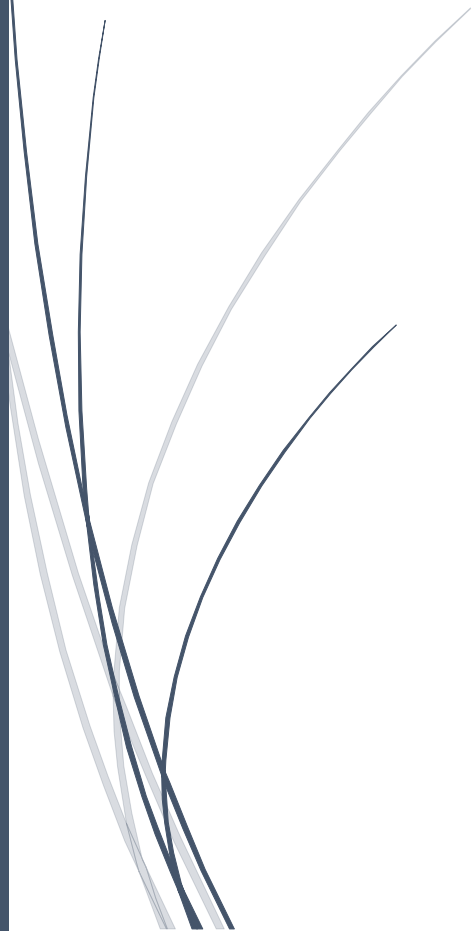
BATTLE OF NEIGHBORHOODS

[Date]

07-05-2020

APPILED DATA SCIENCE PROJECT



Saksham Midha

BATTLE OF NEIGHBORHOODS

1. INTRODUCTION

DELHI IS KNOWN as the food capital of India. The city is famous for a variety of foods, dishes, and restaurants. It was recently named “Best Destination for Food/Drink in India”.

BUSINESS PROBLEM

- As part of my ‘Applied Data Science Capstone’ final project, I shall survey the type of restaurants in different neighborhoods of Delhi to determine the most suitable place to open an Indian restaurant. The use of Data science techniques and machine learning algorithms such as Clustering, shall help us answer the above described business problem.

TARGET AUDIENCE

- The target audience of the above business problem is a restauranteur who is looking for the right neighborhood to open an Indian restaurant.

2.*DATA*

To solve the above business problem, we shall need the following data:-

- A list of neighborhoods in Delhi.
- Latitude and Longitude data of neighborhoods- Latitude and longitude data could be extracted using the Geocoder package.
- Venue-related data- Venue related data could be extracted using the Foursquare API.

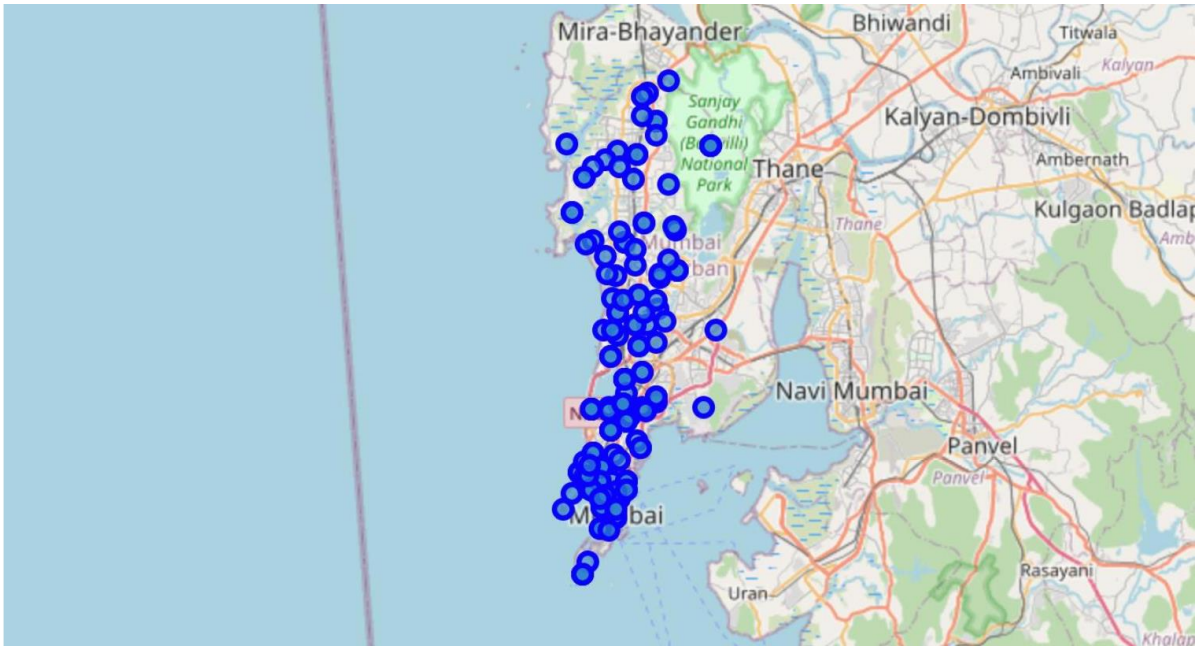
3.METHODOLOGY

We begin by making a data-frame that contains a list of all the neighborhoods in Delhi, along with their latitude and longitude.

Delhi_Neighborhood_Data

	Location	Pincode	State	District	Latitude	Longitude
0	Ashok Vihar	110052	Delhi	North West Delhi	28.688390	77.688390
1	Bawana	110039	Delhi	North West Delhi	28.800000	77.030000
2	Begum Pur	110086	Delhi	North West Delhi	28.727140	77.059480
3	Dhaka	110052	Delhi	North West Delhi	28.707800	77.205700
4	Karala	110081	Delhi	North West Delhi	28.738000	77.039200
5	Keshav Puram	110035	Delhi	North West Delhi	28.688900	77.161600
6	Kingsway Camp	110052	Delhi	North West Delhi	28.415500	77.121600
7	Model Town	110009	Delhi	North West Delhi	28.718000	77.191600
8	Narela	110040	Delhi	North West Delhi	28.850000	77.100000
9	Pitam Pura	110034	Delhi	North West Delhi	28.000000	77.000000
10	Rani Bagh	110034	Delhi	North West Delhi	28.687900	77.134200
11	Rohini	110085	Delhi	North West Delhi	28.732865	77.124180
12	Shalimar Bagh	110088	Delhi	North West Delhi	28.696400	77.127100
13	Chandni Chowk	110006	Delhi	North Delhi	28.656000	77.231000
14	Chawri Bazar	110006	Delhi	North Delhi	28.650402	77.229379

- Then we find the latitude and longitude of Delhi from the Nominatim package, so as to plot the map of Delhi with all the neighborhoods using the Folium package.

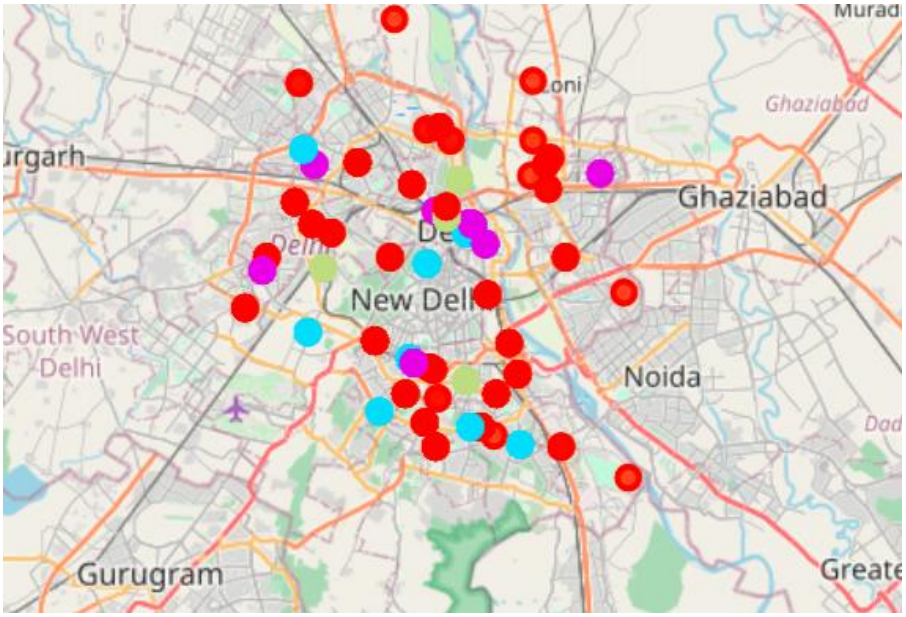


-Then we filtered out the venues that just had 'Indian Restaurant' as the Venue Category, as we just want to take Indian Restaurants into consideration.

	Location	Indian Restaurant
0	Ashok Nagar	0.000000
1	Babarpur	0.000000
2	Badarpur	0.000000
3	Bali Nagar	0.000000
4	Bawana	0.000000
5	Chanakyapuri	0.047619
6	Chandni Chowk	0.357143
7	Chawri Bazar	0.272727
8	Chittaranjan Park	0.000000
9	Civil Lines	0.090909

-After this we used “KMeans” clustering to create clusters, and group the neighborhood in Delhi into different clusters, based on Indian Restaurants for further analysis of each area.

4.***RESULTS***



Cluster 0, Cluster 1, Cluster 2, Cluster 3

The results of the clustering are as follows:-

Cluster 0, has the least number of Indian Restaurants, so it is the most ideal place to open an Indian Restaurant.

Location	Indian Restaurant	Cluster Labels	Pincode	Latitude	Longitude	Ver
0	0	0	0	0	0	0

Cluster 1,2 and 3 have a lot of Indian Restaurants, with Cluster 1 having the most.

5. ***DISCUSSION***

The accuracy of determining the best places to open an Indian Restaurant can be improved by not just taking the number of Indian Restaurants into consideration, but also taking into consideration the demographic of a certain area, and determining the food preferences of the people in a certain neighborhood. Since, that data is not readily available, and can only be gathered through a widespread survey, for the time being only the number of restaurants has been used to find the solution to the posed Business problem. In the future, when such data is available we can update our findings by further incorporating the extra data.

6.***CONCLUSION***

Thus, to conclude the above study, one can say that there is always scope for more improvement, by finding more relevant data and increase its accuracy to narrow down the study, as discussed in the discussion section.