The Battle of Neighborhoods

PROJECT REPORT

Ashish Kumar Singh | Data Science Capstone | Jan/2019

Introduction/Business Problem

Because Bangalore is known as silicon valley of India, So many people from different parts of country moves to Bangalore but India is very big and Diverse every part of a country is different from each other be it language, culture, food or anything else. A large number of people who move to Bangalore are from New Delhi and nearby area (NCR, UP, Haryana, Punjab, Rajasthan). They know New Delhi very well but Bangalore is totally different. So we will try to cluster New Delhi and Bangalore Based on similarity. That is we will cluster New Delhi on the basis of area and cluster Bangalore on the basis of area and will tell which area of New Delhi is similar to which area of Bangalore. This will help not only the new people moving to Bangalore but also those who are already living in Bangalore but are looking for change.

DATA

We will try to get all the data which we can get about New Delhi and Bangalore. we get the list of district of New Delhi and their sub district so district can act as our borough and sub district as Neighborhood. speaking of Bangalore we got a Wikipedia page about Bangalore neighborhood which is just the name of areas of Bangalore as Bangalore is very small as compared to New Delhi which is quite obvious. Moreover we will use foresquare api to get the information about various amenities or facilities available in these places based on which we will cluster theses cities.

We will cluster Delhi then predict on Bangalore, Or we can cluster New Delhi and Bangalore in equal numbers (if possible) then on the bases of similarity we will mark which cluster of Delhi is quite similar to Bangalore This will not only provide the similarity between The Two major cities but also we can individually see which areas are similar to which areas of both cities individually e.g in Delhi is Connaught Place same as Dwarka or in Bangalore is Electronic city similar to J.P. Nagar and so on