Comparison of Toronto to Vancouver Neighborhoods

Daniel Easley

September 29, 2019

1   Introduction

1.1   Background

In this project we will compare two Canadian cities that are highly rated in the Global Livability Ranking study: Toronto and Vancouver. These cities are 2090 miles (3364 km) apart but are ranked as (6th) Vancouver and (7th) Toronto, by the Economist Intelligence Unit (EIU).

Table 1 City Comparison

City	Location	Population (2016 census)	Ethnic Diversity (2016 census)
Vancouver	British Columbia (Pacific Time Zone)	631,486	48.9% European (white) **19.6% Chinese** 12.0% South Asian
Toronto	Ontario (Eastern Time Zone)	2,731,571	47.7% European (white) 12.6% South Asian **11.1% Chinese**

In the 1990's Vancouver had a surge of Chinese immigrants because there was concern as to the 1997 handover of Hong Kong back to China. The NPR article: **Vancouver Has Been Transformed By Chinese Immigrants**, by Jackie Northam, (https://www.npr.org/2019/06/05/726531803/vancouver-has-been-transformed-by-chinese-immigrants) provides more insight into the impacts of the Chinese immigration and how it has affected Vancouver.

1.2   Problem

Given the surge in the Chinese immigrant population, can we define a difference between the Vancouver and Toronto venues each city contains?

1.3   Interest

The restaurant industry may need to shift their products to cater to a more focused Chinese population.

2   Data acquisition and preparation

2.1   Data Sources

The data sources are as follows:

1. The livability rankings are used as part of the reasoning for the comparison.
   Global Livability Ranking: https://en.wikipedia.org/wiki/Global_Liveability_Ranking

2. Population and Ethnic Diversity are used to show the large increase in Chinese immigrants to Vancouver: https://en.wikipedia.org/wiki/Demographics_of_Toronto and https://en.wikipedia.org/wiki/Demographics_of_Metro_Vancouver . Both sources are based on the 2016 Canadian census figures.
3. We will again use postal codes for the centers of the neighborhoods. These will be the center point for the Foursquare queries:
   Toronto: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
   Vancouver: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_V
4. The postal code latitude and longitude will be retrieved by queries to **geocode**, using the provider ArcGis.
5. All venue data will be sourced from the Foursquare REST API.

2.2 Data Cleaning

The Toronto postal code data will be limited in several ways:

1. If the Borough is 'Not assigned', the data will be dropped.
2. If the Neighborhood is 'Not assigned', the Neighborhood will be set to the Borough name.
3. We will focus on the subset of postal codes where the Borough name contains 'Toronto'.
4. All Neighborhoods will be grouped under a single postal code.

The Vancouver postal code data will also be limited:

1. We will limit the postal codes to the V5x and V6x subsets. This will keep us focused on central Vancouver.
2. The data already contains grouped neighborhood data. For multiple neighborhoods the surrounding parenthesis will be removed.

3 Data Analysis

3.1 Postal Codes

The postal codes will be scraped from the wiki pages listed as Data Sources (#3). The pages are mainly composed of HTML tables. The scraping will use **requests** and **pandas** libraries.

To begin the analysis, we have 38 postal codes for Toronto and 40 postal codes for Vancouver. During our venue queries we find that Vancouver **V5C** does not have venue data and is subsequently dropped.

3.2 Geocoding

Each postal code will use **geocoder** to acquire the latitude and longitude needed for the venue queries. The provider **ArcGis** will be used as the source for the geocoding.

3.3 Venue Information

Venue information will be queried from the **Foursquare REST API**. The query will be limited to 100 venues with a radius of 500 meters.

During the collection of the Vancouver venue data one postal code was not returning data for our specified parameters. The **V5C** postal code was dropped from the analysis.

3.4   Clustering

The Toronto and Vancouver venue information will be clustered separately. This will use a K-Means clustering with 10 clusters each. After clustering, we will look at the largest cluster for each city and compare the mix of venues.
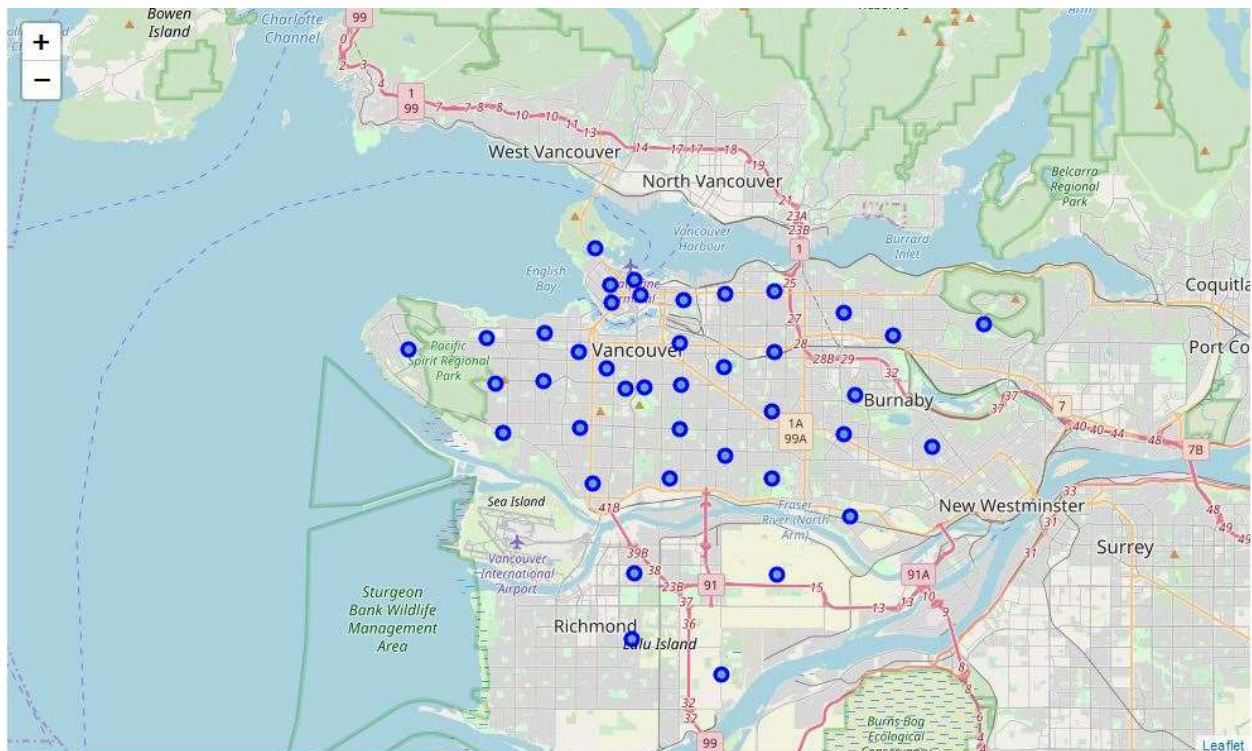
4   Results

4.1   Postal Codes

The maps of the postal code show a similar density within the downtown areas of each city.

Figure 1 Toronto Postal Codes

Figure 2 Vancouver Postal Codes

4.2 Venues

Toronto was able to find 1748 venues. There were 210 distinct venue categories for these venues.

Vancouver was able to find 854 venues. There were 187 distinct venue categories for these venues.

Further analysis of the venue sets was applied to see what venues were in common and unique to each city. The Vancouver unique venues contained several restaurants:

```
Australian Restaurant
Cantonese Restaurant
Filipino Restaurant
Hawaiian Restaurant
Himalayan Restaurant
Japanese Curry Restaurant
Kosher Restaurant
Lebanese Restaurant
Portuguese Restaurant
Shanghai Restaurant
South Indian Restaurant
```

A quick review shows that several could be classified as like 'Chinese Restaurant' or 'Asian Restaurant'. In this analysis we will focus on the 'Chinese Restaurant' venue since it was common to both cities.
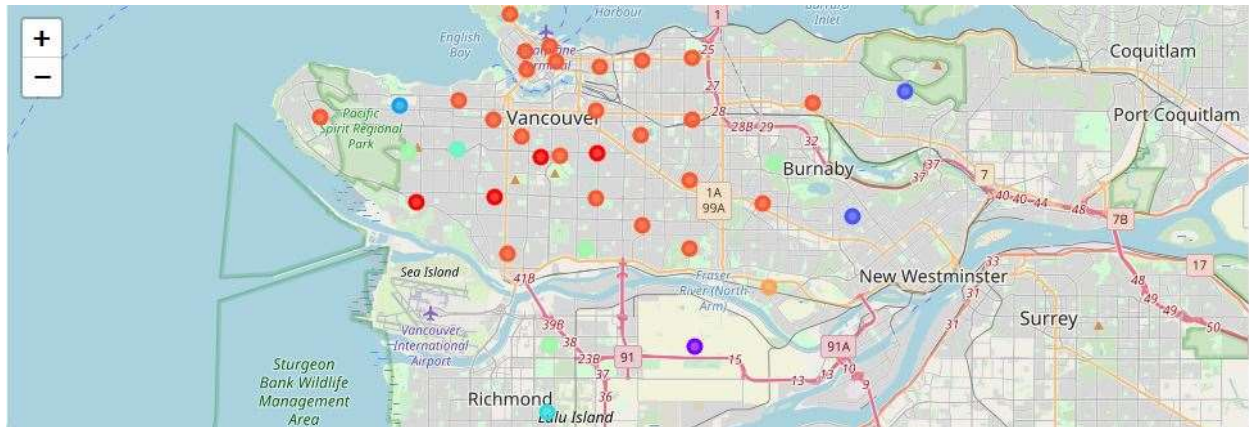
4.3 Clustering

The Toronto and Vancouver both showed the central areas having been clustered together while the fringes of the areas being separated. This is due to the change of the neighborhoods from commercial areas to residential ones. This is evident on the clustering maps.

Figure 3 Toronto Clusters

Figure 4 Vancouver Clusters

Reviewing the largest cluster for each city, we find that these clusters contain many restaurant locations. To compare these clusters, I then focused on the top 4 venues for each cluster to see if they contained a 'Chinese Restaurant'. The results are as follows:

Toronto

> Postal codes within the cluster: **28**
> Boroughs that have 'Chinese Restaurant' in the top 4 most common venues: **0**

Vancouver

> Postal codes within the cluster: **23**
> Boroughs that have 'Chinese Restaurant' in the top 4 most common venues: **7**

Based on these results we can say that the Chinese immigration has heavily affected the restaurant venues in Vancouver. The restaurant industry in Vancouver should shift their product mix towards the Chinese market.