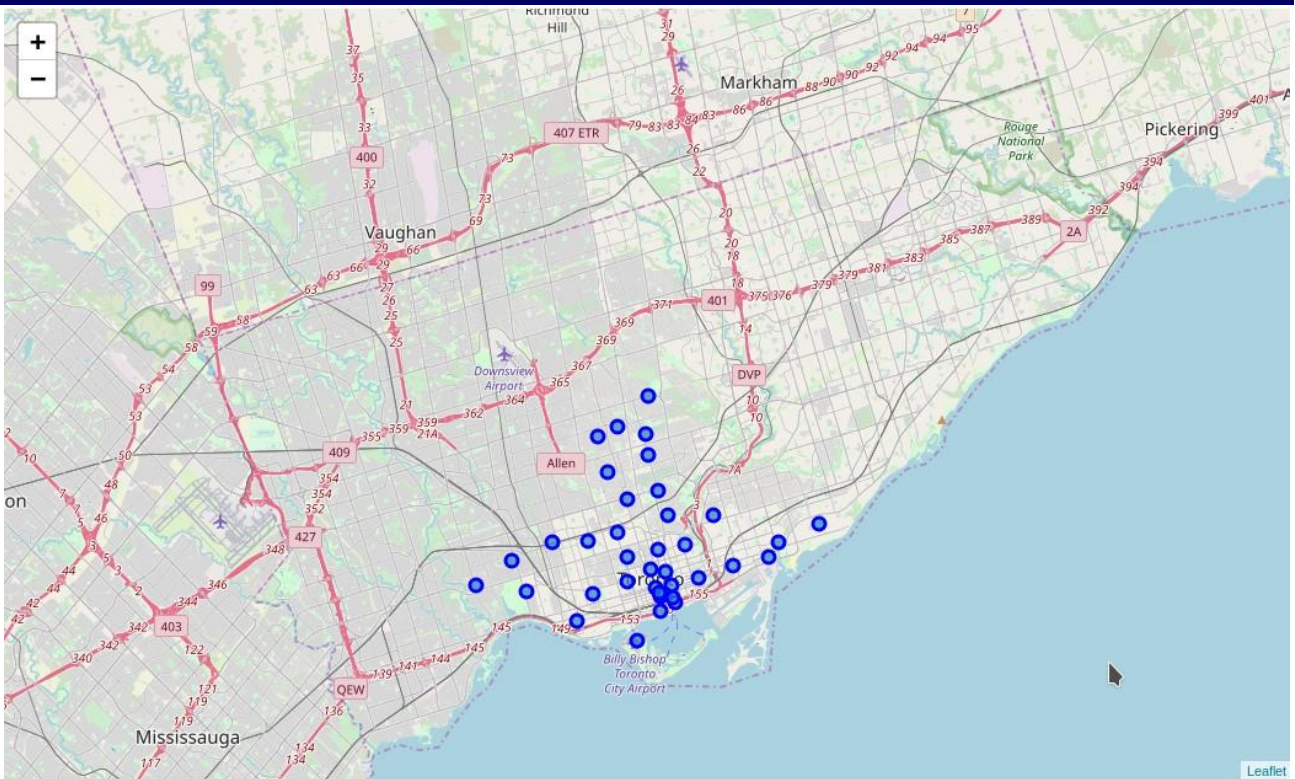


# Selecting a neighbourhood and features of a food venue in Toronto



## COURSERA CAPSTONE WEEK 5

# Contents

---

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Background .....	3
1.2	Problem .....	3
1.3	Interest .....	3
<b>2</b>	<b>Data</b>	<b>3</b>
2.1	Data sources .....	3
2.2	Data cleaning .....	4
2.2.1	Wikipedia table .....	4
2.2.2	Counting venues in a neighbourhood .....	4
2.2.3	Individual venue attributes .....	4
<b>3</b>	<b>Analysis</b>	<b>4</b>
3.1	Selection of candidate neighbourhoods .....	4
3.2	Classification of food venues .....	5
<b>4</b>	<b>Results</b>	<b>6</b>
<b>5</b>	<b>Conclusion</b>	<b>7</b>
5.0.1	Future extensions .....	7

# **1 INTRODUCTION**

## **1.1 BACKGROUND**

The restaurant and coffee shop business is very competitive and highly dynamic. Selecting the right location is a critical first step. But even in a good location, the selection of cuisine and other features can make the difference between success and failure.

One strategy is to provide a familiar cuisine with some feature to brand your venue as new or different to make it stand out.

## **1.2 PROBLEM**

The project is to identify a small number of candidate locations, with the associated cuisines and novel combination of features to maximise interest and traffic through your food venue.

The neighbourhoods in Toronto were examined and the neighbourhoods with the highest density of food venues identified. For the neighbourhood with the highest density of food venues, the cuisine price, outdoor seating and wifi were profiled.

The aim of the analysis is to provide a short list of options to use as a basis for the next decision stage.

## **1.3 INTEREST**

Anyone interested in investing in a new food venue or updating an existing food venue in the Toronto area and looking for good candidate ideas.

# **2 DATA**

## **2.1 DATA SOURCES**

The list of Toronto postcode and the associated neighbourhoods is available from Wikipedia and can be combined with the geolocation of each postcode available from the Week 3 assignment.

With the location of each neighbourhood, the Foursquare API can be used to determine the density of food venues in each neighbourhood and profile the categories (i.e. cuisine) of each food venue.

The Foursquare API can examine each food venue individually and return the attributes (features) price, wifi and outdoor setting.

## **2.2 DATA CLEANING**

### **2.2.1 WIKIPEDIA TABLE**

The table in Wikipedia had several issues to address.

1. Not all postcodes lists were allocated to boroughs, in this case, these postcodes were ignored
2. Some postcodes have boroughs allocated, but no neighbourhood - in this case, the neighbourhood is given the same name as the borough
3. Some postcodes were duplicated and allocated to multiple neighbourhoods - in this case, the neighbourhood names were combined and separated by a comma

### **2.2.2 COUNTING VENUES IN A NEIGHBOURHOOD**

The search API function supports search in a region for only food venues, but it is limited to return a maximum of 50 entries. This means for very high-density suburbs they would all return 50 venues in a given radius with no way to differentiate between them.

To address this the search radius was reduced till a maximum of one neighbourhood would return 50 venues, meaning the density of food venues could be ranked effectively.

### **2.2.3 INDIVIDUAL VENUE ATTRIBUTES**

When collecting the data about individual venues from the Foursquare API, not all attributes (e.g. wifi) were enumerated for each venue. In the case an attribute was not given for a venue, the attribute was considered not present.

Additionally, the venue API calls is a premium API endpoint with only a limited number of calls available in a 24 hour period. To make this project feasible, the number of calls made using this endpoint was kept as low as possible. In the case of a paid access to the Foursquare API more extensive use of the venue API could be made.

### 3 ANALYSIS

#### 3.1 SELECTION OF CANDIDATE NEIGHBOURHOODS

The top five neighbourhoods for food venues are given in table 1.

The neighbourhood Commerce Court, Victoria Hotel was selected for further analysis.

**TABLE 1: TOP FIVE NEIGHBOURHOOD BY NUMBER OF FOOD VENUES**

Neighbourhood	Number of food venues within 200m
Commerce Court, Victoria Hotel	50
Central Bay Street	48
First Canadian Place, Underground city	48
Chinatown, Grange Park, Kensington Market	46
Harbourfront East, Toronto Islands, Union Station	42

#### 3.2 CLASSIFICATION OF FOOD VENUES

The top 3 food venue categories in Commerce Court, Victoria Hotel are given in table 2.

**TABLE 2: TOP 3 FOOD VENUE CATEGORIES IN COMMERCE COURT, VICTORIA HOTEL**

Category	Number of food venues within 200m
Coffee Shop	12
Restaurant	3
Japanese Restaurant	3

The most common version of food venue for each of the categories is given in table

3. The pricing tier is given in four levels, 1 being the cheapest, 4 the most expensive.

**TABLE 3: MOST COMMON VERSION OF FOOD VENUE FOR EACH CATEGORY**

Category	Price	Outdoor seating	Wifi
Coffee Shop	1	Yes	Yes

Restaurant	2	Yes	No
Fastfood	1	Yes	Yes
Restaurant			

When determining the underexploited options with regards to price tier, the tier on either side of the most common price tier for the category was suggested. The higher price tier represents an upmarket, more luxurious version of the standard offering, and the price tier below represents a cheap and cheerful alternative to the standard options. In the case, the price tier is either the lowest (1) or highest (4) only one price tier alternative was presented.

When determining the underexploited option with regards to outdoor seating and wifi, if the majority of venues in a category were offering an option, then not offering the option was presented as an alternative. For example, if wifi was being offered in all venues, not providing wifi could be an alternative - "Unplug from the world". In some cases the number of venues in a category with and without an option were similar. For example, six venues might have wifi and five not. In this situation, both options are well represented and neither could be considered underutilised. So when determining if an option was underexploited or not, a threshold of 35% was applied. That is, if an option was in more 35% of the venues in the category it was not underexploited.

## 4 RESULTS

The optimal neighbourhood for a food venue is Commerce Court, Victoria Hotel.

The options for food venue are given in table 4. The nickname column is a shorthand name to describe the configuration.

**TABLE 4: POTENTIALLY UNDEREXPLOITED ALTERNATIVES IN THE FOOD VENUES IN COMMERCE COURT, VICTORIA HOTEL**

Category	Price	Outdoor seating	Wifi	Nickname
Coffee Shop	1,	No	Yes	Coffee anytime - even when it is raining
Coffee Shop	2	Yes	Yes	Upmarket coffee and treats
Restaurant	1	Yes	No	Cheap and cheerful dining experience

Restaurant	3	Yes	No	Good meals, done right
Fastfood Restaurant	1	Yes	Yes	Everything you expect AND wifi!
Fastfood Restaurant	1	No	No	This place is about two things: food and fast
Fastfood Restaurant	2	Yes	No	Fast food that is cut above the rest

For the restaurant category, there were enough examples of with and without outdoor settings and wifi that it would not be a distinguishing feature, so the only alternative available in this neighbourhood is to vary the price point.

For the coffee shop category, there were few enough venues without outdoor seating to make it a potential feature as well.

In the case of the fast food restaurant category, there is little experimentation with the outdoor seating and wifi features, meaning there are potentially underexploited niches available.

## 5 CONCLUSION

---

In this study, the optimal neighbourhoods for a new food venue were identified, and for the top ranking neighbourhood, the best cuisine and combination of pricing, outdoor seating and wifi were identified.

### 5.0.1 FUTURE EXTENSIONS

The model could be improved with the attributes of more venues included in the analysis, unfortunately because it is a premium call it was possible for this project.

An interesting extension would to segment out trending food venues and identify common elements to those venues. The key challenge in that analysis is the trending results are consistently being updated and a longitudinal analysis would be required to avoid artefacts.