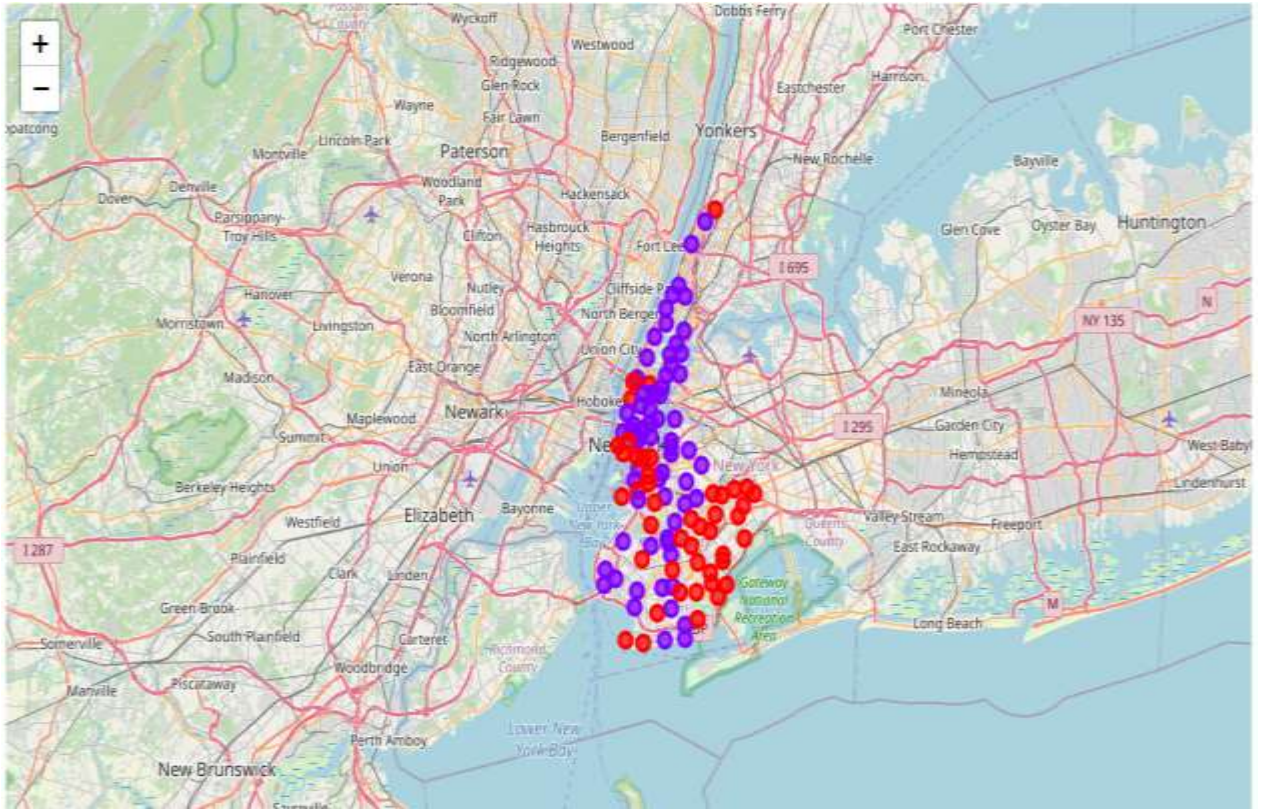
Slide2:

- Now using Forsquare's API we will fetch the data of nearby venues.
- Then using Folium and geopy we create a map of NYC .
- This map shows the different venues nearby Brooklyn and Manhattan.



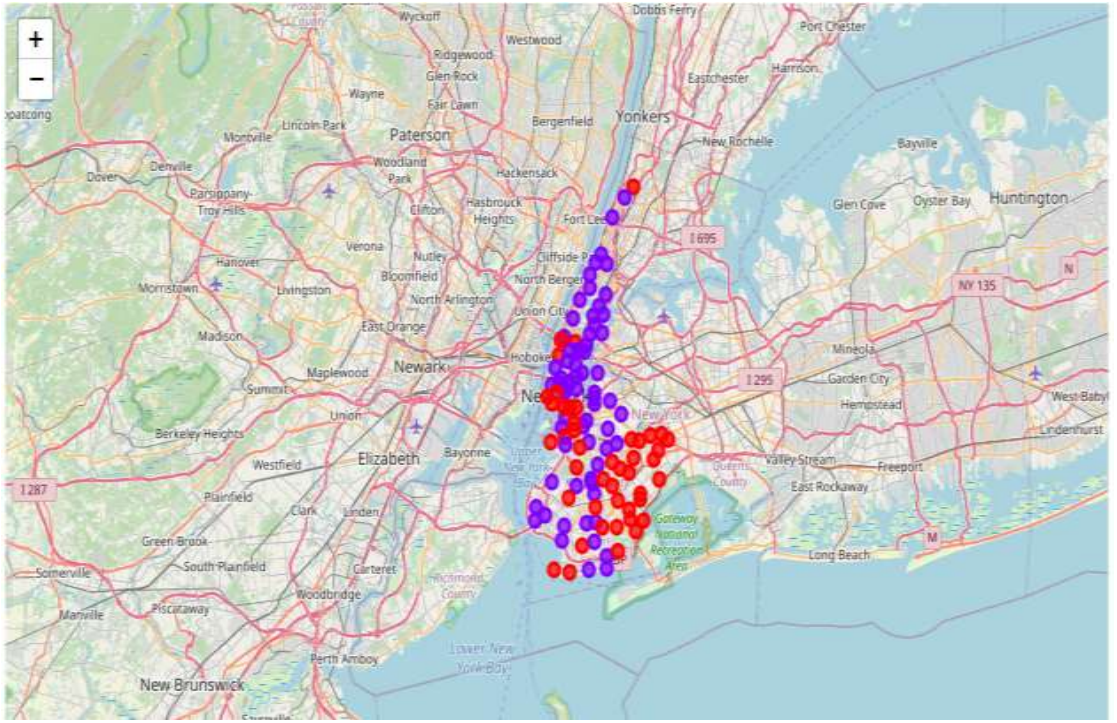
Slide 3:

- We filter the Restaurant data from the rest of the venues .
- After filtering them, we use KMeans clustering algorithm to segment different resaturants in the neighborhoods.
- Finally using Folium library we visualise the clusters.



Result:

- **Segmenting and Clustering the neighborhoods.**
- CLUSTER 0: The total sum of the cluster is low, hence shows the market is unsaturated (lesser number of restaurants).
- CLUSTER 1: The total sum of the cluster is high, hence shows that the market is saturated (significant number of restaurants).



Discussion:

- The major purpose of this project, is to suggest a better neighborhood in Brooklyn and Manhattan for starting a restaurant business here.
- The scope of the project can also be increased by taking into consideration Connectivity to the airport, bus stand, city centre, markets and other daily needs things nearby.

Conclusion:

- Amongst all the neighborhoods in Brooklyn and Manhattan we can suggest few neighborhoods that have lesser number of restaurants .
- For e.g. East Flatbush, East New York, Mill Island etc.
- In these neighborhoods, there is lesser competition with lower risk factor, hence it's good to invest .