

Capstone Project - The Battle of Neighborhoods

Table of contents:

[Introduction/Business Problem](#)

[Data](#)

[Methodology](#)

[Analysis](#)

[Visualization](#)

[Discussion](#)

[Future directions](#)

Introduction/Business Problem

Moving to a different neighborhood or even to another city can be an intimidating experience. You want as much information about potential new neighborhoods as possible. However, maps do not tell you much about what it is like to live in a specific area.

The target group of this project are people who would like to move to or within in Vancouver, BC and for whom one important decision making criteria for selecting their future neighborhood is walking distance to important locations. It is hard for this group of people to get an overview of which of Vancouver's neighborhoods would be best suited for them in terms of their walkability needs. Furthermore, it is hard to know what the next best neighborhood is, in case they cannot afford housing in the area that would be their first choice.

May the distance between the new home and the following locations be the most important factor for this target group:

- nearest beach
- nearest school
- nearest coffee shop
- nearest restaurant
- nearest park

Data

The city of Vancouver publishes all its block numbers and corresponding location data on its Open Data Portal: [link](#). In total, the database includes geospatial data for 5809 blocks numbers. The quality of this data is high. However, some physical blocks have multiple block numbers.

We use the Foursquare API to gather geospatial information on the locations that are of interest (beaches, schools, coffee shops, restaurants, parks). In addition to the information that we are looking for, the Foursquare API returns venues that do not fall into the categories of interest (e.g. if the name of a bus station includes the name of the venue category that we are looking for). Manual filtering of the categories returned was performed to maximise the quality of the data used. After this filtering process, 1072 venues remained in our database for further analysis.

Methodology

In the first step we will collect all the required data: Location of each block within Vancouver, BC's city boundaries and the location of all venues of interest.

In order to identify the location of all venues of interest without exceeding the Foursquare API quota, we will take the following approach:

- 1.1) Identify one central location point per neighborhood (centroid)
- 1.2) Request all "Coffee Shop", "Restaurant", "School", "Beach", "Park" around each centroid
- 1.3) Merge with results of other neighborhoods into one DataFrame and remove duplicates

The second step in our analysis is calculation and exploration of the distance between each block and the nearest venues of interest: Beach, Coffee Shop, Restaurant, School, Park. We go about this as follows:

- 2.1) Develop function to calculate the distance of the closest venue[i] for each block
- 2.2) Calculate the distance of the closest venue for all venue types for each block

In the third and last step, we will cluster Vancouver's blocks using k-means. We compare the means of the features of each cluster to understand the different characteristics of the areas within Vancouver as far as walking distances are concerned.

Analysis

We cluster Vancouver's blocks based on nearest distance to venues of interest.

cluster	Distance to nearest Beach	Distance to nearest Coffee Shop	Distance to nearest Restaurant	Distance to nearest School	Distance to nearest Park
0	1608.130602	296.579779	266.787993	367.357096	310.580154
1	1831.407743	725.483793	544.035172	441.267086	371.192606
2	1953.724817	1282.119105	689.000175	724.043411	497.749299

The mean of the features of all clusters created by K-means ascends with each cluster. The mean of each feature of cluster 1 is higher than the corresponding mean of cluster 0 i.e. the walking distance to all venues of interest are shorter for blocks that are in cluster 0 than cluster 1.

In other words, we can characterize the clusters as follows:

Cluster 0: Close to all venues of interest

Cluster 1: Medium distance to venues of interest

Cluster 2: Furthest away from venues of interest

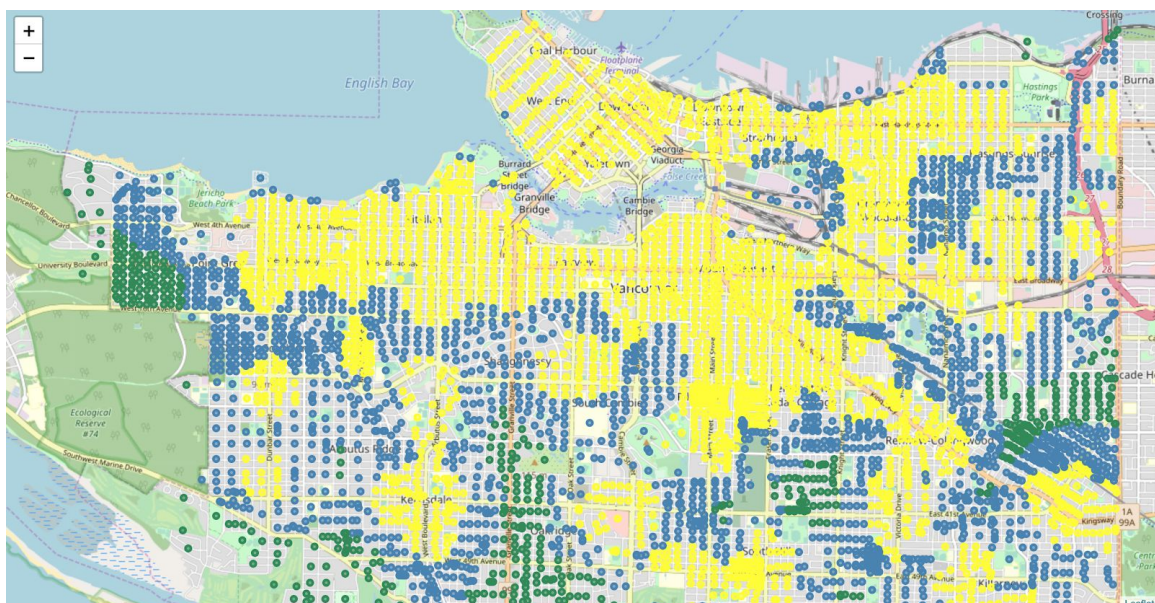
Visualization

Each geolocation of a block has been superimposed on a map of Vancouver. Each colored point represents a block number. The color of the points represents a cluster:

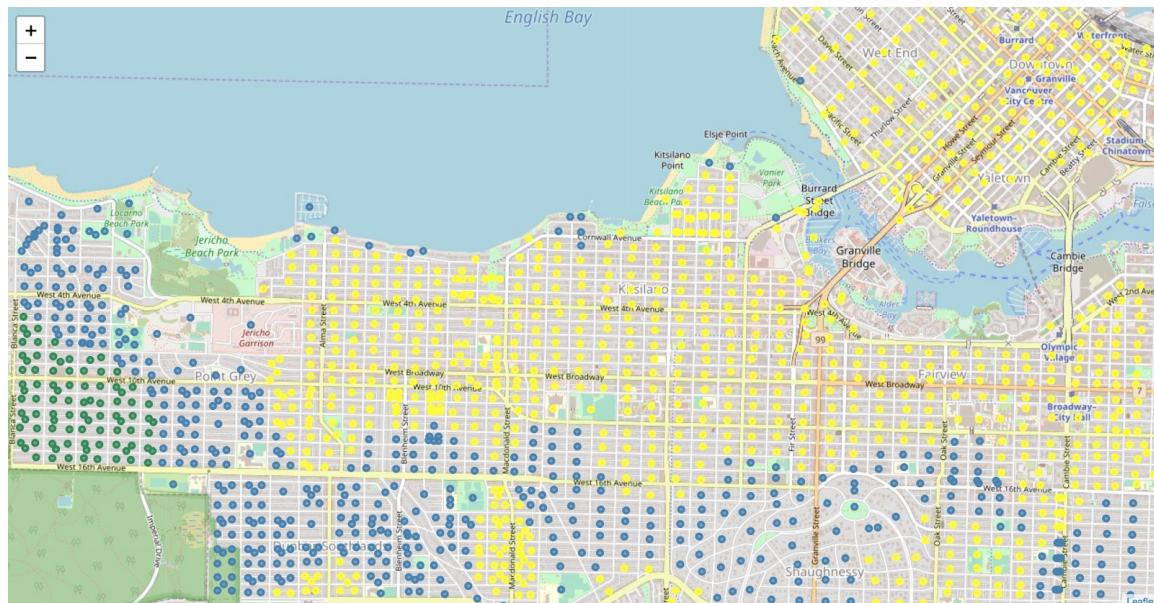
Cluster 0: Close to all venues of interest

Cluster 1: Medium distance to venues of interest

Cluster 2: Furthest away from venues of interest



Screenshot of Vancouver, BC (overview)



Screenshot of Vancouver, BC (Kitsilano area)

Discussion

The goal of this project was to give an overview of the different characteristics of Vancouver's blocks and neighborhoods in terms of walkability. The analysis assesses blocks based on walking distance to the nearest venues of interest (beach, coffee shop, restaurant, school, park). Each type of venue is weighted evenly.

The analysis shows that there are pockets of highly walkable areas with e.g. restaurants, parks, and schools dotted all over Vancouver. At the same time the analysis shows that moving to an extremely pricey area of the city does not necessarily mean that you have all the venues of interest close by e.g. the area south of Locarno Beach Park (Kitsilano) lacks restaurants, schools, coffee shops.

Future directions

When it comes to identifying a neighborhood in Vancouver that reflects your needs as best as possible, this analysis can only be considered one of many data points. This analysis is very limited in terms of the number of features considered. Additionally, all features are weighted evenly which may or may not reflect your personal preferences.

Additional features that may be of interest for selecting a neighborhood include e.g. distance to grocery shops, crime rates, land value, access to public transit, average commute. In order to be able to reflect personal preferences, each of the features needs to be weighed individually (e.g. you may feel that a short walking distance to the beach outweighs the lack of access to coffee shops and restaurants close by).