Bengali Restaurant in New York City

APPLIED DATA SCIENCE CAPSTONE

IBM DATA SCIENCE

Business Problem

▶ The location of food service business will impact its success nearly as much as the menu. If the restaurant is in the wrong place, we cannot attract the number of customers we will need in order to stay in business. The same is true if our location is inundated with competitors, or has poor visibility, or is hard to find. There are many things that must be considered as we look for a location in which to open our business.

▶ The objective of the Capstone project is to analyze and select the best locations in the city of New York to open a new Bengali/Bangladeshi restaurant, using data science methodology and instruments such as data analysis and visualization.

Data

- To solve the problem, we will need the following data:
  - New York City data containing the neighborhoods and boroughs.
  - Latitude and longitude coordinates of those neighborhoods.
  - Venue data, particularly data related to restaurants.
- New York City data containing the neighborhoods and boroughs will be obtained from the open data source: **https://cocl.us/new_york_dataset**.
- Next, we will get the geographical coordinates of the neighborhoods (latitude and longitude) using the **Python Geocoder** package.
- We will use the **Foursquare API** to get the venue data for the neighborhoods defined at the previous step. Foursquare has one of the largest databases of 105+ million places, 900+ venue categories, 30+ attribute fields, and over 150,000 developers use this application. Among the many Foursquare API provided categories of the venue data, we are particularly interested in the restaurant data to solve the business problem defined above.

Methodology

- Collect data from https://cocl.us/new_york_dataset.
- Clean and process data into a dataframe.
- Use Foursquare to locate all venues.
- Filter by Bangladeshi restaurants.
- Sort data based on rankings.
- Visually assess data using graphing from Python libraries.

Results

▶ In Foursquare database, we got 0 Bangladeshi Restaurants across NYC. This should be quite good for business as no pure Bangladeshi restaurant is available to cater to the increasing Bangladeshi population. For further analysis, we will check for Indian restaurants which are similar to the Bangladeshi cuisine.

▶ Manhattan has the least number of neighborhoods, but it does have the highest number of Indian restaurants.

▶ Sutton Place in Manhattan has the highest number of Indian Restaurants with a total count of 3.

▶ Manhattan has the highest average rating for Indian Restaurants. Brooklyn goes right behind the Manhattan with the average rating of its Indian restaurants only a little lower than Manhattan.

▶ Two neighborhoods with the highest average rating for Indian Restaurants are Tribeca and Chelsea (both located in Manhattan).

Discussion

▶ Based on the above information, we can state that Manhattan and Brooklyn are comparatively better locations for Bangladeshi cuisine in NYC.

▶ We can open a Bangladeshi restaurant in Brooklyn which has multiple neighborhoods with average ratings exceeding 8.0 on a scale of 1.0 to 10.0 and has a smaller number of Indian restaurants than Manhattan, making competition easier.

▶ Brooklyn is also close to Jackson Heights, Queens which the main hub for Bangladeshi population. Moreover, the real estate prices in Brooklyn are much cheaper than in Manhattan.

Conclusion

▶ All analysis in the report is depended on the accuracy of Foursquare data.

▶ A more comprehensive analysis would need to incorporate data from other external databases like area-specific population demographics, real estate prices etc.