Coursera Capstone Project : The Battle of Neighbourhoods

Problem Statement:

The following capstone project comprises of segmenting the restaurants based near the vicinity of beaches in tourist destination in India. The chosen Indian state is Goa and all the beaches are considered. The target audience would be the tourists as well as those who are interested in having their own restaurants nearby beaches.

Problem Approach:

For the above problem statement the data of beaches is taken from Wikipedia page and scrapped using beautiful soup package. After the data of all the beaches is gathered, I used geopy package the get the latitude and longitude of all the beaches available. These coordinates then were further made into a single data of beaches having correct latitude and longitude.

Now using foursquare all the venues associated with these beaches were gathered by supplying client Id and client secret. The information about all the unique venue and their category were grouped together. Now only the restaurants data were gathered for the further analysis as it is the problem statement as discussed above. In order to implement the KMeans clustering unsupervised Machine learning model all the categorical data was converted to numerical data. Finally, the model was implemented and 3 segmentation was done.

Conclusion

Using the KMeans clustering model the segmentation of restaurants were done in 3 categories