# 1. Description of the problem

Cynthia would like to open a new clothing store in the city of Toronto. More specifically, she is interested in the area of downtown Toronto, so she wants to explore where are located the already existing stores so as to decide accordingly.

# 2. Description of the Data

The data that will be used are:

- **Toronto neighborhood data:**
  The following Wikipedia page (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M) will be used to pull information such as post codes, boroughs and neighborhoods of Canada.

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools
What links here
Related changes
Upload file
Special pages

Article Talk

List of postal codes of Canada: M

From Wikipedia, the free encyclopedia

This is a list of postal codes in Canada where the first letter is M. Postal codes beginning with characters are listed, corresponding to the Forward Sortation Area.

Canada Post provides a free postal code look-up tool on its website,[1] via its applications for s CD-ROMs. Many vendors also sell validation tools, which allow customers to properly match ac offices, and some libraries.

Toronto - FSAs [ edit ]

Note: There are no rural FSAs in Toronto, hence no postal codes should start with M0, howev Mississauga, suggesting that Canada Post may be allocating the M0 FSA for high volume addr

| Postcode | Borough | Neighbourhood |
|---|---|---|
| M1A | Not assigned | Not assigned |
| M2A | Not assigned | Not assigned |
| M3A | North York | Parkwoods |
| M4A | North York | Victoria Village |
| M5A | Downtown Toronto | Harbourfront |

- **Venues in Toronto:** Foursquare API will help us to get venues by category, location and the relevant business names.
  More specifically, venues and location JSON data can be found on Foursquare API call.
  https://foursquare.com

For example:


```
{'meta': {'code': 200, 'requestId': '5e62584840a7ea001b97ed28'},
 'response': {'suggestedFilters': {'header': 'Tap to show:',
   'filters': [{'name': 'Open now', 'key': 'openNow'}]},
  'headerLocation': 'Bay Street Corridor',
  'headerFullLocation': 'Bay Street Corridor, Toronto',
  'headerLocationGranularity': 'neighborhood',
  'totalResults': 128,
  'suggestedBounds': {'ne': {'lat': 43.6586737045, 'lng': -79.37460365419369},
   'sw': {'lat': 43.6496736955, 'lng': -79.38701963607448}},
  'groups': [{'type': 'Recommended Places',
    'name': 'recommended',
    'items': [{'reasons': {'count': 0,
       'items': [{'summary': 'This spot is popular',
         'type': 'general',
         'reasonName': 'globalInteractionReason'}]},
      'venue': {'id': '57eda381498ebe0e6ef40972',
       'name': 'UNIQLO ユニクロ',
       'location': {'address': '220 Yonge St',
        'crossStreet': 'at Dundas St W',
```


# 3. Methodology

- Data pulling of Toronto neighborhood data from Wikipedia using Beautiful soup package.
- Convert data pulled to a data frame.

| | Postcode | Borough | Neighbourhood |
|---|---|---|---|
| 2 | M3A | North York | Parkwoods |
| 3 | M4A | North York | Victoria Village |
| 4 | M5A | Downtown Toronto | Harbourfront |
| 5 | M6A | North York | Lawrence Heights |
| 6 | M6A | North York | Lawrence Manor |

- Create a new data frame based on the previous one, resetting the index and grouping by post code and borough.

| | Postcode | Borough | Neighbourhood |
|---|---|---|---|
| 0 | M3A | North York | Parkwoods |
| 1 | M4A | North York | Victoria Village |
| 2 | M5A | Downtown Toronto | Harbourfront |
| 3 | M6A | North York | Lawrence Heights, Lawrence Manor |
| 4 | M7A | Downtown Toronto | Queen's Park |
| 5 | M9A | Etobicoke | Islington Avenue |
| 6 | M1B | Scarborough | Rouge, Malvern |
| 7 | M3B | North York | Don Mills North |

- Create a new data frame keeping only borough that is downtown Toronto and resetting index.

| | Postcode | Borough | Neighbourhood |
|---|---|---|---|
| 0 | M5A | Downtown Toronto | Harbourfront |
| 1 | M7A | Downtown Toronto | Queen's Park |
| 2 | M5B | Downtown Toronto | Ryerson, Garden District |
| 3 | M5C | Downtown Toronto | St. James Town |
| 4 | M5E | Downtown Toronto | Berczy Park |

- Find latitude and longitude of Downtown Toronto

```
The geograpical coordinate of Downtown Toronto are 43.6541737, -79.38081164513409.
```

- Call foursquare API and get JSON data

```
{'meta': {'code': 200, 'requestId': '5e62bfc9006dce001bce2f60'},
 'response': {'suggestedFilters': {'header': 'Tap to show:',
   'filters': [{'name': 'Open now', 'key': 'openNow'}]},
  'headerLocation': 'Bay Street Corridor',
  'headerFullLocation': 'Bay Street Corridor, Toronto',
  'headerLocationGranularity': 'neighborhood',
  'totalResults': 128,
  'suggestedBounds': {'ne': {'lat': 43.6586737045, 'lng': -79.37460365419369},
   'sw': {'lat': 43.6496736955, 'lng': -79.38701963607448}},
  'groups': [{'type': 'Recommended Places',
    'name': 'recommended',
    'items': [{'reasons': {'count': 0,
       'items': [{'summary': 'This spot is popular',
         'type': 'general',
         'reasonName': 'globalInteractionReason'}]}
```

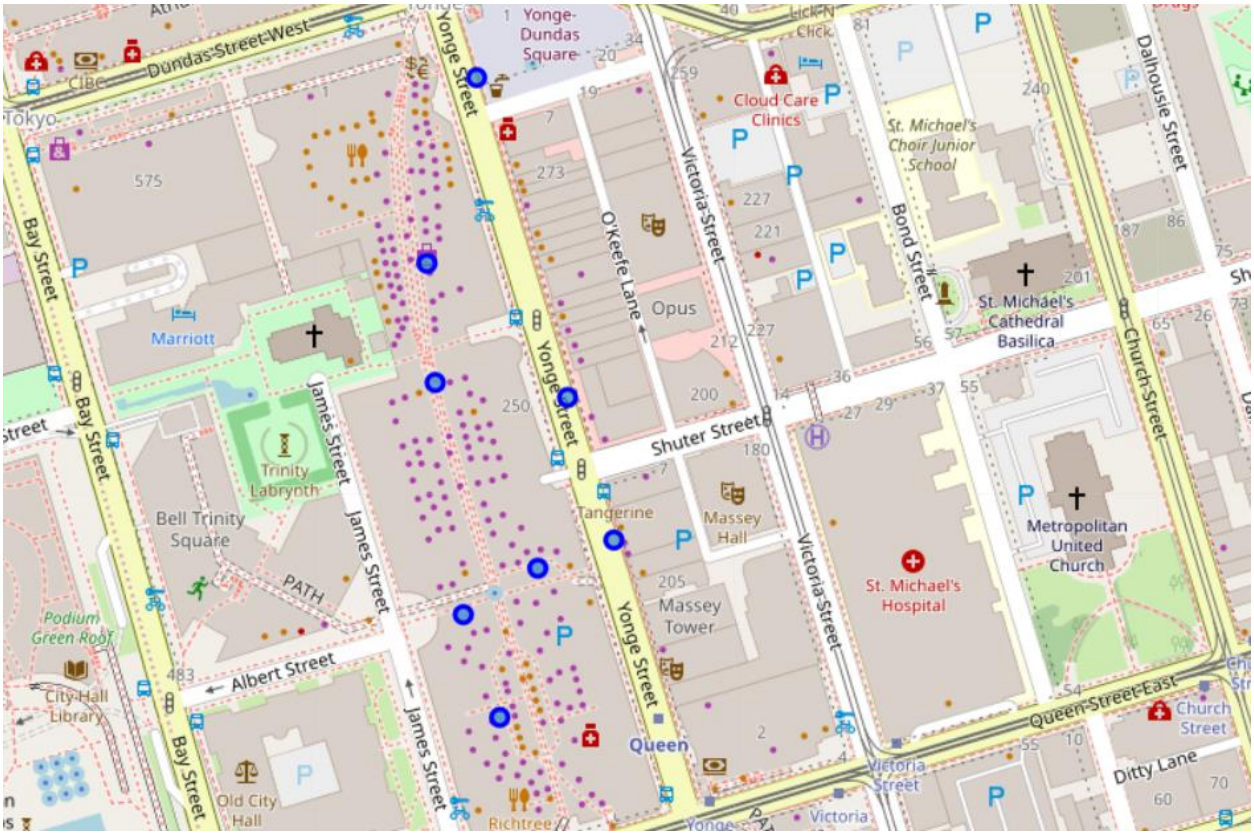

- Clean pulled JSON data and create a new data frame including the required columns

| | postalCode | name | categories | lat | lng |
|---|---|---|---|---|---|
| **0** | M5B 2H1 | UNIQLO ユニクロ | Clothing Store | 43.655910 | -79.380641 |
| **1** | M5B 2H1 | Elgin And Winter Garden Theatres | Theater | 43.653394 | -79.378507 |
| **2** | M5B 2H1 | LUSH | Cosmetics Shop | 43.653557 | -79.380400 |
| **3** | M5B 1V8 | Ed Mirvish Theatre | Theater | 43.655102 | -79.379768 |
| **4** | M5B 2H1 | Indigo | Bookstore | 43.653515 | -79.380696 |
| **5** | M5B 2H1 | CF Toronto Eaton Centre | Shopping Mall | 43.654540 | -79.380677 |
| **6** | M5B 2R8 | Yonge-Dundas Square | Plaza | 43.656054 | -79.380495 |
| **7** | M5G 2C9 | Eggspectation Bell Trinity Square | Breakfast Spot | 43.653144 | -79.381980 |

- Create a new data frame keeping only Clothing Stores

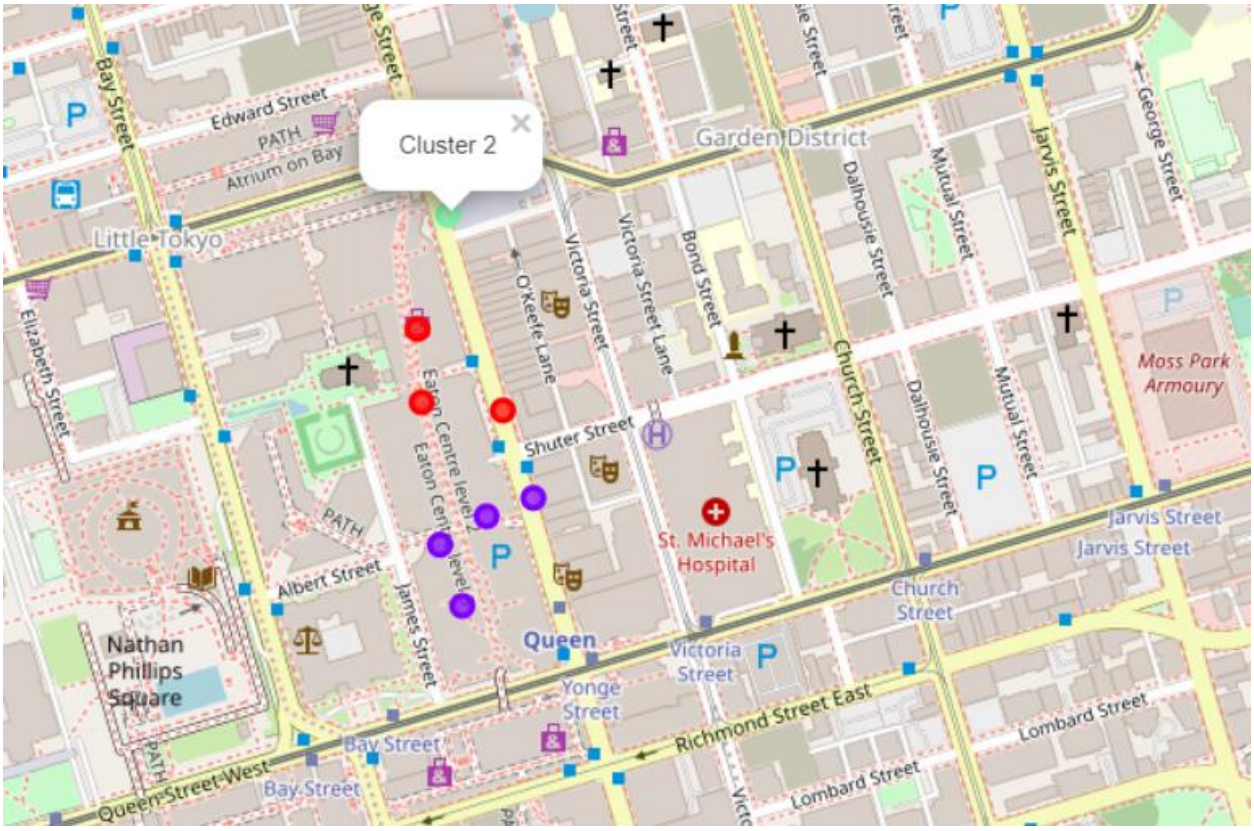
| | postalCode | name | categories | lat | lng |
|---|---|---|---|---|---|
| **0** | M5B 2H1 | UNIQLO ユニクロ | Clothing Store | 43.655910 | -79.380641 |
| **13** | M5B 2L9 | Nordstrom | Clothing Store | 43.655041 | -79.380966 |
| **19** | NaN | Magic Tailor | Clothing Store | 43.653742 | -79.379745 |
| **20** | M5B 2H1 | Roots | Clothing Store | 43.653613 | -79.380244 |
| **27** | M5B 2H1 | Hollister Co. | Clothing Store | 43.654480 | -79.380914 |
| **38** | M5B 2H1 | Abercrombie & Fitch | Clothing Store | 43.652915 | -79.380495 |
| **45** | M5B 2H1 | lululemon athletica | Clothing Store | 43.653394 | -79.380722 |
| **92** | M5B 1N8 | Urban Outfitters | Clothing Store | 43.654411 | -79.380055 |

- Create a folium map to show where the clothing stores are in Downtown Toronto




- Proceed to modeling using k-means clustering method. This method will help us to identify how already existing clothing stores are split in clusters. That way we will be able to conclude on where it would make more sense to open a new clothing store in downtown Toronto area.

- Create a new folium map that shows us the clusters of clothing stores




# 4. Results

We identified that the downtown Toronto clothing stores are located very close each other on a specific area. Only one clothing store is a bit far from the others. Stores are split on 3 clusters. Cluster 2 consists only from “UNIQLO ユニクロ” store which is the one that is more far compared to the others.
As a result, if Cynthia wants to open a new clothing store she should consider opening it in the area near to Cluster 2 since there the competition will be limited due to inexistence of other clothing stores.

# 5. Conclusion

This project helps a person to get a better understanding of the neighborhoods and the existing venues per neighborhood and more specifically for a neighborhood of his choice. So, if someone wants for example, to open a new store in a specific area, he can apply this method with his choice of area so that he gets a view of the locations of existing stores and how those are grouped.

Locations that show there is limited existence of stores will lead to the decision of opening a new store in this area.