

# Exploring Toronto Neighborhoods

## Data Sources

a) Different neighborhoods in Toronto are represented by Postal Codes. I'm using "List of Postal code of Canada: M"

([https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)) wiki page to get all the information about the neighborhoods present in Toronto. This page has the postal code, borough & the name of all the neighborhoods present in Toronto.

b) However, we also need the coordinates of these postal codes to search for competing businesses. The "[https://cocl.us/Geospatial\\_data](https://cocl.us/Geospatial_data)" csv file contains all the geographical coordinates of the neighborhoods. We will need to merge the postal code data with the coordinates to create a dataset of all postal code, names of the neighborhood and their respective coordinates.

## Part 1

**Scraping Toronto Neighborhoods Table from Wikipedia "List of Postal code of Canada: M" in order to obtain the data about the Toronto & the Neighborhoods in it.**

Assumptions made to attain the below DataFrame:

- DataFrame will consist of three columns: PostalCode, Borough, and Neighborhood
- Only the cells that have an assigned borough will be processed. Borough that is not assigned are ignored.
- More than one neighborhood can exist in one postal code area. For example, in the table on the Wikipedia page, you will notice that M5A is listed twice and has two neighborhoods: Harbourfront and Regent Park. These two rows will be combined into one row with the neighborhoods separated with a comma as shown in row 11 in the above table.
- If a cell has a borough but a Not assigned neighborhood, then the neighborhood will be the same as the borough.

wikipedia - package is used to scrape the data from wiki.

```
In [1]: import pandas as pd
import numpy as np
```

```
In [2]: url = 'https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M'
df = pd.read_html(url)

# the above code captures all the 'tables' found in the webpage and return them
type(df)
```

```
Out[2]: list
```

```
In [3]: # if you go through the items on the list, you will see the postal code table is  
df[0]
```

Out[3]:

	0	1	2	3	4	5	
0	M1A Not assigned	M2A Not assigned	M3A North York (Parkwoods)	M4A North York (Victoria Village)	M5A Downtown Toronto (Regent Park / Harbourfront)	M6A North York (Lawrence Manor / Lawrence Heig...	C
1	M1B Scarborough (Malvern / Rouge)	M2B Not assigned	M3B North York (Don Mills) North	M4B East York (Parkview Hill / Woodbine Gardens)	M5B Downtown Toronto (Garden District, Ryerson)	M6B North York (Glencairn)	
2	M1C Scarborough (Rouge Hill / Port Union / Hig...	M2C Not assigned	M3C North York (Don Mills) South (Flemingdon P...	M4C East York (Woodbine Heights)	M5C Downtown Toronto (St. James Town)	M6C York (Humewood-Cedarvale)	
3	M1E Scarborough (Guildwood / Morningside / Wes...	M2E Not assigned	M3E Not assigned	M4E East Toronto (The Beaches)	M5E Downtown Toronto (Berczy Park)	M6E York (Caledonia-Fairbanks)	
4	M1G Scarborough (Woburn)	M2G Not assigned	M3G Not assigned	M4G East York (Leaside)	M5G Downtown Toronto (Central Bay Street)	M6G Downtown Toronto (Christie)	
5	M1H Scarborough (Cedarbrae)	M2H North York (Hillcrest Village)	M3H North York (Bathurst Manor / Wilson Height...	M4H East York (Thornccliffe Park)	M5H Downtown Toronto (Richmond / Adelaide / King)	M6H West Toronto (Dufferin / Dovercourt Village)	
6	M1J Scarborough (Scarborough Village)	M2J North York (Fairview / Henry Farm / Oriole)	M3J North York (Northwood Park / York University)	M4J East York East Toronto (The Danforth East)	M5J Downtown Toronto (Harbourfront East / Unio...	M6J West Toronto (Little Portugal / Trinity)	
7	M1K Scarborough (Kennedy Park / Ionview / East...	M2K North York (Bayview Village)	M3K North York (Downsview) East (CFB Toronto)	M4K East Toronto (The Danforth West / Riverdale)	M5K Downtown Toronto (Toronto Dominion Centre ...	M6K West Toronto (Brockton / Parkdale Village ...	

	0	1	2	3	4	5
8	M1L Scarborough (Golden Mile / Clairlea / Oakr...	M2L North York (York Mills / Silver Hills)	M3L North York (Downsview) West	M4L East Toronto (India Bazaar / The Beaches W...	M5L Downtown Toronto (Commerce Court / Victori...	M6L North York (North Park / Maple Leaf Park /...
9	M1M Scarborough (Cliffside / Cliffcrest / Scar...	M2M North York (Willowdale / Newtonbrook)	M3M North York (Downsview) Central	M4M East Toronto (Studio District)	M5M North York (Bedford Park / Lawrence Manor ...	M6M York (Del Ray / Mount Dennis / Keelsdale a...
10	M1N Scarborough (Birch Cliff / Cliffside West)	M2N North York (Willowdale) South	M3N North York (Downsview) Northwest	M4N Central Toronto (Lawrence Park)	M5N Central Toronto (Roselawn)	M6N York (Runnymede / The Junction North)
11	M1P Scarborough (Dorset Park / Wexford Heights...	M2P North York (York Mills West)	M3P Not assigned	M4P Central Toronto (Davisville North)	M5P Central Toronto (Forest Hill North & West)	M6P West Toronto (High Park / The Junction South)
12	M1R Scarborough (Wexford / Maryvale)	M2R North York (Willowdale) West	M3R Not assigned	M4R Central Toronto (North Toronto West)	M5R Central Toronto (The Annex / North Midtown...	M6R West Toronto (Parkdale / Roncesvalles)
13	M1S Scarborough (Agincourt)	M2S Not assigned	M3S Not assigned	M4S Central Toronto (Davisville)	M5S Downtown Toronto (University of Toronto / ...	M6S West Toronto (Runnymede / Swansea)
14	M1T Scarborough (Clarks Corners / Tam O'Shante...	M2T Not assigned	M3T Not assigned	M4T Central Toronto (Moore Park / Summerhill E...	M5T Downtown Toronto (Kensington Market / Chin...	M6T Not assigned
15	M1V Scarborough (Milliken / Agincourt North / ...	M2V Not assigned	M3V Not assigned	M4V Central Toronto (Summerhill West / Rathnel...	M5V Downtown Toronto (CN Tower / King and Spad...	M6V Not assigned
16	M1W Scarborough (Steeles West / L'Amoreaux West)	M2W Not assigned	M3W Not assigned	M4W Downtown Toronto (Rosedale)	M5W Downtown Toronto Stn A PO Boxes 25 The Esp...	M6W Not assigned

	0	1	2	3	4	5
17	M1X Scarborough (Upper Rouge)	M2X Not assigned	M3X Not assigned	M4X Downtown Toronto (St. James Town / Cabbage...	M5X Downtown Toronto (First Canadian Place / U...	M6X Not assigned
18	M1Y Not assigned	M2Y Not assigned	M3Y Not assigned	M4Y Downtown Toronto (Church and Wellesley)	M5Y Not assigned	M6Y Not assigned
19	M1Z Not assigned	M2Z Not assigned	M3Z Not assigned	M4Z Not assigned	M5Z Not assigned	M6Z Not assigned

In [4]: *# Flatten the DataFrame into one List (a Numpy array)*

```
data = df[0].values.flatten()
data
```

```

Out[4]: array(['M1A Not assigned', 'M2A Not assigned',
               'M3A North York (Parkwoods)', 'M4A North York (Victoria Village)',
               'M5A Downtown Toronto (Regent Park / Harbourfront)',
               'M6A North York (Lawrence Manor / Lawrence Heights)',
               'M7A Queen's Park (Ontario Provincial Government)',
               'M8A Not assigned', 'M9A Etobicoke (Islington Avenue)',
               'M1B Scarborough (Malvern / Rouge)', 'M2B Not assigned',
               'M3B North York (Don Mills) North',
               'M4B East York (Parkview Hill / Woodbine Gardens)',
               'M5B Downtown Toronto (Garden District, Ryerson)',
               'M6B North York (Glencairn)', 'M7B Not assigned',
               'M8B Not assigned',
               'M9B Etobicoke (West Deane Park / Princess Gardens / Martin Grove / Isli
ngton / Cloverdale)',
               'M1C Scarborough (Rouge Hill / Port Union / Highland Creek)',
               'M2C Not assigned',
               'M3C North York (Don Mills) South (Flemingdon Park)',
               'M4C East York (Woodbine Heights)',
               'M5C Downtown Toronto (St. James Town)',
               'M6C York (Humewood-Cedarvale)', 'M7C Not assigned',
               'M8C Not assigned',
               'M9C Etobicoke (Eringate / Bloordale Gardens / Old Burnhamthorpe / Markl
and Wood)',
               'M1E Scarborough (Guildwood / Morningside / West Hill)',
               'M2E Not assigned', 'M3E Not assigned',
               'M4E East Toronto (The Beaches)',
               'M5E Downtown Toronto (Berczy Park)',
               'M6E York (Caledonia-Fairbanks)', 'M7E Not assigned',
               'M8E Not assigned', 'M9E Not assigned', 'M1G Scarborough (Woburn)',
               'M2G Not assigned', 'M3G Not assigned', 'M4G East York (Leaside)',
               'M5G Downtown Toronto (Central Bay Street)',
               'M6G Downtown Toronto (Christie)', 'M7G Not assigned',
               'M8G Not assigned', 'M9G Not assigned',
               'M1H Scarborough (Cedarbrae)',
               'M2H North York (Hillcrest Village)',
               'M3H North York (Bathurst Manor / Wilson Heights / Downsview North)',
               'M4H East York (Thornccliffe Park)',
               'M5H Downtown Toronto (Richmond / Adelaide / King)',
               'M6H West Toronto (Dufferin / Dovercourt Village)',
               'M7H Not assigned', 'M8H Not assigned', 'M9H Not assigned',
               'M1J Scarborough (Scarborough Village)',
               'M2J North York (Fairview / Henry Farm / Oriole)',
               'M3J North York (Northwood Park / York University)',
               'M4J East York East Toronto (The Danforth East)',
               'M5J Downtown Toronto (Harbourfront East / Union Station / Toronto Islan
ds)',
               'M6J West Toronto (Little Portugal / Trinity)', 'M7J Not assigned',
               'M8J Not assigned', 'M9J Not assigned',
               'M1K Scarborough (Kennedy Park / Ionview / East Birchmount Park)',
               'M2K North York (Bayview Village)',
               'M3K North York (Downsview) East (CFB Toronto)',
               'M4K East Toronto (The Danforth West / Riverdale)',
               'M5K Downtown Toronto (Toronto Dominion Centre / Design Exchange)',
               'M6K West Toronto (Brockton / Parkdale Village / Exhibition Place)',
               'M7K Not assigned', 'M8K Not assigned', 'M9K Not assigned',
               'M1L Scarborough (Golden Mile / Clairlea / Oakridge)',
               'M2L North York (York Mills / Silver Hills)',
               'M3L North York (Downsview) West',
               'M4L East Toronto (India Bazaar / The Beaches West)',
               'M5L Downtown Toronto (Commerce Court / Victoria Hotel)',

```

'M6L North York (North Park / Maple Leaf Park / Upwood Park)',  
 'M7L Not assigned', 'M8L Not assigned',  
 'M9L North York (Humber Summit)',  
 'M1M Scarborough (Cliffside / Cliffcrest / Scarborough Village West)',  
 'M2M North York (Willowdale / Newtonbrook)',  
 'M3M North York (Downsview) Central',  
 'M4M East Toronto (Studio District)',  
 'M5M North York (Bedford Park / Lawrence Manor East)',  
 'M6M York (Del Ray / Mount Dennis / Keelsdale and Silverthorn)',  
 'M7M Not assigned', 'M8M Not assigned',  
 'M9M North York (Humberlea / Emery)',  
 'M1N Scarborough (Birch Cliff / Cliffside West)',  
 'M2N North York (Willowdale) South',  
 'M3N North York (Downsview) Northwest',  
 'M4N Central Toronto (Lawrence Park)',  
 'M5N Central Toronto (Roselawn)',  
 'M6N York (Runnymede / The Junction North)', 'M7N Not assigned',  
 'M8N Not assigned', 'M9N York (Weston)',  
 'M1P Scarborough (Dorset Park / Wexford Heights / Scarborough Town Centre)',  
 'M2P North York (York Mills West)', 'M3P Not assigned',  
 'M4P Central Toronto (Davisville North)',  
 'M5P Central Toronto (Forest Hill North & West)',  
 'M6P West Toronto (High Park / The Junction South)',  
 'M7P Not assigned', 'M8P Not assigned',  
 'M9P Etobicoke (Westmount)',  
 'M1R Scarborough (Wexford / Maryvale)',  
 'M2R North York (Willowdale) West', 'M3R Not assigned',  
 'M4R Central Toronto (North Toronto West)',  
 'M5R Central Toronto (The Annex / North Midtown / Yorkville)',  
 'M6R West Toronto (Parkdale / Roncesvalles)',  
 'M7R Mississauga Canada Post Gateway Processing Centre (Enclave of L4 W)',  
 'M8R Not assigned',  
 'M9R Etobicoke (Kingsview Village / St. Phillips / Martin Grove Gardens / Richview Gardens)',  
 'M1S Scarborough (Agincourt)', 'M2S Not assigned',  
 'M3S Not assigned', 'M4S Central Toronto (Davisville)',  
 'M5S Downtown Toronto (University of Toronto / Harbord)',  
 'M6S West Toronto (Runnymede / Swansea)', 'M7S Not assigned',  
 'M8S Not assigned', 'M9S Not assigned',  
 'M1T Scarborough (Clarks Corners / Tam O'Shanter / Sullivan)',  
 'M2T Not assigned', 'M3T Not assigned',  
 'M4T Central Toronto (Moore Park / Summerhill East)',  
 'M5T Downtown Toronto (Kensington Market / Chinatown / Grange Park)',  
 'M6T Not assigned', 'M7T Not assigned', 'M8T Not assigned',  
 'M9T Not assigned',  
 'M1V Scarborough (Milliken / Agincourt North / Steeles East / L'Amoreaux East)',  
 'M2V Not assigned', 'M3V Not assigned',  
 'M4V Central Toronto (Summerhill West / Rathnelly / South Hill / Forest Hill SE / Deer Park)',  
 'M5V Downtown Toronto (CN Tower / King and Spadina / Railway Lands / Harbourfront West / Bathurst Quay / South Niagara / Island airport)',  
 'M6V Not assigned', 'M7V Not assigned',  
 'M8V Etobicoke (New Toronto / Mimico South / Humber Bay Shores)',  
 'M9V Etobicoke (South Steeles / Silverstone / Humbergate / Jamestown / Mount Olive / Beaumont Heights / Thistletown / Albion Gardens)',  
 'M1W Scarborough (Steeles West / L'Amoreaux West)',  
 'M2W Not assigned', 'M3W Not assigned',

```

'M4W Downtown Toronto (Rosedale)',
'M5W Downtown Toronto Stn A PO Boxes 25 The Esplanade (Enclave of M5E)',
'M6W Not assigned', 'M7W Not assigned',
'M8W Etobicoke (Alderwood / Long Branch)',
'M9W Etobicoke Northwest (Clairville / Humberwood / Woodbine Downs / West Humber / Kipling Heights / Rexdale / Elms / Tandridge / Old Rexdale)',
'M1X Scarborough (Upper Rouge)', 'M2X Not assigned',
'M3X Not assigned',
'M4X Downtown Toronto (St. James Town / Cabbagetown)',
'M5X Downtown Toronto (First Canadian Place / Underground city)',
'M6X Not assigned', 'M7X Not assigned',
'M8X Etobicoke (The Kingsway / Montgomery Road / Old Mill North)',
'M9X Not assigned', 'M1Y Not assigned', 'M2Y Not assigned',
'M3Y Not assigned', 'M4Y Downtown Toronto (Church and Wellesley)',
'M5Y Not assigned', 'M6Y Not assigned',
'M7Y East Toronto Business reply mail Processing Centre 969 Eastern (Enclave of M4L)',
'M8Y Etobicoke (Old Mill South / King's Mill Park / Sunnylea / Humber Bay / Mimico NE / The Queensway East / Royal York South East / Kingsway Park South East)",
'M9Y Not assigned', 'M1Z Not assigned', 'M2Z Not assigned',
'M3Z Not assigned', 'M4Z Not assigned', 'M5Z Not assigned',
'M6Z Not assigned', 'M7Z Not assigned',
'M8Z Etobicoke (Mimico NW / The Queensway West / South of Bloor / Kingsway Park South West / Royal York South West)',
'M9Z Not assigned'], dtype=object)

```

```

In [5]: # Split the strings in each cell into Postcode, Borough, and Neighbourhood

# hint: use for loop to loop through each item, use string slicing create a list
# Initialize a list to store the structured data
structured_data = []

# Loop through each element in 'data'
for item in data:
    # Extract Postcode (everything before the first space)
    first_space = item.find(" ")
    postcode = item[:first_space] # Postcode is the first part

    # Extract the remaining part of the string after Postcode
    remaining = item[first_space + 1:]

    # Find the start of the parentheses (if present)
    paren_start = remaining.find("(")

    if paren_start != -1:
        # If parentheses are present, split Borough and Neighbourhood
        borough = remaining[:paren_start].strip() # Everything before '(' is the
        neighbourhood = remaining[paren_start + 1:-1].strip() # Everything inside
    else:
        # If no parentheses, Borough is the rest, and Neighbourhood is "Not assigned"
        borough = remaining.strip()
        neighbourhood = "Not assigned"

    # Append the structured data as a list
    structured_data.append([postcode, borough, neighbourhood])

# Print the structured data to verify
for row in structured_data:
    print(row)

```



['M1A', 'Not assigned', 'Not assigned']  
['M2A', 'Not assigned', 'Not assigned']  
['M3A', 'North York', 'Parkwoods']  
['M4A', 'North York', 'Victoria Village']  
['M5A', 'Downtown Toronto', 'Regent Park / Harbourfront']  
['M6A', 'North York', 'Lawrence Manor / Lawrence Heights']  
['M7A', 'Queen's Park', 'Ontario Provincial Government']  
['M8A', 'Not assigned', 'Not assigned']  
['M9A', 'Etobicoke', 'Islington Avenue']  
['M1B', 'Scarborough', 'Malvern / Rouge']  
['M2B', 'Not assigned', 'Not assigned']  
['M3B', 'North York', 'Don Mills) Nort']  
['M4B', 'East York', 'Parkview Hill / Woodbine Gardens']  
['M5B', 'Downtown Toronto', 'Garden District, Ryerson']  
['M6B', 'North York', 'Glencairn']  
['M7B', 'Not assigned', 'Not assigned']  
['M8B', 'Not assigned', 'Not assigned']  
['M9B', 'Etobicoke', 'West Deane Park / Princess Gardens / Martin Grove / Islington / Cloverdale']  
['M1C', 'Scarborough', 'Rouge Hill / Port Union / Highland Creek']  
['M2C', 'Not assigned', 'Not assigned']  
['M3C', 'North York', 'Don Mills) South (Flemingdon Park']  
['M4C', 'East York', 'Woodbine Heights']  
['M5C', 'Downtown Toronto', 'St. James Town']  
['M6C', 'York', 'Humewood-Cedarvale']  
['M7C', 'Not assigned', 'Not assigned']  
['M8C', 'Not assigned', 'Not assigned']  
['M9C', 'Etobicoke', 'Eringate / Bloordale Gardens / Old Burnhamthorpe / Markland Wood']  
['M1E', 'Scarborough', 'Guildwood / Morningside / West Hill']  
['M2E', 'Not assigned', 'Not assigned']  
['M3E', 'Not assigned', 'Not assigned']  
['M4E', 'East Toronto', 'The Beaches']  
['M5E', 'Downtown Toronto', 'Berczy Park']  
['M6E', 'York', 'Caledonia-Fairbanks']  
['M7E', 'Not assigned', 'Not assigned']  
['M8E', 'Not assigned', 'Not assigned']  
['M9E', 'Not assigned', 'Not assigned']  
['M1G', 'Scarborough', 'Woburn']  
['M2G', 'Not assigned', 'Not assigned']  
['M3G', 'Not assigned', 'Not assigned']  
['M4G', 'East York', 'Leaside']  
['M5G', 'Downtown Toronto', 'Central Bay Street']  
['M6G', 'Downtown Toronto', 'Christie']  
['M7G', 'Not assigned', 'Not assigned']  
['M8G', 'Not assigned', 'Not assigned']  
['M9G', 'Not assigned', 'Not assigned']  
['M1H', 'Scarborough', 'Cedarbrae']  
['M2H', 'North York', 'Hillcrest Village']  
['M3H', 'North York', 'Bathurst Manor / Wilson Heights / Downsview North']  
['M4H', 'East York', 'Thorncliffe Park']  
['M5H', 'Downtown Toronto', 'Richmond / Adelaide / King']  
['M6H', 'West Toronto', 'Dufferin / Dovercourt Village']  
['M7H', 'Not assigned', 'Not assigned']  
['M8H', 'Not assigned', 'Not assigned']  
['M9H', 'Not assigned', 'Not assigned']  
['M1J', 'Scarborough', 'Scarborough Village']  
['M2J', 'North York', 'Fairview / Henry Farm / Oriole']  
['M3J', 'North York', 'Northwood Park / York University']  
['M4J', 'East York East Toronto', 'The Danforth East']

['M5J', 'Downtown Toronto', 'Harbourfront East / Union Station / Toronto Islands']  
['M6J', 'West Toronto', 'Little Portugal / Trinity']  
['M7J', 'Not assigned', 'Not assigned']  
['M8J', 'Not assigned', 'Not assigned']  
['M9J', 'Not assigned', 'Not assigned']  
['M1K', 'Scarborough', 'Kennedy Park / Ionview / East Birchmount Park']  
['M2K', 'North York', 'Bayview Village']  
['M3K', 'North York', 'Downsview) East (CFB Toronto']  
['M4K', 'East Toronto', 'The Danforth West / Riverdale']  
['M5K', 'Downtown Toronto', 'Toronto Dominion Centre / Design Exchange']  
['M6K', 'West Toronto', 'Brockton / Parkdale Village / Exhibition Place']  
['M7K', 'Not assigned', 'Not assigned']  
['M8K', 'Not assigned', 'Not assigned']  
['M9K', 'Not assigned', 'Not assigned']  
['M1L', 'Scarborough', 'Golden Mile / Clairlea / Oakridge']  
['M2L', 'North York', 'York Mills / Silver Hills']  
['M3L', 'North York', 'Downsview) Wes']  
['M4L', 'East Toronto', 'India Bazaar / The Beaches West']  
['M5L', 'Downtown Toronto', 'Commerce Court / Victoria Hotel']  
['M6L', 'North York', 'North Park / Maple Leaf Park / Upwood Park']  
['M7L', 'Not assigned', 'Not assigned']  
['M8L', 'Not assigned', 'Not assigned']  
['M9L', 'North York', 'Humber Summit']  
['M1M', 'Scarborough', 'Cliffside / Cliffcrest / Scarborough Village West']  
['M2M', 'North York', 'Willowdale / Newtonbrook']  
['M3M', 'North York', 'Downsview) Centra']  
['M4M', 'East Toronto', 'Studio District']  
['M5M', 'North York', 'Bedford Park / Lawrence Manor East']  
['M6M', 'York', 'Del Ray / Mount Dennis / Keelsdale and Silverthorn']  
['M7M', 'Not assigned', 'Not assigned']  
['M8M', 'Not assigned', 'Not assigned']  
['M9M', 'North York', 'Humberlea / Emery']  
['M1N', 'Scarborough', 'Birch Cliff / Cliffside West']  
['M2N', 'North York', 'Willowdale) Sout']  
['M3N', 'North York', 'Downsview) Northwes']  
['M4N', 'Central Toronto', 'Lawrence Park']  
['M5N', 'Central Toronto', 'Roselawn']  
['M6N', 'York', 'Runnymede / The Junction North']  
['M7N', 'Not assigned', 'Not assigned']  
['M8N', 'Not assigned', 'Not assigned']  
['M9N', 'York', 'Weston']  
['M1P', 'Scarborough', 'Dorset Park / Wexford Heights / Scarborough Town Centre']  
['M2P', 'North York', 'York Mills West']  
['M3P', 'Not assigned', 'Not assigned']  
['M4P', 'Central Toronto', 'Davisville North']  
['M5P', 'Central Toronto', 'Forest Hill North & West']  
['M6P', 'West Toronto', 'High Park / The Junction South']  
['M7P', 'Not assigned', 'Not assigned']  
['M8P', 'Not assigned', 'Not assigned']  
['M9P', 'Etobicoke', 'Westmount']  
['M1R', 'Scarborough', 'Wexford / Maryvale']  
['M2R', 'North York', 'Willowdale) Wes']  
['M3R', 'Not assigned', 'Not assigned']  
['M4R', 'Central Toronto', 'North Toronto West']  
['M5R', 'Central Toronto', 'The Annex / North Midtown / Yorkville']  
['M6R', 'West Toronto', 'Parkdale / Roncesvalles']  
['M7R', 'Mississauga Canada Post Gateway Processing Centre', 'Enclave of L4W']  
['M8R', 'Not assigned', 'Not assigned']  
['M9R', 'Etobicoke', 'Kingsview Village / St. Phillips / Martin Grove Gardens / R

ichview Gardens']  
 ['M1S', 'Scarborough', 'Agincourt']  
 ['M2S', 'Not assigned', 'Not assigned']  
 ['M3S', 'Not assigned', 'Not assigned']  
 ['M4S', 'Central Toronto', 'Davisville']  
 ['M5S', 'Downtown Toronto', 'University of Toronto / Harbord']  
 ['M6S', 'West Toronto', 'Runnymede / Swansea']  
 ['M7S', 'Not assigned', 'Not assigned']  
 ['M8S', 'Not assigned', 'Not assigned']  
 ['M9S', 'Not assigned', 'Not assigned']  
 ['M1T', 'Scarborough', "Clarks Corners / Tam O'Shanter / Sullivan"]  
 ['M2T', 'Not assigned', 'Not assigned']  
 ['M3T', 'Not assigned', 'Not assigned']  
 ['M4T', 'Central Toronto', 'Moore Park / Summerhill East']  
 ['M5T', 'Downtown Toronto', 'Kensington Market / Chinatown / Grange Park']  
 ['M6T', 'Not assigned', 'Not assigned']  
 ['M7T', 'Not assigned', 'Not assigned']  
 ['M8T', 'Not assigned', 'Not assigned']  
 ['M9T', 'Not assigned', 'Not assigned']  
 ['M1V', 'Scarborough', "Milliken / Agincourt North / Steeles East / L'Amoreaux East"]  
 ['M2V', 'Not assigned', 'Not assigned']  
 ['M3V', 'Not assigned', 'Not assigned']  
 ['M4V', 'Central Toronto', 'Summerhill West / Rathnelly / South Hill / Forest Hill SE / Deer Park']  
 ['M5V', 'Downtown Toronto', 'CN Tower / King and Spadina / Railway Lands / Harbourfront West / Bathurst Quay / South Niagara / Island airport']  
 ['M6V', 'Not assigned', 'Not assigned']  
 ['M7V', 'Not assigned', 'Not assigned']  
 ['M8V', 'Etobicoke', 'New Toronto / Mimico South / Humber Bay Shores']  
 ['M9V', 'Etobicoke', 'South Steeles / Silverstone / Humbergate / Jamestown / Mount Olive / Beaumont Heights / Thistletown / Albion Gardens']  
 ['M1W', 'Scarborough', "Steeles West / L'Amoreaux West"]  
 ['M2W', 'Not assigned', 'Not assigned']  
 ['M3W', 'Not assigned', 'Not assigned']  
 ['M4W', 'Downtown Toronto', 'Rosedale']  
 ['M5W', 'Downtown Toronto Stn A PO Boxes 25 The Esplanade', 'Enclave of M5E']  
 ['M6W', 'Not assigned', 'Not assigned']  
 ['M7W', 'Not assigned', 'Not assigned']  
 ['M8W', 'Etobicoke', 'Alderwood / Long Branch']  
 ['M9W', 'Etobicoke Northwest', 'Clairville / Humberwood / Woodbine Downs / West Humber / Kipling Heights / Rexdale / Elms / Tandridge / Old Rexdale']  
 ['M1X', 'Scarborough', 'Upper Rouge']  
 ['M2X', 'Not assigned', 'Not assigned']  
 ['M3X', 'Not assigned', 'Not assigned']  
 ['M4X', 'Downtown Toronto', 'St. James Town / Cabbagetown']  
 ['M5X', 'Downtown Toronto', 'First Canadian Place / Underground city']  
 ['M6X', 'Not assigned', 'Not assigned']  
 ['M7X', 'Not assigned', 'Not assigned']  
 ['M8X', 'Etobicoke', 'The Kingsway / Montgomery Road / Old Mill North']  
 ['M9X', 'Not assigned', 'Not assigned']  
 ['M1Y', 'Not assigned', 'Not assigned']  
 ['M2Y', 'Not assigned', 'Not assigned']  
 ['M3Y', 'Not assigned', 'Not assigned']  
 ['M4Y', 'Downtown Toronto', 'Church and Wellesley']  
 ['M5Y', 'Not assigned', 'Not assigned']  
 ['M6Y', 'Not assigned', 'Not assigned']  
 ['M7Y', 'East Toronto Business reply mail Processing Centre 969 Eastern', 'Enclave of M4L']  
 ['M8Y', 'Etobicoke', "Old Mill South / King's Mill Park / Sunnylea / Humber Bay /

```
Mimico NE / The Queensway East / Royal York South East / Kingsway Park South East"]
['M9Y', 'Not assigned', 'Not assigned']
['M1Z', 'Not assigned', 'Not assigned']
['M2Z', 'Not assigned', 'Not assigned']
['M3Z', 'Not assigned', 'Not assigned']
['M4Z', 'Not assigned', 'Not assigned']
['M5Z', 'Not assigned', 'Not assigned']
['M6Z', 'Not assigned', 'Not assigned']
['M7Z', 'Not assigned', 'Not assigned']
['M8Z', 'Etobicoke', 'Mimico NW / The Queensway West / South of Bloor / Kingsway
Park South West / Royal York South West']
['M9Z', 'Not assigned', 'Not assigned']
```

if you are doing this correctly, you shall have a list of lists that looks like this:

```
[['M1A', 'Not assigned', 'Not assigned'],
 ['M2A', 'Not assigned', 'Not assigned'],
 ['M3A', 'North York', 'Parkwoods'],
 ['M4A', 'North York', 'Victoria Village'],
 ['M5A', 'Downtown Toronto', 'Regent Park / Harbourfront'],
 ['M6A', 'North York', 'Lawrence Manor / Lawrence Heights'],
 ['M7A', 'Queen's Park', 'Ontario Provincial Government'],
 ['M8A', 'Not assigned', 'Not assigned'],
 ['M9A', 'Etobicoke', 'Islington Avenue'],
 ['M1B', 'Scarborough', 'Malvern / Rouge'],
 ['M2B', 'Not assigned', 'Not assigned'],
 ['M3B', 'North York', 'Don MillsNorth'],
 ['M4B', 'East York', 'Parkview Hill / Woodbine Gardens'],
 ['M5B', 'Downtown Toronto', 'Garden District, Ryerson'],
 ['M6B', 'North York', 'Glencairn'],
 ['M7B', 'Not assigned', 'Not assigned'],
 ['M8B', 'Not assigned', 'Not assigned'],
```

```
In [6]: # Create a DataFrame from the cleaned data
columns = ['Postcode', 'Borough', 'Neighbourhood']
df = pd.DataFrame(structured_data, columns=columns)

# Handle multiple neighbourhoods (you can choose to keep them in the same row or
# Assuming you want to handle entries with multiple neighbourhoods by leaving th
df['Neighbourhood'] = df['Neighbourhood'].str.replace(' / ', ', ')

# Display the cleaned DataFrame
# Simply use this to display the DataFrame in your local environment
df.head()
```

```
Out[6]:
```

	Postcode	Borough	Neighbourhood
0	M1A	Not assigned	Not assigned
1	M2A	Not assigned	Not assigned
2	M3A	North York	Parkwoods
3	M4A	North York	Victoria Village
4	M5A	Downtown Toronto	Regent Park, Harbourfront

```
In [7]: print(df.columns)
```

```
Index(['Postcode', 'Borough', 'Neighbourhood'], dtype='object')
```

```
In [8]: #Only process the cells that have an assigned borough. Ignore cells with a borou
df = df[df.Borough != 'Not assigned']
df = df.rename(columns={'Postcode': 'Postalcode'})

df
```

```
Out[8]:
```

	Postalcode	Borough	Neighbourhood
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	Queen's Park	Ontario Provincial Government
...	...	...	...
160	M8X	Etobicoke	The Kingsway, Montgomery Road, Old Mill North
165	M4Y	Downtown Toronto	Church and Wellesley
168	M7Y	East Toronto Business reply mail Processing Ce...	Enclave of M4L
169	M8Y	Etobicoke	Old Mill South, King's Mill Park, Sunnylea, Hu...
178	M8Z	Etobicoke	Mimico NW, The Queensway West, South of Bloor,...

103 rows × 3 columns

```
In [9]: # Sort data and reset index
toronto_DF = df.groupby(['Borough', 'Postalcode'])['Neighbourhood'].apply(list).
toronto_DF
```

Out[9]:

	Borough	Postalcode	Neighbourhood
0	Central Toronto	M4N	Lawrence Park
1	Central Toronto	M4P	Davisville North
2	Central Toronto	M4R	North Toronto West
3	Central Toronto	M4S	Davisville
4	Central Toronto	M4T	Moore Park, Summerhill East
...	...	...	...
98	York	M6C	Humewood-Cedarvale
99	York	M6E	Caledonia-Fairbanks
100	York	M6M	Del Ray, Mount Dennis, Keelsdale and Silverthorn
101	York	M6N	Runnymede, The Junction North
102	York	M9N	Weston

103 rows × 3 columns

## Part 2

### Adding geographical coordinates to the neighborhoods

Next important step is adding the geographical coordinates to these neighborhoods. To do so I'm extracting the data present in the Geospatial Data csv file and I'm combining it with the existing neighborhood dataframe by merging them both based on the postal code.

In [10]: *#Reading the Latitude & Longitude data from CSV file*

```
import io
import requests

url = "https://cocl.us/Geospatial_data"
lat_long = requests.get(url).text
lat_long_df=pd.read_csv(io.StringIO(lat_long))
lat_long_df.head()
```

Out[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

In [11]: *# renaming the columns to match the existing dataframe*

```
lat_long_df = lat_long_df.rename(columns={'Postal Code': 'Postalcode'})
```

```
lat_long_df.head()
```

```
Out[11]:
```

	Postalcode	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

```
In [12]: # merging both the dataframe into one by matching on the postal code.
# Merge the two DataFrames on the 'PostalCode' column
toronto_DF = toronto_DF.merge(lat_long_df, on="Postalcode", how="left")
toronto_DF
```

```
Out[12]:
```

	Borough	Postalcode	Neighbourhood	Latitude	Longitude
0	Central Toronto	M4N	Lawrence Park	43.728020	-79.388790
1	Central Toronto	M4P	Davisville North	43.712751	-79.390197
2	Central Toronto	M4R	North Toronto West	43.715383	-79.405678
3	Central Toronto	M4S	Davisville	43.704324	-79.388790
4	Central Toronto	M4T	Moore Park, Summerhill East	43.689574	-79.383160
...	...	...	...	...	...
98	York	M6C	Humewood-Cedarvale	43.693781	-79.428191
99	York	M6E	Caledonia-Fairbanks	43.689026	-79.453512
100	York	M6M	Del Ray, Mount Dennis, Keelsdale and Silverthorn	43.691116	-79.476013
101	York	M6N	Runnymede, The Junction North	43.673185	-79.487262
102	York	M9N	Weston	43.706876	-79.518188

103 rows × 5 columns

```
In [13]: print('The dataframe has {} boroughs and {} neighborhoods.'.format(
len(toronto_DF['Borough'].unique()),
toronto_DF.shape[0]
))
```

The dataframe has 15 boroughs and 103 neighborhoods.

```
In [14]: toronto_DF.to_csv('Toronto_postal_codes_and_coordinates.csv')
```

If you do this right, you shall produce a table like below:

	Borough	Postalcode	Neighborhood	Latitude	Longitude
0	Central Toronto	M4N	Lawrence Park	43.728020	-79.388790
1	Central Toronto	M4P	Davisville North	43.712751	-79.390197
2	Central Toronto	M4R	North Toronto West	43.715383	-79.405678
3	Central Toronto	M4S	Davisville	43.704324	-79.388790
4	Central Toronto	M4T	Moore Park, Summerhill East	43.689574	-79.383160
...	...	...	...	...	...
98	York	M6C	Humewood-Cedarvale	43.693781	-79.428191
99	York	M6E	Caledonia-Fairbanks	43.689026	-79.453512
100	York	M6M	Del Ray, Mount Dennis, Keelsdale and Silverthorn	43.691116	-79.476013
101	York	M6N	Runnymede, The Junction North	43.673185	-79.487262
102	York	M9N	Weston	43.706876	-79.518188

103 rows × 5 columns

## Very important!

Save the output as a csv file.

Upload both this notebook in PDF with output AND the csv file to your GitHub repo.