Regional Chinese Restaurant Saturation in Greater Vancouver, Canada

Amos Turin

2020

Introduction

- Vancouver has a large Asian community.

- The 2011 Census found that people of Chinese ethnicity made up 27.7% of the city's population.

- Among the city of Vancouver's thousands of restaurants there is a large proportion of Chinese and other Asian restaurants.

- This analysis aims to assist an entrepreneurial restauranteur to identify a potential location for a new Chinese restaurant within the greater city area.

- The research aims to answer the question: Which region has a large number of restaurants (and hence prospective clients) while having a small number of directly competing Chinese restaurants?

Data

- Regional postal code-based data for the province of British Columbia and Vancouver.

- A dataset of 100 top venues within a 1000 meter proximity of each distinct region within the greater Vancouver area from the Foursquare API
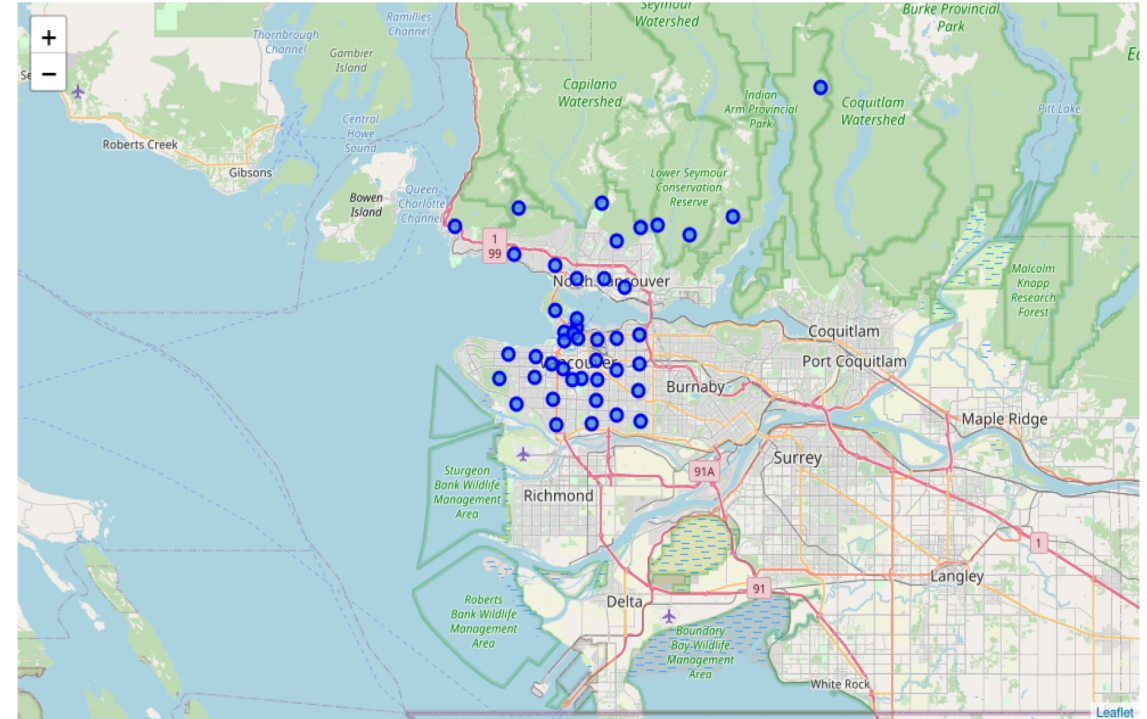

Map of Greater Vancouver, BC, with key regions superimposed in purple markers

Methodology

- Use of K-Means Clustering algorithm. An Unsupervised Learning approach.
- Through Machine Learning, identify patterns and similarities by which city regions can be split into subgroups.
- Key project steps:
  1. Scrape and wrangle region post code dataset and merge with venue data.
  2. One-hot coding to classify venue frequency at the regional level.
  3. Split regions into subgroups of similar characteristics using K-Means Clustering algorithm.

Results

- With K-Means algorithm, where K = 10, 2 largest clusters represent north and south regions of central Vancouver (the Metro area).

- Most common venues in these clusters include restaurants, cafes, and hotels.

- Other clusters appear on city outskirts and less urban areas.

- These clusters display top common venues such as parks, trails, golf courses, and ski areas.
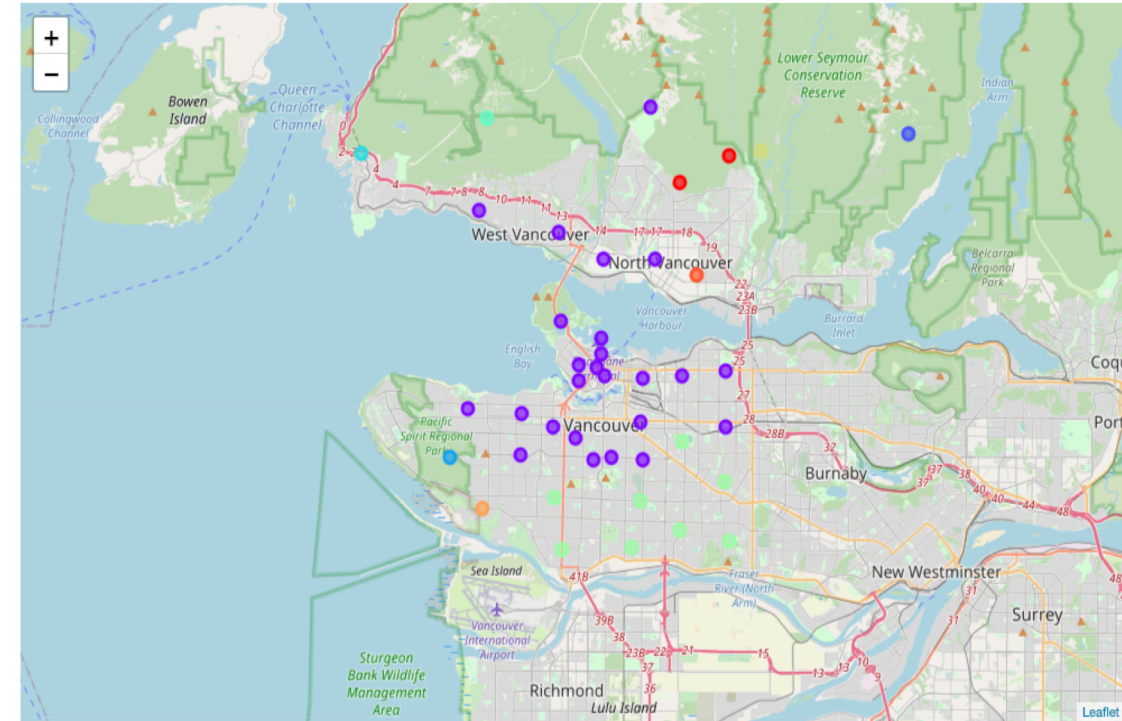
Greater Vancouver with clusters superimposed and color-coded

Discussion

- Results are based on a pre-defined number of K = 10. Additional exploration may identify a more optimal number K.

- Other key factors for each region were ignored and could be accounted for in future research such as:

    1. Commercial rent prices
    2. Average spend on eating out
    3. Local business tax rates
    4. Socio-economic status of residents

Conclusion

- The project successfully narrows potential target regions for a new restaurant to one key subgroup cluster.

- Within the target subgroup, the results also go further and identify regions to avoid due to high saturation of existing Chinese restaurant competitors.

- Prospective entrepreneurs should conduct further due diligence on the remaining target regions taking into account additional factors mentioned above.