IBM - Data Science

Capstone Project

**Introduction**

This report is aimed at stakeholders looking to open an Indian Restaurant Business in British Columbia, Canada.

It involves comparison of localities on the basis of population of the area, frequency of restaurants and restaurants per person. It also divides the Lower Mainland Region into 4 classes based on average household income in the area with category A regions having highest income followed by regions B, C and D respectively.

This report has been made as a submission for IBM Professional Data Science Certification by Japman Singh.

**Business Problem**

Due to great quality of life and better weather conditions than rest of Canada, Vancouver and the rest of Lower Mainland region has become a top location for immigrants to settle. This has led to a high number of businesses opening up limiting the scope for new businesses to open. This project will help to find ideal locations which have best chances of becoming successful.

**Data**

This project uses the following Data Sets

1. BC Postal Codes

2. BC Population

<https://en.wikipedia.org/wiki/List_of_the_100_largest_municipalities_in_Canada_by_population>

https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hlt-fst/pd-pl/Tables/File.cfm?T=1201&SR=1&RPP=9999&PR=0&CMA=0&CSD=0&S=22&O=A&Lang=Eng&OFT=CSV

3. Geocoder API for coordinates of addresses

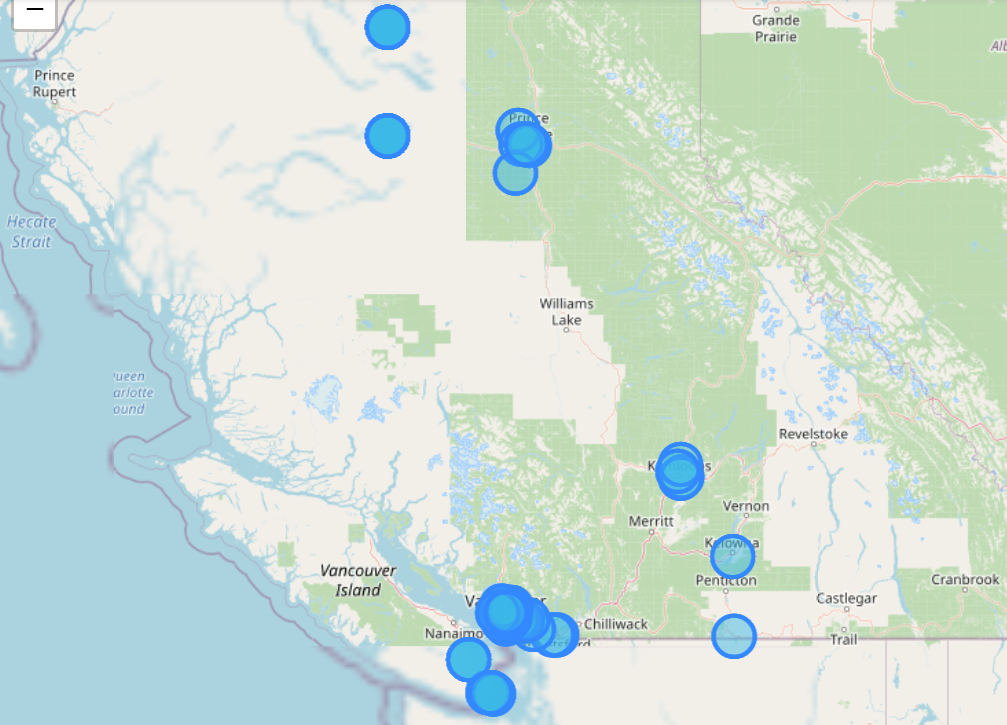
4. Foursquare API to find frequency of Venues

The data in (1) and (2) was scraped using beautifulsoup library in python and was then stored into a separate csv file which has been uploaded in the github repository.

**Methodology**

1.After obtaining the data by scraping, the data was filtered out to remove non-required locations from the analysis.

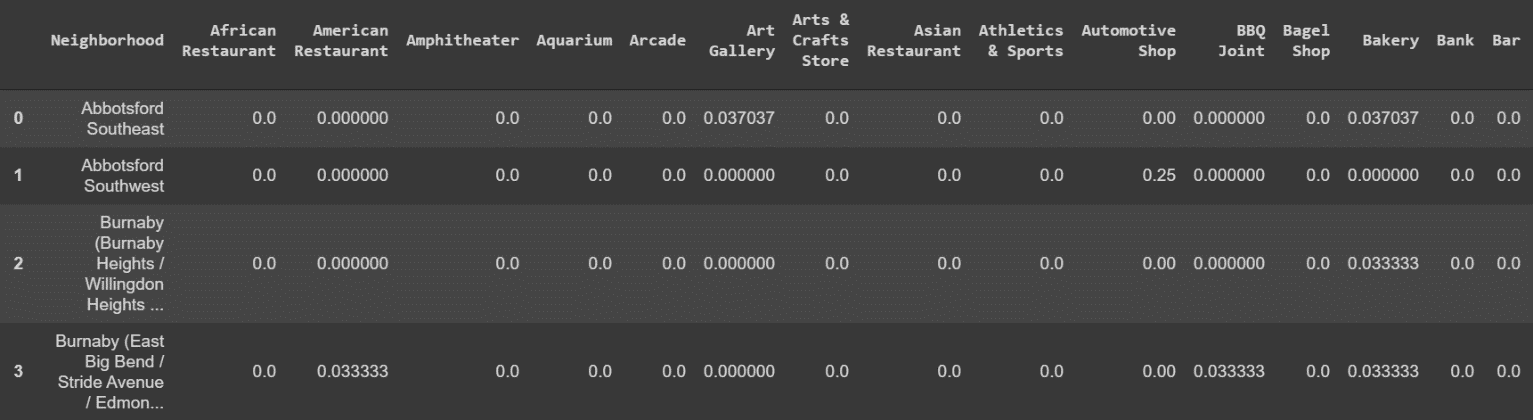
2. The latitudes and longitudes for each location were entered into a dataframe and were plotted using folium.



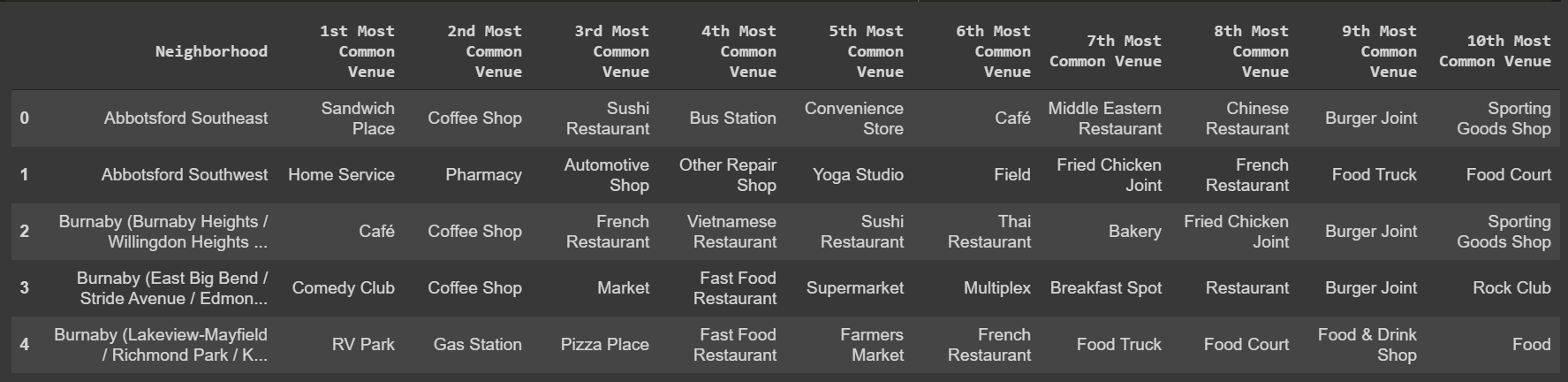
The above cities were used in the analysis.

3. Foursquare API was then used to calculate the venues and their frequencies and an in depth analysis was made for each city

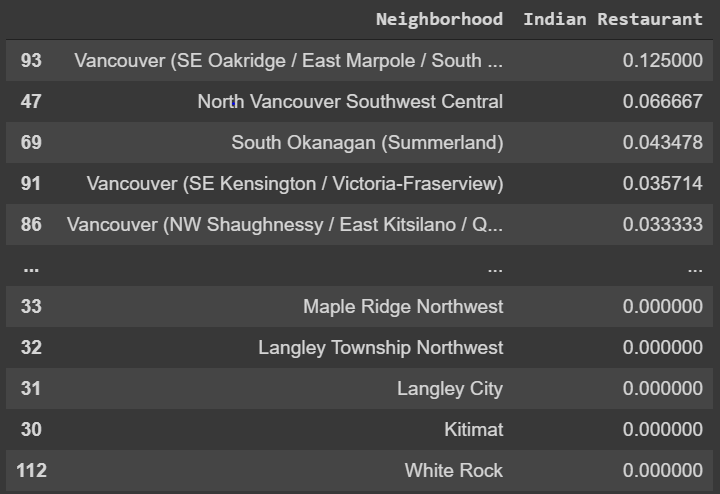




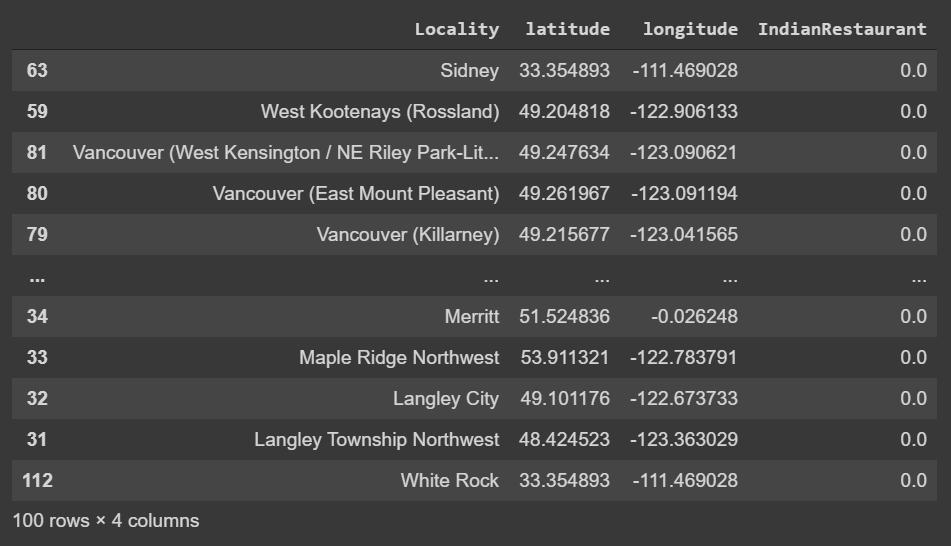
4. The 10 most common venues were listed for each city.



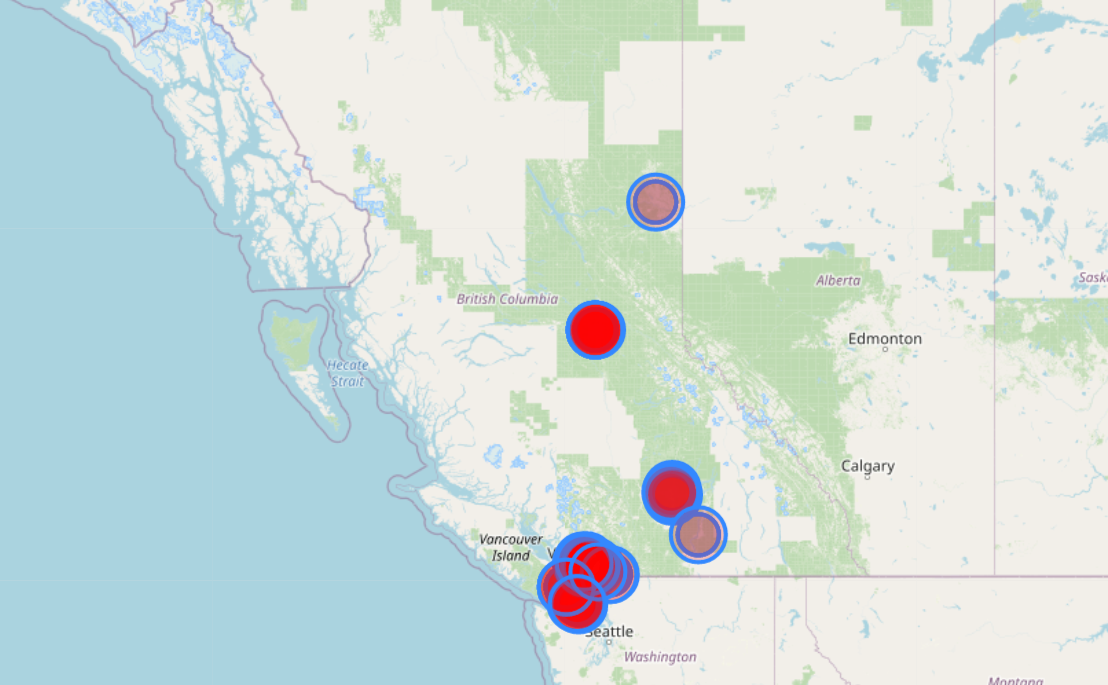
5. Now since we needed to find locations with least number of Indian Restaurants, we analysed Indian Restaurant Frequencies.



6. Now we made a dataframe of locations with Least frequency of Indian Restaurants.



7. To make the list easy to visualise, we plotted these results on a map.



The darker red shaded areas show a greater business opportunity than ligher shaded areas.

**Result**

The places plotted in the above map are suitable for opening an Indian Restaurant. The live map with popups for city name can be found in the JupiterNotebook in the Github repository.

**Observation**

Indian Restaurants are limited to very few locations and have a great scope for expansion.

**Conclusion**

Indian Restaurants can easily be opened in the above red marked locations.