



Predicting Toronto Subway Stations for Coffee Shop Expansion

Bradley Jones

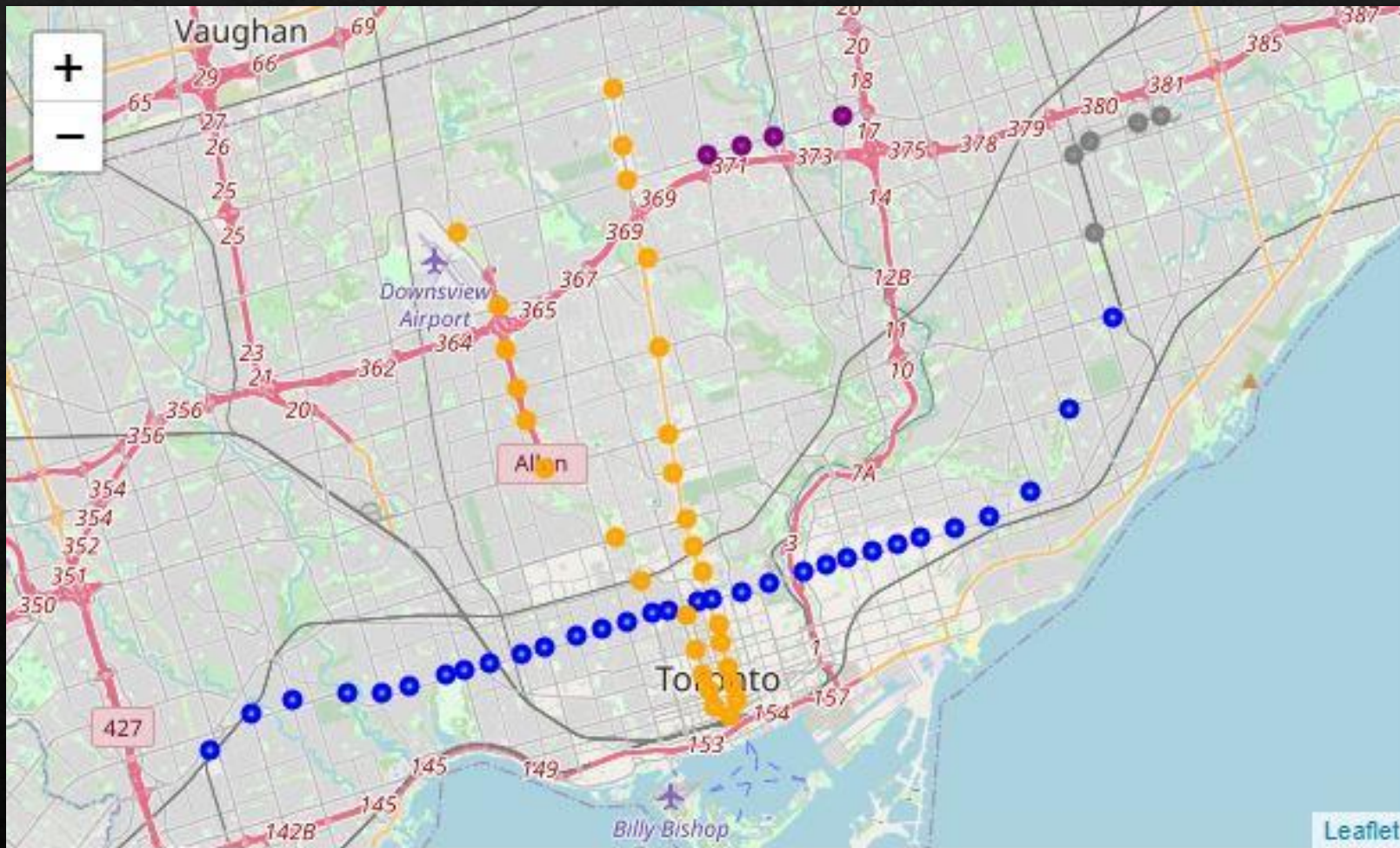
Predicting Expansion Locations is Valuable for Stakeholders in the Coffee Shop Business

- ◈ Expanding to an improper location where coffee shops are not very popular can have great financial consequences on the business
- ◈ Missing out on the opportunity to expand to a good location can have financial consequences as well
- ◈ Expanding to a good location can provide major financial benefits to a business

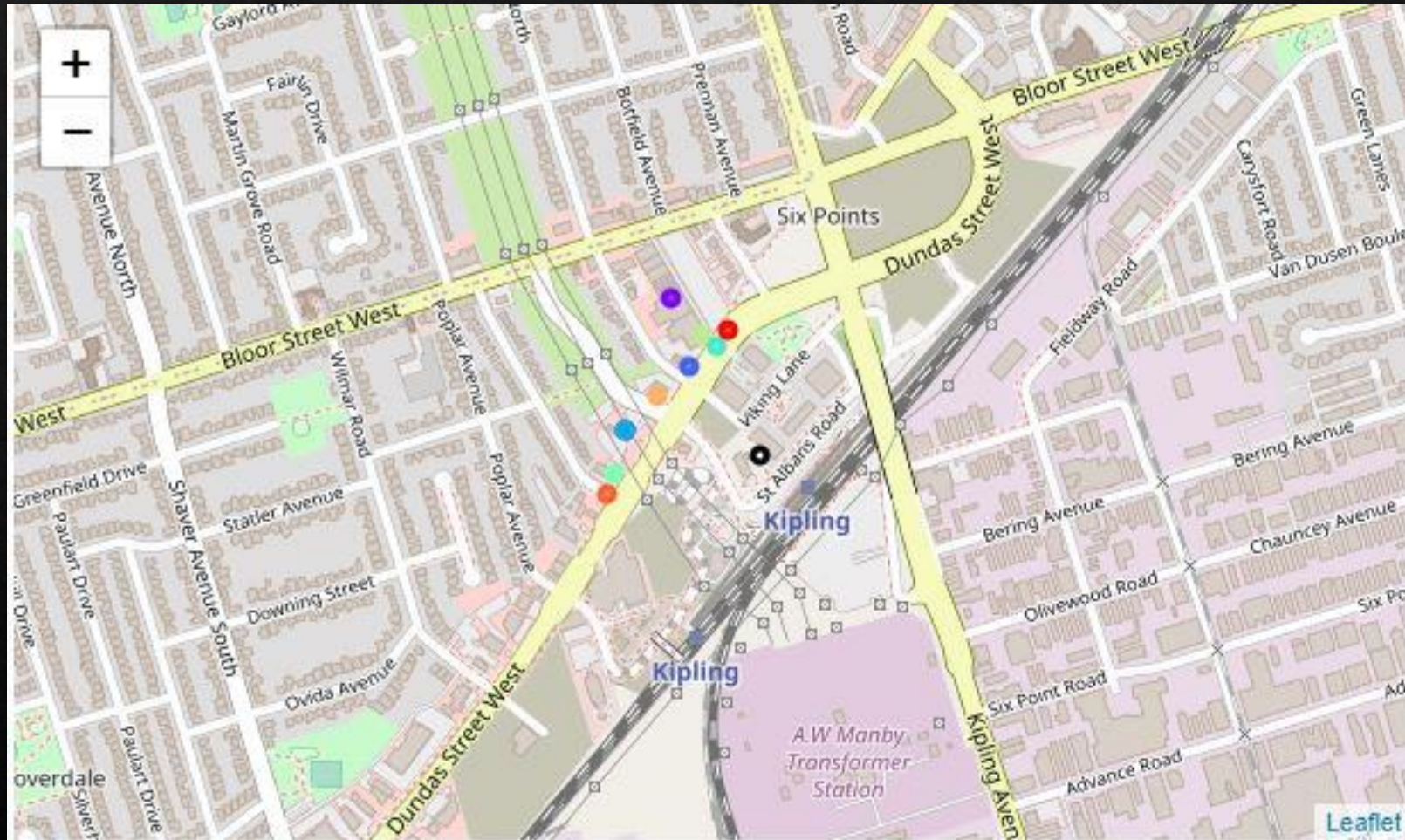
Data Retrieval and Cleaning

- ◆ Subway Station name, latitude and longitude retrieved from <https://scruss.com/blog/2005/12/14/toronto-subway-station-gps-locations/>
- ◆ Venue name, latitude, longitude and category retrieved from Foursquare API <https://developer.foursquare.com/docs/api-reference/venues/explore/>
- ◆ Duplicate values were dropped to ensure data integrity

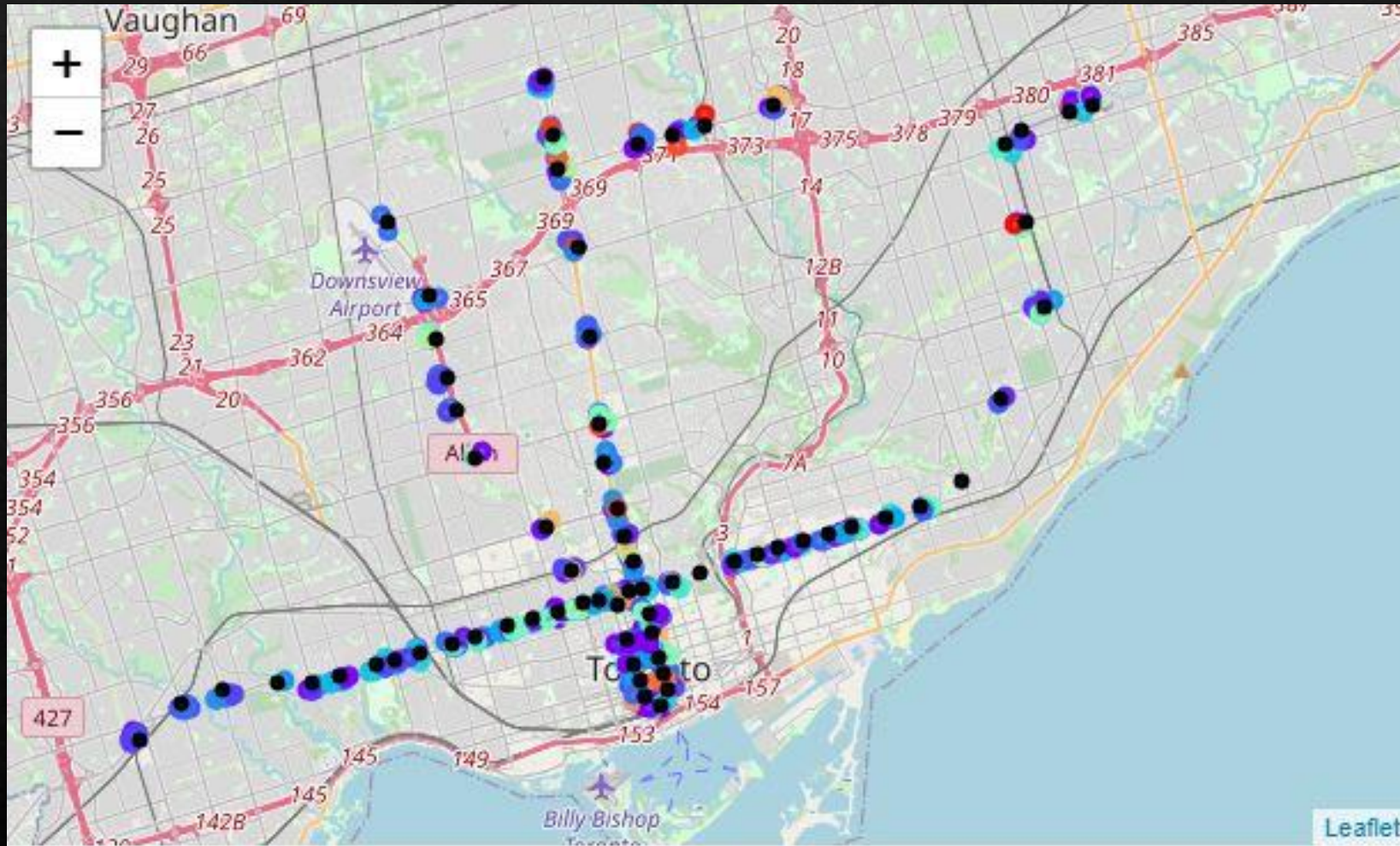
Toronto Subway Lines



Popular Categories Near First Station



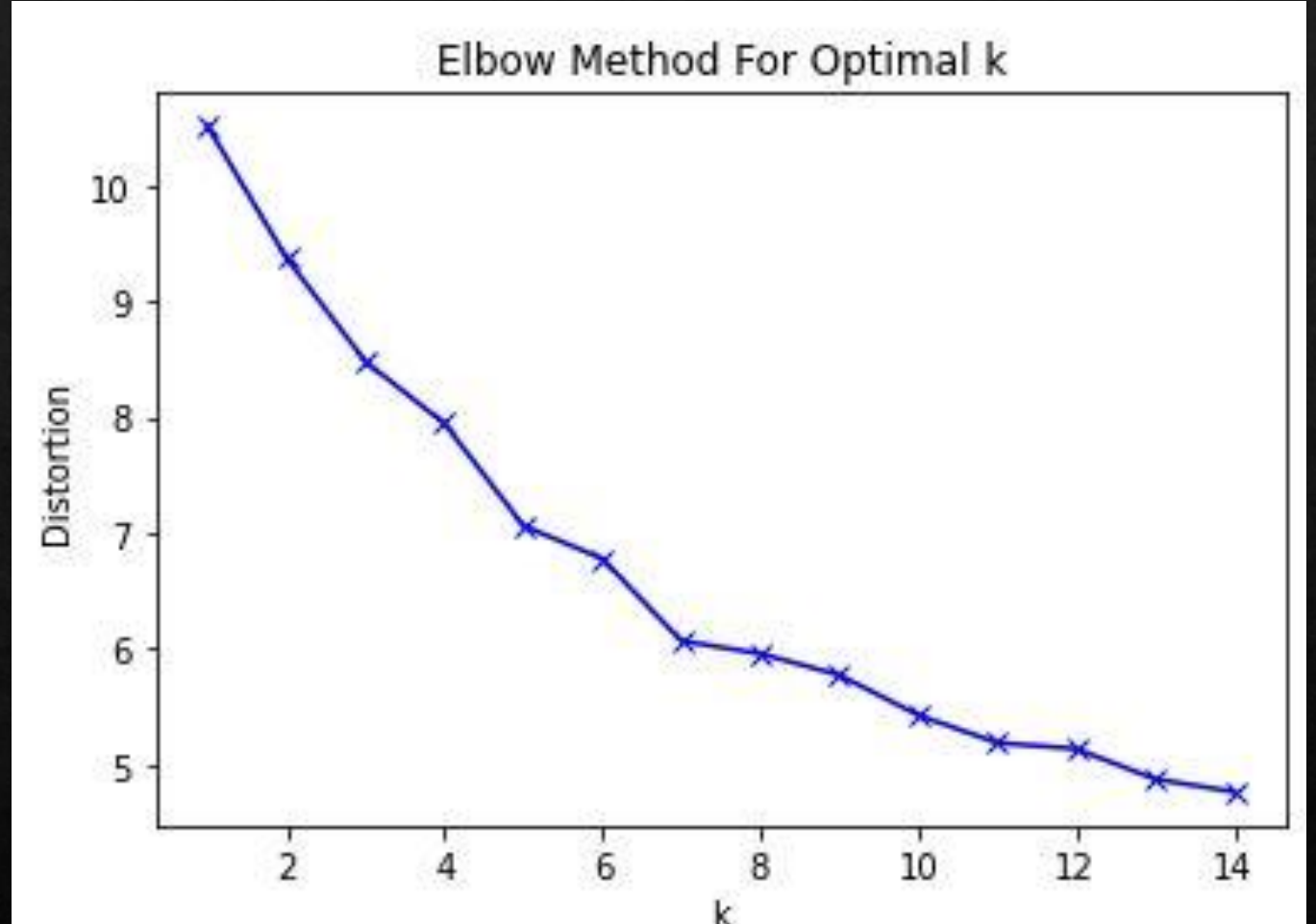
Popular Categories Near All Stations



25 Most Popular Venue Categories Near Toronto Subway Stations

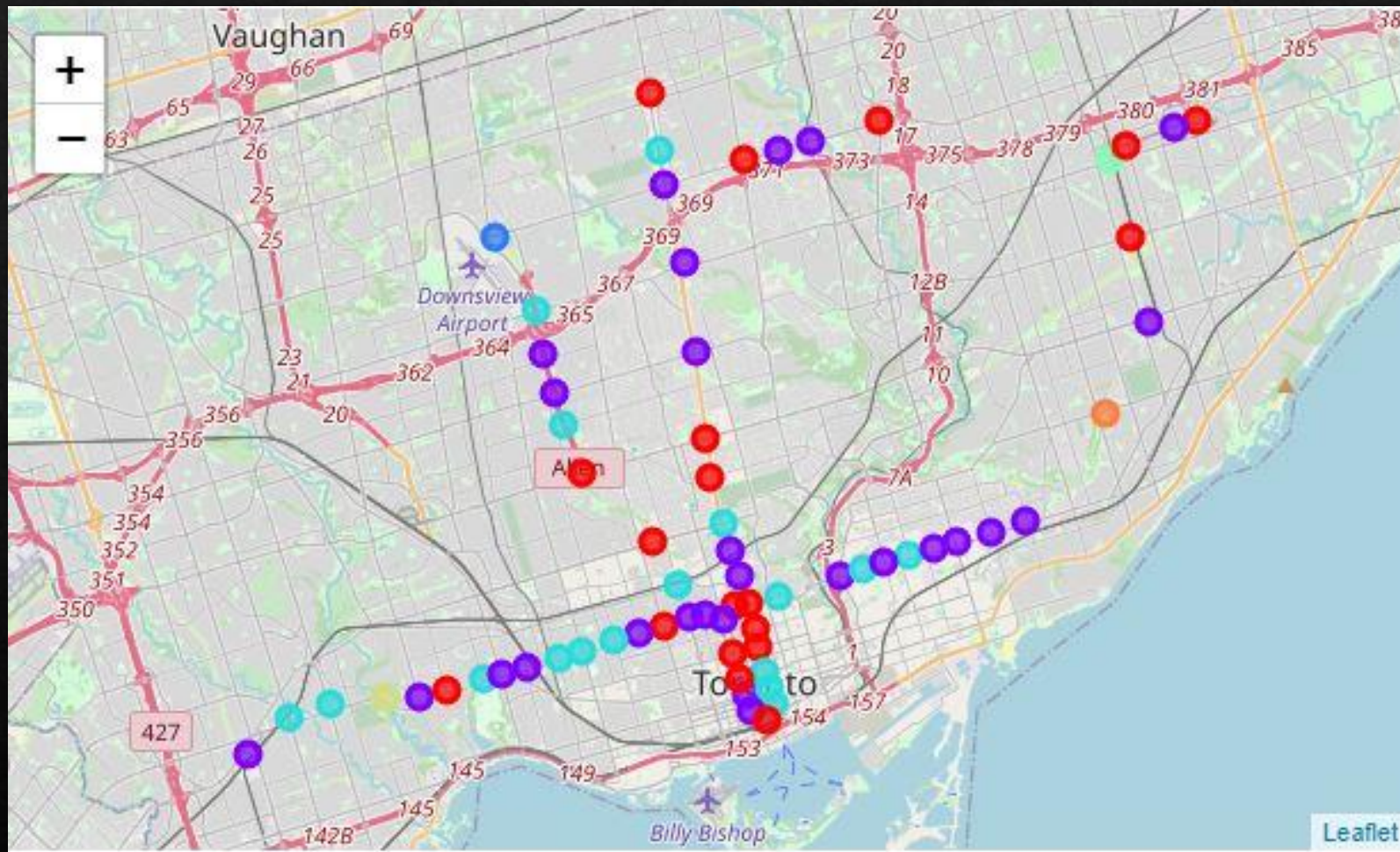
```
----Most Popular Food Venues at Subway Station----
venue  freq
0      Coffee Shop  0.18
1      Café         0.06
2      Fast Food Restaurant  0.05
3      Pizza Place  0.05
4      Restaurant   0.04
5      Bakery       0.04
6      Italian Restaurant  0.03
7      Sandwich Place  0.03
8      Chinese Restaurant  0.02
9      Grocery Store  0.02
10     Japanese Restaurant  0.02
11     Bubble Tea Shop  0.02
12     Sushi Restaurant  0.02
13     Thai Restaurant  0.02
14     Fried Chicken Joint  0.02
15     Food Court      0.02
16     Asian Restaurant  0.01
17     Gastropub       0.01
18     Ice Cream Shop  0.01
19     Indian Restaurant  0.01
20     American Restaurant  0.01
21     Mexican Restaurant  0.01
22     Juice Bar       0.01
23     Korean Restaurant  0.01
24     Mediterranean Restaurant  0.01
```


Elbow Method for K-Means Clustering



Elbow method indicating a K value of 7

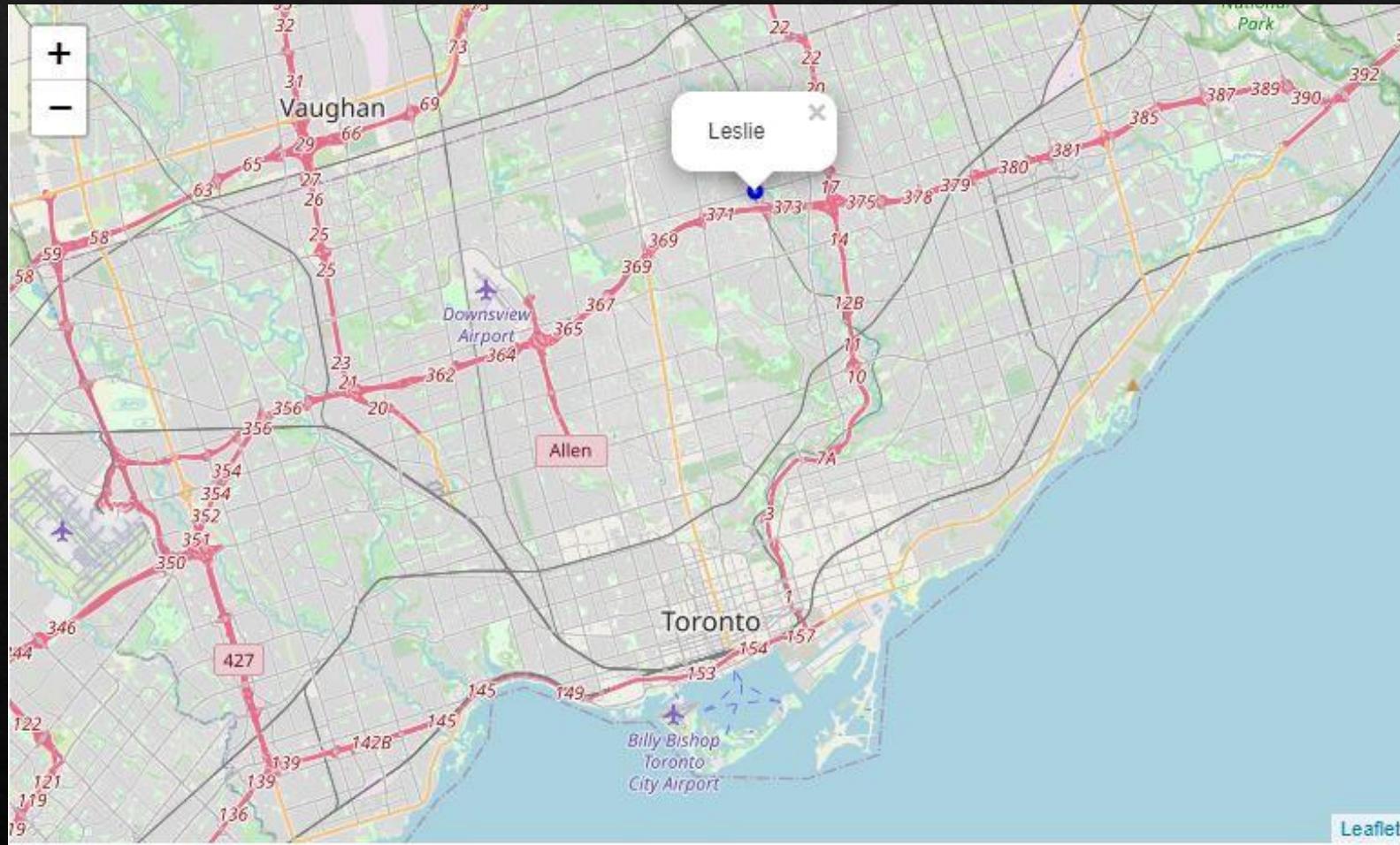
Toronto Subway Stations Grouped by Cluster



Cluster Analysis

- ◆ Upon analyzing the clusters some conclusions could be made:
 - Cluster 1 is already dominated by coffee shops and can be ignored
 - Clusters 3, 5, 6 and 7 only contain outliers and can also be ignored
 - Cluster 2's average rank for coffee shops is better than Cluster 4's average rank
 - Cluster 2 will be the focused cluster
- ◆ Drop stations from Cluster 2 that already contain a coffee shop and map results

Cluster 2 Stations Without a Coffee Shop



Conclusions

- ◆ Built model that clusters subway stations in Toronto based on the popularity of nearby food venues
- ◆ I recommend that a coffee shop business expands to Leslie Station
- ◆ Future modifications include:
 - Adjust the radius value to canvas a wider area
 - Recommend location for different venue type such as Pizza Place