The Battle of the Neighborhoods : Visakhapatnam City, INDIA

Applied Data Science Capstone Course by IBM/Coursera

Shaik Shahid
121722501020@gitam.in

Introduction/Business Problem

Visakhapatnam is the largest city in the Indian state of Andhra Pradesh, serving as headquarters of the Visakhapatnam district. It lies between the Eastern Ghats and the coast of the Bay of Bengal. The city is set to become the executive capital of Andhra Pradesh once a decentralization bill is enacted in the Andhra Pradesh Legislative Assembly. It is one of the four smart cities of Andhra Pradesh selected under Smart Cities Mission. It is the most populous city in the state, with a population of 2,035,922 as of 2011, making it the 11th largest city in the country and the fourth most populous in South India. It is also the ninth most populous metropolitan area in India, with a population of 5,018,000. With an output of $43.5.billion, Visakhapatnam is the ninth largest contributor to India's overall gross domestic product as of 2016.

Whenever people move to any other place, they explore the place and try to fetch as much information as possible about it. It can be the neighborhood, locality, market, price of the place and many more factors including neighborhood analysis. This can be termed as a request for a search algorithm which usually returns the requested features such as population rate, median house price, school ratings, crime rates, weather conditions, recreational facilities etc. It would be beneficial and nice to have an application which could be made easy by considering a comparative analysis between the neighborhoods with provided factors.

This project helps the end user or the stakeholder to achieve the results which will not only be recommended but also saves a lot of time in manual search. This will indeed save the time and money of the user.

This project can be used by the user at the time of rental apartment or buy house in a locality based on the distribution of various facilities available around the neighborhood. As an example, this project would compare 2 randomly picked neighborhoods and analyse the top 10 most common venues in each of those two neighborhoods based on the number of visits by people in each of those places. Also, this project uses K-mean clustering unsupervised machine learning algorithm to cluster the venues based on the place category such as restaurants, park, coffee shop, gym, clubs etc. This would give a better understanding of the similarities and dissimilarities between the two chosen neighborhoods to retrieve more insights and to conclude with ease which neighborhood wins over others.