

INTRODUCTION

- This is a list of postal codes in Canada where the first letter is M. Postal codes beginning with M are located within the city of Toronto in the province of Ontario. Only the first three characters are listed, corresponding to the Forward Sortation Area.
- Canada Post provides a free postal code look-up tool on its website, via its applications for such smartphones as the iPhone and BlackBerry, and sells hard-copy directories and CD-ROMs. Many vendors also sell validation tools, which allow customers to properly match addresses and postal codes. Hard-copy directories can also be consulted in all post offices, and some libraries.

Problem

- In this project we will try to find an optimal location for a restaurant. Specifically, this report will be targeted to stakeholders interested in opening restaurant in Toronto
- Since there are lots of restaurants in Berlin, we will try to detect locations that are not already crowded with restaurants.
- We will use our data science powers to generate a few most promising neighborhoods based on these criteria. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

Data acquisition

- Data is imported from the below link: 'https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M'
- We have to do web scraping in order to get the data from the above website.
- The final dataset which we will get after scraping the website is as below:

Neighborhood\n	Borough\n	Postal code\n	
\n	Not assigned\n	M1A\n	0
\n	Not assigned\n	M2A\n	1
Parkwoods\n	North York\n	M3A\n	2
Victoria Village\n	North York\n	M4A\n	3
Regent Park / Harbourfront\n	Downtown Toronto\n	M5A\n	4

Data Cleaning

- We could see that the data which we got from the website is not clean.
- So, we have to clean the data.
- First, we have to remove '\n' characters after each word.
- Then, we have to rename the column names.
- Remove all the rows in the column 'Borough' which is are assigned to the value 'Not Assigned'. And also, we need to check the null values in the data set and we need to manage them. And also, we need to manage the duplicate values in the dataset.
- As there are 77 rows with Not Assigned value in Borough column so we will remove all the 77 rows. And also, there are no null and duplicate values in the dataset. And also we need to replace the character '/' with ','.
- So, after cleaning the data we could see that there are 103 rows in the dataset

Importing Geospatial Data

	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

Merging the two datasets

	Postal code	Borough	Neighborhood	Latitude	Longitude
0	МЗА	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

Exploring the Clusters

• Cluster 1:

	Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
4	East Toronto	0	Trail	Health Food Store	Pub	Women's Store	Dance Studio	Electronics Store	Eastern European Restaurant	Donut Shop	Doner Restaurant	Dog Run

• Cluster 2:

		-			-							
	Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Downtown Toronto	1	Coffee Shop	Park	Bakery	Pub	Breakfast Spot	Café	Theater	Mexican Restaurant	Restaurant	Chocolate Shop
1	Downtown Toronto	1	Coffee Shop	Sushi Restaurant	Diner	Yoga Studio	Burrito Place	Beer Bar	Italian Restaurant	Juice Bar	Sandwich Place	Burger Joint
2	Downtown Toronto	1	Clothing Store	Coffee Shop	Café	Cosmetics Shop	Restaurant	Japanese Restaurant	Italian Restaurant	Bubble Tea Shop	Middle Eastern Restaurant	Ramen Restaurant
3	Downtown Toronto	1	Café	Coffee Shop	Cocktail Bar	American Restaurant	Gastropub	Hotel	Restaurant	Gym	Italian Restaurant	Department Store
5	Downtown Toronto	1	Coffee Shop	Cocktail Bar	Beer Bar	Cheese Shop	Restaurant	Café	Bakery	Seafood Restaurant	Bistro	Jazz Club
6	Downtown Toronto	1	Coffee Shop	Café	Italian Restaurant	Sandwich Place	Ice Cream Shop	Middle Eastern Restaurant	Bar	Thai Restaurant	Burger Joint	Fried Chicken Joint
7	Downtown Toronto	1	Grocery Store	Café	Park	Candy Store	Athletics & Sports	Restaurant	Italian Restaurant	Diner	Nightclub	Coffee Shop
8	Downtown Toronto	1	Coffee Shop	Café	Restaurant	Thai Restaurant	Deli / Bodega	Gym	Clothing Store	Hotel	Bakery	Steakhouse
9	West Toronto	1	Bakery	Pharmacy	Supermarket	Music Venue	Bank	Brewery	Pool	Recording Studio	Café	Middle Eastern Restaurant
10	Downtown Toronto	1	Coffee Shop	Aquarium	Hotel	Café	Restaurant	Brewery	Sporting Goods Shop	Italian Restaurant	Scenic Lookout	Fried Chicken Joint

Exploring the Clusters (Contd..)

• Cluster 3:

Toronto 2 School Park Landscaping Bus Line Store Restaurant Store European Restaurant 21 Central 2 Park Store Trail Bus Line Sushi Deli / Electronics Eastern European Restaurant 22 Park Store Trail Bus Line Restaurant Bodega Store Eastern European Restaurant 23 Park Store Trail Bus Line Restaurant Bodega Store Eastern Restaurant 24 Park Store Trail Bus Line Restaurant Bodega Store Restaurant 25 Park Store Trail Bus Line Restaurant Bodega Store Restaurant 26 Park Store Restaurant Bodega Store Restaur		Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
21 Central 2 Park Store Trail Bus Line Sushi Dell' Electronics European Restaurant 33 Downtown 2 Park Playground Trail Women's Dance Electronics Eastern European Donut Shop Restaurant 34 Downtown 2 Park Playground Trail Store Studio Store European Donut Shop Restaurant 35 Downtown Doner Doner Doner Doner Doner Restaurant	18		2		Park		Bus Line				European	Donut Shop	Doner Restaurant
33 Downtown 2 Park Playground Trail Store Electronics European Donut Shop Doner Dog Run	21		2	Park		Trail	Bus Line				European	Donut Shop	Doner Restaurant
	33	_	2	Park	Playground	Trail				European	Donut Shop		Dog Run

Cluster 4:

	Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
19	Central Toronto	3	Garden	Women's Store	Deli / Bodega	Event Space	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant	Donut Shop	Doner Restaurant	Dog Run

Exploring the Clusters:

• Cluster 5:

E	Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
29	Central Toronto	4	Park	Women's Store	Deli / Bodega	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant	Donut Shop	Doner Restaurant	Dog Run	Distribution Center

Results and Conclusion

As per the above results, the stakeholders can open the restaurant in Cluster 1, Cluster 4 and Cluster 5.