Capstone Project - The Battle of Neighborhoods
Introduction

IBM Data Science Professional

New York City's demographics show that it is a large and ethnically diverse metropolis. It is the largest city in the United States with a long history of international immigration. New York City was home to nearly 8.5 million people in 2014, accounting for over 40% of the population of New York State and a slightly lower percentage of the New York metropolitan area, home to approximately 23.6 million. Over the last decade the city has been growing faster than the region. The New York region continues to be by far the leading metropolitan gateway for legal immigrants admitted into the United States.

Throughout its history, New York City has been a major point of entry for immigrants; the term "melting pot" was coined to describe densely populated immigrant neighborhoods on the Lower East Side. As many as 800 languages are spoken in New York, making it the most linguistically diverse city in the world. English remains the most widely spoken language, although there are areas in the outer boroughs in which up to 25% of people speak English as an alternate language, and/or have limited or no English language fluency. English is least spoken in neighborhoods such as Flushing, Sunset Park, and Corona.

With it's diverse culture , comes diverse food items. There are many resturants in New york City, each beloning to different categories like Chinese , Indian , French etc.

Data

For this project we need the following data :
- New York City data that contains list Boroughs, Neighborhoods along with their latitude and longitude.
  - Data source : https://cocl.us/new_york_dataset
  - Description : This data set contains the required information. And we will use this data set to explore various neighborhoods of new york city.
- Indian resturants in each neighborhood of new york city.
  - Data source : Fousquare API
  - Description : By using this api we will get all the venues in each neighborhood. We can filter these venues to get only indian resturants.
- GeoSpace data
  - Data source : https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm
  - Description : By using this geo space data we will get the New york Borough boundaries that will help us visualize choropleth map.

Approach
- Collect the new york city data from https://cocl.us/new_york_dataset
- Using FourSquare API we will find all venues for each neighborhood.
- Filter out all venues that are Indian Resturants.
- Find rating , tips and like count for each Indian Resturants using FourSquare API.
- Using rating for each resturant , we will sort that data.
- Visualize the Ranking of neighborhoods using folium library(python)

Questions that can be asked using the above mentioned datasets
- What is best location in New York City for Indian Cuisine ?
- Which areas have potential Indian Resturant Market ?
- Which all areas lack Indian Resturants ?
- Which is the best place to stay if I prefer Indian Cuisine ?

Analysis
We will import the required libraries for python.
- pandas and numpy for handling data.
- request module for using FourSquare API.
- geopy to get co-ordinates of City of New York.
- folium to visualize the results on a map

Discussion

Discussion section where you discuss any observations you noted and any recommendations you can make based on the results.

When I first decided to create this study I was expecting to find clusters of Indian restaurants in certain regions and the final result should meet our expectation. So as part of this project , we will list and visualize all major parts of New York City that has great Indian restaurants.