

# The Battle of Neighborhoods- Open a Kids Mall

Sayali Gholap

June, 2021

# Table of Contents

## Contents

|        |   |    |
|--------|---|----|
| 1.     | Introduction .....                                | 3  |
| 1.1.   | Background .....                                  | 3  |
| 1.2.   | Business Problem .....                            | 3  |
| 1.3.   | Interest.....                                     | 3  |
| 2.     | Data .....  | 3  |
| 2.1.   | Data sources.....                                 | 3  |
| 2.2.   | Data Cleaning .....                               | 3  |
| 2.3.   | Feature Selection .....                           | 4  |
| 3.     | Methodology.....                                  | 4  |
| 3.1.   | Web scraping .....                                | 4  |
| 3.2.   | Exploring Vancouver's Neighborhood .....          | 5  |
| 3.3.   | Cluster Neighborhoods .....                       | 6  |
| 3.3.1. | Elbow method to determine optimal value of K..... | 6  |
| 5.     | Analysis .....                                    | 7  |
| 5.1.   | Examine Clusters.....                             | 7  |
| 6.     | Results and Discussion.....                       | 11 |
| 7.     | Conclusion.....                                   | 13 |
| 8.     | References .....                                  | 13 |

# 1. Introduction

## 1.1. Background

- The purpose of this project is to explore neighborhoods in Vancouver, Canada to open a kids mall. A mall that offers kids accessories, food, clothing, gaming and entertainment. The mall would be purely dedicated to kids entertainment and requirements.
- Finding a kids mall that would suffice all the needs of children is rare. A place that would be convenient especially for parents to buy baby food, clothing, toys, etc. at one stop with no need of traveling to different places for different requirement.
- The mall will provide the residents of Vancouver a place that would fulfil all the requirements of their child.

## 1.2. Business Problem

- The project aims to explore different neighborhoods in Vancouver that would be optimal to open a new Kids mall. The analysis would be based on different features like number of playground, parks, kids store, garden in the neighborhood.

**Business problem:** Explore neighborhoods optimal to open a new kids store.

## 1.3. Interest

**Stakeholders:** Residents of Vancouver especially parents that would get benefit of all the services at one stop.

# 2. Data

## 2.1. Data sources

- For information related to Vancouver city and its postal codes; webscrapping would be done from List\_of\_postal\_codes\_of\_Canada:\_V [1]  
Vancouver's neighborhoods along with postal codes would be extracted.
- Get location co-ordinates of each postal code using- Geopy [2].
- Venue related data such as venue name, type, location, tips via- Foursquare API [3]
- Venue types- garden, parks, playground, kids store would be used for analysis
- Vancouver city official website [4]

## 2.2. Data Cleaning

- The data extracted from List\_of\_postal\_codes\_of\_Canada:\_V [1] had total 180 postal codes. We filtered the data resulting postal codes of only Vancouver City. Vancouver City has 31 postal codes. These 31 postal codes are used further to find the optimal neighborhood for opening new kids mall.
- After web scrapping, the City-neighborhood data column is separated. Any postal code with 2 or more neighborhood is grouped together and separated with commas in a column.

- Location co-ordinates are assigned to each postal code with help of Geocoder. If any postal code isn't assigned with the coordinates then they are dropped.

### 2.3.Feature Selection

- There are total 236 unique venue categories in Vancouver neighborhoods. Out of 236 only 4 Venue categories are selected- 'Park', 'Toy / Game Store', 'Garden', 'Playground'.

## 3. Methodology

**Relevant factors considered in analysis:** As we would be determining a more suitable neighborhood to open a new kid's mall. Some factors like-

1. Number of existing kids store in the neighborhood that would lead to competition.
2. Number of playground, parks, garden in the neighborhood that would attract customers to the Kids mall
3. Demographics of the neighborhood

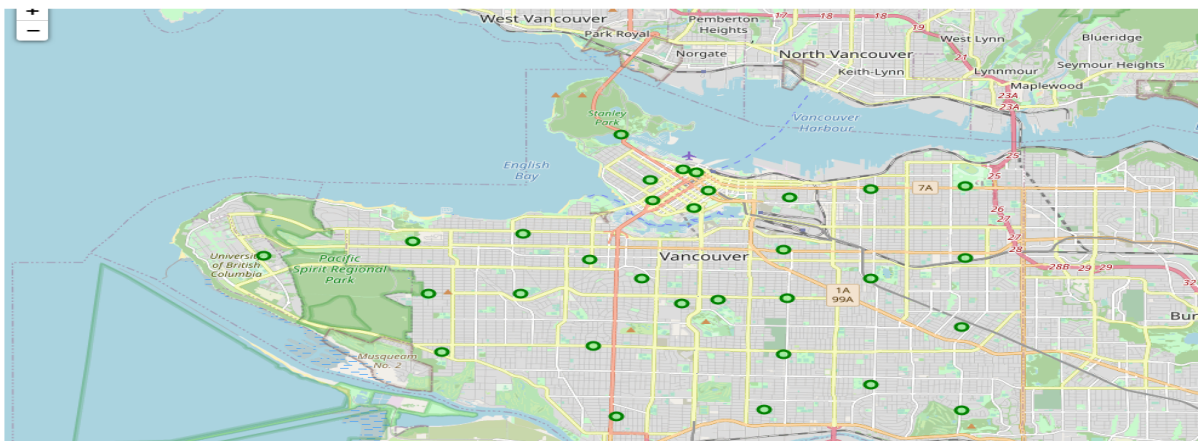
### 3.1.Web scraping

Beautiful Soup and Geopy libraries are used to extract and create a data frame as below-

|   | PostalCode | City      | Neighborhood                                      | Latitude | Longitude  |
|---|------------|-----------|---|----------|------------|
| 0 | V5K        | Vancouver | North Hastings-Sunrise                            | 49.28167 | -123.03998 |
| 1 | V5L        | Vancouver | North Grandview-Woodland                          | 49.28070 | -123.06684 |
| 2 | V5M        | Vancouver | South Hastings-Sunrise , North Renfrew-Colling... | 49.26043 | -123.04009 |
| 3 | V5N        | Vancouver | South Grandview-Woodland , NE Kensington-Cedar... | 49.25442 | -123.06680 |
| 4 | V5P        | Vancouver | SE Kensington-Cedar Cottage , Victoria-Fraserview | 49.22293 | -123.06688 |

The resulting data frame has 31 rows, i.e. 31 postal codes describing its City, Neighborhood[s], Latitude and Longitude.

A map is visualized for these postal codes using folium library. **Map1- Vancouver's Neighborhood**



### 3.2.Exploring Vancouver's Neighborhood

Foursquare API is used to explore Neighborhoods. A limit of total 100 venues and radius of 1000m was set to collect all the data.

11 neighborhoods returned venues up to 100 which are sufficient to explore the neighborhoods.

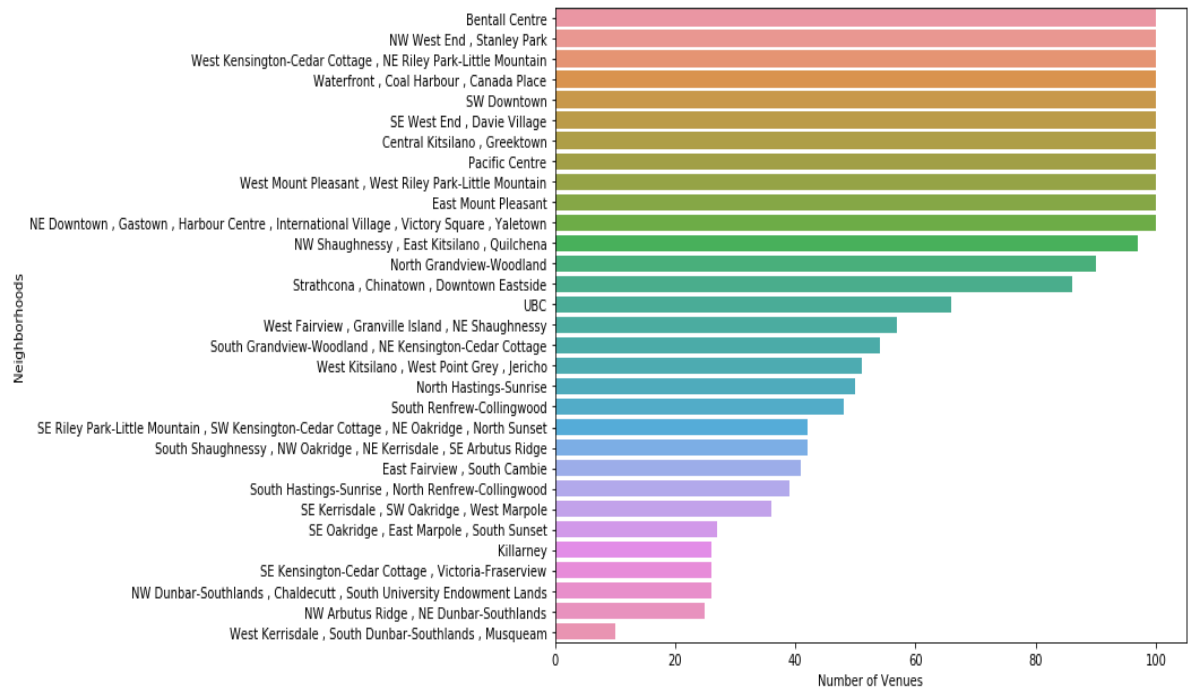


Figure1. Number of Venues per Neighborhood

There are total 236 unique Venue Categories and we have chosen 4 for our analysis.

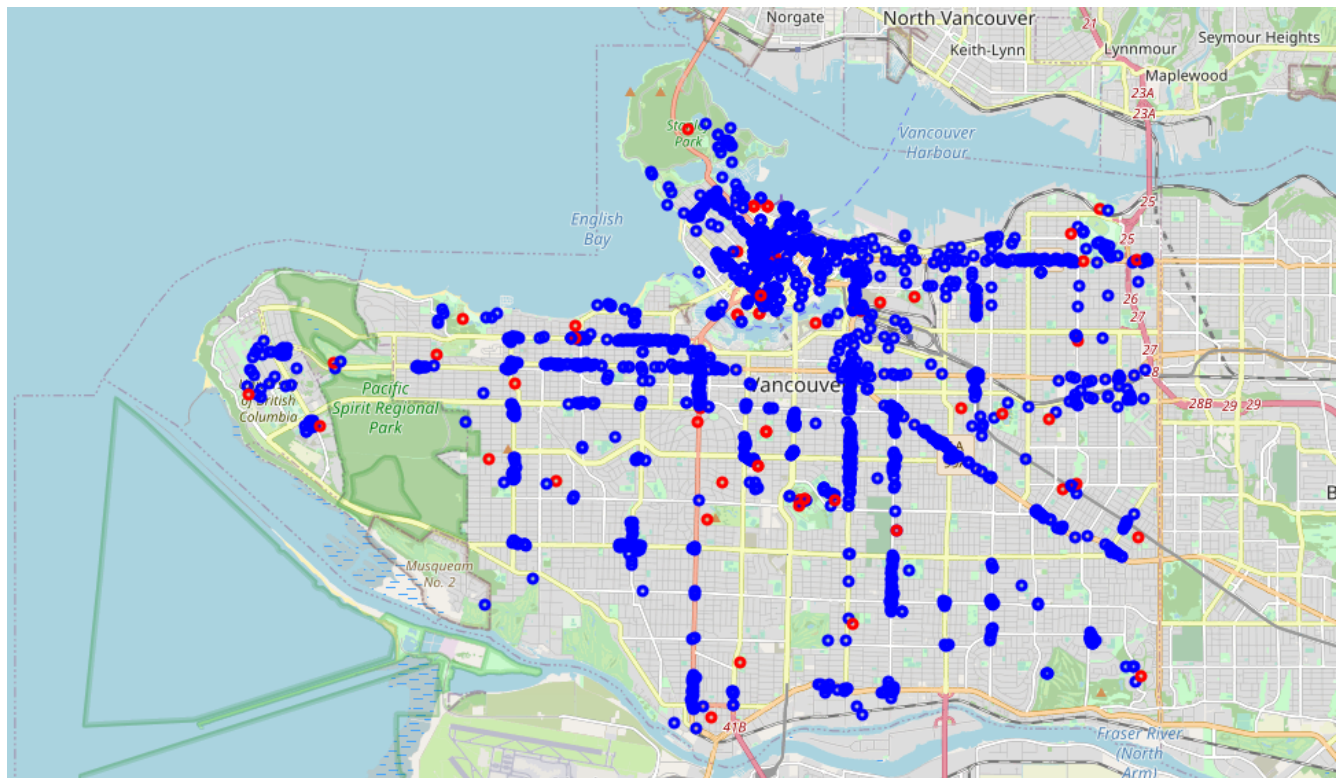
The 4 Venue Categories of our interest are-

1. Park
2. Playground
3. Garden
4. Toy / Game Store

The first 3 are the areas where more customers would be attracted to our Kids Mall. So we would be focusing more on these areas to open a Kids mall near to park/ playground, garden.

The areas with existing Toy/ Game Store are more likely to be a competition but on the other hand they also serve to be an area that is already known for Kids store.

We can visualize it in a map with red marker as our 4 Venue Category and blue marker as the rest of the Venue Categories.



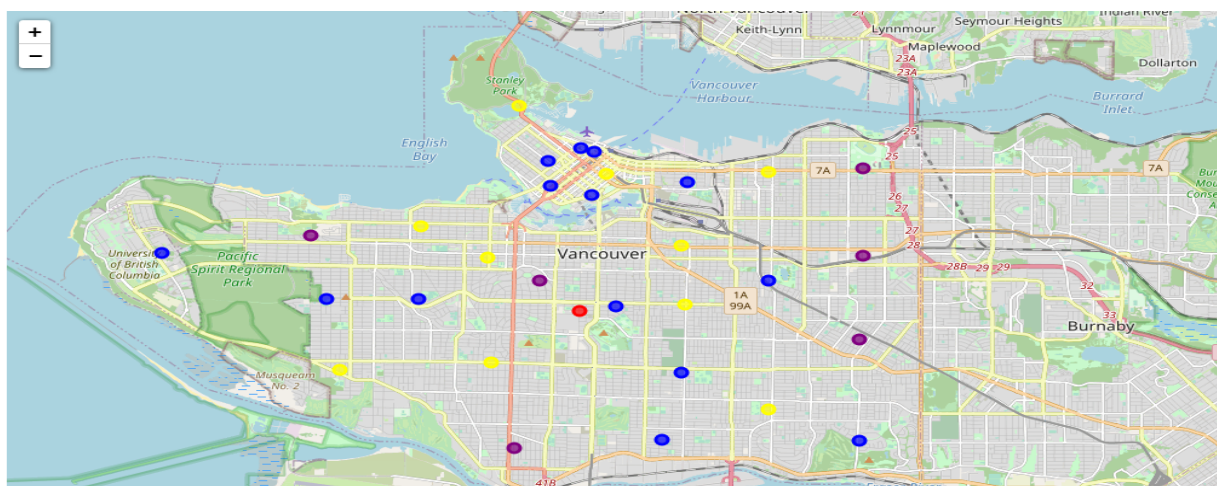
**Map 2:** Venue Categories- [Park, Playground, Garden, Toy / Game Store]-Red marker, [Rest of the Venue Categories] - Blue marker

### 3.3. Cluster Neighborhoods

K-Means algorithm is used for Clustering analysis. Using one hot encoding, categorical data is converted into numerical data where each Neighborhood is turned into frequency of the venue located in each neighborhood. Park, Playground, Garden, Toy / Game Store Venues are used for the analysis.

#### 3.3.1. Elbow method to determine optimal value of K

Using elbow method the optimal value of K is found to be 4. Hence the data would be segmented into 4 clusters. **Map-3 – Cluster analysis of Vancouver Venue**



In the **Map3**, Cluster 0 is shown with Blue marker, Cluster 1 with Yellow marker, Cluster 2 with Red, Cluster 3 with Purple.

## 5. Analysis

### 5.1. Examine Clusters

At first glance, Cluster0 (Blue) is spread throughout the map, Cluster1 (Yellow) is mostly in the central region, Cluster2 (Red) has only 1 postal code in it, Cluster3 (Purple) is mostly on the Outline of the city.

Next, we'll see frequency of different venues per cluster.

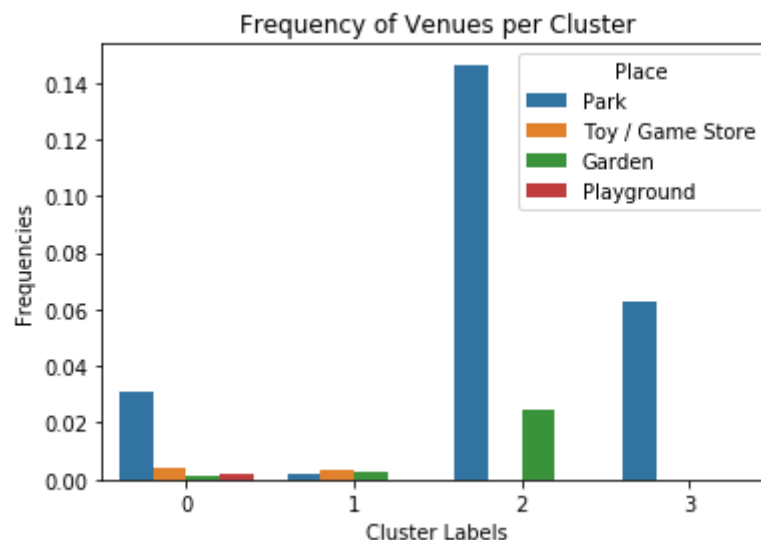


Figure2- Frequency of Venues per Cluster

**Cluster0:** has all types of venues but less in count as compared to Cluster2 and Cluster3.

**Cluster1:** has no playground and has the lowest counts of venues compared to the rest.

**Cluster2:** has the most number of Parks and Garden, but no existing toy store and playground. Also Cluster 2 has only 1 postal code as mentioned earlier. We'll check in detail for that particular postal code.

**Cluster3:** has moderate number of parks but no other venue type.

**Following is the Analysis of Venues per Postal Code-**

#### 1. Examine Venues per Postal code- Cluster0

If we check in Figure3(Venues per Postal code area- Cluster0) The postal code **V5Y** seems to be a good choice to open a new Kids mall as the number of Parks, Garden is moderate; also has a playground and no kids store.

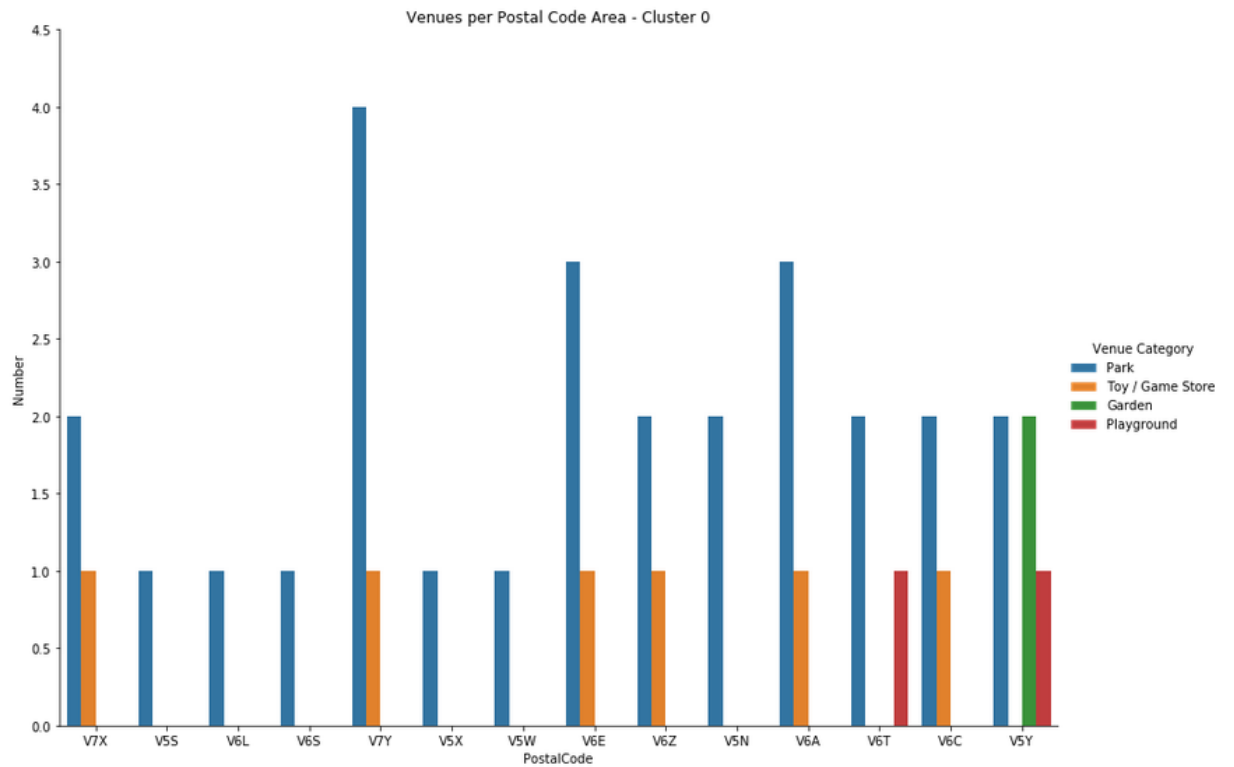


Figure3- Venues per postal code area- Cluster0

## 2. Examine Venues per Postal code- Cluster1

In Cluster1, there are 3 toy stores in 3 different postal code areas. Postal code **V6G, V6M** seems to be better choice than the rest for open a Kids mall as there is at least 1 park/garden to attract kids and no other toy store in competition.

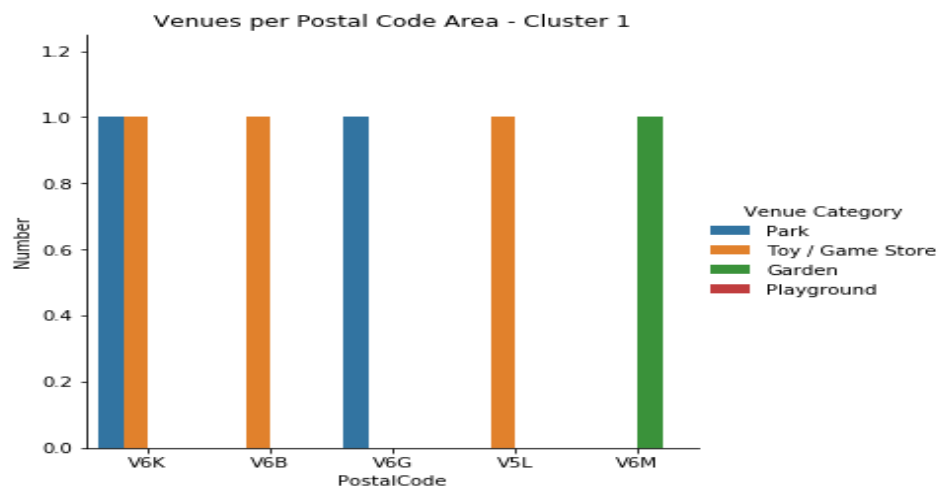


Figure4- Venues per postal code area- Cluster1



3. Examine Venues per Postal code- Cluster2

**V5Z** seems to be a good choice to open a kids store nearby as the number of parks is 6 and garden 1 with no existing toy store.

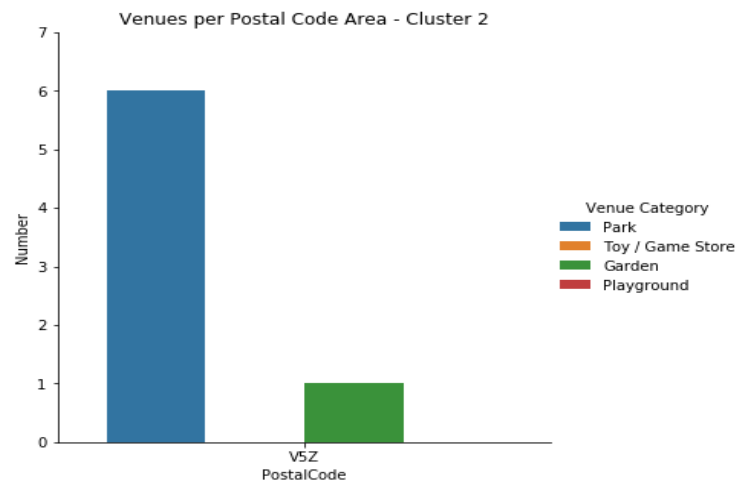


Figure5- Venues per postal code area- Cluster2

4. Examine Venues per Postal code- Cluster3

**V5K** and **V6H** have high number of parks in the Cluster and no existing toy store. Can be considered as a option to open.

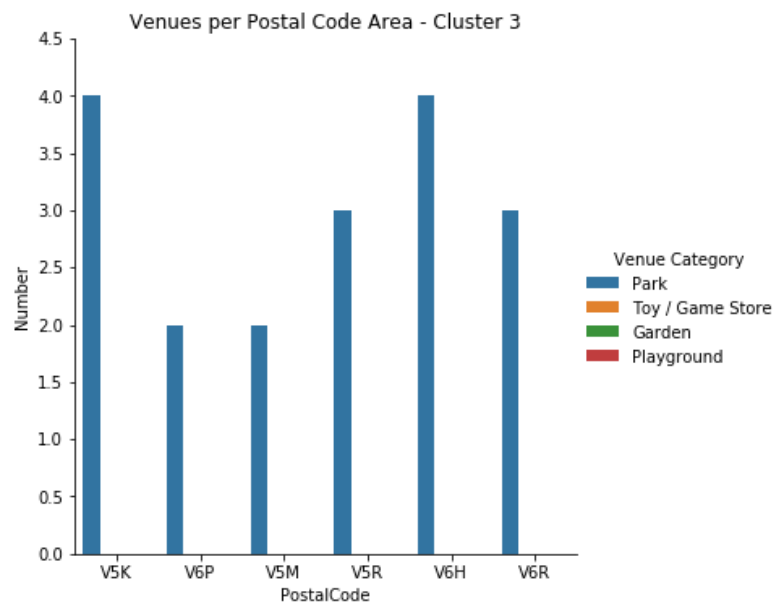
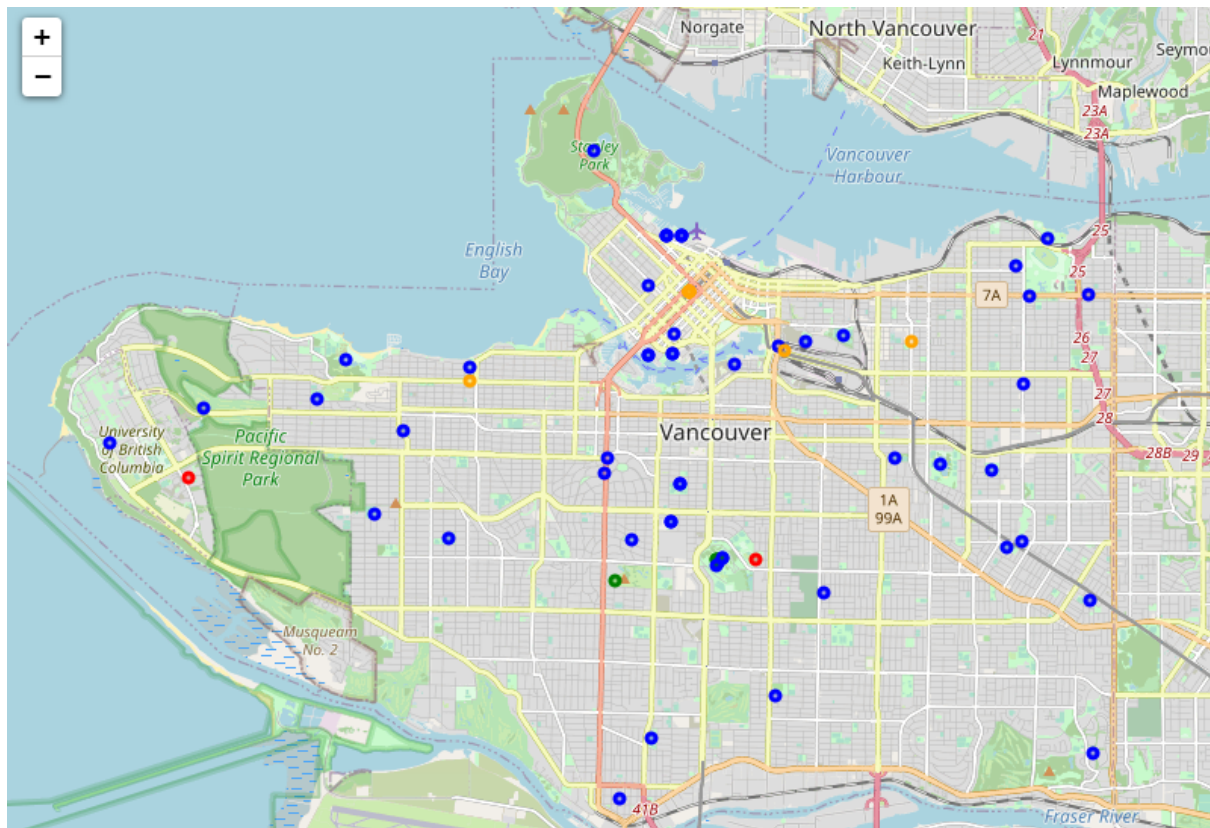
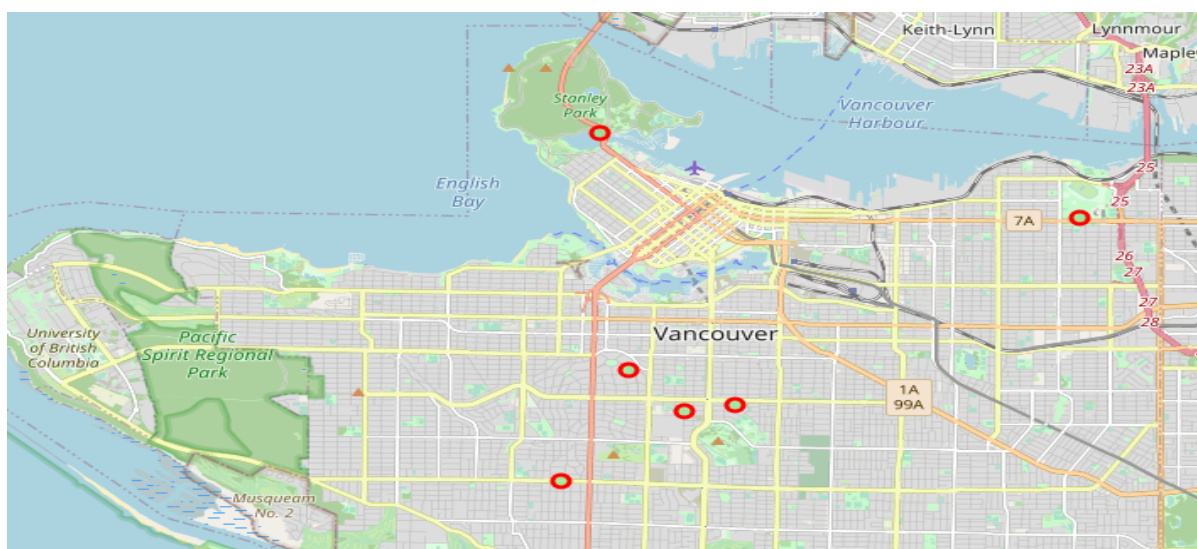


Figure6- Venues per postal code area- Cluster3

**Below Map4:** shows all the Venue types- Park, Playground, Garden, Toy / Game Store distributed over Vancouver City.



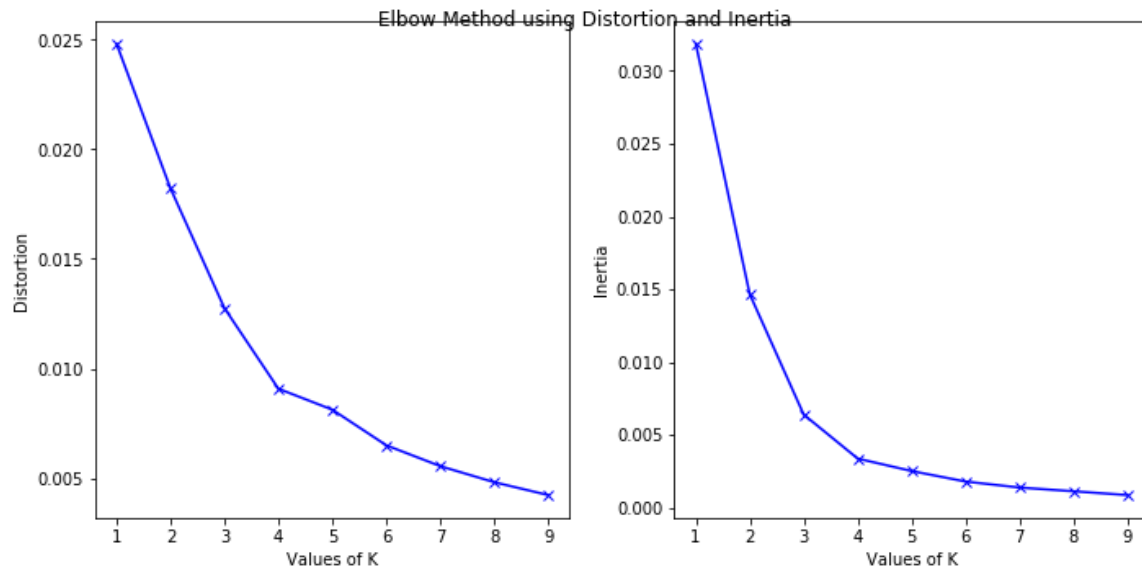
**Below Map5:** shows all the postal codes that are considered to be optimal to open a Kid's Mall after examining the clusters. Postal codes- V5Y, V6G, V6M, V5Z, V5K and V6H.



## 6. Results and Discussion

Using elbow method we determined the optimal value of K for cluster analysis.

The value optimal value of K is 4.



The K-Means algorithm classified the Vancouver's data into 4 clusters.

**Cluster0:** Refer Map-3, Cluster 0 is shown with Blue marker.

Cluster 0 data is scattered over the Vancouver City and dense in the north region. From Figure-2, we can observe that Cluster0 has all types of Venues in it but the total venue count is less than that of Cluster2 and Cluster3.

|    | Neighborhood           | 1st Most Common Venue | 2nd Most Common Venue        | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | 6th Most Common Venue | 7th Most Common Venue | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue | PostalCode | City      | Latitude | Longitude  |
|----|------------------------|-----------------------|------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------|------------|-----------|----------|------------|
| 11 | North Hastings-Sunrise | Park                  | Theme Park Ride / Attraction | Fast Food Restaurant  | Vietnamese Restaurant | Coffee Shop           | Theme Park            | Pharmacy              | Burger Joint          | Chinese Restaurant    | Food Truck             | V5K        | Vancouver | 49.28167 | -123.03998 |

**North Hastings-Sunrise- V5K** is a residential area in Vancouver with most Common Venue as parks and theme park. Many customers could be attracted to Kids mall due to existing parks, garden, playground and no kid's store.

**Cluster1:** has no playground and has the lowest counts of venues compared to the rest.

|   | Neighborhood              | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | 6th Most Common Venue | 7th Most Common Venue | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue | PostalCode | City      | Latitude | Longitude  |
|---|---------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------|------------|-----------|----------|------------|
| 9 | NW West End, Stanley Park | Café                  | Coffee Shop           | Japanese Restaurant   | Dessert Shop          | Noodle House          | Korean Restaurant     | Grocery Store         | Sushi Restaurant      | Hotel                 | Seafood Restaurant     | V6G        | Vancouver | 49.29686 | -123.13759 |

**NW West End , Stanley Park-V6G** is a urban park that offers an escape from busy City life. Its a good place to relax and hike. Its not an ideal place to open a mall but can be a good choice to

open in the same postal area V6G. For this neighborhood the most common restaurants are food places.

|    | Neighborhood                                    | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | 6th Most Common Venue | 7th Most Common Venue | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue | PostalCode | City      | Latitude | Longitude |
|----|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------|------------|-----------|----------|-----------|
| 22 | South Shaughnessy, NW Oakridge, NE Kerrisdal... | Bus Stop              | Chinese Restaurant    | Japanese Restaurant   | Asian Restaurant      | Sushi Restaurant      | Tea Room              | Bank                  | Coffee Shop           | Gym Pool              | Dessert Shop           | V6M        | Vancouver | 49.2344  | -123.1454 |

**V6M-** is a residential area with most common venue as Bus stop and 7th most as Bank. Being a residential area more residents are attracted to Kids mall. It is more suitable than V6G to open a Kids mall

**Cluster2:** has the most number of Parks and Garden, but no existing toy store and playground. Also Cluster 2 has only 1 postal code as mentioned earlier.

|   | Neighborhood                | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | 6th Most Common Venue | 7th Most Common Venue | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue | PostalCode | City      | Latitude | Longitude  |
|---|-----------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------|------------|-----------|----------|------------|
| 2 | East Fairview, South Cambie | Coffee Shop           | Park                  | Bank                  | Chinese Restaurant    | Bubble Tea Shop       | Café                  | Sushi Restaurant      | Sandwich Place        | Garden                | Cantonese Restaurant   | V5Z        | Vancouver | 49.247   | -123.12051 |

With many parks (2<sup>nd</sup> most common venue) and garden it can be considered as a decent place to open a Kids mall

**Cluster3:** has moderate number of parks but no other venue type.

|    | Neighborhood           | 1st Most Common Venue | 2nd Most Common Venue        | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | 6th Most Common Venue | 7th Most Common Venue | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue | PostalCode | City      | Latitude | Longitude  |
|----|------------------------|-----------------------|------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------|------------|-----------|----------|------------|
| 11 | North Hastings-Sunrise | Park                  | Theme Park Ride / Attraction | Fast Food Restaurant  | Vietnamese Restaurant | Coffee Shop           | Theme Park            | Pharmacy              | Burger Joint          | Chinese Restaurant    | Food Truck             | V5K        | Vancouver | 49.28167 | -123.03998 |

**North Hastings-Sunrise** is a residential area with many parks attracting customers to Kids mall.

|    | Neighborhood                                    | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | 6th Most Common Venue | 7th Most Common Venue  | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue | PostalCode | City      | Latitude | Longitude  |
|----|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------|-----------------------|-----------------------|------------------------|------------|-----------|----------|------------|
| 26 | West Fairview, Granville Island, NE Shaughnessy | Park                  | Coffee Shop           | Bank                  | Pizza Place           | Restaurant            | Breakfast Spot        | Furniture / Home Store | Bakery                | Café                  | Bus Stop               | V6H        | Vancouver | 49.25444 | -123.13172 |

Granville Island is a shopping district and if it doesn't have a toy store then would be perfect to open a kids Mall

Based on the analysis I would recommend following places to explore to open a new Kid's Mall-

1. **V5K- North Hastings-Sunrise** -> A perfect neighborhood that offers all venues of our interest with no existing competitor (from Cluster0).
2. **V6M- South Shaughnessy, NW Oakridge, NE Kerrisdal** -> A residential area with more potential than other neighborhoods in same cluster (from Cluster1)
3. **V5Z- East Fairview, South Cambie** -> A new place to explore Kids Mall (from Cluster2)
4. **V5K- North Hastings-Sunrise** ->A residential area with many parks attracting more customers to the store (from Cluster3)
5. **V6H- West Fairview, Granville Island, NE Shaughnessy**-> A shopping district that would definitely offer customers to Kids Mall (from Cluster3)

## 7. Conclusion

In this study, I analysed different neighborhoods of Vancouver City to explore promising locations to open a Kid's Mall. Features like parks, garden, existing kids store/ toy store, playground helped in the analysis of optimal locations. Foursquare API was used to get location details like Venue names, category, etc. K-means clustering algorithm helped to clusters the neighborhoods into 4 clusters. Postal codes from each neighborhood were analysed in detail with help of visualization. As a result top 5 locations can be presented to stakeholders to open a new kid's store.

## 8. References

- [1] List\_of\_postal\_codes\_of\_Canada:\_V -  
[https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_V](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_V)
- [2] Geopy - <https://pypi.org/project/geopy/>
- [3] Foursquare API - <https://developer.foursquare.com/>
- [4] Vancouver city official website - <https://vancouver.ca/>
- [5] Vancouver Wiki- <https://en.wikipedia.org/wiki/Vancouver>