Coursera Capstone

IBM Applied Data Science Capstone

Real Estate Prices & Venues Data Analysis of Winnipeg City

By: Lilivett Ovalle Borrego
November 2020

Problem Statement

Winnipeg is the capital and largest city of the province of Manitoba in Canada. It is centred on the confluence of the Red and Assiniboine rivers, near the longitudinal centre of North America.The region was a trading centre for Indigenous peoples long before the arrival of Europeans. French traders built the first fort on the site in 1738. As of the Canada 2016 Census there were 705,244 people living in Winnipeg proper, with approximately 778,489 living in the Winnipeg Census Metropolitan Area (CMA). Thus, Winnipeg is Manitoba's largest city and Canada's seventh largest. Furthermore, it represents 54.9% of the population of the province of Manitoba, the highest population concentration in one city of any province in Canada.[1]

This boom means a strong interest by buyers and investors in the real estate market. Real estate investors look for locations where house prices are low and facilities (shops, restaurants, parks, etc.) and social spaces are close by. For an individual to find these locations in such a big city, and to gather such information, is a difficult task.

Considering these problems, we seek to create a map and an information chart of the real estate index of Winnipeg where each of its districts are clustered according to venue density.

Data Collection

To solve the problem, we will need the following data:

1. List of areas of Winnipeg with its neighborhoods and postcodes from www.geonames.org .[2]
2. Properties data such as address, price, postcode from the website www.point2homes.com [3]
3. Foursquare API to explore the boroughs of Winnipeg and segment them. [4]

Data Preprocessing

For the first table, the data scraped from www.geonames.org was cleaned and formatted. I removed all the columns with not relevant information and deleted the special characters. I only selected the location and the postcode columns.

For the second table, the data was scraped from www.point2homes.com. I removed all null values and only kept the Neighborhood and Avg. Price columns. The Avg Price column contained strings so it was processed to convert to an integer by removing the '$' special character and comma.

After cleaning these two tables, I performed an inner join and merged the information.Then by using the geocoder library I found the Longitudes and Latitudes of each Location and added new columns for each in my dataframe.

I used the folium python library to visualize geographic details from Winnipeg city and its boroughs, and I created a map of Winnipeg with the boroughs superimposed on top.

Finally, using the Foursquare API in conjunction with the created datasets, a new table with the most common visited venues in Winnipeg neighborhoods is generated.

References

[1] https://en.wikipedia.org/wiki/Winnipeg

[2] https://www.geonames.org/postal-codes/CA/MB/manitoba.html

[3] https://www.point2homes.com/CA/Real-Estate-Listings/MB/Winnipeg.html

[4] https://developer.foursquare.com