

# Analysis of sources of entertainments in the Telangana cities

## Business Problem

Telangana India has been a growing hub for many businesses for many decades. The cities of Telangana have been home to the ever-rising food and entertainment industry. New entrepreneurs are ever eager to invest in Telangana cities but not sure which location or sector would reap maximum profits. So, we will analyse the existing market, come up with valuable insights regarding the demand of various food and entertainment sectors, and help them make an informed decision.

## Data Description

We will be analyzing these 12 cities of the Telangana State. We scraped a webpage ([https://en.wikipedia.org/wiki/List\\_of\\_cities\\_in\\_Telangana\\_by\\_population](https://en.wikipedia.org/wiki/List_of_cities_in_Telangana_by_population)) which provided me with a table of cities of Telangana. We only kept the city name and district column and dropped the remaining columns as they were irrelevant to our analysis. We then used the Foursquare(<https://www.foursquare.com/places>) API to obtain the geographical locations of the cities.

We then generated all the venues in every city using the Foursquare API. These venues were grouped into unique categories. These categories are then one-hot encoded to obtain the highest or the lowest available sectors in the cities of Telangana.

## Data Cleaning

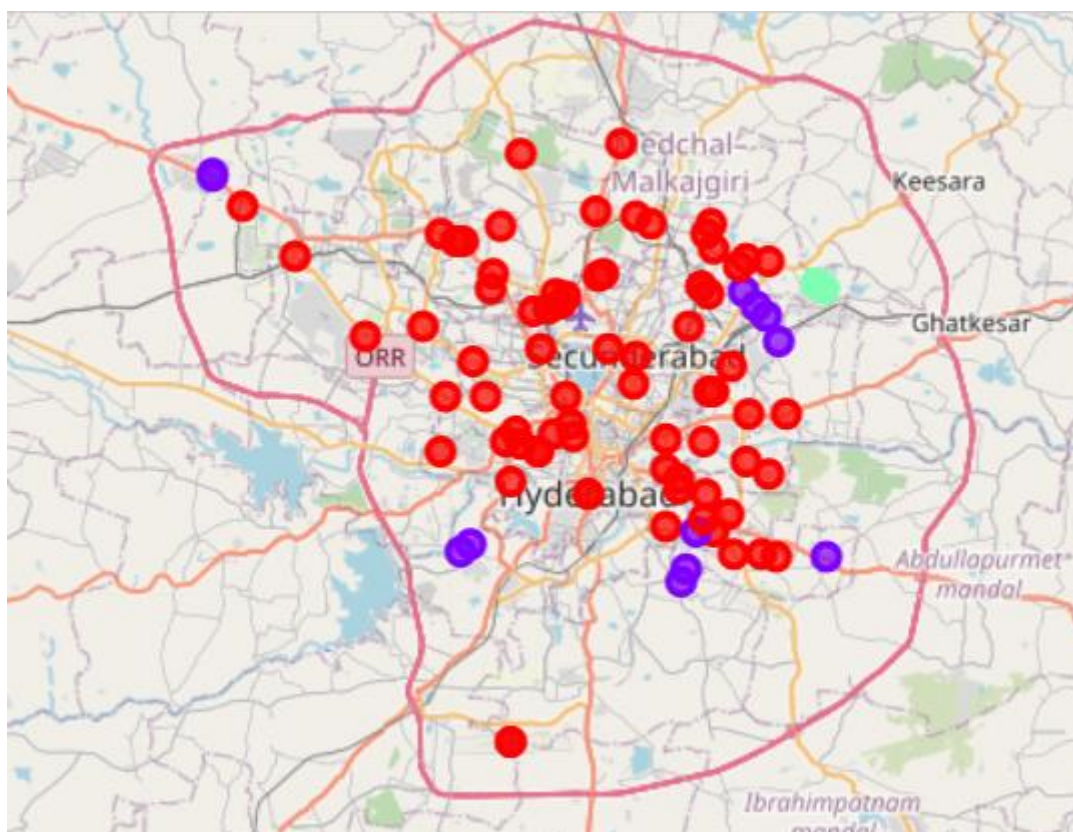
We cleaned the data by checking for null values which might have taken place when there was no data available through the GeoPy API. After generating the venues using foursquare API, the number of venues generated was very less due to the co-ordinates of the cities, which won't cover more considerable distances. So we decided to concentrate our study in Hyderabad, a town in Telangana only.

## Feature Selection

After data scraping, we dropped the columns which were unnecessary to us for our analysis like the area of the city. We did not consider the population of the cities as it was out of scope for this project. After getting the venues of every neighbourhood in Hyderabad, we generated the list of top 10 venue categories in every neighbourhood. All 10 venue categories were used as features for clustering our data into multiple groups.

## Clustering

K-Means unsupervised clustering was done on the data to group the neighbourhoods based on the similarities and dissimilarities between them. We tried different K-values and the best K-value was turning out to be 3. So we divided the neighbourhoods into 3 clusters.



## Conclusion

We observed that the most available venues are Indian Restaurants and Cafes followed by movie theatres and shopping malls in the first cluster. ATMs are the most popular followed by music

venues and Farms in the second cluster. Any investment other than Indian Restaurant and Cafes is better, considering there's a market for it. Movie theatres, Shopping Malls and music venues are lagging behind in the frequency. So investing in them would make a better sense as there's a market for it but the availability is lesser.

## Future Directions

We did not consider the area of the city/neighbourhood, population, income levels of the citizens and the demand for a certain type of venues. In the subsequent analysis, we will be considering all the above data and come up with a much more sophisticated solution to the problem.