Applied Data Science Capstone Project
The Battle of Neighborhoods

-Prerita Kalra

Introduction

- India is a country with varied cuisines. Given the diversity in soil, climate, ethnic groups, cultures and occupations of people in various parts of India, these cuisines vary substantially- making use of locally available spices, herbs, fruits and vegetables.

- Indian cuisines have influenced many other cuisines across the globe. Even the global cuisines that have been brought into India have been given a local twist. This makes India a preferred spot for food connoisseurs around the globe.

- This project was aimed at analysing the venues in the capital cities of India so as to find the cuisines of popular restaurants located there. This would help the food connoisseurs plan their trip across India based on their cuisine preferences and save them the trouble of having to search individual cities for the cuisines served there.

Data Description

The following sources of data were used for the project-

1. The list of the capital cities of India was scraped from the Wikipedia page- "https://en.wikipedia.org/wiki/List_of_state_and_union_territory_capitals_in_India".

States and Union Territories [edit]

No.	State or union territory	Administrative capital	Legislative capital	Judicial capital	Year of establishment	Former capital
1	Andaman and Nicobar Islands †	Port Blair	—	Kolkata	1956	—
2	Andhra Pradesh	Amaravati (*de facto* from 2017)[3][4][a]	Amaravati[3]	Amaravati	1956 2014	Hyderabad[b] (1956–2017)
3	Arunachal Pradesh	Itanagar	Itanagar	Guwahati	1987	—
4	Assam	Dispur	Guwahati	Guwahati	1972	Shillong[c] (1950–1972)
5	Bihar	Patna	Patna	Patna	1950	—
6	Chandigarh †	Chandigarh[d]	—	Chandigarh	1966	—
7	Chhattisgarh	Naya Raipur	Raipur	Bilaspur	2000	—
8	Dadra and Nagar Haveli and Daman and Diu †	Daman	—	Mumbai	2020	—
9	Delhi †	New Delhi	New Delhi	New Delhi	1956	—
10	Goa	Panaji[e]	Porvorim	Mumbai	1987	—
11	Gujarat	Gandhinagar	Gandhinagar	Ahmedabad	1960	Ahmedabad (1960–1970)
12	Haryana	Chandigarh	Chandigarh	Chandigarh	1966	—
13	Himachal Pradesh	Shimla	Shimla (Summer) Dharamsala (Winter)[8]	Shimla	1971 2017	—
14	Jammu and Kashmir †	Srinagar (Summer)[9] Jammu (Winter)	Srinagar (Summer) Jammu (Winter)	Srinagar (Summer) Jammu (Winter)	2019	—
15	Jharkhand	Ranchi	Ranchi	Ranchi	2000	—
16	Karnataka	Bangalore	Bangalore	Bangalore	1956	—
17	Kerala	Thiruvananthapuram	Thiruvananthapuram	Kochi	1956	—
18	Ladakh †	Leh (Summer)		Srinagar (Summer)	2019	

Figure 1. The table on Wikipedia page- "List of state and union territory capitals in India", from which the "Administrative capital" column data was used.

2. Python's Geopy package was used to get the latitudes and longitudes of the capital cities of India.

3. The Foursquare API was used to explore the capital cities of India.

Methodology

The following steps were taken to analyse the venues in the capital cities of India for finding the cuisines of popular restaurants located there-

1. A dataframe containing the list of capital cities of India was formed by scraping the Wikipedia page- List of state and union territory capitals in India (https://en.wikipedia.org/wiki/List_of_state_and_union_territory_capitals_in_India). The data from the "Administrative capital" column of the table shown in Figure 1 was used for the analysis.

2. Python's Geopy package was used to add the latitudes and longitudes of the capital cities of India to the dataframe.

	Capital City	Latitude	Longitude
0	Port Blair	11.664535	92.739045
1	Amaravati	16.574420	80.355608
2	Itanagar	27.097966	93.623729
3	Dispur	26.151308	91.793380
4	Patna	25.609324	85.123525
5	Chandigarh	30.719402	76.764655
6	Naya Raipur	21.161027	81.786441
7	Daman	20.420005	72.863763
8	New Delhi	28.614179	77.202266
9	Panaji	15.498995	73.828214
10	Gandhinagar	23.223288	72.649227
11	Shimla	31.104153	77.170973
12	Srinagar	34.074744	74.820444
13	Jammu	32.718561	74.858092
14	Ranchi	23.370035	85.325013
15	Bangalore	12.979120	77.591300
16	Thiruvananthapuram	8.505891	76.957048
17	Leh	34.164203	77.584813
18	Kargil	34.420894	75.665486
19	Kavaratti	10.559050	72.636569
20	Bhopal	23.258486	77.401989
21	Mumbai	18.938771	72.835335
22	Imphal	24.800609	93.937000

Figure 2. Dataframe created using webscraping and Python's Geopy package.

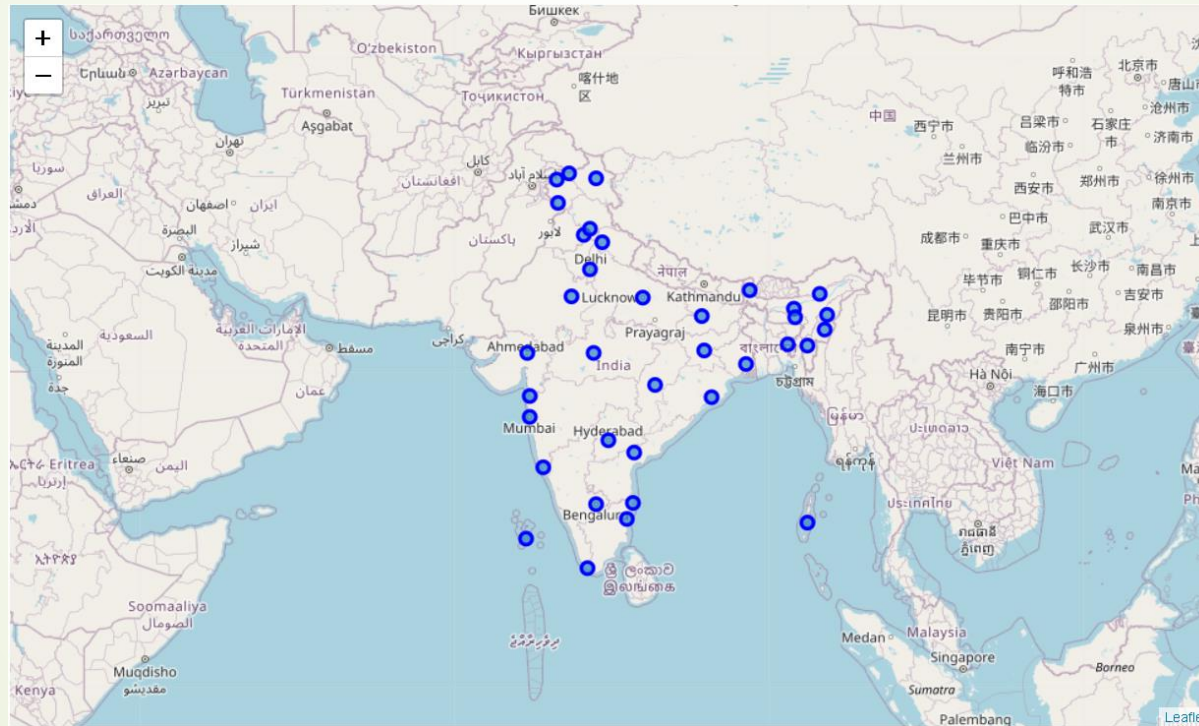3. The Folium library was used to visualise the capital cities of India.

Figure 3. Folium map depicting the capital cities of India.

4. The Foursquare API was used to explore the venues in the capital cities of India.

	Capital City	City Latitude	City Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Port Blair	11.664535	92.739045	MM Tours & Travels	11.664988	92.738922	Tourist Information Center
1	Port Blair	11.664535	92.739045	Sagar Emporium	11.663861	92.735847	Arts & Crafts Store
2	Port Blair	11.664535	92.739045	andaman fried chicken	11.667943	92.737825	Fast Food Restaurant
3	Port Blair	11.664535	92.739045	Wandoor	11.667975	92.737324	Beach
4	Amaravati	16.574420	80.355608	Amaravathi Buddhist Museum	16.575949	80.356569	History Museum
...	...	...	...	...	...	...	...
244	Kolkata	22.545412	88.356775	Shisha Reincarnated	22.546911	88.353341	Nightclub
245	Kolkata	22.545412	88.356775	Cafe Coffee Day	22.545478	88.352501	Café
246	Kolkata	22.545412	88.356775	Pantaloons	22.546955	88.353446	Clothing Store
247	Kolkata	22.545412	88.356775	Urban Desi	22.544248	88.352160	Indian Restaurant
248	Kolkata	22.545412	88.356775	Cafe 4/1	22.546093	88.352437	Café

249 rows × 7 columns

Figure 4. Dataframe containing the coordinates and categories of various venues located in the capital cities of India.

5. The rows of the dataframe shown in Figure 4, whose venue category was a restaurant were extracted to form a list of various types of restaurants in the capital cities of India.

6. k-means clustering was used to cluster the capital cities of India based on the cuisines served by the popular restaurants located there.

Capital City	Restaurant Category
Agartala	Fast Food Restaurant
	Indian Restaurant
Aizawl	Restaurant
Bhubaneswar	Asian Restaurant
	Fast Food Restaurant
Chandigarh	American Restaurant
	Asian Restaurant
	Fast Food Restaurant
	Indian Restaurant
	Italian Restaurant
Dehradun	Fast Food Restaurant
	Indian Restaurant
Dispur	Northeast Indian Restaurant
Gangtok	Dumpling Restaurant
Hyderabad	Afghan Restaurant
	Indian Restaurant
Imphal	Indian Restaurant
Jammu	Asian Restaurant
Kolkata	Indian Restaurant
	Mexican Restaurant
Leh	Himalayan Restaurant
	Indian Restaurant
	Multicuisine Indian Restaurant
	Tibetan Restaurant

Figure 5. List of various types of restaurants in the capital cities of India.

Results

- The capital cities of India were divided into five clusters based on the cuisines served by the popular restaurants located there.

- A part of the dataframe containing the clusters as well as the top ten restaurant types of each capital city of India is given in Figure 6.

	Cluster Labels	Capital City	1st Most Common Restaurant	2nd Most Common Restaurant	3rd Most Common Restaurant	4th Most Common Restaurant	5th Most Common Restaurant	6th Most Common Restaurant	7th Most Common Restaurant	8th Most Common Restaurant	9th Most Common Restaurant	10th Most Common Restaurant
0	1	Agartala	Fast Food Restaurant	Indian Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant	American Restaurant	Asian Restaurant	Chinese Restaurant	Dumpling Restaurant	French Restaurant	Himalayan Restaurant
1	2	Aizawl	Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant	American Restaurant	Asian Restaurant	Chinese Restaurant	Dumpling Restaurant	Fast Food Restaurant	French Restaurant	Himalayan Restaurant
2	4	Bhubaneswar	Asian Restaurant	Fast Food Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant	American Restaurant	Chinese Restaurant	Dumpling Restaurant	French Restaurant	Himalayan Restaurant	Indian Restaurant
3	3	Chandigarh	Italian Restaurant	American Restaurant	Asian Restaurant	Fast Food Restaurant	Indian Restaurant	Vegetarian / Vegan Restaurant	Chinese Restaurant	Dumpling Restaurant	French Restaurant	Himalayan Restaurant
4	1	Dehradun	Fast Food Restaurant	Indian Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant	American Restaurant	Asian Restaurant	Chinese Restaurant	Dumpling Restaurant	French Restaurant	Himalayan Restaurant
5	3	Dispur	Northeast Indian Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant	American Restaurant	Asian Restaurant	Chinese Restaurant	Dumpling Restaurant	Fast Food Restaurant	French Restaurant	Himalayan Restaurant
6	3	Gangtok	Dumpling Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant	American Restaurant	Asian Restaurant	Chinese Restaurant	Fast Food Restaurant	French Restaurant	Himalayan Restaurant	Indian Restaurant
7	0	Hyderabad	Afghan Restaurant	Indian Restaurant	Italian Restaurant	American Restaurant	Asian Restaurant	Chinese Restaurant	Dumpling Restaurant	Fast Food Restaurant	French Restaurant	Himalayan Restaurant
8	0	Imphal	Indian Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant	American Restaurant	Asian Restaurant	Chinese Restaurant	Dumpling Restaurant	Fast Food Restaurant	French Restaurant	Himalayan Restaurant
9	4	Jammu	Asian Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant	American Restaurant	Chinese Restaurant	Dumpling Restaurant	Fast Food Restaurant	French Restaurant	Himalayan Restaurant	Indian Restaurant
10	0	Kolkata	Mexican Restaurant	Indian Restaurant	Italian Restaurant	American Restaurant	Asian Restaurant	Chinese Restaurant	Dumpling Restaurant	Fast Food Restaurant	French Restaurant	Himalayan Restaurant
11	3	Leh	Indian Restaurant	Multicuisine Indian Restaurant	Himalayan Restaurant	Tibetan Restaurant	Vegetarian / Vegan Restaurant	American Restaurant	Asian Restaurant	Chinese Restaurant	Dumpling Restaurant	Fast Food Restaurant

Figure 6. Dataframe containing the clusters as well as the top ten restaurant types of each capital city of India.

Discussion

- From the results obtained, it can be inferred that Indian Cuisine is available in at least one of the popular restaurants in the capital cities falling under cluster 0 and Fast Food is available in at least one of the popular restaurants in the capital cities falling under cluster 1.

- The only popular restaurant in the capital cities falling under cluster 2 is a general restaurant while a variety of popular restaurants like- American, Indian, Italian, Northeast Indian, Dumpling, Himalayan, Tibetan, Mughlai, Punjabi, Chinese etc. can be visited by the food connoisseurs visiting the cities in cluster 3.

- As for the capital cities falling under cluster 4, Asian Cuisine can be found in at least one of the popular restaurants located there.

Conclusion

- The map shown in Figure 7 was obtained using the Folium library and can be used to visualize the clusters of capital cities of India based on the cuisines served by the popular restaurants located there.

- The red dots represent the capital cities lying in cluster 0 while the purple dots represent the capital cities lying in cluster 1.

- The blue dots represent the capital cities lying in cluster 2.

- The capital cities lying in cluster 3 are marked in green while those in cluster 4 are marked in orange.

- The food connoisseurs can use this map as a reference to plan their trip across India based on their cuisine preferences, thereby avoiding the trouble of having to search individual cities for the cuisines served there.
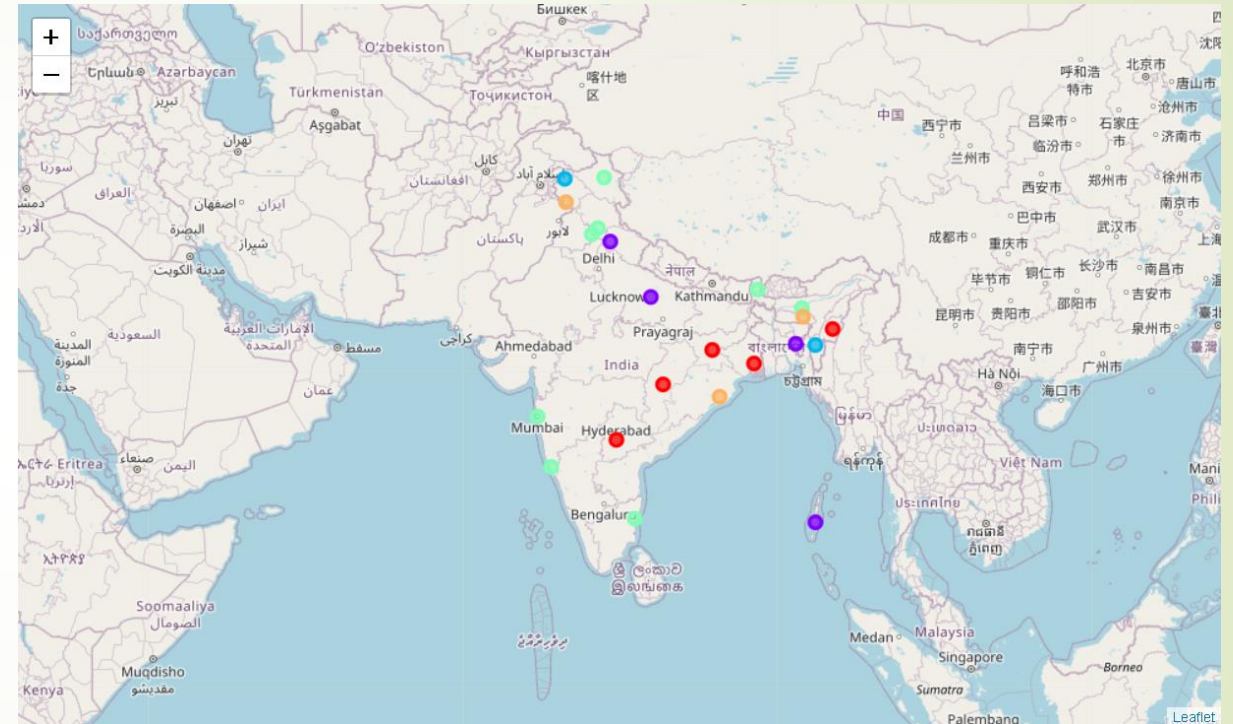

Figure 7. Folium map depicting the clusters of capital cities of India based on the cuisines served by the popular restaurants located there.

Thank You!