Capstone Project - The Battle of the Neighborhoods

Introduction: Business Problem of knowledge of neighborhood to explore venues

Mumbai and Delhi are two metro cities of India. These two cities are crowded and always jam packed with people. These cities are one of the best tourist attractions. So, we will explore these two cities and try to find out which city has more venues to visit.

Nobody can remember or know all venues in Mumbai and Delhi area and so cannot promote all venues and categories which can found through Foursquare API. We would like to provide information near these neighborhoods to tell the tourists and people who are unaware. We would also cluster the similar venues and categorize them to quickly find out which category of venues and areas are unique and have good parks and cafeterias. This could make the difference for tourists who are unaware of the places and it will provide information that can be even crucial when families deciding where they are going to move or buy new home.

Data

Based on definition of our problem, factors that will help:

- All venues of neighborhood
- Top venue categories in neighborhood
- Overall style for example cafes and parks

The following data sources will be needed to generate the required information:

- Data found from the data.gov.in about the post offices in India
- I extracted the data about Mumbai and Delhi
- And using this data I changed office names as neighborhoods
- All venues or neighborhood area through Foursquare API
- After cleaning the data, I used geopy library
- Geopy library to get coordinates of neighborhoods

We will use the **explore** function to get the most common venue categories in each neighborhood of Mumbai and Delhi. We will also cluster neighborhoods to give similarity information to end customer. And in the end, which city is the best to live in according to the end customers.