Suggesting Best Restaurant Around

Pankhuri Arora

June 9, 2020

1. Introduction

1.1 Problem

There are thousands of restaurants are available these days in a city but not all restaurants are good to try when you just visit on a vacation of day or two. This project aims to suggest best restaurant to try based on locality or cuisine.

1.2 Interest

This is useful for the localities as well as tourists. As the project suggests the best restaurant for them to try out and enjoy their time.

2. Data acquisition and cleaning

2.1 Data source

Restaurant ID, name and their location is available in the dataset which is available on kaggle. The data contains almost all the restaurant names and id that are present in New Delhi. Later used foursquare API which is used to gather detailed information including reviews about the restaurant.

2.2 Data cleaning

Firstly, dataset I get from Kaggle was World data. I filtered it down to New Delhi restaurants only. Later, eliminated all the restaurants that has 0 rating.

3. Exploratory Data Analysis

3.1 Calculation of target variable

Firstly, plotted all the restaurants in the dataset to visualize the area they are in. Later, dataset is filtered by looking for rows with same locality which has high rating.

4. Methodology

I used dataset for restaurant available on kaggle along with Foursqaure API for this project.
After cleaning the data and bringing it to the required form, k-mean clustering is used to form groups of restaurants. To get the optimal K-mean, k-mean elbow method is used.

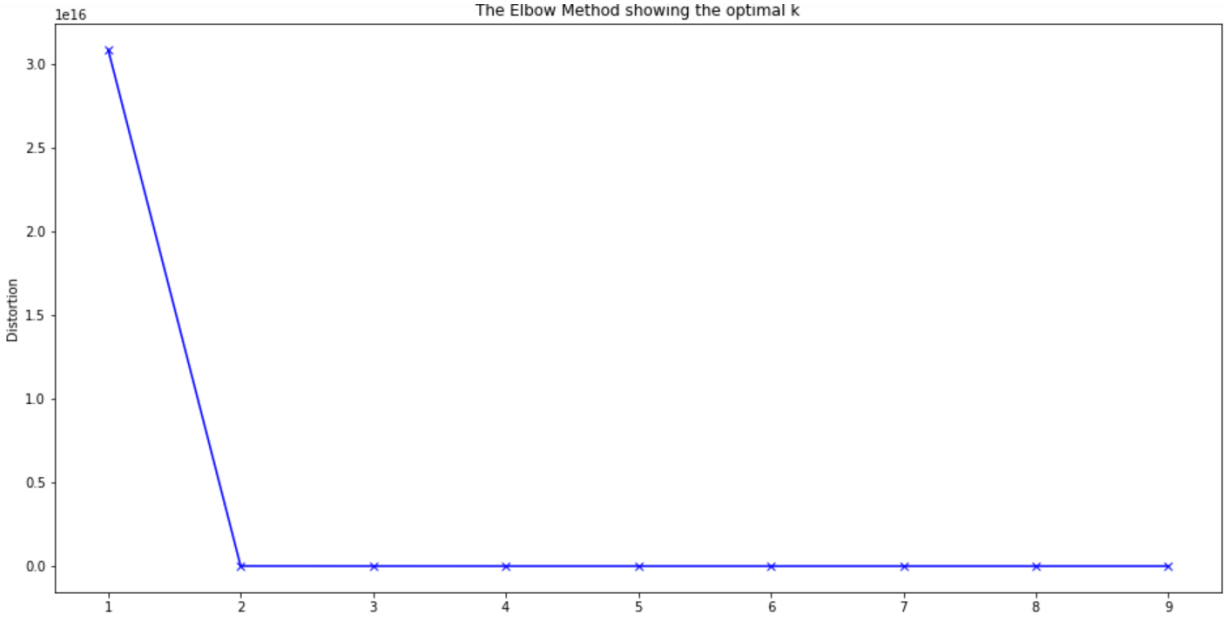

Fig 1.1

I used python folium library to visualize geographic details of all the restaurants that have high ratings in every area inside New Delhi. I used the longitude and latitude values to get the visual as below:

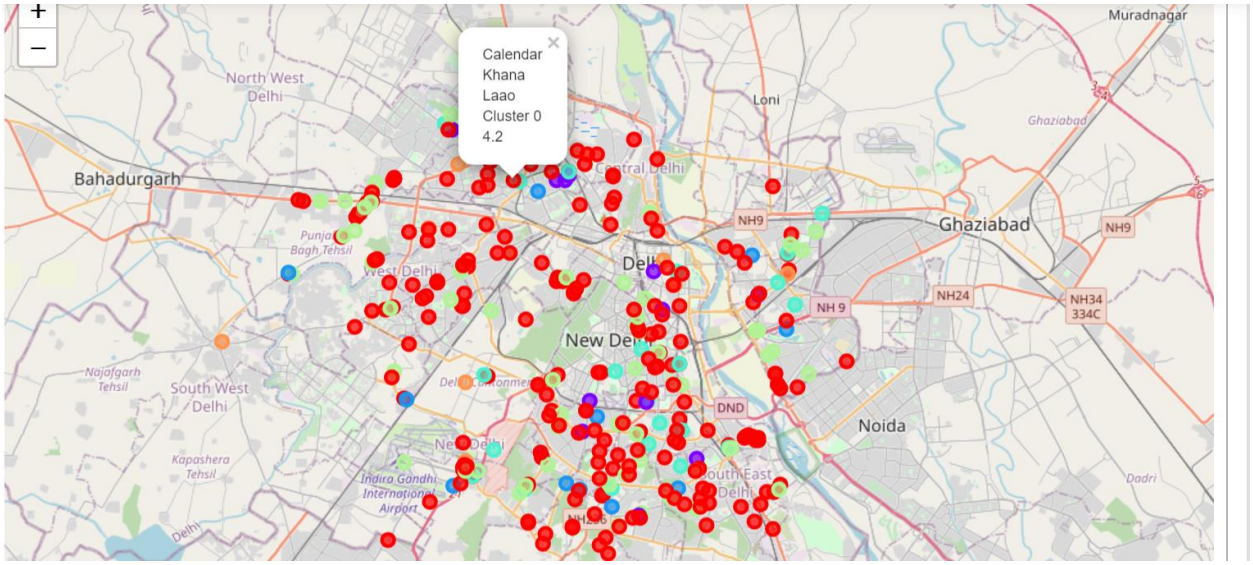

Fig 1.2

Repeated the same process to get best restaurant for different cuisine in New Delhi and visualized it to get the visual as below:

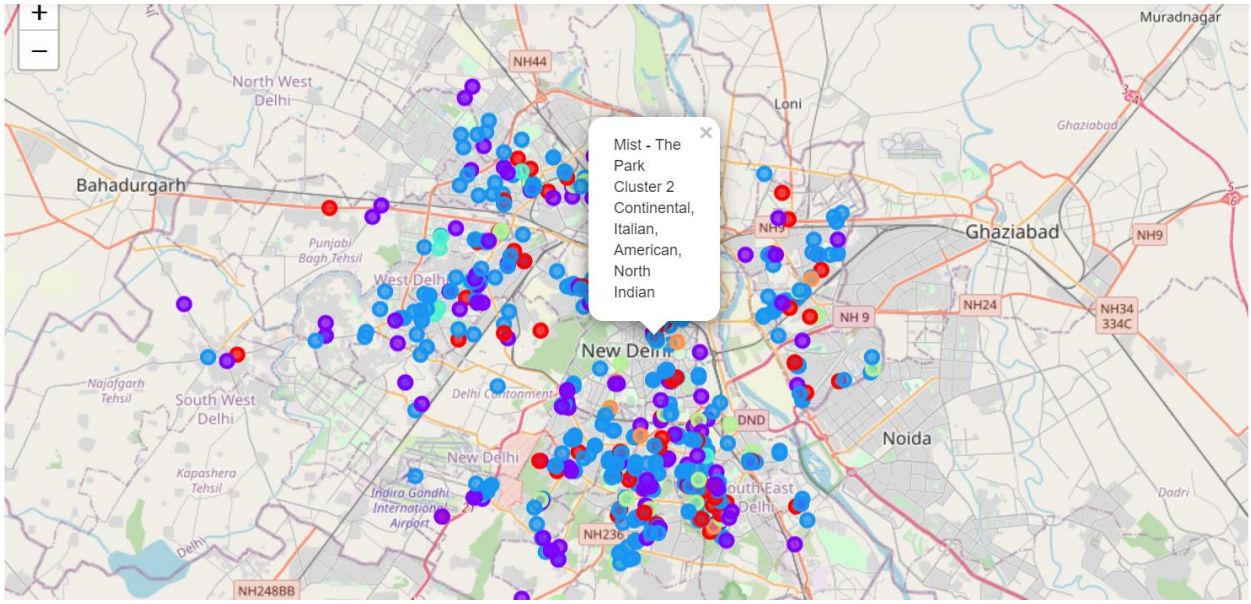

Fig 1.3

After visualizing, I used drop downs to select the cuisine and locality which will result in the name and address of the restaurant best suggested.

Fig 1.4

These dropdown results in the restaurant name and address which are highest rated for.

In the end, I used foursquare API to get reviews and detailed information about the restaurant that are resulted in dropdown results.

5. Results

This model results in the best restaurant per locality per cuisine. Using dropdowns one can easily get the name and address of the restaurant which is best rated for selected cuisine in selected locality.

6. Discussion

As I mentioned before, New Delhi have 5000+ restaurants and a population of 2.18 crores. As people have become more interested in eating out and enjoying the time model like these helps them to suggest the best pick for them. As there is such a complexity, very different approaches can be tried in clustering and classification studies. Moreover, it is obvious that not every classification method can yield the same high-quality results for this metropole.

I used K mean algorithm as part of this clustering study. When I tested the Elbow method, I set the optimum K value to 3. I tried to visualize as many restaurants as possible. For more detailed and accurate guidance, dataset can be expanded, and thorough details of restaurant can be drilled.
I ended the study by visualizing the data and clustering information on the New Delhi map. In future studies, restaurants with most positive reviews and based on services can be suggested.

7. Conclusion

As a result, restaurant business is at peak as people have started to visit it like they used to visit wonders. Also, restaurants are expanding in numbers day by day so it is required to have a model that can list down the restaurants available in the locality.

8. References

8.1 https://en.wikipedia.org/wiki/New_Delhi

8.2 https://developer.foursquare.com/