# An analysis of optimal retail store location based in Waterfront station, Vancouver

By Hua-Wei Huang

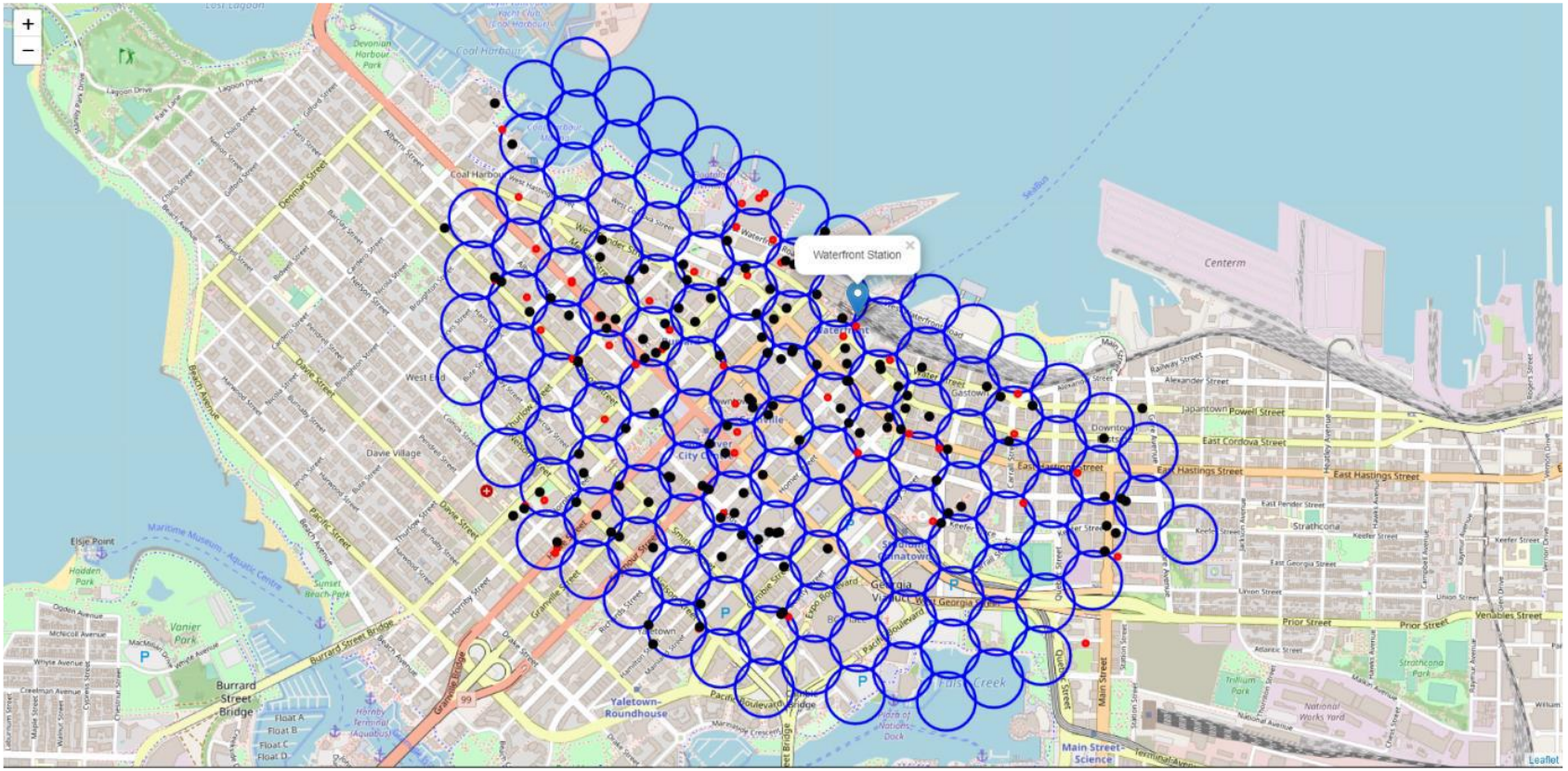
# Introduction

- Problem description
  - Aims at predicting the customer rating of a coffee shop in downtown Vancouver, Canada
  - Investor is interesting in opening a high rating store to increase his/her brand exposure

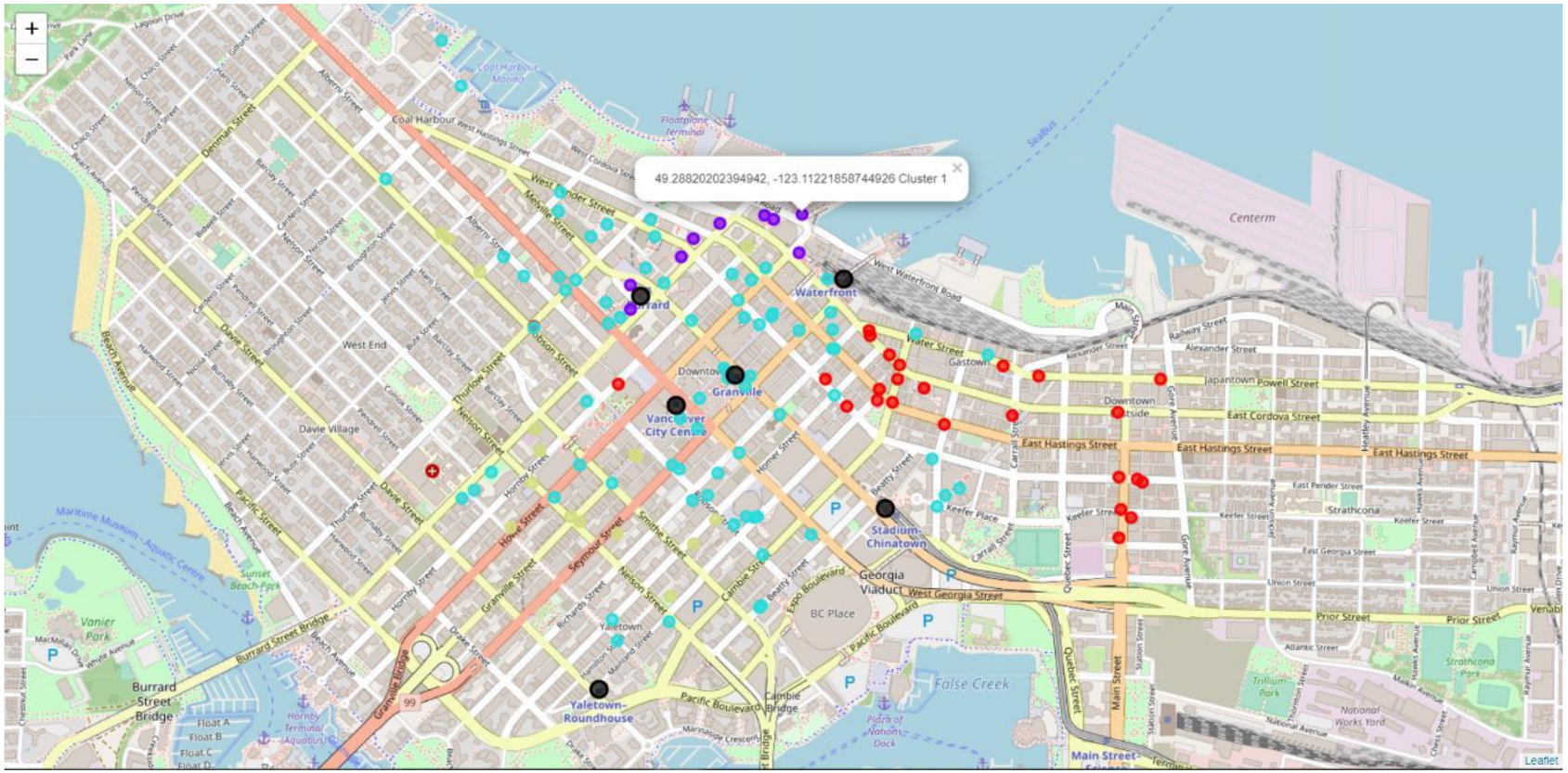
- Data description
  - Google API
  - Foursquare API
  - Statistic Canada
  - censusmapper.ca

# Methodology

- 1. Define neighbourhoods in the targeted area, 1.5km within Waterfront Station, see figure 1.
- 2. Retrieve venue information from Foursquare
- 3. Explore coffee shops through clustering
- 4. Implementing machine learning techniques to predict customer rating
- 5. Perform prediction on the selected potential locations based on the model with the lowest Root Mean Square Prediction Error (RMSE)


Waterfront Station
Centerm
Yaletown-Roundhouse
Burrard Street Bridge
Granville Bridge
Main Street-Science
Leaflet

# Exploratory analysis



# Model comparison

Table 1. Model comparison: prediction error

| | Test RMSE | Training RMSE |
|---|---|---|
| Linear | 0.896 | 0.703 |
| Stepwise | 1.209 | 1.063 |
| LASSO | 0.975 | 0.745 |
| Bagging | 0.883 | 0.339 |
| Boosting | 0.895 | 0.303 |
| Random forest | 0.895 | 0.303 |

# Prediction and potential locations

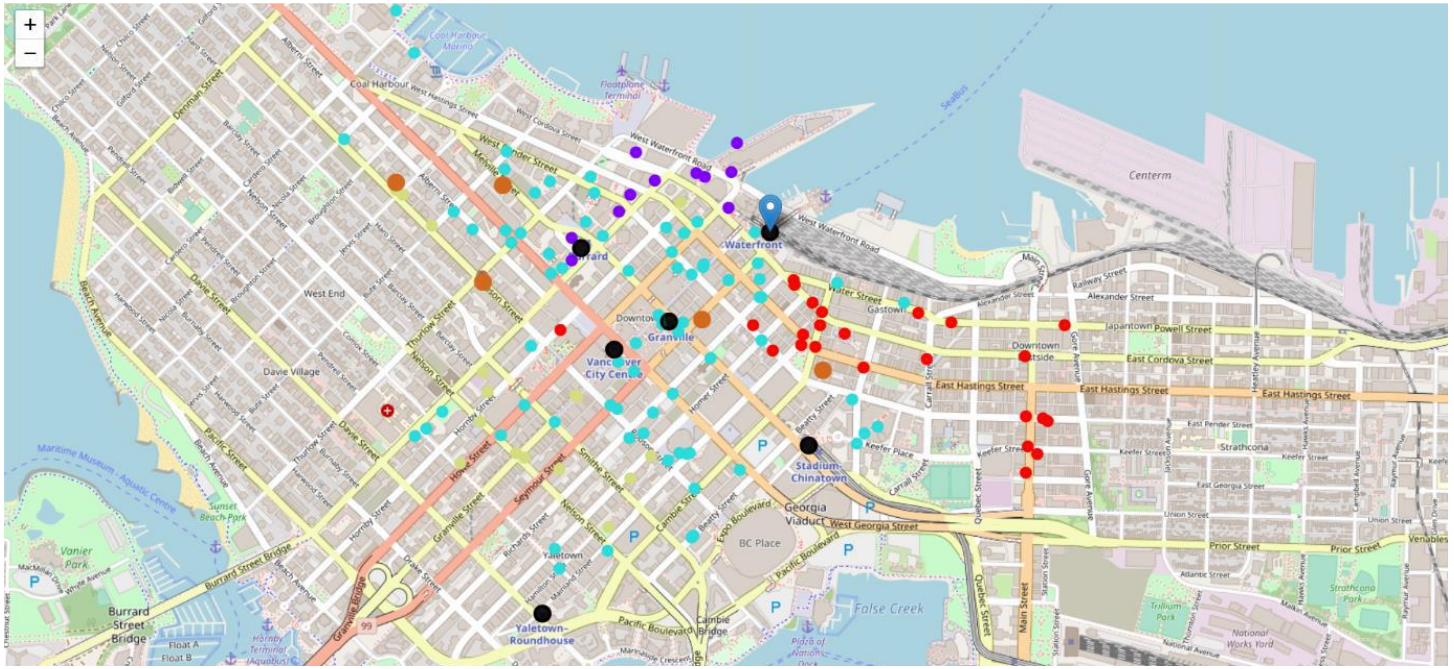



Table 2. Customer rating prediction on potential locations

| | Address | Postal Code | Customer Rating |
|---|---|---|---|
| 1 | 789 Jervis St, Vancouver, BC | V6E 2B1 | 7.34 |
| 2 | 560 Seymour St, Vancouver, BC | V6B 3H7 | 6.62 |
| 3 | 1160 Melville St, Vancouver, BC | V6E 2S8 | 7.07 |
| 4 | 150 W Hastings St, Vancouver, BC | V6B 1R3 | 7.22 |
| 5 | 1098-1008 Robson St, Vancouver, | V6E 1A7 | 7.19 |

# Conclusion and limitation

- The highest prediction customer rating by using bagging model is 7.34
- Certain limitations in this study
  - From a business perspective, an investor may want to know the expected profit or revenue he could earn from opening a new retail store.
- An alternative way of measuring revenue is
  - Using user mobility or traffic flow
  - The time and amount of people staying a an area could potentially transform into spending in stores in that area.