**Introduction to data**

In this section the data related to hotels and resorts that are in Visakhapatnam city and the surrounding neighbourhoods within the limits of 10 miles are considered for the study. By clustering the hotels and resorts based on the concentration of the location and neighbourhoods.

**Data**

**The following data is considered for this problem to solve:**

1. Since the scope of this project is confines to Visakhapatnam city, considered list of neighbourhoods in and around Visakhapatnam city proposed executive capital Andhra Pradesh state in India.
2. In order to building the meaningful visuals using python the latitude and longitudes of the Visakhapatnam city are considered.
3. Also, the data related to the hotels and resorts that are already exist in the city are considered to build a meaningful cluster using machine learning

**Data source & Methodology**

The sources of the above mentioned data is Wikipedia and the data related to Visakhapatnam city is fetched from URL <https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Visakhapatnam> contains data for 121 neighbourhoods in Visakhapatnam. Web scraping is used to extract the data from the Wikipedia using python and beautiful soup package. For geographical coordinates data Python Geocoder package is used to get the latitude and longitude for all the neighbourhoods.

The Foursquare API the largest locations data base with more than 100 million places used to get the information and details about the places of interest (Hotels and Resorts in Visakhapatnam). Post gathering the data related to places of interest applied K means clustering to create the homogeneous groups of the places of interest and analyse each of these clusters and provide recommendations.