Venues/Business Analysis in India

Introduction

In today's world where it is becoming challenging for earning good money from a job. So, people are mostly thinking of starting a business, but they are always confused and ask me to tell me some solution which can give them some idea about the location and business to start in a location. This project is about doing a business and strategizing the location and business. Sometimes, it is challenging for the investors to invest the money. It will help in doing the analysis for them and help in taking better business decisions.

Introduction/Business Problem

Whenever people are thinking about doing a business it is always challenging to work from scratch and find out a location and the type of business. Most of the time they heard from people about the business and location without doing any analysis which i am afraid to say that people used to fail multiple times only because of this reason.

Nowadays, Investors are looking forward to investing their money in good start-ups which help both to get succeeded but they may take wrong decisions as they don't know the competition in that city or area. This project will give a good insight to all the people who want to explore and study more about the existing businesses along with help them to do analysis first further take any decision.

It will provide the information about the similar cities in an area to determine the same kind of opportunities in different locations.

Data section

For this purpose, this project contains the data of India and based on the requirements it will analysis the state and give the result in the form of Data Frames and maps with proper location. Here, one example of state i.e. Andhra Pradesh has been taken. It will generate all the businesses present in Andhra Pradesh state with city information as well.

Data will analyse the top 10 venues/businesses which are present in the respective city and classify them using K-Means algorithm to give the better recommendation of the businesses which a person can think about it. It will prove the complete report about the cities and will help to understand the similar cities in terms of business.

Here, Folium library is used for plotting the maps and Foursquare API is used to get the venues information and extracting the features related to venues. We can explore more on Foursquare API to get the results about the items on a venue along with the ratings and comments etc.
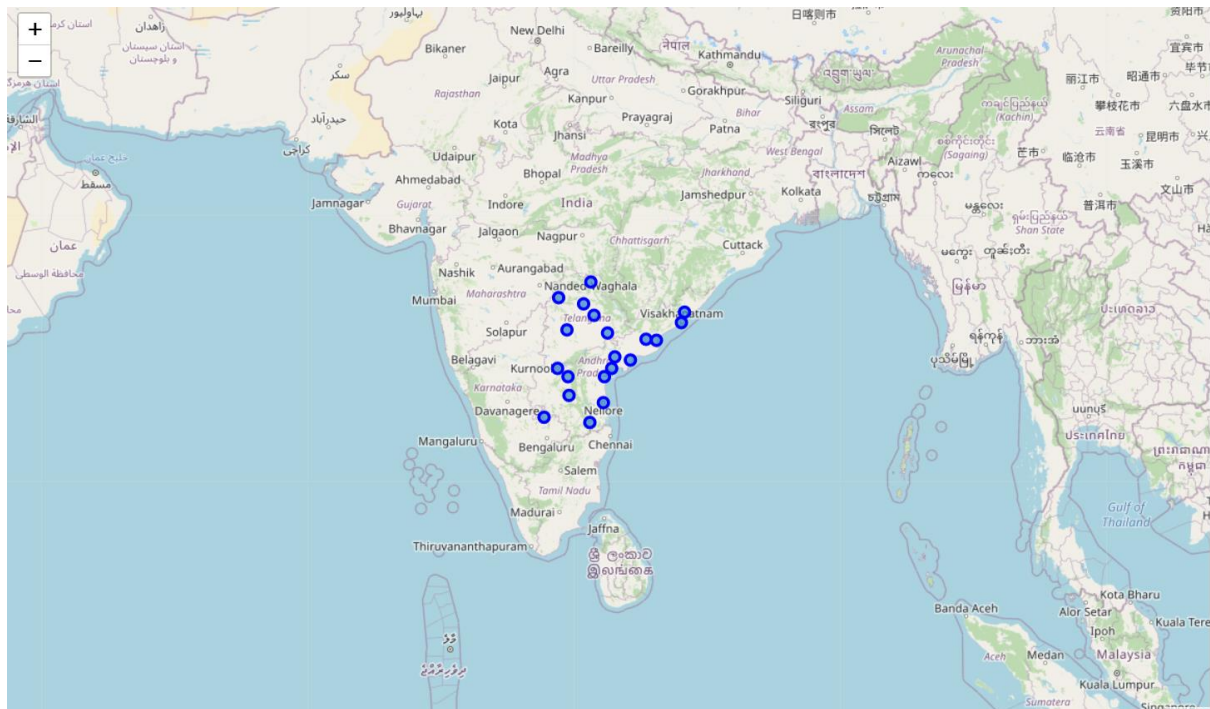
For this project, Data is gathered from https://simplemaps.com/data/in-cities which provides the information about the major Cities, States, Population, Latitude and Longitude information of India. This information will be used to analyse existing business in a state and help to determine the scope of more business areas. Data is related to all the states but for better understanding only one state I have considered for now to showcase and understand one part of it. Later, it can be used to analyse the entire country or some specific state or city. It depends on the person how and for what purpose they want to use it.

Data pre-processing step is done where the data will contain only the useful information's. Post this step the data is in the below mentioned format.

State	City	Latitude	Longitude	Population
Andaman and Nicobar Islands	Port Blair	11.666667	92.750000	127562.0
Andhra Pradesh	Proddatūr	14.750200	78.548129	197451.0
Andhra Pradesh	Hindupur	13.828065	77.491425	168312.0
Andhra Pradesh	Khammam	17.247672	80.143682	290839.0
Andhra Pradesh	Guntūr	16.299737	80.457293	530577.0
...	...	...	...	...
West Bengal	Haldia	22.025278	88.058333	200762.0
West Bengal	Hāora	22.576882	88.318566	4841638.0
West Bengal	Alīpur Duār	26.483500	89.522855	127342.0
West Bengal	Kolkata	22.562627	88.363044	14787000.0
West Bengal	Barddhamān	23.255716	87.856906	301725.0

Data has been displayed in the maps form for better understanding about the location of the business and the distance between them to do better analysis and making good decision.

This figure contains the information about all the states of Andhra Pradesh which has been used as example in this project. For getting the information about all the businesses, Foursquare API is been used which will give the information about the cities we have and existing venues/business.

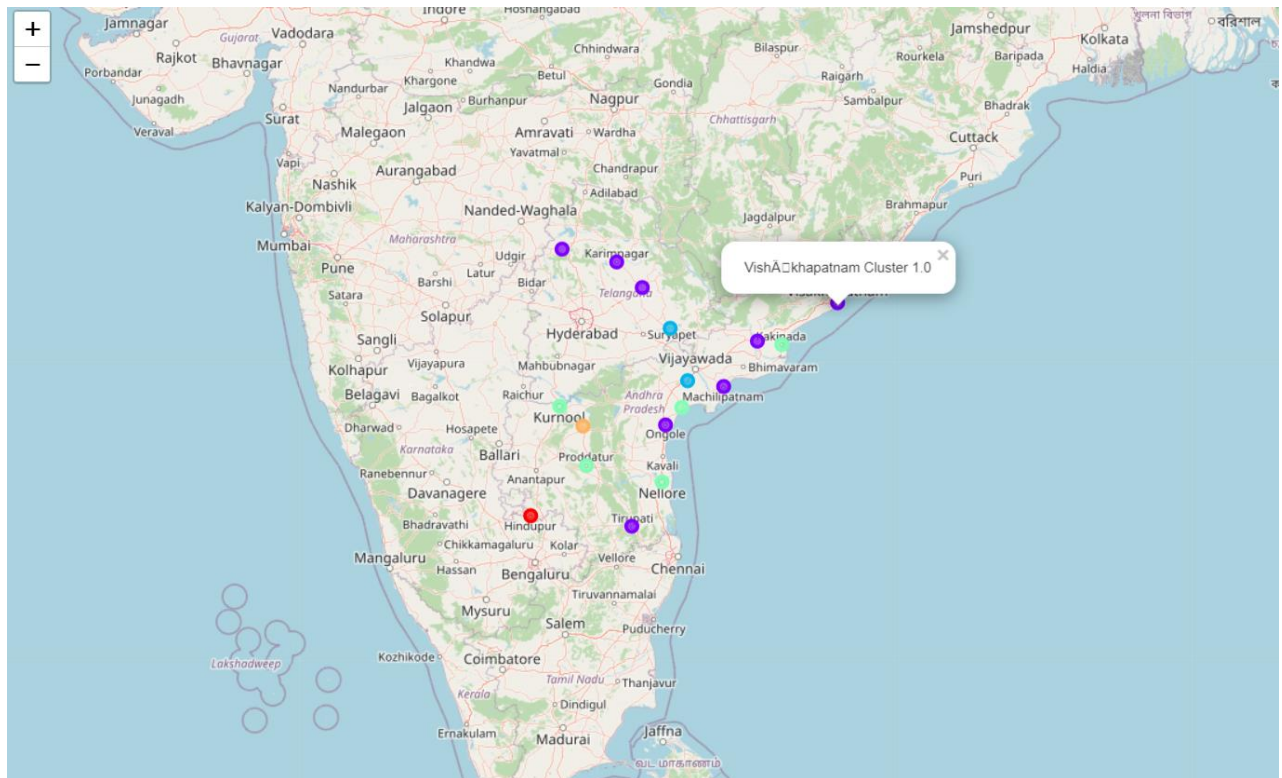| City	City Latitude	City Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Proddatūr	14.750200	78.548129	Cine Hub	14.750211	78.551526	Multiplex
Proddatūr	14.750200	78.548129	ibaco	14.751263	78.548719	Ice Cream Shop
Hindupur	13.828065	77.491425	Sundaram Soda	13.830011	77.492712	Juice Bar
Khammam	17.247672	80.143682	Khammam Bus Station	17.250140	80.142381	Bus Station
Khammam	17.247672	80.143682	Hotel Vishnu Residency	17.250873	80.142643	Hotel

Data is formatted again in a tabular format with the required columns. Later analysis is done on this data using folium map's and data frames to get the top 10 venues/businesses in a city to compare and take the decision about their similarity and the most commonly used businesses in a city.

City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
Chīrāla	Vegetarian / Vegan Restaurant	Clothing Store	Gift Shop	Breakfast Spot	Fried Chicken Joint	Food Court	Fast Food Restaurant	Electronics Store	Deli / Bodega	Watch Shop
Guntūr	Bus Station	Café	Fried Chicken Joint	Food Court	Fast Food Restaurant	Electronics Store	Deli / Bodega	Clothing Store	Watch Shop	Hotel
Hindupur	Juice Bar	Watch Shop	Hotel	Fried Chicken Joint	Food Court	Fast Food Restaurant	Electronics Store	Deli / Bodega	Clothing Store	Café
Karīmnagar	ATM	Tourist Information Center	Shoe Store	Pharmacy	Asian Restaurant	Bakery	Arts & Crafts Store	Boat or Ferry	Breakfast Spot	Gift Shop
Khammam	Bus Station	Fried Chicken Joint	Electronics Store	Hotel	Café	Food Court	Fast Food Restaurant	Deli / Bodega	Clothing Store	Watch Shop

Here is an example of the data we construct to get the top 10 business city wise.

Results

Using k-means algorithm and clustering the data is combined and the analysis is done. Using K-means the data is split into 5 clusters. Some cities we are having very few businesses and some of the city are having a good number of businesses. It will provide the complete insight about all of them.

Here the 5 clusters are representing in 5 different colours and cluster 1.0 cities are having the maximum number of business opportunities.

The result will be displayed as city name and the business which can be done in that area. It will also provide the information about the similar cities where same kind of business can be setup based on the financial cost and other benefits.

```
List of businesses which are most common in ['Hindupur'] are as:-
                    0
             Juice Bar
            Watch Shop
                 Hotel
    Fried Chicken Joint
            Food Court
   Fast Food Restaurant
       Electronics Store
          Deli / Bodega
         Clothing Store
                  Café
```

It will also provide the information about the similar cities where same kind of business can be setup based on the financial cost and other benefits.

```
List of businesses which are most common in ['Machilīpatnam' 'Rājahmundry' 'Karīmnagar' 'Vishākhapatnam' 'Tirupati'
 'Warangal' 'Nizāmābād' 'Vizianagaram' 'Ongole'] are as:-

                          0
                  Pharmacy
             Movie Theater
                       ATM
               Snack Place
          Indian Restaurant
             Train Station
        Arts & Crafts Store
 Tourist Information Center
         Indie Movie Theater
              Boat or Ferry
         Fried Chicken Joint
                  Multiplex
             Shopping Mall
             Clothing Store
                 Shoe Store
        Fast Food Restaurant
           Asian Restaurant
           Electronics Store
                 Restaurant
                 Watch Shop
                Bus Station
                     Buffet
                   Mountain
                  Gift Shop
                       Café
                      Hotel
              Deli / Bodega
                 Food Court
                     Bakery
             Breakfast Spot
```

Conclusion

The results it displayed are correct and have been confirmed by the people who want to use it. Here it is specific to Andhra Pradesh, but it can be used to compare two cities, two states etc. It can provide more information about the other cities as well which are a part of India.