

# Venue Characteristics Near the Rapid Transit Stations and Their Relationship to the Ridership

Siyu Lin

December 27, 2020

## 1 Introduction

Metro Vancouver is expected to accommodate 1 million population and 500,000 jobs by 2040. If all the new residents adopt a car-oriented lifestyle, the regional road network will be heavily congested. The congestion translates into direct economic loss because people are wasting time sitting in the traffic. A resilient transit network is the solution - it provides an affordable and equitable mobility option to all the residents. Compared to a private car, a bus is able to move more people to their destination. In addition, the rapid transit network such as SkyTrain and West Coast Express (WCE, the commuter rail) with dedicated right of way - this can reduce the transit rider's travel time significantly compared to the auto driver during the rush hour.

TransLink, the transit agency serving Metro Vancouver area, relies on the fare revenue to maintain financial sustainability and to reserve fund for transit infrastructure investments. In particular, the SkyTrain guideway and stations are expensive - a recent business case for Surrey-Langley SkyTrain Project estimated a cost \$1.63 billion to extend the guideway by seven kilometres and construct four new stations.

To learn from the existing operations, it would be helpful to characterize the existing rapid transit stations and understand the factors contributing to the high-ridership. This project will provide transit agency insights into correlation between the venue characteristics surrounding transit stops and whether the venue types have an impact on the transit ridership. The transit agency can then develop transit oriented development policies to foster an business environment that stimulate the ridership growth.

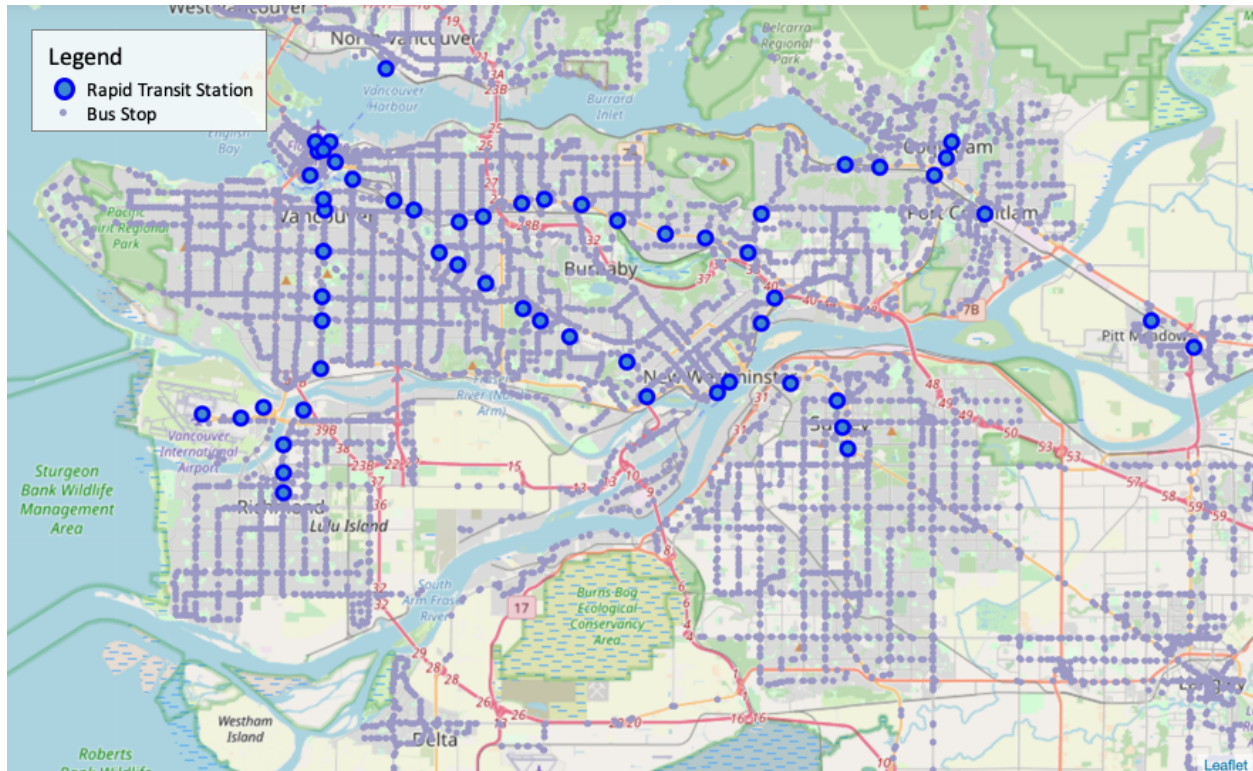
## 2 Data

### 2.1 Rapid Transit Station Locations

Metro Vancouver is serviced by three types of rapid transit systems: SkyTrain, SeaBus and WCE. The rapid transit services are operating on their own dedicated right of way and the users are paying a premium fare for the high capacity and service frequency. The rapid transit network is

supported by the bus network – a feeder system that moves people between the rapid transit station and their destination (home).

The transit stops location data (latitude and longitude) are loaded from the General Transit Feed Specification (GTFS) database published by TransLink<sup>1</sup>. The transit station locations are mapped below – this project will focus on the rapid transit stop locations highlighted in blue.



## 2.2 Foursquare Venue Data

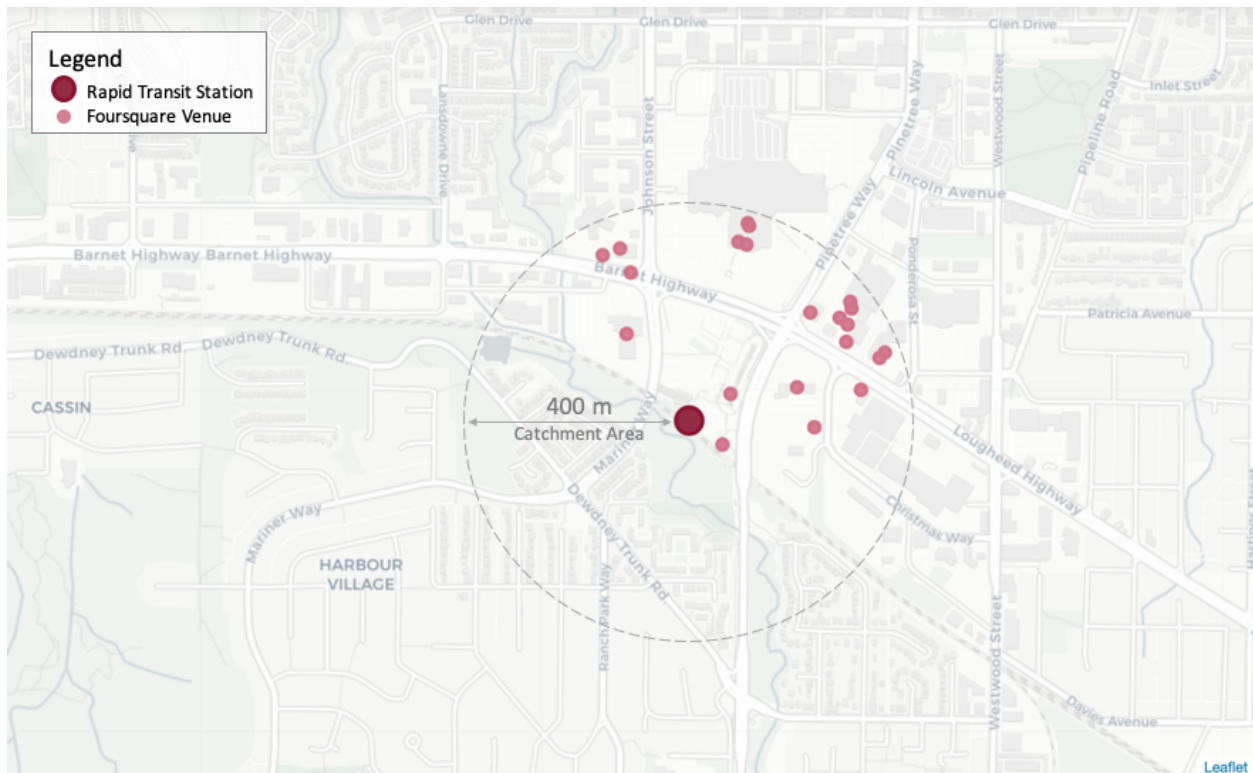
The venue database, generated through the Foursquare API, includes the venue within the catchment area of the transit stops. The catchment area is set to be 400-metre radius, equivalent to five-minute walk, from the station. Following information are registered into the venue database:

- Transit Stop (the query location)
- Venue ID
- Venue name
- Venue type
- Venue location (latitude, longitude)

The map below illustrates the 21 venues scattering around Coquitlam Central Station. More than one-third of the venues are located in the Pinetree Village, a plaza to northeast corner of the catchment area. Around this Coquitlam Central Station, the top three venue categories are Vietnamese restaurant, coffee shop and clothing store.

<sup>1</sup> <https://developer.translink.ca/servicesgtfs/gtfsdata>

The venue database assembles a total of 1,520 unique venues in 236 categories surrounding the rapid transit stations in Metro Vancouver. On average, each rapid transit station captures 30 venues.



## 2.3 Transit Ridership

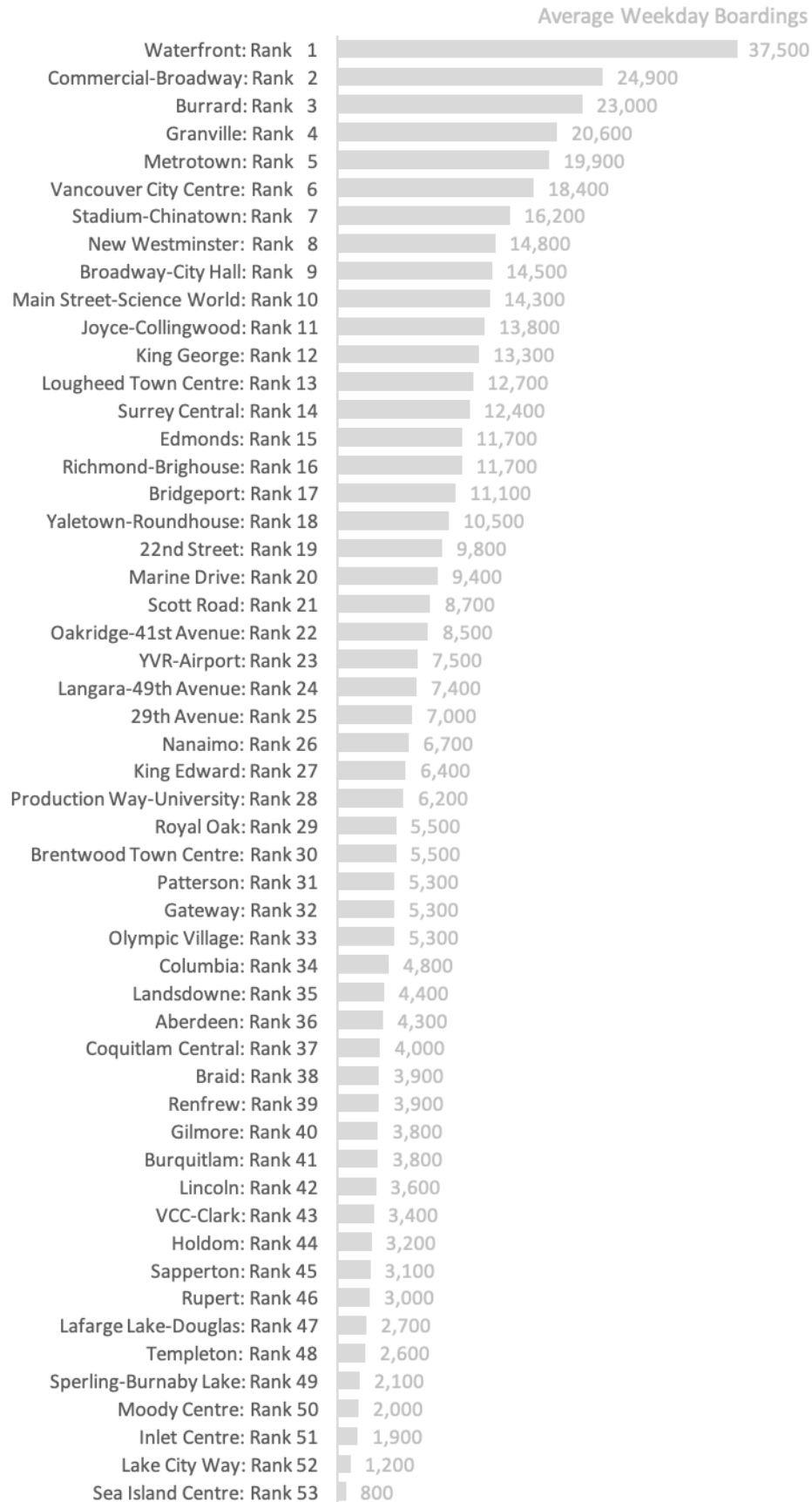
The transit ridership is measured by average daily boarding at 53 rapid transit station – the data was published by TransLink annually in the Transit Service and Performance Review Report<sup>2</sup>.

The Waterfront Station (rank 1), with an average weekday boardings of 37,500 people, is busiest station in the region. The station is located near tourist attractions such as Gas Town and Canada Place in the city centre. Vancouver cruise ship terminal and the seaplane airport are just steps away. In addition, it is a transportation hub that houses the SkyTrain, SeaBus and WCE terminals.

In contrast, Sea Island Centre Station (rank 53) attracted less than 1,000 boardings a day. The main travel market captured by this station is the Vancouver Airport Staff who park their cars in the staff parking lot nearby and took the free SkyTrain service between the parking lot and the airport.

This project will focus on the rank of the rapid transit stations and explores the factors contributing to a station cluster's ridership performance.

<sup>2</sup> [https://www.translink.ca/-/media/translink/documents/plans-and-projects/managing-the-transit-network/tspr/2017\\_tspr\\_summary.pdf](https://www.translink.ca/-/media/translink/documents/plans-and-projects/managing-the-transit-network/tspr/2017_tspr_summary.pdf) (Page 30/31)



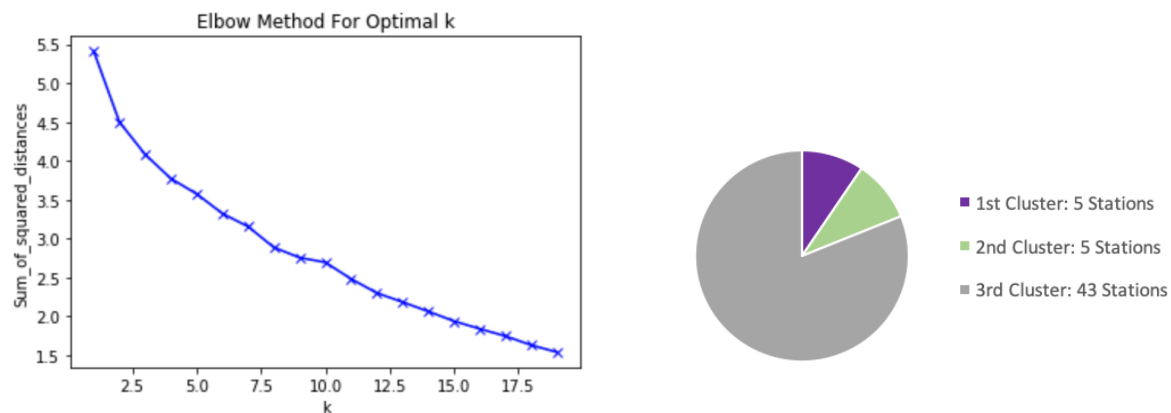
### 3 Methodology

In this project we will disseminate the rapid transit stations based on the types of venues surrounding the station and examine the station category's ridership performance. In the first step, we have gathered the station ridership ranking and a database the Foursquare venues within the station catchment area. The second step is creating clusters of rapid transit stations surrounded by similar type of venues. We will limit the venue types to ten most common ones in the catchment area. The third step is examining the correlation results: we will label the clusters on the ridership ranking chart and analyze their geographic distribution on the map.

#### 3.1 Creating Rapid Transit Station Clusters

For each station, we use one-hot encoding to convert the venue database into a quantitative attribute describing one venue occurrence. The venue frequencies are normalized at each station and only the top ten most common venue types are retained for the analysis.

We then use the elbow method to find the optimal number of clusters. The elbow is not obvious in this case – it seems to be at  $k=2$ , or you may argue it is at  $k=8$ . We picked three clusters for the initial analysis, because the incremental error being eliminated by additional cluster is very even. This suggests the additional clusters are capturing very small number of stations, which might make the cluster prone to the outlier venue types and making the result less meaningful. This results in two clusters of five rapid transit stations and one cluster of 43 stations.



#### 3.2 The First Cluster

The keywords in the first cluster are bus station or bus stop as the most common venue. These stations are located near the park, industrial area or residential neighbourhood with very low density; hence there is not much commercial activities around the station. The ridership at these five stations would depend on the bus routes feeding into the stations.

	stop_name	Cluster Labels	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
8880	Sperling-Burnaby Lake Station	1	Metro Station	Bus Station	Ethiopian Restaurant	Food	Flower Shop	Flea Market	Fish Market	Fish & Chips Shop	Financial or Legal Service	Filipino Restaurant
8894	Nanaimo Station	1	Bus Station	Bus Stop	Bike Trail	Park	Bus Line	Metro Station	Yoga Studio	Farmers Market	Food	Flower Shop
8895	29th Avenue Station	1	Bus Stop	Park	Metro Station	Plaza	Flower Shop	Bus Station	Bar	Fabric Shop	Flea Market	Fish Market
8901	22nd Street Station	1	Bus Stop	Sandwich Place	Vietnamese Restaurant	Bus Station	Light Rail Station	Yoga Studio	Flea Market	Fish Market	Fish & Chips Shop	Financial or Legal Service
8905	Braid Station	1	Bus Station	Bus Stop	Metro Station	Storage Facility	Laser Tag	Fabric Shop	Food	Flower Shop	Flea Market	Fish Market

### 3.3 The Second Cluster

The second clusters include stations surrounded by parks and light rail stations. One thing that makes these five station stands out might be the Ethiopian Restaurant near the station – this type of restaurant is not as common in Vancouver compared to Chinese restaurant or fast food restaurant. Similar to the first cluster, these restaurants also located in the low-density part of the city.

	stop_name	Cluster Labels	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
8866	Langara-49th Avenue Station	2	Park	Bus Stop	Track	Light Rail Station	Golf Course	Yoga Studio	Ethiopian Restaurant	Flea Market	Fish Market	Fish & Chips Shop
8868	King Edward Station	2	Chinese Restaurant	Bus Stop	Coffee Shop	Light Rail Station	Juice Bar	Dessert Shop	Bubble Tea Shop	Park	Financial or Legal Service	Filipino Restaurant
8881	Lake City Way Station	2	Sporting Goods Shop	Park	Bus Stop	Light Rail Station	Yoga Studio	Ethiopian Restaurant	Flower Shop	Flea Market	Fish Market	Fish & Chips Shop
8900	Edmonds Station	2	Park	Coffee Shop	Gym / Fitness Center	Bus Station	Garden	Trail	Café	Fast Food Restaurant	Filipino Restaurant	Ethiopian Restaurant
8906	Scott Road Station	2	Sporting Goods Shop	Lounge	Light Rail Station	Sandwich Place	Yoga Studio	Ethiopian Restaurant	Flower Shop	Flea Market	Fish Market	Fish & Chips Shop

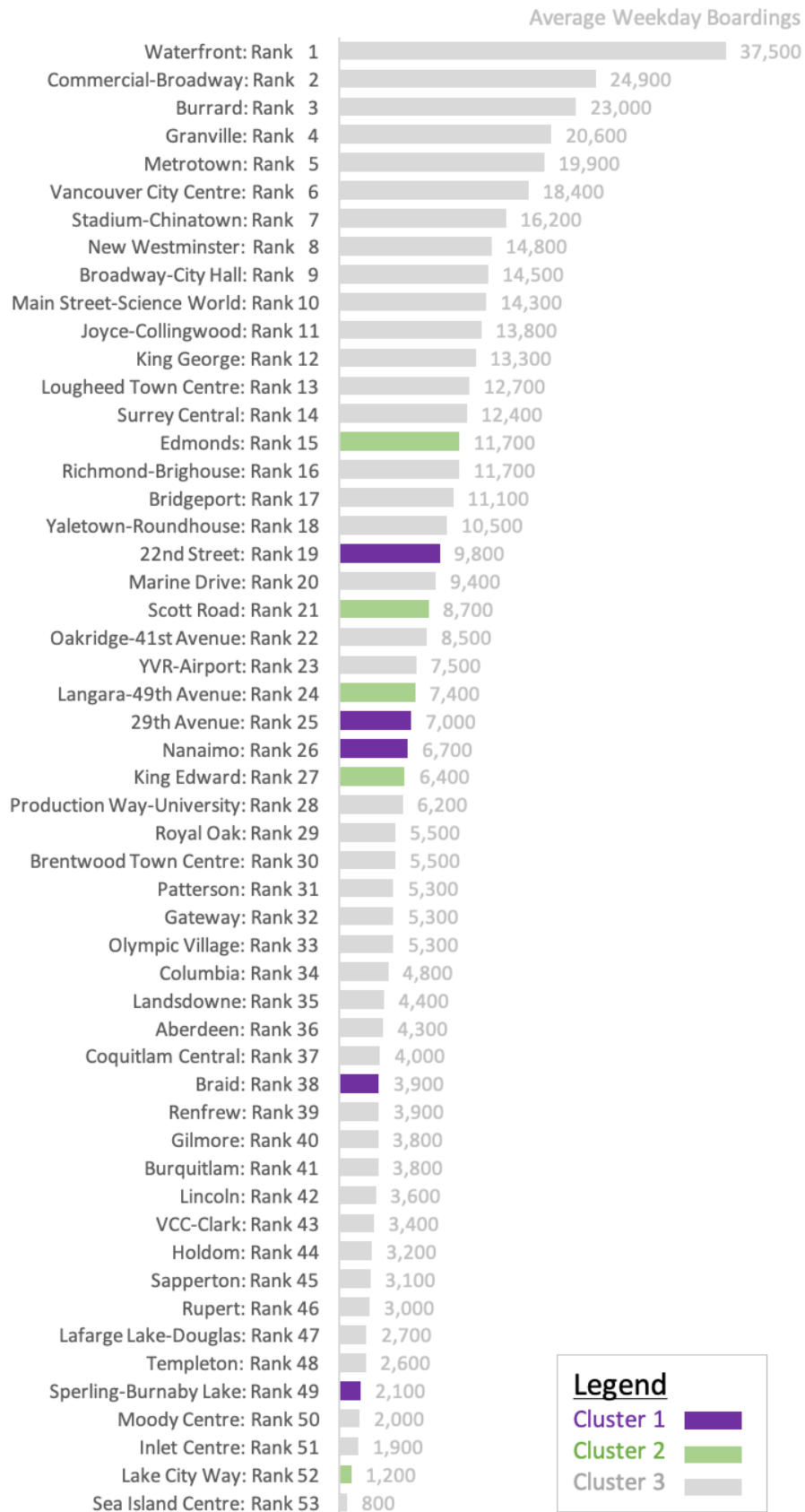
### 3.4 The Last Cluster

The last cluster of 43 stations is essentially “everything else”. The surrounding venue characteristics are dominated by the coffee shops, various stores and restaurants. There are significant commercial activities within the walking distance of the station.

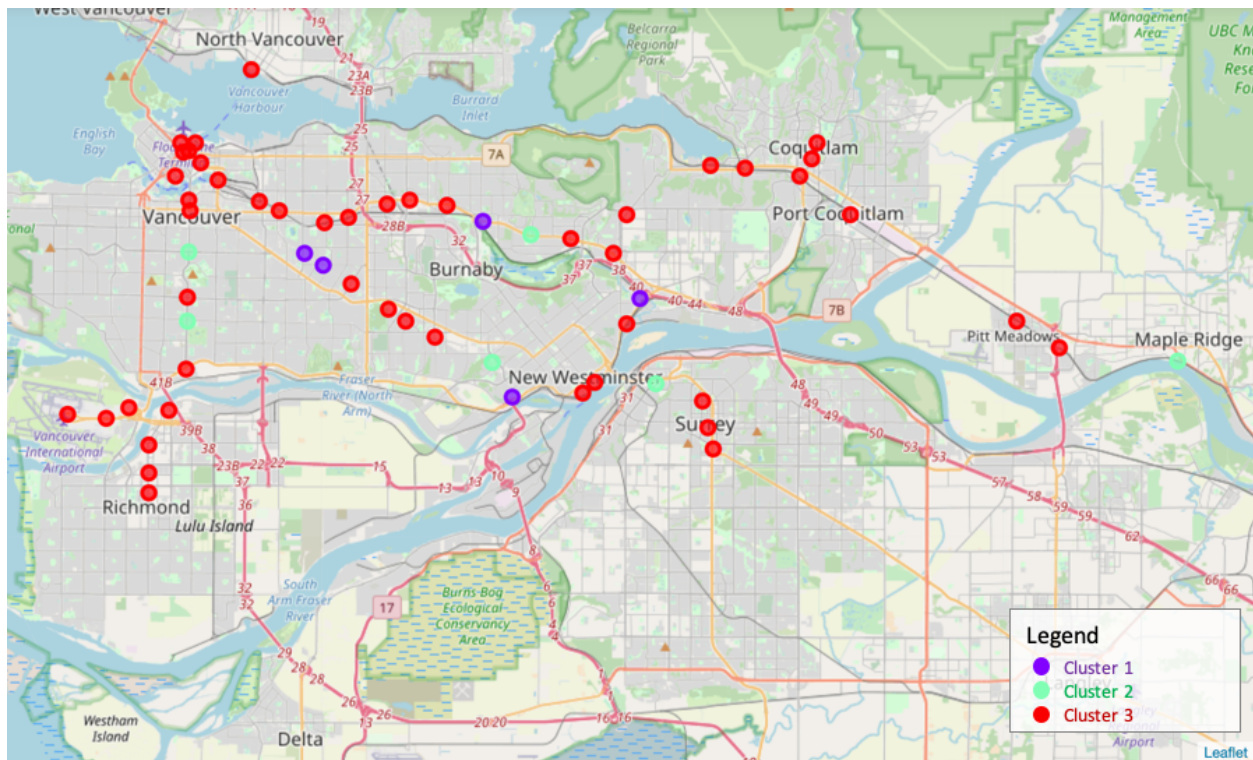
## 4 Results

The three rapid transit station clusters are labelled onto the station ridership chart below. The first cluster captures stations from rank 19 to 49 and the second cluster captures stations from rank 15 to 52; the wide range of station ranking suggests there is little correlation between the type of the business near the station and the ridership ranking. The stations located in less-populated area can also attract boardings from kiss-and-ride and feeder bus routes.





The stations in the first two clusters are scattering around the transit systems. There is no obvious pattern based in the geographic distribution, but it is safe to point out they are all located outside the Metro Vancouver's urban centres – this means the business is less concentrated in these area hence making the two clusters stand out.



## 5 Discussion

While we are hopeful that the clustering algorithm can help us highlighting a cluster of stations with exceptionally good or poor ridership, the analysis results let us down. The result shows there is no correlation between the types of business near the rapid transit station and the ridership performance.

The purpose of the transit system is moving people from A to B, hence both sides of the trip ends are important to the riders. By clustering the venue characteristics to one station, we are ignoring the activities on the other end of the trips. This suggests the transit oriented development policy that bringing in business around transit station will not necessarily stimulate the ridership growth.

## 6 Conclusion

The purpose of this project is to explore the impact of venue types surrounding a rapid transit station on the ridership. The analysis shows there is no correlation between the types of the business nearby and the station ridership performance. This project suggests the transit oriented development does not mean immediate ridership growth. The business near the



transit station might be an intermediate stop but not an trip end – the transit agency should put more weights on trip ends (home/work locations)