

The Battle of neighborhoods - Toronto borough

Author: Abishek Venkatakrishnan

Date: August 30, 2020

Introduction:

- ▶ Canada has one of the hottest real estate markets in the world. Cities like Vancouver and Toronto lead the charts in terms of expensive housing in the country. Toronto attracts a lot of people from different part of the country and as well as the world as its technological hub is constantly growing and thriving.
- ▶ With ever increasing population in the Greater Toronto Area, the real estate market is having a hard time meeting the demand. This has resulted in the housing prices skyrocketing in the last few years.
- ▶ Given the limited supply of housing and ever-increasing prices, it is imperative that people make an informed decision while buying a property.
- ▶ This project aims to help the general public such as buyers and real estate agents to make such an informed decision.

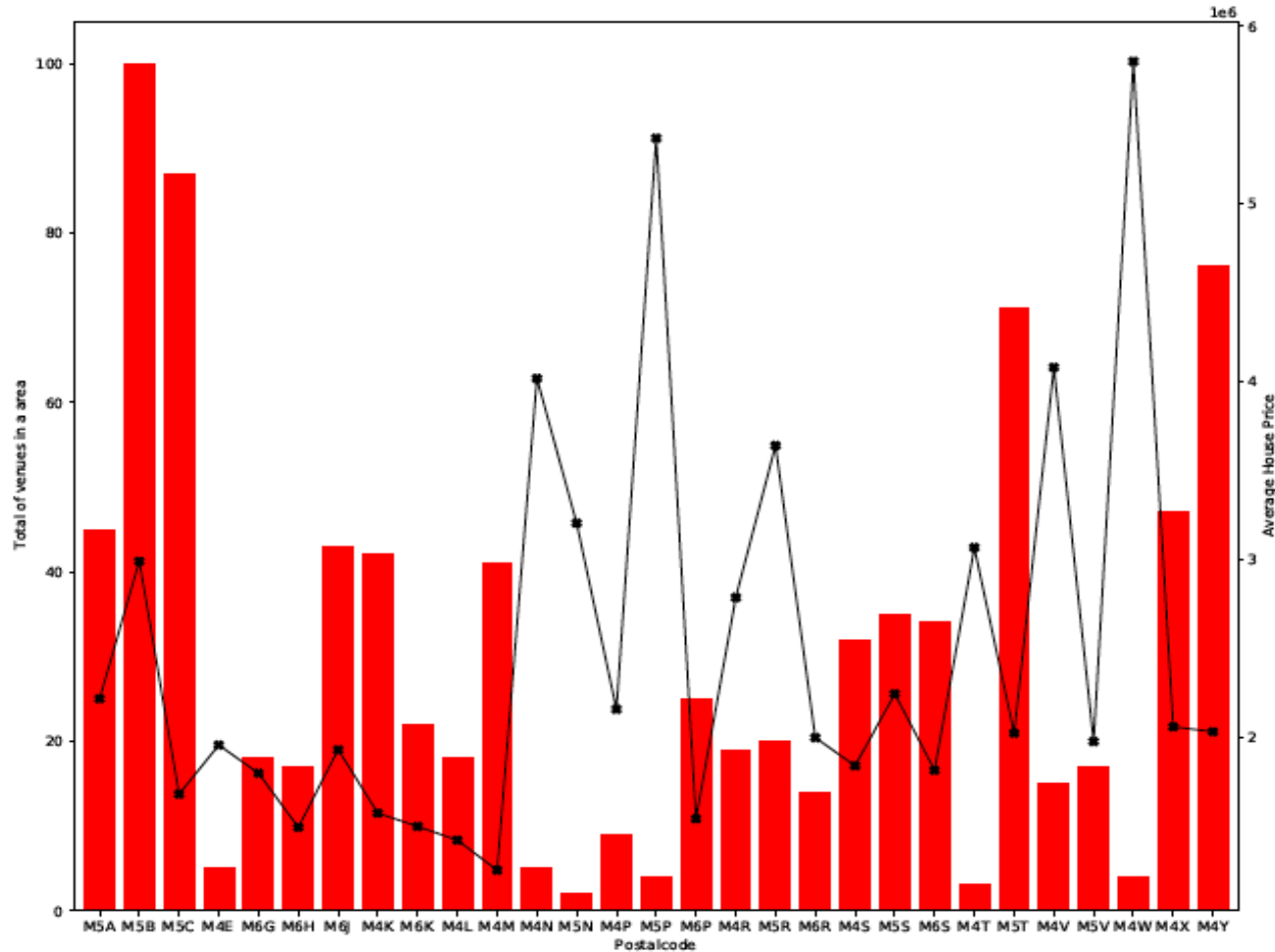
Acquiring data:

- ▶ We need to obtain data from different sources that include Wikipedia, www.housepricehub.com and Foursquare.
- ▶ All the postal code, borough and neighborhood information can be obtained from Wikipedia.
- ▶ Average house prices for each postal code/neighborhood are found at www.housepricehub.com.
- ▶ Venue information and venue categories can be fetched from Foursquare using API.

Methodology:

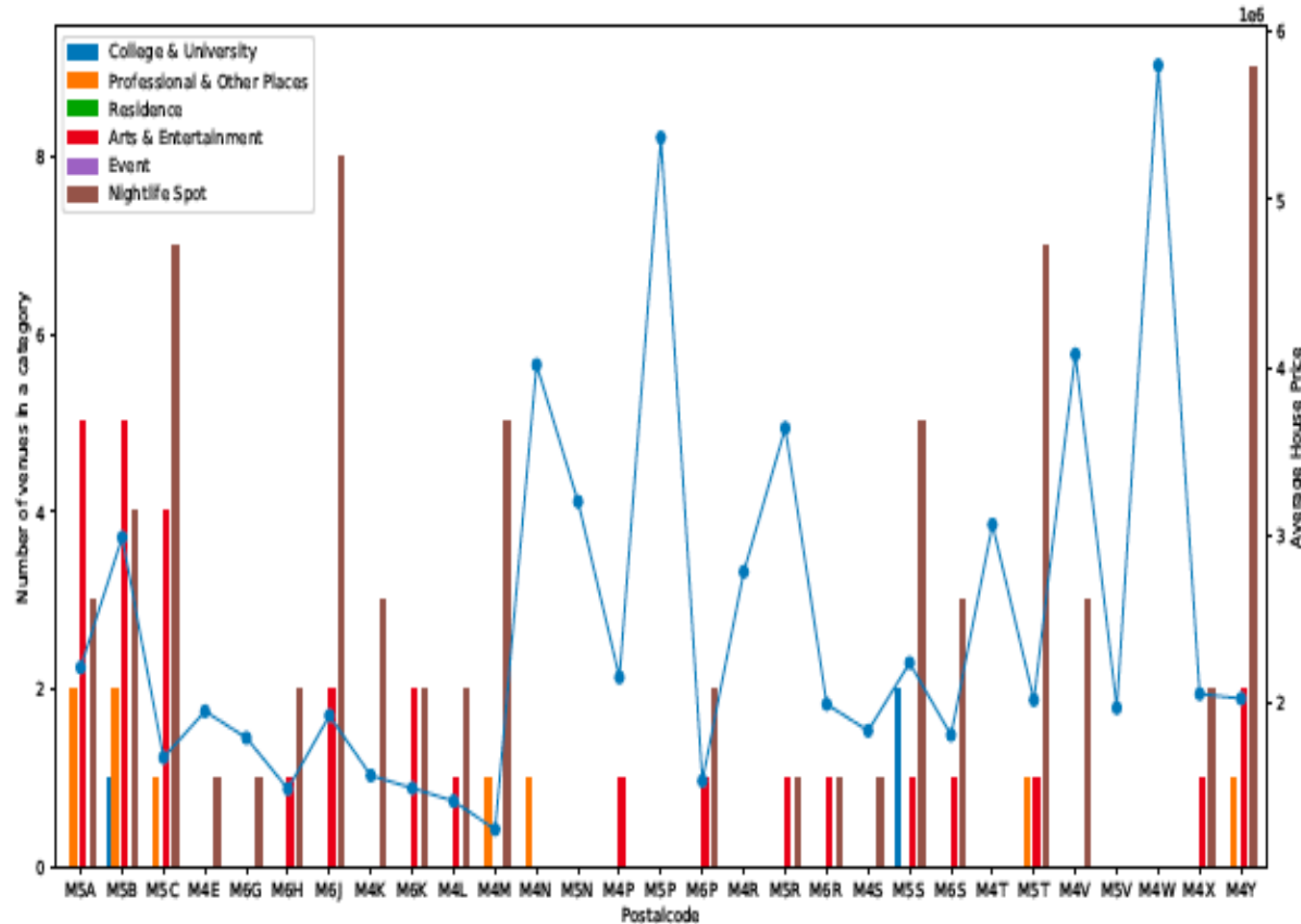
- ▶ Analyze the impact of venue categories and number of venues in each neighborhood on average house prices in that area.
- ▶ Use k-means clustering to cluster house prices, neighborhoods and use that information to pick our sample space to be used in the detailed analysis.
- ▶ Assign main category type to each venue entries based on the category hierarchical level tree obtained from Foursquare website. This "main category" will help us better understand and evaluate data.
- ▶ Count the number of each type of venues in an area and store that information in the dataset. This data will be used to compare against house prices to study their impact.

Results:



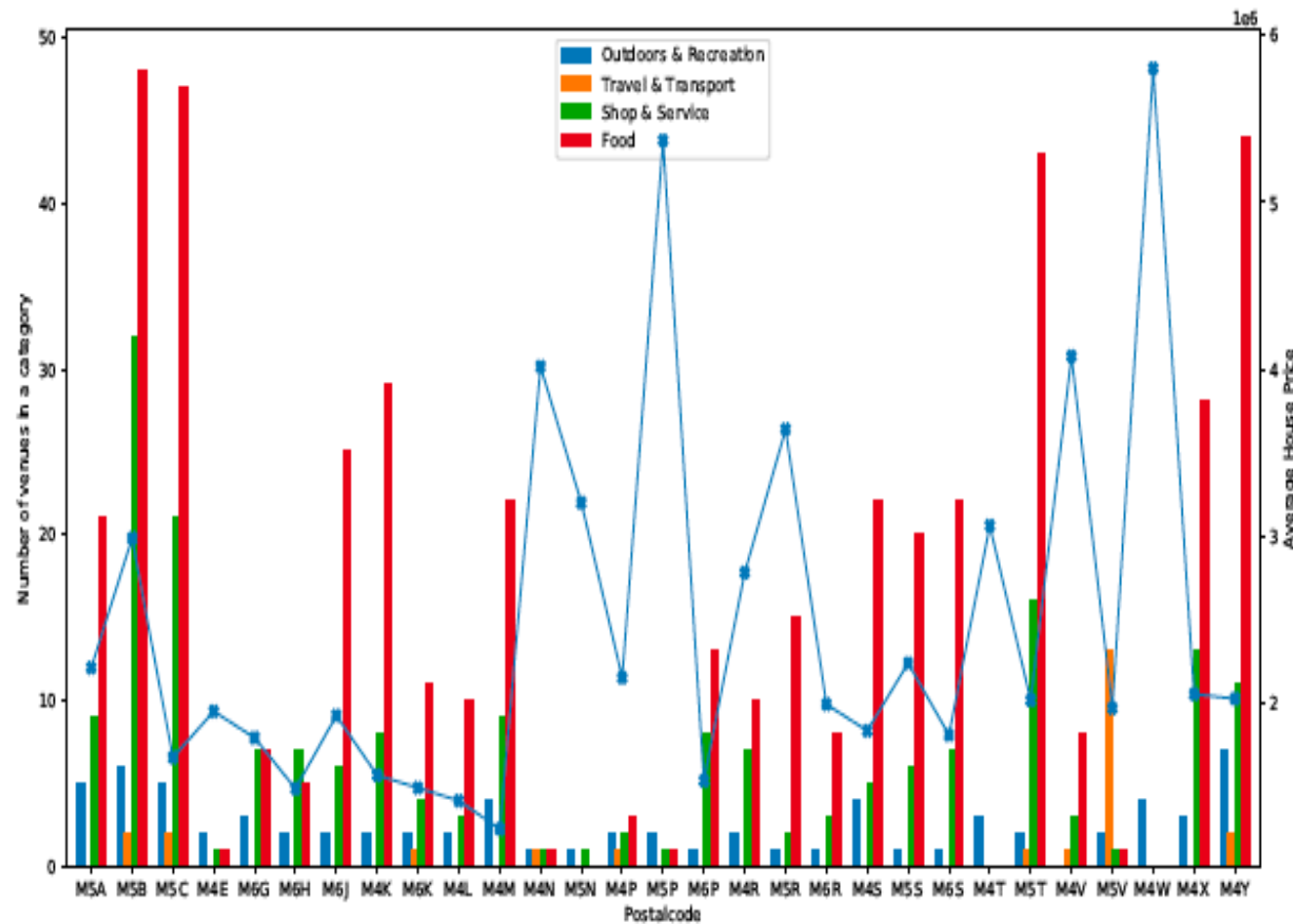
- Plot shows total number of venues in an area for different neighborhoods (bar plot). The line plots shows average house prices.
- Based on the plot, there seems to be no direct impact of total venues on house prices.
- There does seem to be few scenarios where the number of venues seems to have an opposite effect on house prices i.e. lower the number of venues in an area, the higher the house prices.
- However, there are also scenarios where the higher number of venues in an area, the lower the house prices.

Results:



- The “Nightlife Spot” category has a clear trend in the plot shown above. Areas with higher count of nightlife spots seem to have lower average house prices.
- The areas with highest house prices don’t have any or very few nightlife spots in the above plot.

Results:



- One category that seems to have a trend is Food. For the most part, the higher the count for food venues in an area, the lower the house prices.
- The top 3-4 areas that have the highest prices have a very low count of food venues.

Discussion:

- ▶ Based upon the findings in the results section, it is evident that there exists a co-relation between number of venues in an area, number of different venue categories and average house prices.
- ▶ Findings from this study can be used to make an informed decision while trying to decide which neighborhood would be the best to buy a house in.
- ▶ This study can be further extended in the future by including data that contains information about schools, their rankings and proximity, etc. I believe that will add another valuable layer to this project.

Conclusion:

- ▶ In this project, we analyzed the house prices in various neighborhoods in Toronto and studied the impact of various venue categories on house prices. This study can be used by the general public such as buyers and real estate agents to make informed decisions.