ANALYSIS OF POPULOUS CITIES OF INDIA TO FIND CITIES WITH SIMILAR RESTAURANT CULTURE

P N M MRUDU

Data science project

HOW THIS HELPS RESTAURANT CHAINS?

- Finding similar cities in terms of Restaurant venue category helps restaurant chains to narrow down on certain cities where the restaurant culture is inclined towards their venue category.

- They can get an insight on the most common venue categories prevalent in India.

- New businesses can look for similar cities to initially start their restaurant.

- Existing businesses can look for cities where the probability of low acceptance and inflow of customers is reduced.
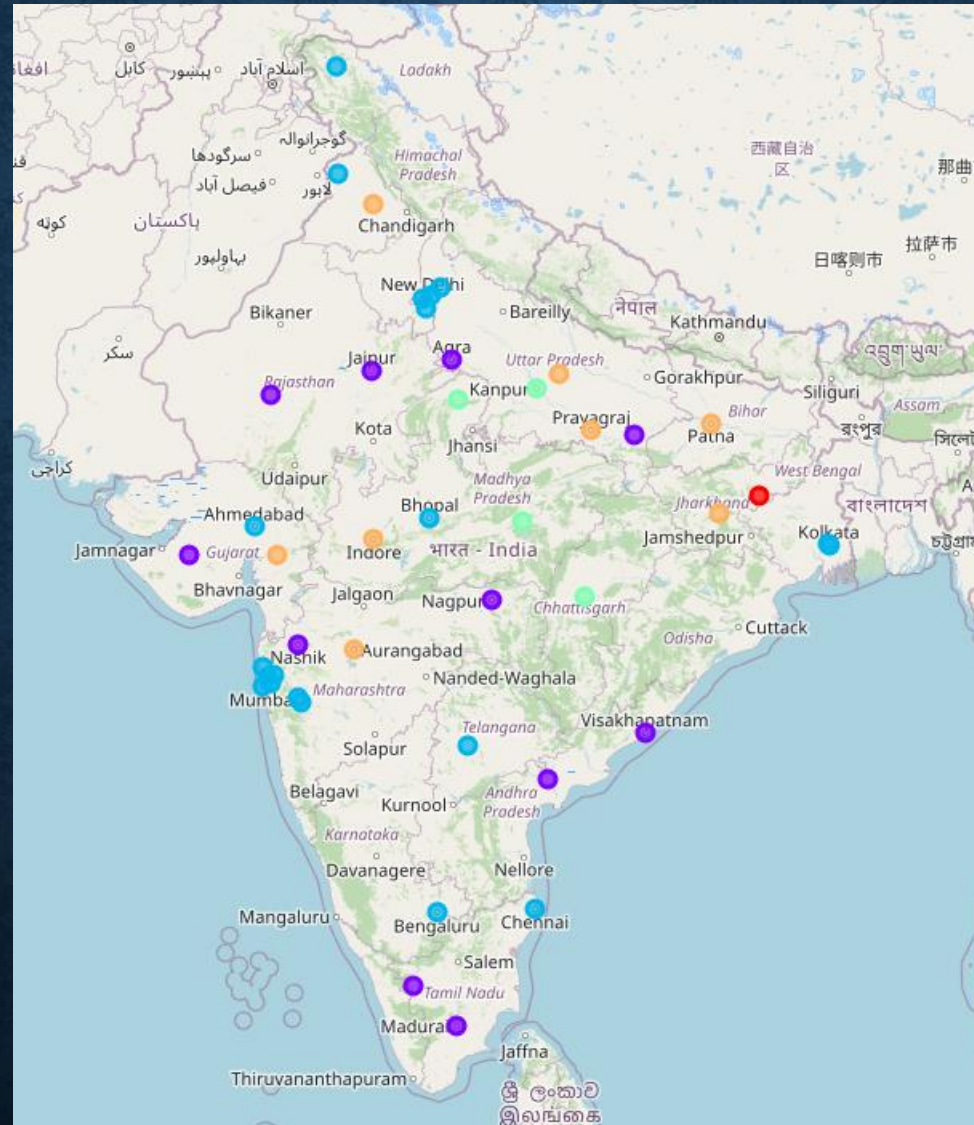
DATA, EXTRACTION AND CLEANING

- Populous cities of India dataset scrapped from Wikipedia. (Link: https://en.wikipedia.org/wiki/List_of_cities_in_India_by_population)

- Original data contains a list of 300 cities, top 45 cities has been utilized for this project.

- Venue and venue category for every city has been collected from Foursquare API.

- Venue limit has been placed at 100 and city radius at 60,000 meters.

- Both the datasets has been combined to form a dataset with all 45 cities and their corresponding 20 most common venue categories.

ALGORITHM APPLIED

- Since we wanted the cities to be grouped into similar clusters, K-means algorithm has been applied to the final dataset.

- As a result the dataset has been clustered to 5 different clusters.

Cluster Number	Number of cities
Cluster 1	1
Cluster 2	11
Cluster 3	21
Cluster 4	4
Cluster 5	8

VISUALIZATION OF THE CLUSTERS ON MAP

LIST OF CLUSTER 1 CITIES:

	latitude	City	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
32	23.795281	Dhanbad	23.795281	86.430964	0	Pizza Place	Indian Restaurant	Asian Restaurant

LIST OF CLUSTER 2 CITIES:

	latitude	City	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	26.916194	Jaipur	26.916194	75.820349	1	Café	Indian Restaurant	Bakery
1	21.149813	Nagpur	21.149813	79.082056	1	Indian Restaurant	Fast Food Restaurant	Sandwich Place
2	17.723128	Visakhapatnam	17.723128	83.301284	1	Indian Restaurant	Café	Restaurant
3	27.175255	Agra	27.175255	78.009816	1	Indian Restaurant	Café	Fast Food Restaurant
4	20.011247	Nashik	20.011247	73.790236	1	Indian Restaurant	Café	Pizza Place
5	22.305199	Rajkot	22.305199	70.802834	1	Indian Restaurant	Fast Food Restaurant	Pizza Place
6	25.335649	Varanasi	25.335649	83.007629	1	Indian Restaurant	Café	Pizza Place
7	11.001812	Coimbatore	11.001812	76.962842	1	Indian Restaurant	Vegetarian / Vegan Restaurant	Café
8	16.508759	Vijayawada	16.508759	80.618510	1	Indian Restaurant	Fast Food Restaurant	Café
9	26.296772	Jodhpur	26.296772	73.035143	1	Indian Restaurant	Restaurant	Café
10	9.926115	Madurai	9.926115	78.114098	1	Indian Restaurant	Café	Restaurant

LIST OF CLUSTER 3 CITIES:

	latitude	City	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	18.938771	Mumbai	18.938771	72.835335	2	Indian Restaurant	Restaurant	Café
1	28.651718	Delhi	28.651718	77.221939	2	Indian Restaurant	Café	Italian Restaurant
2	12.979120	Bangalore	12.979120	77.591300	2	Indian Restaurant	Café	Burger Joint
3	17.388786	Hyderabad	17.388786	78.461065	2	Indian Restaurant	Bakery	Café
4	23.021624	Ahmedabad	23.021624	72.579707	2	Indian Restaurant	Café	Fast Food Restaurant
5	13.080172	Chennai	13.080172	80.283833	2	Indian Restaurant	Café	Seafood Restaurant
6	22.545412	Kolkata	22.545412	88.356775	2	Chinese Restaurant	Indian Restaurant	Café
7	45.938300	Surat	45.938300	3.255300	2	French Restaurant	Restaurant	Italian Restaurant
8	18.521428	Pune	18.521428	73.854454	2	Café	Indian Restaurant	Fast Food Restaurant
9	19.194329	Thane	19.194329	72.970178	2	Indian Restaurant	Restaurant	Seafood Restaurant
10	23.253092	Bhopal	23.253092	77.396272	2	Indian Restaurant	Pizza Place	Bakery
11	18.627929	Pimpri-Chinchwad	18.627929	73.800983	2	Café	Fast Food Restaurant	Indian Restaurant
12	28.711241	Ghaziabad	28.711241	77.444537	2	Indian Restaurant	Café	Italian Restaurant
13	28.402837	Faridabad	28.402837	77.308563	2	Indian Restaurant	Café	Italian Restaurant
14	28.916667	Meerut	28.916667	77.683333	2	Indian Restaurant	Café	Fast Food Restaurant
15	19.233487	Kalyan-Dombivli	19.233487	73.118246	2	Indian Restaurant	Restaurant	Café
16	19.425879	Vasai-Virar	19.425879	72.822490	2	Indian Restaurant	Restaurant	Seafood Restaurant
17	34.074744	Srinagar	34.074744	74.820444	2	Café	Indian Restaurant	Fried Chicken Joint
18	31.634308	Amritsar	31.634308	74.873679	2	Café	Pakistani Restaurant	Fast Food Restaurant
19	19.030826	Navi Mumbai	19.030826	73.019854	2	Indian Restaurant	Fast Food Restaurant	Restaurant
20	22.588222	Howrah	22.588222	88.323139	2	Indian Restaurant	Chinese Restaurant	Café

LIST OF CLUSTER 4 CITIES:

	latitude	City	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	26.460914	Kanpur	26.460914	80.321759	3	Fast Food Restaurant	Café	Pizza Place
1	26.203725	Gwalior	26.203725	78.157363	3	Fast Food Restaurant	Café	Pizza Place
2	23.160894	Jabalpur	23.160894	79.949770	3	Café	Chinese Restaurant	Asian Restaurant
3	21.237947	Raipur	21.237947	81.633683	3	Café	Pizza Place	Fast Food Restaurant

LIST OF CLUSTER 5 CITIES:

	latitude	City	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	26.838100	Lucknow	26.838100	80.934600	4	Indian Restaurant	Fast Food Restaurant	Café
1	22.720362	Indore	22.720362	75.868200	4	Indian Restaurant	Fast Food Restaurant	Café
2	25.609324	Patna	25.609324	85.123525	4	Café	Pizza Place	Indian Restaurant
3	22.297314	Vadodara	22.297314	73.194257	4	Indian Restaurant	Café	Fast Food Restaurant
4	30.909016	Ludhiana	30.909016	75.851601	4	Fast Food Restaurant	Indian Restaurant	Café
5	19.877263	Aurangabad	19.877263	75.339024	4	Café	Indian Restaurant	Pizza Place
6	25.438130	Allahabad	25.438130	81.833800	4	Fast Food Restaurant	Pizza Place	Indian Restaurant
7	23.370035	Ranchi	23.370035	85.325013	4	Indian Restaurant	Café	Pizza Place

RESULTS

- From all the results, we can observe that Indian Restaurants, Cafes and Fast Food Restaurants are the most common Venues in India. Venturing into any one of these or a combination of these venue categories is advisable for starting a new Restaurant chain.

- Existing Restaurants who want to expand their business to other cities can look up the cluster data:

- Restaurants already in Indian restaurant category can expand to cities in cluster 2 or to cluster 3 cities if they want to start Café venue category along with Indian restaurants category.

- Pizza seems to be very common venue in Dhanbad, if any pizza restaurant which doesn't have a branch in Dhanbad can definitely try to start one.

- From the Data we can observe that Chinese, seafood and French restaurants are not so common in India.

- Restaurants with a combination of Fast food and café venue category can target cluster 4 cities.

CONCLUSION AND FUTURE DIRECTION

- . All this data and insights could be utilized by both new and existing restaurant chains. For example, Indian restaurant category and cafés appears to be the common restaurant culture in India, therefore, venturing into any one of these venue categories reduces the possibility of poor in-flow of customers.

- This analysis could be further enhanced by adding:
  (i) reviews and ratings data of each venue category.(can help in understanding user acceptance).

  (ii) venue menus.