

Analyzing the Neighborhoods in Visakhapatnam for setting up a new Restaurant

IBM Applied Data Science Capstone Project

by : Swarna Naga Sri Tejaswi

Data Collection

The data required for this project is as follows and has been collected from multiple sources. The following data is required for the project:

- 1) Neighborhood data of Visakhapatnam
- 2) Geographical Coordinates of Visakhapatnam and all neighborhoods in Visakhapatnam
- 3) Venue data for neighborhoods in Visakhapatnam

Neighborhood Data

The data of the neighborhoods in Visakhapatnam was scraped from https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Visakhapatnam. We will use web scraping techniques to extract the data from the Wikipedia page, with the help of Python requests and BeautifulSoup packages. Then we will get the geographical coordinates of the neighborhoods using Python Geocoder package which will give us the latitude and longitude coordinates of the neighborhoods. After that, we will use Foursquare API to get the venue data for those neighborhoods. Foursquare API will provide many categories of the venue data, we are particularly interested in the Restaurant category in order to help us to solve the business problem put forward. This is a project that will make use of many data science skills, from web scraping (Wikipedia), working with API (Foursquare), data cleaning, data wrangling, to machine learning (K-means clustering) and map visualization (Folium).