Battle Of Neighbourhood

Introduction

This is my first post as part of my IBM Professional Data Science Certification. This project involving the Foursquare API and various data tools such as wikipedia etc. The project allowed me to pull together many of the tools\libraries I learned in the course:

- Jupyter notebook
- pandas
- folium
- json
- requests
- bs4

Background and Problem Statement

Indians are more welcoming now a days to drink in public rather than in their hideouts or home. Drinking is becoming norm at social gatherings. India has population of 1.3Billion people but still not enough bars to cater such a large customer base.

One of the report suggest that Alcohol consumption in India amounted to about 5.4 billion liters in 2016 and was estimated to reach about 6.5 billion liters by 2020. The steady increase in consuming these beverages can be attributed to multiple factors including the rising levels of disposable income and a growing urban population among others

Looking at the above facts I think opening a bar would be a good business proposition. Mumbai is a matropolitan city with huge population and high average income so seems to be a good candidate city. Finding the right place to open any restaurent or bar is always the key to business. I am trying to build a model to forecast best locality to open the next bar in Mumbai(India).

Data acquision

Finding suitable data and quality data is always a challenge. I have used wikipedia which has emerged as one of the golden source for reliable data so i have used wikipedia to source neighbour hood data in mumbai. I have BS4 and Request libraries in python to scrub that data.

I have used foursquaer to get the venue data to build attibutes of a particular neighbour hoods. I have also used foursquare to collect the data about number of bars in a particular neighbourhood.

DataSource

- Wikipedia
- FourSquare

Challenges With Data Acquision

I wanted to add another dimension of per capita income in each region but i failed to find some good reliable source to find average income in each regio or tax collection in each neighbourhood etc and hence didnt include that in my analysis.

Conclusion

Afte careful analysis and using clusturing technique i have identified 'Breach Candy' and 'Kemps Corner' in south mumbai are 2 neighbourhoods that suitable to open next bar.