State Capitals Comparision

Shubham Mahidhar

Urbanization

Urbanization is important for the benefits it brings and due to increase in commercial activity that creates demand and growth for many more industries.Thus analyzing the state capitals will help to track how diverse each capital is compared to other.
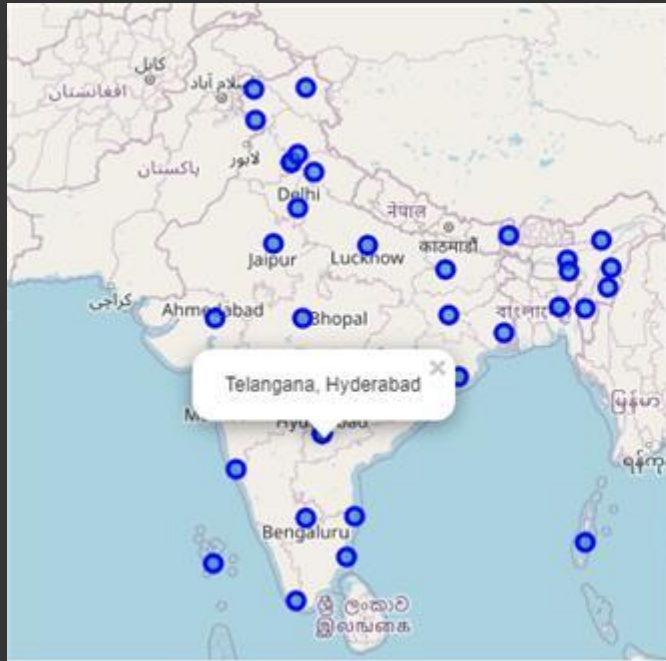
1. Data Description

To solve the problem we need listed items.

➔ **State and their capitals**

➔ **Latitude and Longitude of capitals**
   Python geopy package to fetch details

Data Exploration

Population in each capital
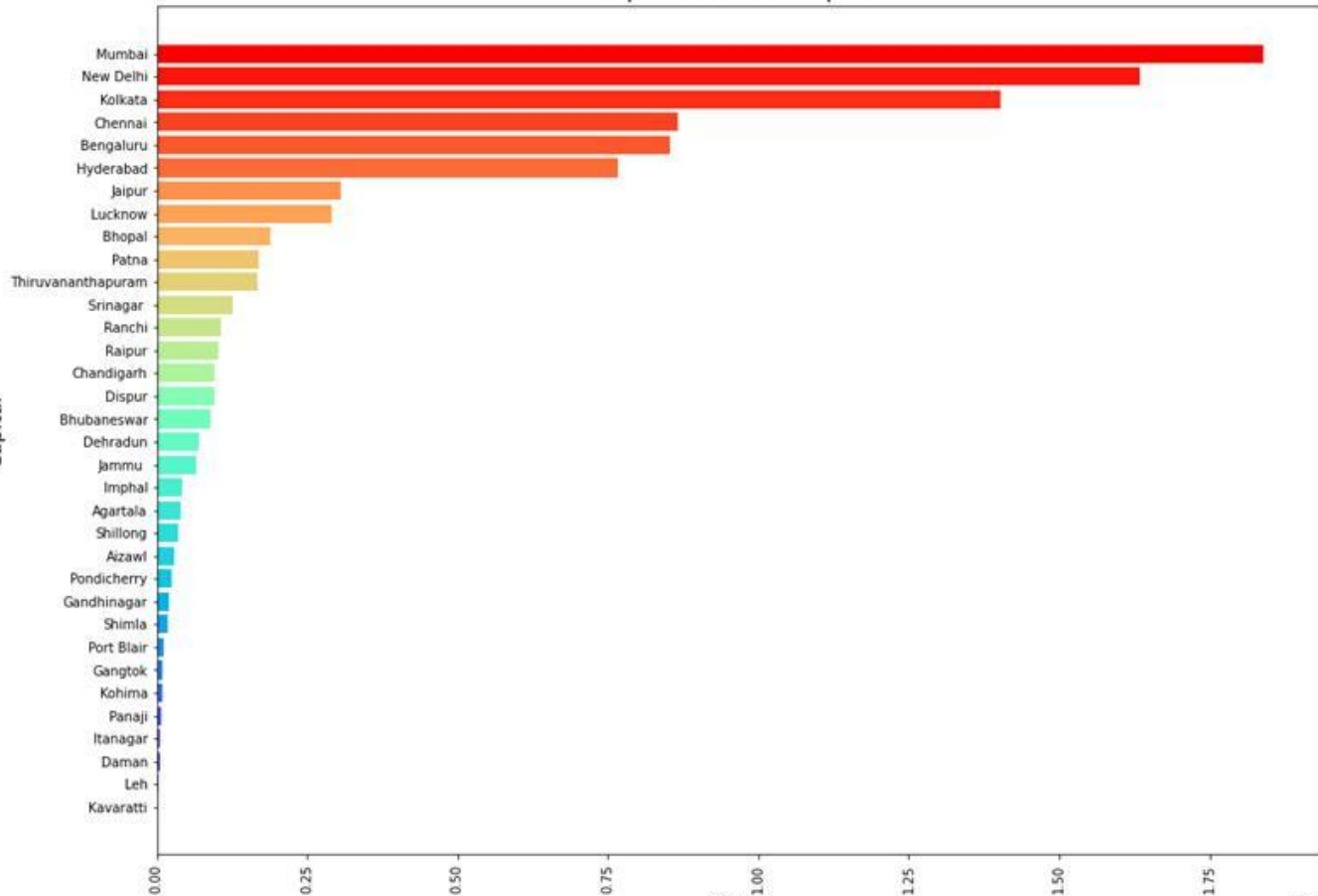
```
tourist_venues.head()
```

Out[24]:

	Capital	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
4	Leh	34.164203	77.584813	Leh Palace	34.166287	77.586390	Palace
11	Leh	34.164203	77.584813	Shey Palace	34.159815	77.633323	Palace
13	Daman	20.420005	72.863763	Daman	20.421682	72.839065	Beach
14	Daman	20.420005	72.863763	The Deltin	20.412848	72.856197	Casino
17	Daman	20.420005	72.863763	Jetty	20.412949	72.832805	Arcade

```
business_venues.head()
```

Out[25]:

|	Capital	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Kavaratti	10.559050	72.636569	kavaratti	10.549820	72.626635	Boat or Ferry
1	Leh	34.164203	77.584813	The Grand Dragon	34.156612	77.580351	Hotel
3	Leh	34.164203	77.584813	Rendezvous Cafe	34.165095	77.584020	Bakery
6	Leh	34.164203	77.584813	Coffee Culture	34.165123	77.584145	Coffee Shop
7	Leh	34.164203	77.584813	Pangong Hotel	34.161511	77.582676	Hotel

In [29]:
```
print('The recors obtained for tourist venue are - ' + str(tourist_venues.shape[0]))
print('The recors obtained for business venue are - ' + str(business_venues.shape[0]))
```

```
The recors obtained for tourist venue are - 111
The recors obtained for business venue are - 966
```

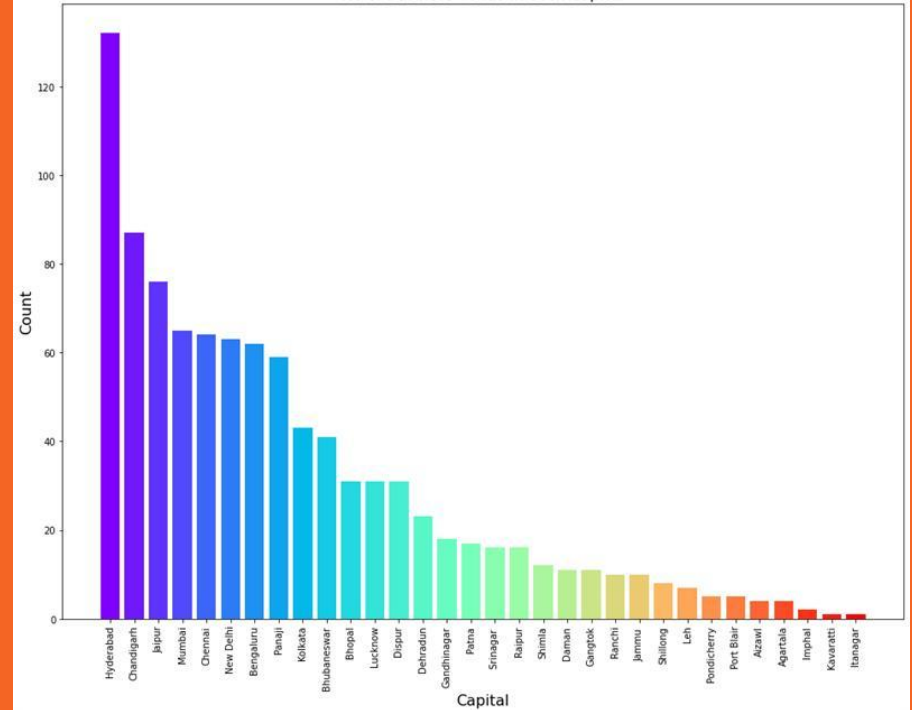We got 111 tourist and 966 tourist spots for all the capitals
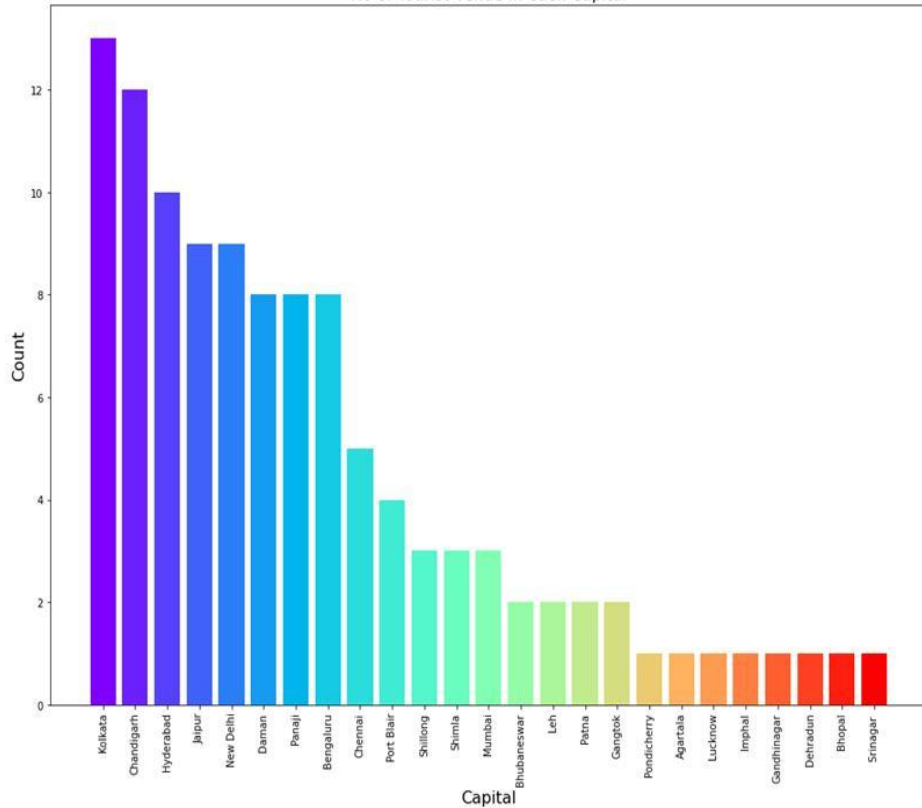
While Kolkata had more tourist spots than any other capital

Hyderabad was spotted having more business venue

2. Analysis

➜ **Tourist spots**
What kind of tourist spot each capital has and does any other capital depict same behavior?
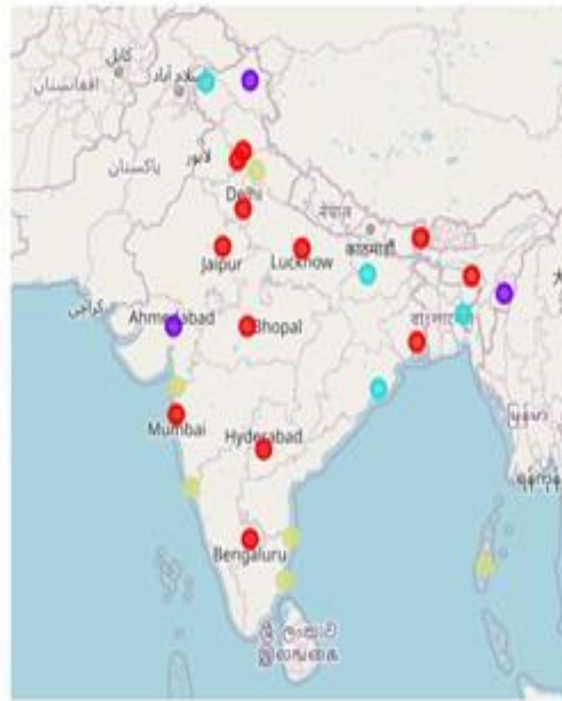
➜ **Business**
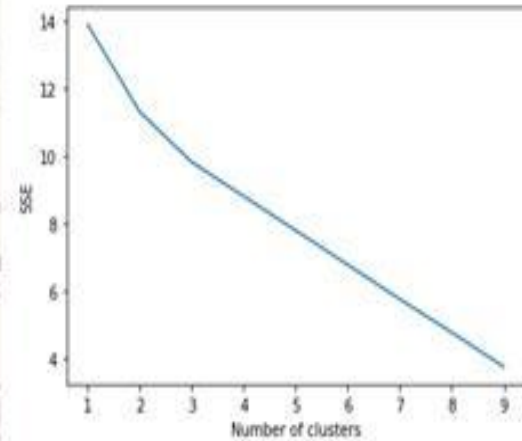Business the capital support does other capitals have same model .

- Perform one hot encoding to obtain famous tourist and business spots around capital.

- Use K-Means algorithm to cluster the capitals.

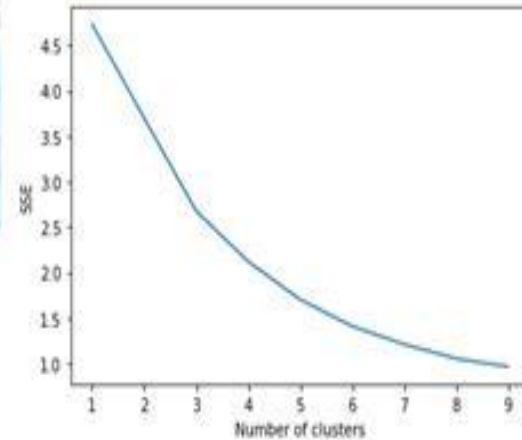- Elbow method to find optimal value for k

Capital Cluster Using Business Data

Capital Cluster Using Tourist Spot Data

Elbow Graph Using Tourist Spot

Elbow Graph Using Business Spot

Business data suggest majority capital look alike.

Tourist cluster look more diverse

Food joints is key business for cluster zero cities per business data.

Beach is key tourist spot for cluster 3 cities per tourist data.

Future Directions

Per capita income ,Housing demands etc can help understand which cities are showing growth thus helping in improving the Indian Economy.