Indian Restaurant Analysis in NJ and NY

Data Science Capstone - IBM Data Science Professional Certificate on Coursera

Abstract

Analyzing Indian Restaurants and their density for setting up a new restaurant.

Indian Restaurant Analysis in NJ and NY cities.

Data Science Capstone - IBM Data Science Professional Certificate on Coursera

Introduction

A business owner of successful Indian chain restaurants wants to open his new branch in NY and NJ areas. He is not aware of the area and he wants to identify the best city in these two states to open the restaurants by using the density of already existing Indian restaurants in the area.

Data Section

For this project I will be using Four Square API. I will be scouting for the Indian restaurants in New York City, Newark, Edison, Elizabeth, Secaucus. These are some of the cities in that geographical area with highest number of Indian communities and Indian restaurants.

Project Plan

For this project I will be using data from Foursquare API to assess the density of Indian restaurants through venues in the cities mentioned in Data Section. From Four Square developer website I got the category of Indian Restaurants - 4bf58dd8d48988d10f941735.

I will be using the coordinates of these restaurants to map them on the folium maps imported for visual inspection. Next, I will also try to find the centroid of these restaurants in every city. Once I have the centroid Coordinates for each city, I will be calculating the distance of each restaurant from the centroid.

Python Packages I will be using for this project are

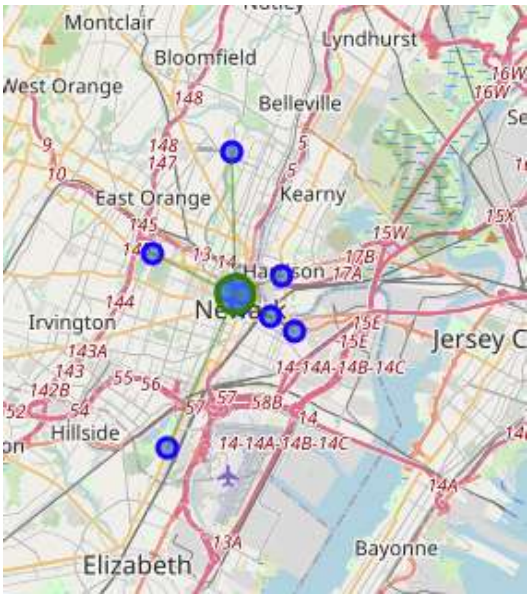
- Numpy
- Pandas
- Json
- folium

Results

After importing and plotting the data from Four Square API, on maps using geoplot by folium here are the pictures.

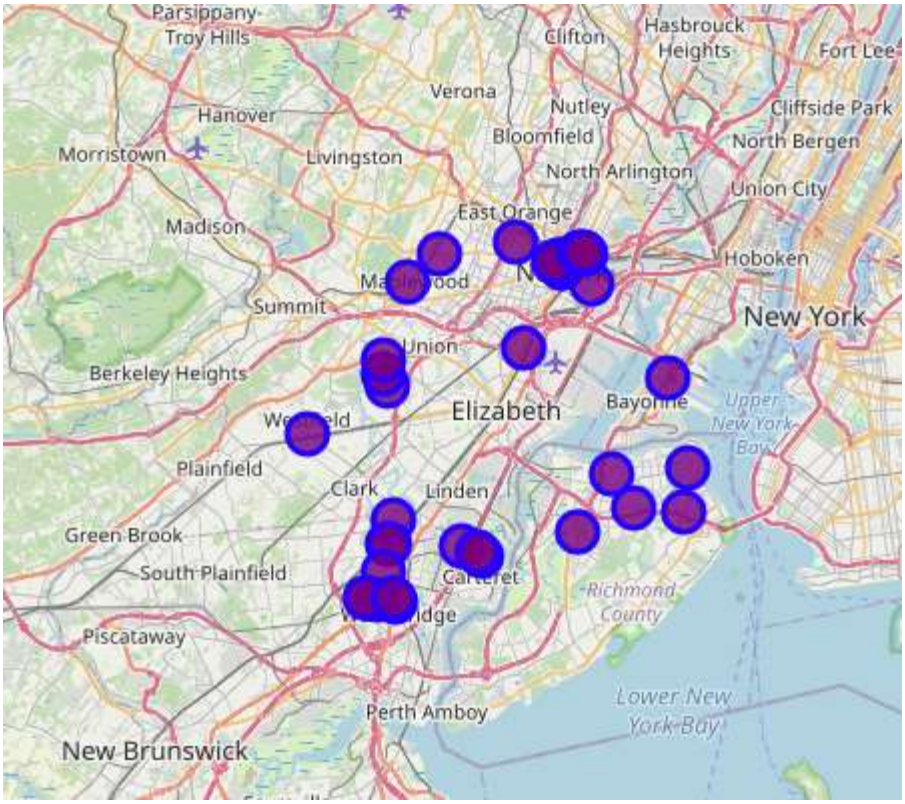
New York NY:

Newark NJ:

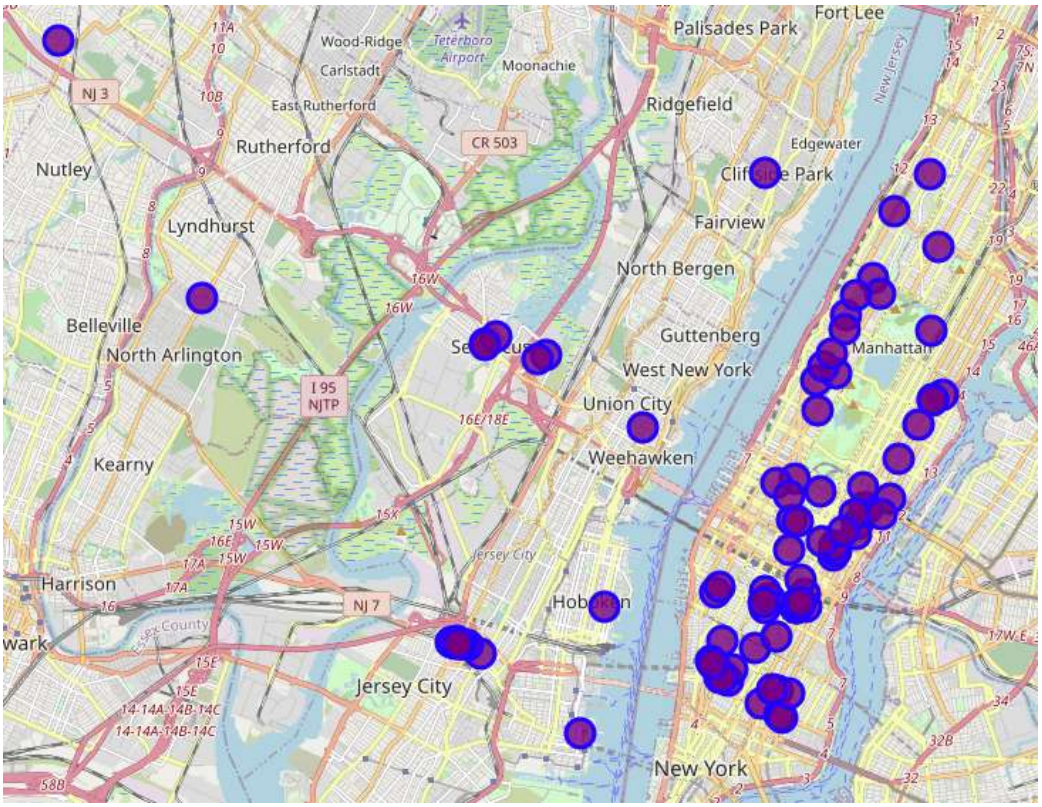


Edision NJ:

Elizabeth NJ:



Secaucus NJ:



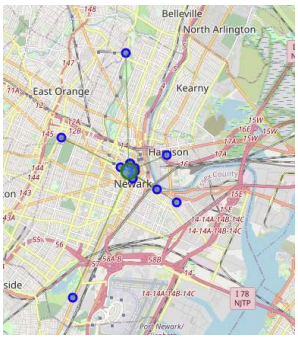
Observation:

Here we have identified that in Four Square data almost all the restaurants in Secaucus are also present in New York NY. So, we can ignore Secaucus from the Analysis.

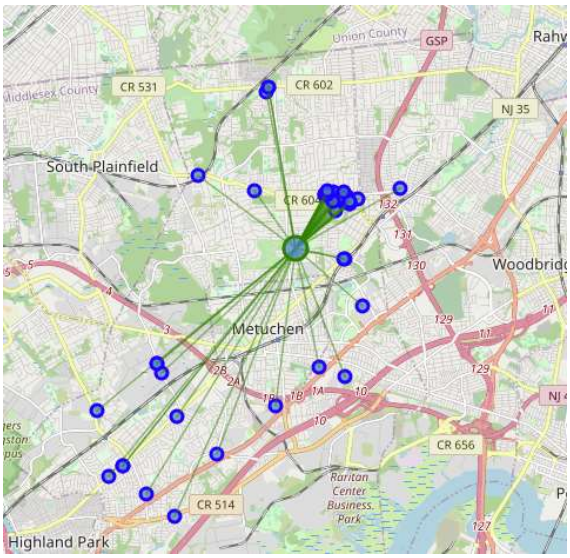
Now let us calculate the density of restaurants in Each City.

New York NY:

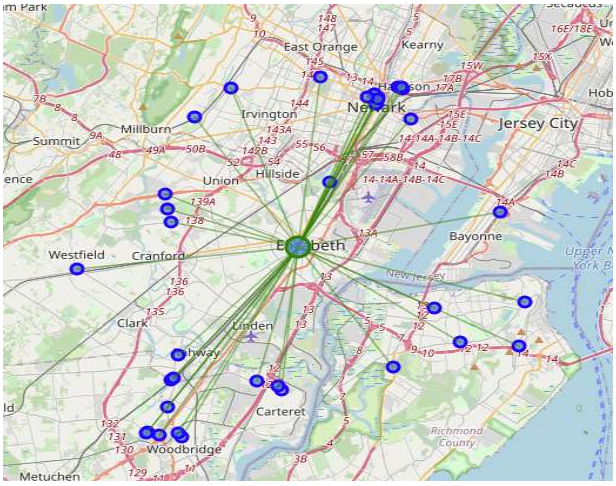
Newark NJ:



Edison NJ:



Elizabeth NJ:



Discussion & Conclusion:

Indian Restaurants in New York, NY = 254
Indian Restaurants in Newark, NJ = 13
Indian Restaurants in Edison, NJ = 60
Indian Restaurants in Elizabeth, NJ = 37
Indian Restaurants in Secaucus, NJ = 222

New York, NY Mean Distance from Mean coordinates 0.025204772613466226
Newark, NJ Mean Distance from Mean coordinates 0.015400081373466852
Edison, NJ Mean Distance from Mean coordinates 0.026978970644616223
Elizabeth, NJ Mean Distance from Mean coordinates 0.09812715045889935
Secaucus, NJ Mean Distance from Mean coordinates 0.042743460153164846

If We compare the total number of restaurants, New York stands on the top. But by the Density Newark has highly dense restaurants. But the total number of restaurants in Newark is 13. So, it's safe to ignore Newark. The Next Dense Ranking and Number of restaurants go to Edison.

As New York is already crowded, its better if the Restaurant is opened in Edison NJ

By adding additional data points like - Asian Indian Population Density, Ratings of the existing Restaurants etc. this project's scope can be further expanded.