**Battle of Neighborhoods - Presentation**

**Business Problem :**

We want to explore, segment, and cluster the neighborhoods in Downtown Toronto, to find neighborhoods that would be profitable to open a Yoga Studio. However, unlike New York, the neighborhood data is not readily available on the internet. What is interesting about the field of data science is that each project can be challenging in its unique way, so you need to learn to be agile and refine the skill to learn new libraries and tools quickly depending on the project.

For the Toronto neighborhood data, a Wikipedia page exists that has all the information we need to explore and cluster the neighborhoods in Toronto : [**https://en.wikipedia.org/wiki/List\_of\_postal\_codes\_of\_Canada:\_M**](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)

It will be required to scrape the Wikipedia page and wrangle the data, clean it, and then read it into a pandas dataframe so that it is in a structured format.

The question is in which neighborhood, would it be recommended to open a Yoga Studio. The straight answer would be to open one in a neighborhood, that doesn’t have any.

**Data to be used :**

1.Scrape the Wikipedia page [**https://en.wikipedia.org/wiki/List\_of\_postal\_codes\_of\_Canada:\_M**](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)into a dataframe :

df=pd.read\_html('https://en.wikipedia.org/wiki/List\_of\_postal\_codes\_of\_Canada:\_M')[0]

df.head(10)

|  | **Postal Code** | **Borough** | **Neighborhood** |
| --- | --- | --- | --- |
| **0** | M1A | Not assigned | NaN |
| **1** | M2A | Not assigned | NaN |
| **2** | M3A | North York | Parkwoods |
| **3** | M4A | North York | Victoria Village |
| **4** | M5A | Downtown Toronto | Regent Park, Harbourfront |
| **5** | M6A | North York | Lawrence Manor, Lawrence Heights |
| **6** | M7A | Downtown Toronto | Queen's Park, Ontario Provincial Government |
| **7** | M8A | Not assigned | NaN |
| **8** | M9A | Etobicoke | Islington Avenue |
| **9** | M1B | Scarborough | Malvern, Rouge |

2. Drop the rows where a borough is not assigned

3. Read the geospatial co-ordinates of each postal code into a dataframe

df1=pd.read\_csv('http://cocl.us/Geospatial\_data/Geospatial\_Coordinates.csv')

df1.head(10)

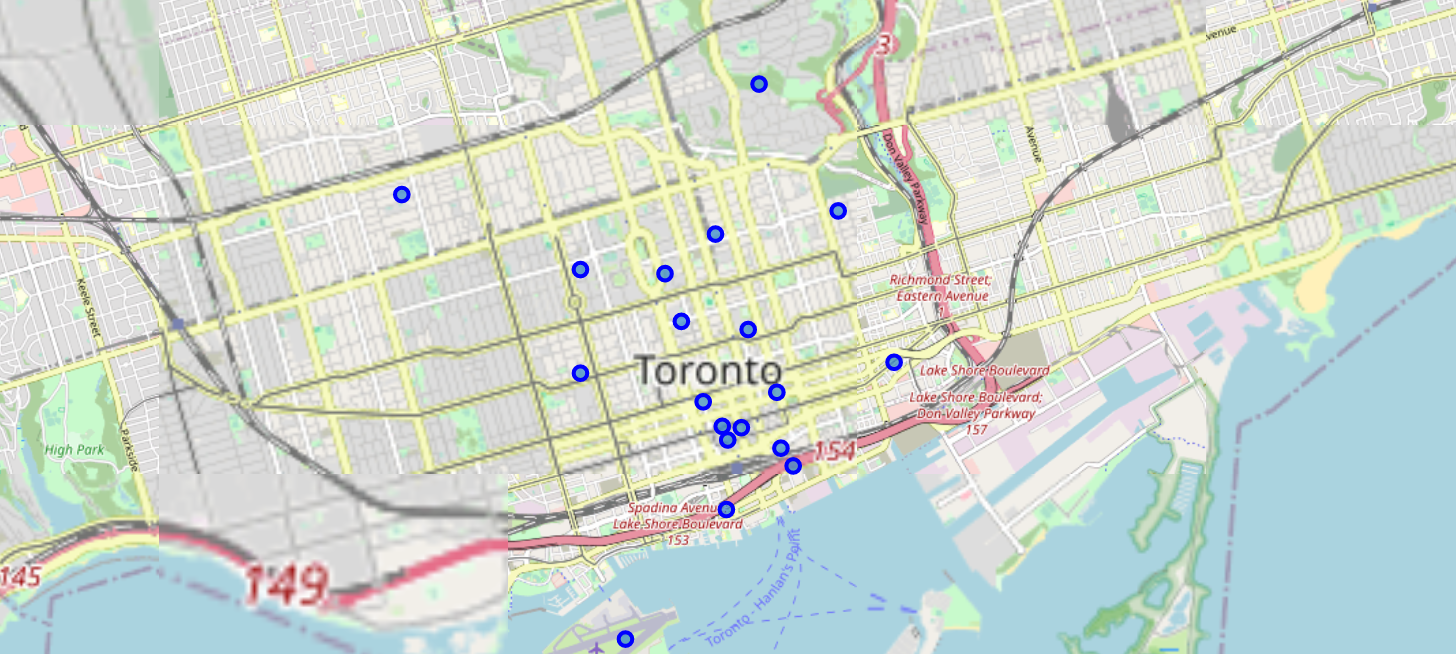
|  | **Postal Code** | **Latitude** | **Longitude** |
| --- | --- | --- | --- |
| **0** | M1B | 43.806686 | -79.194353 |
| **1** | M1C | 43.784535 | -79.160497 |
| **2** | M1E | 43.763573 | -79.188711 |
| **3** | M1G | 43.770992 | -79.216917 |
| **4** | M1H | 43.773136 | -79.239476 |
| **5** | M1J | 43.744734 | -79.239476 |
| **6** | M1K | 43.727929 | -79.262029 |
| **7** | M1L | 43.711112 | -79.284577 |
| **8** | M1M | 43.716316 | -79.239476 |
| **9** | M1N | 43.692657 | -79.264848 |

4. Merge the geospatial codes into the dataframe which has the neighborhoods and boroughs based on the postal codes

|  | **Postal Code** | **Borough** | **Neighborhood** | **Latitude** | **Longitude** |
| --- | --- | --- | --- | --- | --- |
| **0** | M3A | North York | Parkwoods | 43.753259 | -79.329656 |
| **1** | M4A | North York | Victoria Village | 43.725882 | -79.315572 |
| **2** | M5A | Downtown Toronto | Regent Park, Harbourfront | 43.654260 | -79.360636 |
| **3** | M6A | North York | Lawrence Manor, Lawrence Heights | 43.718518 | -79.464763 |
| **4** | M7A | Downtown Toronto | Queen's Park, Ontario Provincial Government | 43.662301 | -79.389494 |
| **5** | M9A | Etobicoke | Islington Avenue | 43.667856 | -79.532242 |
| **6** | M1B | Scarborough | Malvern, Rouge | 43.806686 | -79.194353 |
| **7** | M3B | North York | Don Mills | 43.745906 | -79.352188 |
| **8** | M4B | East York | Parkview Hill, Woodbine Gardens | 43.706397 | -79.309937 |
| **9** | M5B | Downtown Toronto | Garden District, Ryerson | 43.657162 | -79.378937 |
| **10** | M6B | North York | Glencairn | 43.709577 | -79.445073 |

**Methodology :**

1. Filter data for borough of Downtown Toronto and then create map of Toronoto and mark Neighborhoods using Folium.



1. Explore the neighborhoods using Foursquare API to get the different venues and venue category in each neighborhood. Snippet of Venue and Venue Category for a neighborhood.

|  | **Neighborhood** | **Neighborhood Latitude** | **Neighborhood Longitude** | **Venue** | **Venue Latitude** | **Venue Longitude** | **Venue Category** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | Regent Park, Harbourfront | 43.65426 | -79.360636 | Roselle Desserts | 43.653447 | -79.362017 | Bakery |
| **1** | Regent Park, Harbourfront | 43.65426 | -79.360636 | Tandem Coffee | 43.653559 | -79.361809 | Coffee Shop |
| **2** | Regent Park, Harbourfront | 43.65426 | -79.360636 | Morning Glory Cafe | 43.653947 | -79.361149 | Breakfast Spot |
| **3** | Regent Park, Harbourfront | 43.65426 | -79.360636 | Cooper Koo Family YMCA | 43.653249 | -79.358008 | Distribution Center |
| **4** | Regent Park, Harbourfront | 43.65426 | -79.360636 | Body Blitz Spa East | 43.654735 | -79.359874 | Spa |

1. Filter the data to check which neighborhoods have Venue Category as Yoga Studio, shown below :

['Regent Park, Harbourfront',

"Queen's Park, Ontario Provincial Government",

'Central Bay Street',

'University of Toronto, Harbord',

'Stn A PO Boxes',

'Church and Wellesley']

1. The remaining neighborhoods, are the ones, to open a Yoga Studio

**Results/Conclusion :**

1. The neighborhoods identified in Step 4 above, are the ones where it would be benefical/profitable to open a Yoga Studio

|  | **Postal Code** | **Borough** | **Neighborhood** | **Latitude** | **Longitude** |
| --- | --- | --- | --- | --- | --- |
| **2** | M5B | Downtown Toronto | Garden District, Ryerson | 43.657162 | -79.378937 |
| **3** | M5C | Downtown Toronto | St. James Town | 43.651494 | -79.375418 |
| **4** | M5E | Downtown Toronto | Berczy Park | 43.644771 | -79.373306 |
| **6** | M6G | Downtown Toronto | Christie | 43.669542 | -79.422564 |
| **7** | M5H | Downtown Toronto | Richmond, Adelaide, King | 43.650571 | -79.384568 |
| **8** | M5J | Downtown Toronto | Harbourfront East, Union Station, Toronto Islands | 43.640816 | -79.381752 |
| **9** | M5K | Downtown Toronto | Toronto Dominion Centre, Design Exchange | 43.647177 | -79.381576 |
| **10** | M5L | Downtown Toronto | Commerce Court, Victoria Hotel | 43.648198 | -79.379817 |
| **12** | M5T | Downtown Toronto | Kensington Market, Chinatown, Grange Park | 43.653206 | -79.400049 |
| **13** | M5V | Downtown Toronto | CN Tower, King and Spadina, Railway Lands, Har... | 43.628947 | -79.394420 |
| **14** | M4W | Downtown Toronto | Rosedale | 43.679563 | -79.377529 |
| **16** | M4X | Downtown Toronto | St. James Town, Cabbagetown | 43.667967 | -79.367675 |
| **17** | M5X | Downtown Toronto | First Canadian Place, Underground city | 43.648429 | -79.382280 |