The Battle of the Neighborhoods: Visakhapatnam City, INDIA

Applied Data Science Capstone by IBM/Coursera

Introduction: Business Problem

- To recommend the best neighborhood to live, to buy a house, to rent an apartment or build a restaurant etc in Visakhapatnam, India .
- To understand the similarities and differences between the neighborhoods using Unsupervised K-Mean Clustering Algorithm

Objective

- Collecting the top trending venues in the using Foursquare API(Beautiful Soup, http request)
- Forming neighborhood clusters based on venue categories using unsupervised k-mean clustering algorithm(sklearn)
- Identifying and understanding the similarities and differences between two chosen neighborhoods to retrieve more insights and to conclude which neighborhood wins over other.

Data Description

The data about Boroughs, PinCode, and Neighbourhood was scraped from a url (https://www.indiatvnews.com/pincode/andhra-pradesh/visakhapatnam) to create our own dataset.

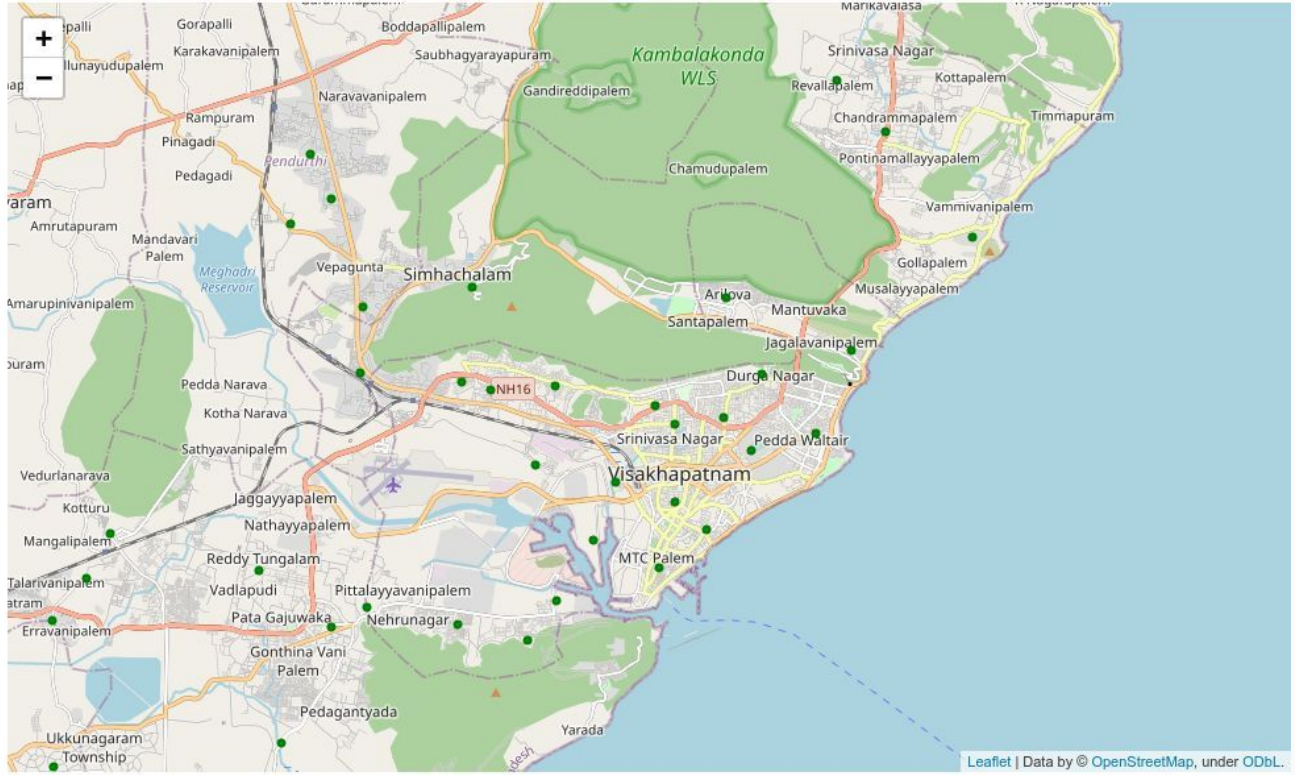
HTTP requests would be made to this Foursquare API server using pin codes of the Visakhapatnam city neighborhoods to pull the location information (Latitude and Longitude).

Pincode	Neighbourhood	Borough	Latitude	Longitude
530001	Fortward	Visakhapatnam Urban	17.6998	83.2961
530002	D.C. Buildings	Visakhapatnam Urban	17.7092	83.3081
530003	A U Engg College	Visakhapatnam Urban	17.7288	83.3201
530004	Gnanapuram	Visakhapatnam Urban	17.7207	83.2847
530005	Gandhigram Visakhapatnam	Pedagantyada	17.6815	83.2618
530007	Industrial Estate Visakhapatnam	Visakhapatnam Urban	17.7447	83.2686
530008	IRSD Area	Visakhapatnam Urban	17.7253	83.2637
530009	Marripalem Vuda Colony	Visakhapatnam Urban	17.7459	83.2446
530011	Malkapuram	Visakhapatnam Urban	17.6855	83.2433
530012	Autonagar	Visakhapatnam Urban	17.6991	83.1917
530013	P & T Colony VM	Visakhapatnam Urban	17.7368	83.3127

Data Frame Glimpses

Methodology

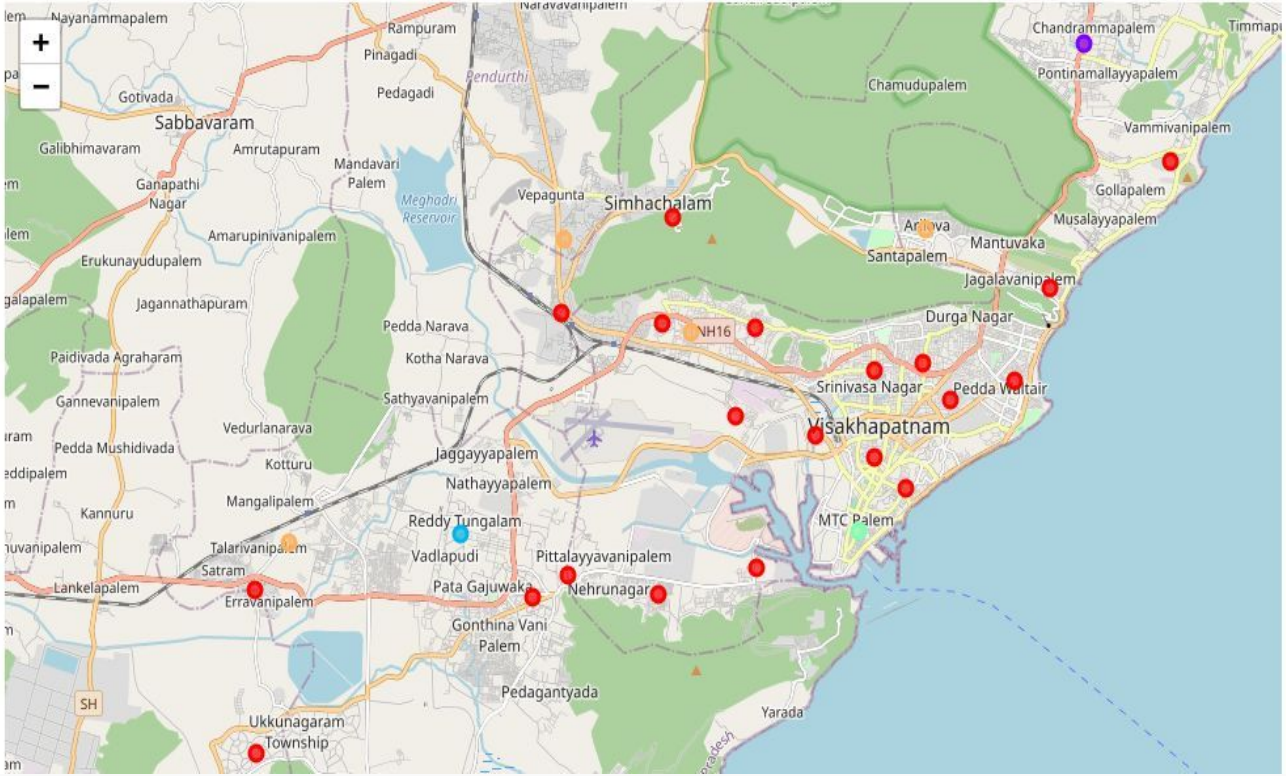
- The data about Boroughs, PinCode, and Neighbourhood was extracted using web scraping from the url to create our own dataset.
- The data consisted of 650+ rows.
- Feature Selection and cleaning was carried out on the data.
- Map for all the Neighbourhoods in Visakhapatnam City was plotted.
- Foursquare API was used to find all the venues near each Neighbourhood.
- OneHotEncoding was used to create dummy variables for the VenueCategory as it is a categorical variable.
- The venues were clustered using KMeans ML algorithm.
- The plot of all the Clustered venues and the Neighbourhood was made on the map of Visakhapatnam.



Neighbourhoods in Visakhapatnam City

Results and Discussions

- Most of the neighborhoods fall into Cluster 1 which are mostly business areas with cafes, restaurants, supermarkets etc.
- Mostly business areas in Cluster 5 are ATMs, Pet services and Pharmacies.
- Most business areas in Cluster 4 are just Harbor / Marina.
- Mostly business areas in Cluster 3 are Business Service and Warehouse Store.
- Mostly business areas in Cluster 2 are just Sculpture Gardens.



Clustered Neighbourhoods in Visakhapatnam City

Conclusion

With this we can conclude that the people who are looking for new apartment or home for rent or purchase can select the Neighborhoods that belong to Cluster 1 as there are many facilities and business areas.

Thanks!