Segmenting and Clustering Toronto Neighbourhoods

Objective: Segment and cluster the Toronto neighbourhoods based on post codes

Download the table of post codes for neighbourhoods in Toronto from Wikipedia. https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)

I have Excel 2016, which can hold 1,048,576 rows. The post code table has only 287 rows. I tried using BeautifulSoup first. But when I was half-way through and dealing with all those issues of cleaning the data, I realized that for such a small table, directlying downloading it into Excel, then uploading it into Jupyter Notebook will be much easier and faster as well. So, I scraped all my BeautifulSoup codes.

In [65]:
```python
import pandas as pd
postcode_df=pd.read_excel('Toronto Post Codes.xlsx')
postcode_df.head()
```

Out[65]:

	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	Harbourfront

In [66]:
```python
# size of the table
postcode_df.shape
```

Out[66]:
```
(287, 3)
```

In [67]:
```python
# number of unique Boroughs
import numpy as np
postcode_df["Borough"].value_counts()
```

Out[67]:
```
Not assigned        77
Etobicoke           45
North York          38
Scarborough         37
Downtown Toronto    37
Central Toronto     17
West Toronto        13
York                 9
East Toronto         7
East York            6
Mississauga          1
Name: Borough, dtype: int64
```

Okay. So there are 77 "Not assigned" in the Borough column. I could use drop.duplicates() to drop those "Not assigned". However, that will drop the other boroughs as well, since there are boroughs that have more than one neighbourhoods assigned to it. So, drop by "Borough" won't work.

Let's check the Neighbourhood column. The neighbourhoods should be unique, except for the "Not assigned".

```
In [68]:  postcode_df["Neighbourhood"].value_counts()
```

```
Out[68]:  Not assigned                                77
          St. James Town                               2
          Runnymede                                    2
          Mount Dennis                                 1
          York Mills West                              1
          The Queensway East                           1
          Royal York South East                        1
          Martin Grove                                 1
          Long Branch                                  1
          Wexford                                      1
          East Birchmount Park                         1
          Adelaide                                     1
          Woodbine Gardens                             1
          The Queensway West                           1
          Humber Summit                                1
          Woburn                                       1
          Clarks Corners                               1
          Cloverdale                                   1
          Downsview North                              1
          CN Tower                                     1
          Old Mill South                               1
          Steeles East                                 1
          Thistletown                                  1
          Toronto Dominion Centre                      1
          Cliffcrest                                   1
          Harbourfront East                            1
          Studio District                             1
          Eringate                                     1
          Dorset Park                                  1
          Roselawn                                     1
          Design Exchange                              1
          Richview Gardens                             1
          West Hill                                    1
          Highland Creek                               1
          Rouge                                        1
          Davisville                                   1
          Islington                                    1
          Railway Lands                                1
          Princess Gardens                             1
          East Toronto                                 1
          Scarborough Village                          1
          Don Mills South                              1
          Port Union                                   1
          Downsview Northwest                          1
          Silverstone                                  1
          Roncesvalles                                 1
          New Toronto                                  1
          Canada Post Gateway Processing Centre        1
          Berczy Park                                  1
          Kingsway Park South West                     1
          Keelesdale                                   1
          Chinatown                                    1
          The Beaches                                  1
          Willowdale South                             1
          Trinity                                      1
          Harbourfront                                 1
          Forest Hill SE                               1
          Forest Hill North                            1
          Rathnelly                                    1
          Birch Cliff                                  1
          Garden District                             1
          Ionview                                      1
          Hillcrest Village                            1
          Silverthorn                                  1
          Willowdale West                              1
          Woodbine Heights                             1
          Parkview Hill                                1
          University of Toronto                        1
          Wilson Heights                               1
          Downsview Central                            1
          Agincourt North                              1
          Lawrence Heights                             1
          Tam O'Shanter                                1
          Moore Park                                   1
          Emery                                        1
          Oakridge                                     1
          Parkdale                                     1
          Victoria Hotel                               1
          Lawrence Manor                               1
          Underground city                            1
          Summerhill East                             1
          Thorncliffe Park                             1
          Brockton                                     1
```

Humberlea	1
Markland Wood	1
Islington Avenue	1
Alderwood	1
Scarborough Town Centre	1
Riverdale	1
The Junction North	1
Del Ray	1
Weston	1
The Kingsway	1
Scarborough Village West	1
Parkdale Village	1
Yorkville	1
Island airport	1
Upwood Park	1
Golden Mile	1
Albion Gardens	1
The Annex	1
South Niagara	1
Central Bay Street	1
Maryvale	1
King's Mill Park	1
Royal York South West	1
Northwest	1
Kingsview Village	1
Grange Park	1
Glencairn	1
Milliken	1
Willowdale	1
High Park	1
Harbourfront West	1
The Junction South	1
Malvern	1
Rouge Hill	1
Victoria Village	1
Richmond	1
Silver Hills	1
Northwood Park	1
Kensington Market	1
Bathurst Manor	1
Downsview West	1
L'Amoreaux East	1
Martin Grove Gardens	1
Caledonia-Fairbanks	1
Summerhill West	1
Downsview	1
Guildwood	1
North Midtown	1
Henry Farm	1
Exhibition Place	1
Cliffside	1
Flemingdon Park	1
Jamestown	1
Union Station	1
Humbergate	1
Beaumond Heights	1
Davisville North	1
West Deane Park	1
Mimico NW	1
Dufferin	1
Clairlea	1
Sunnylea	1
Wexford Heights	1
Commerce Court	1
Stn A PO Boxes 25 The Esplanade	1
Little Portugal	1
St. Phillips	1
King	1
South Steeles	1
Church and Wellesley	1
Newtonbrook	1
Mimico South	1
York Mills	1
Agincourt	1
Bathurst Quay	1
Fairview	1
Kingsway Park South East	1
North Park	1
Cliffside West	1
Deer Park	1
First Canadian Place	1
Upper Rouge	1
Morningside	1
Parkwoods	1
Cabbagetown	1
Mount Olive	1

```
Montgomery Road                                   1
Downsview East                                    1
South Hill                                        1
Old Burnhamthorpe                                 1
Lawrence Manor East                               1
Harbord                                           1
Cedarbrae                                         1
York University                                   1
Humber Bay                                        1
Queen's Park                                      1
Oriole                                            1
Bedford Park                                      1
North Toronto West                                1
King and Spadina                                  1
Toronto Islands                                   1
Old Mill North                                    1
Bayview Village                                   1
The Danforth West                                 1
Dovercourt Village                                1
CFB Toronto                                       1
Sullivan                                          1
Swansea                                           1
Mimico NE                                         1
Forest Hill West                                  1
Christie                                          1
Kennedy Park                                      1
India Bazaar                                      1
Ryerson                                           1
The Beaches West                                  1
Business Reply Mail Processing Centre 969 Eastern 1
Lawrence Park                                     1
Humber Bay Shores                                 1
Humewood-Cedarvale                                1
Leaside                                           1
Westmount                                         1
Rosedale                                          1
South of Bloor                                    1
Bloordale Gardens                                 1
L'Amoreaux West                                   1
Don Mills North                                   1
Name: Neighbourhood, dtype: int64
```

All the neighbourhoods are unique, except for 1) the "Not assigned" and 2) "Runnymede" and "St. James Town" where there are two copies.

In [69]:
```
#Remove the "Not assigned" from the Borough column.
postcode_df.drop_duplicates(subset="Neighbourhood",keep=False, inplace=True)
postcode_df
```

Out[69]:

	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
7	M7A	Downtown Toronto	Queen's Park
9	M9A	Etobicoke	Islington Avenue
10	M1B	Scarborough	Rouge
11	M1B	Scarborough	Malvern
13	M3B	North York	Don Mills North
14	M4B	East York	Woodbine Gardens
15	M4B	East York	Parkview Hill
16	M5B	Downtown Toronto	Ryerson
17	M5B	Downtown Toronto	Garden District
18	M6B	North York	Glencairn
21	M9B	Etobicoke	Cloverdale
22	M9B	Etobicoke	Islington
23	M9B	Etobicoke	Martin Grove
24	M9B	Etobicoke	Princess Gardens
25	M9B	Etobicoke	West Deane Park
26	M1C	Scarborough	Highland Creek
27	M1C	Scarborough	Rouge Hill
28	M1C	Scarborough	Port Union
30	M3C	North York	Flemingdon Park
31	M3C	North York	Don Mills South
32	M4C	East York	Woodbine Heights
34	M6C	York	Humewood-Cedarvale
37	M9C	Etobicoke	Bloordale Gardens
38	M9C	Etobicoke	Eringate
39	M9C	Etobicoke	Markland Wood
40	M9C	Etobicoke	Old Burnhamthorpe
41	M1E	Scarborough	Guildwood
42	M1E	Scarborough	Morningside
43	M1E	Scarborough	West Hill
46	M4E	East Toronto	The Beaches
47	M5E	Downtown Toronto	Berczy Park
48	M6E	York	Caledonia-Fairbanks
52	M1G	Scarborough	Woburn
55	M4G	East York	Leaside
56	M5G	Downtown Toronto	Central Bay Street
57	M6G	Downtown Toronto	Christie
61	M1H	Scarborough	Cedarbrae
62	M2H	North York	Hillcrest Village
63	M3H	North York	Bathurst Manor
64	M3H	North York	Downsview North
65	M3H	North York	Wilson Heights
66	M4H	East York	Thorncliffe Park
67	M5H	Downtown Toronto	Adelaide
68	M5H	Downtown Toronto	King
69	M5H	Downtown Toronto	Richmond
70	M6H	West Toronto	Dovercourt Village

	Postcode	Borough	Neighbourhood
71	M6H	West Toronto	Dufferin
75	M1J	Scarborough	Scarborough Village
76	M2J	North York	Fairview
77	M2J	North York	Henry Farm
78	M2J	North York	Oriole
79	M3J	North York	Northwood Park
80	M3J	North York	York University
81	M4J	East York	East Toronto
82	M5J	Downtown Toronto	Harbourfront East
83	M5J	Downtown Toronto	Toronto Islands
84	M5J	Downtown Toronto	Union Station
85	M6J	West Toronto	Little Portugal
86	M6J	West Toronto	Trinity
90	M1K	Scarborough	East Birchmount Park
91	M1K	Scarborough	Ionview
92	M1K	Scarborough	Kennedy Park
93	M2K	North York	Bayview Village
94	M3K	North York	CFB Toronto
95	M3K	North York	Downsview East
96	M4K	East Toronto	The Danforth West
97	M4K	East Toronto	Riverdale
98	M5K	Downtown Toronto	Design Exchange
99	M5K	Downtown Toronto	Toronto Dominion Centre
100	M6K	West Toronto	Brockton
101	M6K	West Toronto	Exhibition Place
102	M6K	West Toronto	Parkdale Village
106	M1L	Scarborough	Clairlea
107	M1L	Scarborough	Golden Mile
108	M1L	Scarborough	Oakridge
109	M2L	North York	Silver Hills
110	M2L	North York	York Mills
111	M3L	North York	Downsview West
112	M4L	East Toronto	The Beaches West
113	M4L	East Toronto	India Bazaar
114	M5L	Downtown Toronto	Commerce Court
115	M5L	Downtown Toronto	Victoria Hotel
116	M6L	North York	Downsview
117	M6L	North York	North Park
118	M6L	North York	Upwood Park
121	M9L	North York	Humber Summit
122	M1M	Scarborough	Cliffcrest
123	M1M	Scarborough	Cliffside
124	M1M	Scarborough	Scarborough Village West
125	M2M	North York	Newtonbrook
126	M2M	North York	Willowdale
127	M3M	North York	Downsview Central
128	M4M	East Toronto	Studio District
129	M5M	North York	Bedford Park
130	M5M	North York	Lawrence Manor East
131	M6M	York	Del Ray
132	M6M	York	Keelesdale
133	M6M	York	Mount Dennis
134	M6M	York	Silverthorn
137	M9M	North York	Emery
138	M9M	North York	Humberlea

	Postcode	Borough	Neighbourhood
139	M1N	Scarborough	Birch Cliff
140	M1N	Scarborough	Cliffside West
141	M2N	North York	Willowdale South
142	M3N	North York	Downsview Northwest
143	M4N	Central Toronto	Lawrence Park
144	M5N	Central Toronto	Roselawn
145	M6N	York	The Junction North
149	M9N	York	Weston
150	M1P	Scarborough	Dorset Park
151	M1P	Scarborough	Scarborough Town Centre
152	M1P	Scarborough	Wexford Heights
153	M2P	North York	York Mills West
155	M4P	Central Toronto	Davisville North
156	M5P	Central Toronto	Forest Hill North
157	M5P	Central Toronto	Forest Hill West
158	M6P	West Toronto	High Park
159	M6P	West Toronto	The Junction South
162	M9P	Etobicoke	Westmount
163	M1R	Scarborough	Maryvale
164	M1R	Scarborough	Wexford
165	M2R	North York	Willowdale West
167	M4R	Central Toronto	North Toronto West
168	M5R	Central Toronto	The Annex
169	M5R	Central Toronto	North Midtown
170	M5R	Central Toronto	Yorkville
171	M6R	West Toronto	Parkdale
172	M6R	West Toronto	Roncesvalles
173	M7R	Mississauga	Canada Post Gateway Processing Centre
175	M9R	Etobicoke	Kingsview Village
176	M9R	Etobicoke	Martin Grove Gardens
177	M9R	Etobicoke	Richview Gardens
178	M9R	Etobicoke	St. Phillips
179	M1S	Scarborough	Agincourt
182	M4S	Central Toronto	Davisville
183	M5S	Downtown Toronto	Harbord
184	M5S	Downtown Toronto	University of Toronto
186	M6S	West Toronto	Swansea
190	M1T	Scarborough	Clarks Corners
191	M1T	Scarborough	Sullivan
192	M1T	Scarborough	Tam O'Shanter
195	M4T	Central Toronto	Moore Park
196	M4T	Central Toronto	Summerhill East
197	M5T	Downtown Toronto	Chinatown
198	M5T	Downtown Toronto	Grange Park
199	M5T	Downtown Toronto	Kensington Market
204	M1V	Scarborough	Agincourt North
205	M1V	Scarborough	L'Amoreaux East
206	M1V	Scarborough	Milliken
207	M1V	Scarborough	Steeles East
210	M4V	Central Toronto	Deer Park
211	M4V	Central Toronto	Forest Hill SE
212	M4V	Central Toronto	Rathnelly
213	M4V	Central Toronto	South Hill
214	M4V	Central Toronto	Summerhill West
215	M5V	Downtown Toronto	CN Tower

	Postcode	Borough	Neighbourhood
216	M5V	Downtown Toronto	Bathurst Quay
217	M5V	Downtown Toronto	Island airport
218	M5V	Downtown Toronto	Harbourfront West
219	M5V	Downtown Toronto	King and Spadina
220	M5V	Downtown Toronto	Railway Lands
221	M5V	Downtown Toronto	South Niagara
224	M8V	Etobicoke	Humber Bay Shores
225	M8V	Etobicoke	Mimico South
226	M8V	Etobicoke	New Toronto
227	M9V	Etobicoke	Albion Gardens
228	M9V	Etobicoke	Beaumond Heights
229	M9V	Etobicoke	Humbergate
230	M9V	Etobicoke	Jamestown
231	M9V	Etobicoke	Mount Olive
232	M9V	Etobicoke	Silverstone
233	M9V	Etobicoke	South Steeles
234	M9V	Etobicoke	Thistletown
235	M1W	Scarborough	L'Amoreaux West
238	M4W	Downtown Toronto	Rosedale
239	M5W	Downtown Toronto	Stn A PO Boxes 25 The Esplanade
242	M8W	Etobicoke	Alderwood
243	M8W	Etobicoke	Long Branch
244	M9W	Etobicoke	Northwest
245	M1X	Scarborough	Upper Rouge
248	M4X	Downtown Toronto	Cabbagetown
250	M5X	Downtown Toronto	First Canadian Place
251	M5X	Downtown Toronto	Underground city
254	M8X	Etobicoke	The Kingsway
255	M8X	Etobicoke	Montgomery Road
256	M8X	Etobicoke	Old Mill North
261	M4Y	Downtown Toronto	Church and Wellesley
264	M7Y	East Toronto	Business Reply Mail Processing Centre 969 Eastern
265	M8Y	Etobicoke	Humber Bay
266	M8Y	Etobicoke	King's Mill Park
267	M8Y	Etobicoke	Kingsway Park South East
268	M8Y	Etobicoke	Mimico NE
269	M8Y	Etobicoke	Old Mill South
270	M8Y	Etobicoke	The Queensway East
271	M8Y	Etobicoke	Royal York South East
272	M8Y	Etobicoke	Sunnylea
281	M8Z	Etobicoke	Kingsway Park South West
282	M8Z	Etobicoke	Mimico NW
283	M8Z	Etobicoke	The Queensway West
284	M8Z	Etobicoke	Royal York South West
285	M8Z	Etobicoke	South of Bloor

All the "Not assigned" in the Neighbourhood column has been dropped. However, since the parameter "keep" was set to False. The process above dropped the two neighbourhoods "Runnemede" and "St. James Town" as well. But I would like to keep those two neighbourhoods, since each copy belong to two different boroughs.

Put the two neighbourhoods back.

```
In [70]: append1=pd.DataFrame({"Postcode":["M5C","M6N","M6S","M4X"],"Borough":["Downtown Toronto","York","West Toronto","Downtown Tor
         postcode_df.append(append1,ignore_index=False)
```

Out[70]:

	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
7	M7A	Downtown Toronto	Queen's Park
9	M9A	Etobicoke	Islington Avenue
10	M1B	Scarborough	Rouge
11	M1B	Scarborough	Malvern
13	M3B	North York	Don Mills North
14	M4B	East York	Woodbine Gardens
15	M4B	East York	Parkview Hill
16	M5B	Downtown Toronto	Ryerson
17	M5B	Downtown Toronto	Garden District
18	M6B	North York	Glencairn
21	M9B	Etobicoke	Cloverdale
22	M9B	Etobicoke	Islington
23	M9B	Etobicoke	Martin Grove
24	M9B	Etobicoke	Princess Gardens
25	M9B	Etobicoke	West Deane Park
26	M1C	Scarborough	Highland Creek
27	M1C	Scarborough	Rouge Hill
28	M1C	Scarborough	Port Union
30	M3C	North York	Flemingdon Park
31	M3C	North York	Don Mills South
32	M4C	East York	Woodbine Heights
34	M6C	York	Humewood-Cedarvale
37	M9C	Etobicoke	Bloordale Gardens
38	M9C	Etobicoke	Eringate
39	M9C	Etobicoke	Markland Wood
40	M9C	Etobicoke	Old Burnhamthorpe
41	M1E	Scarborough	Guildwood
42	M1E	Scarborough	Morningside
43	M1E	Scarborough	West Hill
46	M4E	East Toronto	The Beaches
47	M5E	Downtown Toronto	Berczy Park
48	M6E	York	Caledonia-Fairbanks
52	M1G	Scarborough	Woburn
55	M4G	East York	Leaside
56	M5G	Downtown Toronto	Central Bay Street
57	M6G	Downtown Toronto	Christie
61	M1H	Scarborough	Cedarbrae
62	M2H	North York	Hillcrest Village
63	M3H	North York	Bathurst Manor
64	M3H	North York	Downsview North
65	M3H	North York	Wilson Heights
66	M4H	East York	Thorncliffe Park
67	M5H	Downtown Toronto	Adelaide
68	M5H	Downtown Toronto	King
69	M5H	Downtown Toronto	Richmond
70	M6H	West Toronto	Dovercourt Village
71	M6H	West Toronto	Dufferin

	Postcode	Borough	Neighbourhood
75	M1J	Scarborough	Scarborough Village
76	M2J	North York	Fairview
77	M2J	North York	Henry Farm
78	M2J	North York	Oriole
79	M3J	North York	Northwood Park
80	M3J	North York	York University
81	M4J	East York	East Toronto
82	M5J	Downtown Toronto	Harbourfront East
83	M5J	Downtown Toronto	Toronto Islands
84	M5J	Downtown Toronto	Union Station
85	M6J	West Toronto	Little Portugal
86	M6J	West Toronto	Trinity
90	M1K	Scarborough	East Birchmount Park
91	M1K	Scarborough	Ionview
92	M1K	Scarborough	Kennedy Park
93	M2K	North York	Bayview Village
94	M3K	North York	CFB Toronto
95	M3K	North York	Downsview East
96	M4K	East Toronto	The Danforth West
97	M4K	East Toronto	Riverdale
98	M5K	Downtown Toronto	Design Exchange
99	M5K	Downtown Toronto	Toronto Dominion Centre
100	M6K	West Toronto	Brockton
101	M6K	West Toronto	Exhibition Place
102	M6K	West Toronto	Parkdale Village
106	M1L	Scarborough	Clairlea
107	M1L	Scarborough	Golden Mile
108	M1L	Scarborough	Oakridge
109	M2L	North York	Silver Hills
110	M2L	North York	York Mills
111	M3L	North York	Downsview West
112	M4L	East Toronto	The Beaches West
113	M4L	East Toronto	India Bazaar
114	M5L	Downtown Toronto	Commerce Court
115	M5L	Downtown Toronto	Victoria Hotel
116	M6L	North York	Downsview
117	M6L	North York	North Park
118	M6L	North York	Upwood Park
121	M9L	North York	Humber Summit
122	M1M	Scarborough	Cliffcrest
123	M1M	Scarborough	Cliffside
124	M1M	Scarborough	Scarborough Village West
125	M2M	North York	Newtonbrook
126	M2M	North York	Willowdale
127	M3M	North York	Downsview Central
128	M4M	East Toronto	Studio District
129	M5M	North York	Bedford Park
130	M5M	North York	Lawrence Manor East
131	M6M	York	Del Ray
132	M6M	York	Keelesdale
133	M6M	York	Mount Dennis
134	M6M	York	Silverthorn
137	M9M	North York	Emery
138	M9M	North York	Humberlea
139	M1N	Scarborough	Birch Cliff

	Postcode	Borough	Neighbourhood
140	M1N	Scarborough	Cliffside West
141	M2N	North York	Willowdale South
142	M3N	North York	Downsview Northwest
143	M4N	Central Toronto	Lawrence Park
144	M5N	Central Toronto	Roselawn
145	M6N	York	The Junction North
149	M9N	York	Weston
150	M1P	Scarborough	Dorset Park
151	M1P	Scarborough	Scarborough Town Centre
152	M1P	Scarborough	Wexford Heights
153	M2P	North York	York Mills West
155	M4P	Central Toronto	Davisville North
156	M5P	Central Toronto	Forest Hill North
157	M5P	Central Toronto	Forest Hill West
158	M6P	West Toronto	High Park
159	M6P	West Toronto	The Junction South
162	M9P	Etobicoke	Westmount
163	M1R	Scarborough	Maryvale
164	M1R	Scarborough	Wexford
165	M2R	North York	Willowdale West
167	M4R	Central Toronto	North Toronto West
168	M5R	Central Toronto	The Annex
169	M5R	Central Toronto	North Midtown
170	M5R	Central Toronto	Yorkville
171	M6R	West Toronto	Parkdale
172	M6R	West Toronto	Roncesvalles
173	M7R	Mississauga	Canada Post Gateway Processing Centre
175	M9R	Etobicoke	Kingsview Village
176	M9R	Etobicoke	Martin Grove Gardens
177	M9R	Etobicoke	Richview Gardens
178	M9R	Etobicoke	St. Phillips
179	M1S	Scarborough	Agincourt
182	M4S	Central Toronto	Davisville
183	M5S	Downtown Toronto	Harbord
184	M5S	Downtown Toronto	University of Toronto
186	M6S	West Toronto	Swansea
190	M1T	Scarborough	Clarks Corners
191	M1T	Scarborough	Sullivan
192	M1T	Scarborough	Tam O'Shanter
195	M4T	Central Toronto	Moore Park
196	M4T	Central Toronto	Summerhill East
197	M5T	Downtown Toronto	Chinatown
198	M5T	Downtown Toronto	Grange Park
199	M5T	Downtown Toronto	Kensington Market
204	M1V	Scarborough	Agincourt North
205	M1V	Scarborough	L'Amoreaux East
206	M1V	Scarborough	Milliken
207	M1V	Scarborough	Steeles East
210	M4V	Central Toronto	Deer Park
211	M4V	Central Toronto	Forest Hill SE
212	M4V	Central Toronto	Rathnelly
213	M4V	Central Toronto	South Hill
214	M4V	Central Toronto	Summerhill West
215	M5V	Downtown Toronto	CN Tower
216	M5V	Downtown Toronto	Bathurst Quay

	Postcode	Borough	Neighbourhood
217	M5V	Downtown Toronto	Island airport
218	M5V	Downtown Toronto	Harbourfront West
219	M5V	Downtown Toronto	King and Spadina
220	M5V	Downtown Toronto	Railway Lands
221	M5V	Downtown Toronto	South Niagara
224	M8V	Etobicoke	Humber Bay Shores
225	M8V	Etobicoke	Mimico South
226	M8V	Etobicoke	New Toronto
227	M9V	Etobicoke	Albion Gardens
228	M9V	Etobicoke	Beaumond Heights
229	M9V	Etobicoke	Humbergate
230	M9V	Etobicoke	Jamestown
231	M9V	Etobicoke	Mount Olive
232	M9V	Etobicoke	Silverstone
233	M9V	Etobicoke	South Steeles
234	M9V	Etobicoke	Thistletown
235	M1W	Scarborough	L'Amoreaux West
238	M4W	Downtown Toronto	Rosedale
239	M5W	Downtown Toronto	Stn A PO Boxes 25 The Esplanade
242	M8W	Etobicoke	Alderwood
243	M8W	Etobicoke	Long Branch
244	M9W	Etobicoke	Northwest
245	M1X	Scarborough	Upper Rouge
248	M4X	Downtown Toronto	Cabbagetown
250	M5X	Downtown Toronto	First Canadian Place
251	M5X	Downtown Toronto	Underground city
254	M8X	Etobicoke	The Kingsway
255	M8X	Etobicoke	Montgomery Road
256	M8X	Etobicoke	Old Mill North
261	M4Y	Downtown Toronto	Church and Wellesley
264	M7Y	East Toronto	Business Reply Mail Processing Centre 969 Eastern
265	M8Y	Etobicoke	Humber Bay
266	M8Y	Etobicoke	King's Mill Park
267	M8Y	Etobicoke	Kingsway Park South East
268	M8Y	Etobicoke	Mimico NE
269	M8Y	Etobicoke	Old Mill South
270	M8Y	Etobicoke	The Queensway East
271	M8Y	Etobicoke	Royal York South East
272	M8Y	Etobicoke	Sunnylea
281	M8Z	Etobicoke	Kingsway Park South West
282	M8Z	Etobicoke	Mimico NW
283	M8Z	Etobicoke	The Queensway West
284	M8Z	Etobicoke	Royal York South West
285	M8Z	Etobicoke	South of Bloor
0	M5C	Downtown Toronto	St. James Town
1	M6N	York	Runnymede
2	M6S	West Toronto	Runnymede
3	M4X	Downtown Toronto	St. James Town

As shown above,the original table has 287 rows in total and 77 "Not assigned" in the Borough column. After the 77 "Not assigned" rows are dropped, the cleaned table should have 287-77=210 rows. The result of the above code block shows 210 rows.

Let's confirm that there is no more "Not assigned" in Borough column.

In [73]:
```python
postcode_df["Borough"].value_counts()
```

Out[73]:
```
Etobicoke           45
North York          38
Scarborough         37
Downtown Toronto    35
Central Toronto     17
West Toronto        12
York                 8
East Toronto         7
East York            6
Mississauga          1
Name: Borough, dtype: int64
```

Cool. There is no more "Not assigned" in the Borough column.

The table is cleaned now. I can start working with it. The next thing I am going to do is to put the neighbourhoods that have the same postcode in the same row. Let's see how many unique postcodes are in the cleaned table.

```
In [75]:   count=postcode_df["Postcode"].value_counts()
```

Out[75]:   M8Y   8
           M9V   8
           M5V   7
           M4V   5
           M8Z   5
           M9B   5
           M1V   4
           M6M   4
           M9R   4
           M9C   4
           M6K   3
           M8X   3
           M1T   3
           M8V   3
           M5T   3
           M1K   3
           M6L   3
           M1C   3
           M1L   3
           M5R   3
           M1E   3
           M1P   3
           M2J   3
           M1M   3
           M5J   3
           M3H   3
           M5H   3
           M6P   2
           M1R   2
           M6J   2
           M6H   2
           M4K   2
           M1N   2
           M6A   2
           M6R   2
           M5P   2
           M5X   2
           M5S   2
           M2M   2
           M4T   2
           M5B   2
           M3J   2
           M8W   2
           M9M   2
           M5M   2
           M4L   2
           M4B   2
           M3C   2
           M3K   2
           M5L   2
           M2L   2
           M1B   2
           M5K   2
           M4Y   1
           M4S   1
           M5A   1
           M3A   1
           M4E   1
           M4C   1
           M7Y   1
           M4W   1
           M4G   1
           M6C   1
           M4A   1
           M6N   1
           M1J   1
           M4R   1
           M7A   1
           M1H   1
           M1X   1
           M1S   1
           M1G   1
           M3M   1
           M3N   1
           M9W   1
           M2H   1
           M4M   1
           M7R   1
           M4N   1
           M5G   1
           M9N   1
           M9L   1
           M4P   1
```

```
M9P     1
M9A     1
M5N     1
M4J     1
M3L     1
M3B     1
M6G     1
M2R     1
M4H     1
M5W     1
M2N     1
M4X     1
M2P     1
M6E     1
M6B     1
M2K     1
M6S     1
M5E     1
M1W     1
Name: Postcode, dtype: int64
```

There are 102 unique postcodes. Most of the postcodes has only 1 neighbourhoods associated with it. However, postcodeS "M8Y" and "M9V" have 8 neighbourhoods associated with it, "M5V" has 7, "M4V" and "M8Z" have 5. I need to concatenate all those neighbourhoods under the postcode they are associated with, in one row.

```
In [76]:   postcode_df=pd.DataFrame(postcode_df.groupby(["Postcode","Borough"])["Neighbourhood"].apply(lambda x: ','.join(x)))
           postcode_df
```

Out[76]:

Postcode	Borough	Neighbourhood
M1B	Scarborough	Rouge,Malvern
M1C	Scarborough	Highland Creek,Rouge Hill,Port Union
M1E	Scarborough	Guildwood,Morningside,West Hill
M1G	Scarborough	Woburn
M1H	Scarborough	Cedarbrae
M1J	Scarborough	Scarborough Village
M1K	Scarborough	East Birchmount Park,Ionview,Kennedy Park
M1L	Scarborough	Clairlea,Golden Mile,Oakridge
M1M	Scarborough	Cliffcrest,Cliffside,Scarborough Village West
M1N	Scarborough	Birch Cliff,Cliffside West
M1P	Scarborough	Dorset Park,Scarborough Town Centre,Wexford He...
M1R	Scarborough	Maryvale,Wexford
M1S	Scarborough	Agincourt
M1T	Scarborough	Clarks Corners,Sullivan,Tam O'Shanter
M1V	Scarborough	Agincourt North,L'Amoreaux East,Milliken,Steel...
M1W	Scarborough	L'Amoreaux West
M1X	Scarborough	Upper Rouge
M2H	North York	Hillcrest Village
M2J	North York	Fairview,Henry Farm,Oriole
M2K	North York	Bayview Village
M2L	North York	Silver Hills,York Mills
M2M	North York	Newtonbrook,Willowdale
M2N	North York	Willowdale South
M2P	North York	York Mills West
M2R	North York	Willowdale West
M3A	North York	Parkwoods
M3B	North York	Don Mills North
M3C	North York	Flemingdon Park,Don Mills South
M3H	North York	Bathurst Manor,Downsview North,Wilson Heights
M3J	North York	Northwood Park,York University
M3K	North York	CFB Toronto,Downsview East
M3L	North York	Downsview West
M3M	North York	Downsview Central
M3N	North York	Downsview Northwest
M4A	North York	Victoria Village
M4B	East York	Woodbine Gardens,Parkview Hill
M4C	East York	Woodbine Heights
M4E	East Toronto	The Beaches
M4G	East York	Leaside
M4H	East York	Thorncliffe Park
M4J	East York	East Toronto
M4K	East Toronto	The Danforth West,Riverdale
M4L	East Toronto	The Beaches West,India Bazaar
M4M	East Toronto	Studio District
M4N	Central Toronto	Lawrence Park
M4P	Central Toronto	Davisville North
M4R	Central Toronto	North Toronto West
M4S	Central Toronto	Davisville
M4T	Central Toronto	Moore Park,Summerhill East
M4V	Central Toronto	Deer Park,Forest Hill SE,Rathnelly,South Hill,...
M4W	Downtown Toronto	Rosedale

Postcode	Borough	Neighbourhood
M4X	Downtown Toronto	Cabbagetown
M4Y	Downtown Toronto	Church and Wellesley
M5A	Downtown Toronto	Harbourfront
M5B	Downtown Toronto	Ryerson,Garden District
M5E	Downtown Toronto	Berczy Park
M5G	Downtown Toronto	Central Bay Street
M5H	Downtown Toronto	Adelaide,King,Richmond
M5J	Downtown Toronto	Harbourfront East,Toronto Islands,Union Station
M5K	Downtown Toronto	Design Exchange,Toronto Dominion Centre
M5L	Downtown Toronto	Commerce Court,Victoria Hotel
M5M	North York	Bedford Park,Lawrence Manor East
M5N	Central Toronto	Roselawn
M5P	Central Toronto	Forest Hill North,Forest Hill West
M5R	Central Toronto	The Annex,North Midtown,Yorkville
M5S	Downtown Toronto	Harbord,University of Toronto
M5T	Downtown Toronto	Chinatown,Grange Park,Kensington Market
M5V	Downtown Toronto	CN Tower,Bathurst Quay,Island airport,Harbourf...
M5W	Downtown Toronto	Stn A PO Boxes 25 The Esplanade
M5X	Downtown Toronto	First Canadian Place,Underground city
M6A	North York	Lawrence Heights,Lawrence Manor
M6B	North York	Glencairn
M6C	York	Humewood-Cedarvale
M6E	York	Caledonia-Fairbanks
M6G	Downtown Toronto	Christie
M6H	West Toronto	Dovercourt Village,Dufferin
M6J	West Toronto	Little Portugal,Trinity
M6K	West Toronto	Brockton,Exhibition Place,Parkdale Village
M6L	North York	Downsview,North Park,Upwood Park
M6M	York	Del Ray,Keelesdale,Mount Dennis,Silverthorn
M6N	York	The Junction North
M6P	West Toronto	High Park,The Junction South
M6R	West Toronto	Parkdale,Roncesvalles
M6S	West Toronto	Swansea
M7A	Downtown Toronto	Queen's Park
M7R	Mississauga	Canada Post Gateway Processing Centre
M7Y	East Toronto	Business Reply Mail Processing Centre 969 Eastern
M8V	Etobicoke	Humber Bay Shores,Mimico South,New Toronto
M8W	Etobicoke	Alderwood,Long Branch
M8X	Etobicoke	The Kingsway,Montgomery Road,Old Mill North
M8Y	Etobicoke	Humber Bay,King's Mill Park,Kingsway Park Sout...
M8Z	Etobicoke	Kingsway Park South West,Mimico NW,The Queensw...
M9A	Etobicoke	Islington Avenue
M9B	Etobicoke	Cloverdale,Islington,Martin Grove,Princess Gar...
M9C	Etobicoke	Bloordale Gardens,Eringate,Markland Wood,Old B...
M9L	North York	Humber Summit
M9M	North York	Emery,Humberlea
M9N	York	Weston
M9P	Etobicoke	Westmount
M9R	Etobicoke	Kingsview Village,Martin Grove Gardens,Richvie...
M9V	Etobicoke	Albion Gardens,Beaumond Heights,Humbergate,Jam...
M9W	Etobicoke	Northwest

All the neighbourhoods are now concatenated under the postcode they are associated with. However, instead of regular index, the table has "Postcode" and "Borough" as its multi-column index. I am going to fix it below.

```
postcode_df.reset_index(inplace=True)
postcode_df
```

	Postcode	Borough	Neighbourhood
0	M1B	Scarborough	Rouge,Malvern
1	M1C	Scarborough	Highland Creek,Rouge Hill,Port Union
2	M1E	Scarborough	Guildwood,Morningside,West Hill
3	M1G	Scarborough	Woburn
4	M1H	Scarborough	Cedarbrae
5	M1J	Scarborough	Scarborough Village
6	M1K	Scarborough	East Birchmount Park,Ionview,Kennedy Park
7	M1L	Scarborough	Clairlea,Golden Mile,Oakridge
8	M1M	Scarborough	Cliffcrest,Cliffside,Scarborough Village West
9	M1N	Scarborough	Birch Cliff,Cliffside West
10	M1P	Scarborough	Dorset Park,Scarborough Town Centre,Wexford He...
11	M1R	Scarborough	Maryvale,Wexford
12	M1S	Scarborough	Agincourt
13	M1T	Scarborough	Clarks Corners,Sullivan,Tam O'Shanter
14	M1V	Scarborough	Agincourt North,L'Amoreaux East,Milliken,Steel...
15	M1W	Scarborough	L'Amoreaux West
16	M1X	Scarborough	Upper Rouge
17	M2H	North York	Hillcrest Village
18	M2J	North York	Fairview,Henry Farm,Oriole
19	M2K	North York	Bayview Village
20	M2L	North York	Silver Hills,York Mills
21	M2M	North York	Newtonbrook,Willowdale
22	M2N	North York	Willowdale South
23	M2P	North York	York Mills West
24	M2R	North York	Willowdale West
25	M3A	North York	Parkwoods
26	M3B	North York	Don Mills North
27	M3C	North York	Flemingdon Park,Don Mills South
28	M3H	North York	Bathurst Manor,Downsview North,Wilson Heights
29	M3J	North York	Northwood Park,York University
30	M3K	North York	CFB Toronto,Downsview East
31	M3L	North York	Downsview West
32	M3M	North York	Downsview Central
33	M3N	North York	Downsview Northwest
34	M4A	North York	Victoria Village
35	M4B	East York	Woodbine Gardens,Parkview Hill
36	M4C	East York	Woodbine Heights
37	M4E	East Toronto	The Beaches
38	M4G	East York	Leaside
39	M4H	East York	Thorncliffe Park
40	M4J	East York	East Toronto
41	M4K	East Toronto	The Danforth West,Riverdale
42	M4L	East Toronto	The Beaches West,India Bazaar
43	M4M	East Toronto	Studio District
44	M4N	Central Toronto	Lawrence Park
45	M4P	Central Toronto	Davisville North
46	M4R	Central Toronto	North Toronto West
47	M4S	Central Toronto	Davisville
48	M4T	Central Toronto	Moore Park,Summerhill East
49	M4V	Central Toronto	Deer Park,Forest Hill SE,Rathnelly,South Hill,...
50	M4W	Downtown Toronto	Rosedale
51	M4X	Downtown Toronto	Cabbagetown

	Postcode	Borough	Neighbourhood
52	M4Y	Downtown Toronto	Church and Wellesley
53	M5A	Downtown Toronto	Harbourfront
54	M5B	Downtown Toronto	Ryerson,Garden District
55	M5E	Downtown Toronto	Berczy Park
56	M5G	Downtown Toronto	Central Bay Street
57	M5H	Downtown Toronto	Adelaide,King,Richmond
58	M5J	Downtown Toronto	Harbourfront East,Toronto Islands,Union Station
59	M5K	Downtown Toronto	Design Exchange,Toronto Dominion Centre
60	M5L	Downtown Toronto	Commerce Court,Victoria Hotel
61	M5M	North York	Bedford Park,Lawrence Manor East
62	M5N	Central Toronto	Roselawn
63	M5P	Central Toronto	Forest Hill North,Forest Hill West
64	M5R	Central Toronto	The Annex,North Midtown,Yorkville
65	M5S	Downtown Toronto	Harbord,University of Toronto
66	M5T	Downtown Toronto	Chinatown,Grange Park,Kensington Market
67	M5V	Downtown Toronto	CN Tower,Bathurst Quay,Island airport,Harbourf...
68	M5W	Downtown Toronto	Stn A PO Boxes 25 The Esplanade
69	M5X	Downtown Toronto	First Canadian Place,Underground city
70	M6A	North York	Lawrence Heights,Lawrence Manor
71	M6B	North York	Glencairn
72	M6C	York	Humewood-Cedarvale
73	M6E	York	Caledonia-Fairbanks
74	M6G	Downtown Toronto	Christie
75	M6H	West Toronto	Dovercourt Village,Dufferin
76	M6J	West Toronto	Little Portugal,Trinity
77	M6K	West Toronto	Brockton,Exhibition Place,Parkdale Village
78	M6L	North York	Downsview,North Park,Upwood Park
79	M6M	York	Del Ray,Keelesdale,Mount Dennis,Silverthorn
80	M6N	York	The Junction North
81	M6P	West Toronto	High Park,The Junction South
82	M6R	West Toronto	Parkdale,Roncesvalles
83	M6S	West Toronto	Swansea
84	M7A	Downtown Toronto	Queen's Park
85	M7R	Mississauga	Canada Post Gateway Processing Centre
86	M7Y	East Toronto	Business Reply Mail Processing Centre 969 Eastern
87	M8V	Etobicoke	Humber Bay Shores,Mimico South,New Toronto
88	M8W	Etobicoke	Alderwood,Long Branch
89	M8X	Etobicoke	The Kingsway,Montgomery Road,Old Mill North
90	M8Y	Etobicoke	Humber Bay,King's Mill Park,Kingsway Park Sout...
91	M8Z	Etobicoke	Kingsway Park South West,Mimico NW,The Queensw...
92	M9A	Etobicoke	Islington Avenue
93	M9B	Etobicoke	Cloverdale,Islington,Martin Grove,Princess Gar...
94	M9C	Etobicoke	Bloordale Gardens,Eringate,Markland Wood,Old B...
95	M9L	North York	Humber Summit
96	M9M	North York	Emery,Humberlea
97	M9N	York	Weston
98	M9P	Etobicoke	Westmount
99	M9R	Etobicoke	Kingsview Village,Martin Grove Gardens,Richvie...
100	M9V	Etobicoke	Albion Gardens,Beaumond Heights,Humbergate,Jam...
101	M9W	Etobicoke	Northwest

In [78]:
```python
print(postcode_df.shape)
```

```
(102, 3)
```

Next, find the latitude and longitude coordinates for each postcode.

In [79]:
```python
# Load the "Geospatial_Coordinates.csv" dataset
Geo_df=pd.read_csv("Geospatial_Coordinates.csv")
Geo_df.head()
```

Out[79]:

	Postcode	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 [80]:  # Merge "postcode_df" and "Geo_df".
          # Use "left" join so that all records from "postcode_df" and matched records from "Geo_df" are returned.
          Geomerge_df=pd.merge(postcode_df,Geo_df, how="left")
          Geomerge_df
```

Out[80]:

	Postcode	Borough	Neighbourhood	Latitude	Longitude
0	M1B	Scarborough	Rouge,Malvern	43.806686	-79.194353
1	M1C	Scarborough	Highland Creek,Rouge Hill,Port Union	43.784535	-79.160497
2	M1E	Scarborough	Guildwood,Morningside,West Hill	43.763573	-79.188711
3	M1G	Scarborough	Woburn	43.770992	-79.216917
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476
5	M1J	Scarborough	Scarborough Village	43.744734	-79.239476
6	M1K	Scarborough	East Birchmount Park,Ionview,Kennedy Park	43.727929	-79.262029
7	M1L	Scarborough	Clairlea,Golden Mile,Oakridge	43.711112	-79.284577
8	M1M	Scarborough	Cliffcrest,Cliffside,Scarborough Village West	43.716316	-79.239476
9	M1N	Scarborough	Birch Cliff,Cliffside West	43.692657	-79.264848
10	M1P	Scarborough	Dorset Park,Scarborough Town Centre,Wexford He...	43.757410	-79.273304
11	M1R	Scarborough	Maryvale,Wexford	43.750072	-79.295849
12	M1S	Scarborough	Agincourt	43.794200	-79.262029
13	M1T	Scarborough	Clarks Corners,Sullivan,Tam O'Shanter	43.781638	-79.304302
14	M1V	Scarborough	Agincourt North,L'Amoreaux East,Milliken,Steel...	43.815252	-79.284577
15	M1W	Scarborough	L'Amoreaux West	43.799525	-79.318389
16	M1X	Scarborough	Upper Rouge	43.836125	-79.205636
17	M2H	North York	Hillcrest Village	43.803762	-79.363452
18	M2J	North York	Fairview,Henry Farm,Oriole	43.778517	-79.346556
19	M2K	North York	Bayview Village	43.786947	-79.385975
20	M2L	North York	Silver Hills,York Mills	43.757490	-79.374714
21	M2M	North York	Newtonbrook,Willowdale	43.789053	-79.408493
22	M2N	North York	Willowdale South	43.770120	-79.408493
23	M2P	North York	York Mills West	43.752758	-79.400049
24	M2R	North York	Willowdale West	43.782736	-79.442259
25	M3A	North York	Parkwoods	43.753259	-79.329656
26	M3B	North York	Don Mills North	43.745906	-79.352188
27	M3C	North York	Flemingdon Park,Don Mills South	43.725900	-79.340923
28	M3H	North York	Bathurst Manor,Downsview North,Wilson Heights	43.754328	-79.442259
29	M3J	North York	Northwood Park,York University	43.767980	-79.487262
30	M3K	North York	CFB Toronto,Downsview East	43.737473	-79.464763
31	M3L	North York	Downsview West	43.739015	-79.506944
32	M3M	North York	Downsview Central	43.728496	-79.495697
33	M3N	North York	Downsview Northwest	43.761631	-79.520999
34	M4A	North York	Victoria Village	43.725882	-79.315572
35	M4B	East York	Woodbine Gardens,Parkview Hill	43.706397	-79.309937
36	M4C	East York	Woodbine Heights	43.695344	-79.318389
37	M4E	East Toronto	The Beaches	43.676357	-79.293031
38	M4G	East York	Leaside	43.709060	-79.363452
39	M4H	East York	Thorncliffe Park	43.705369	-79.349372
40	M4J	East York	East Toronto	43.685347	-79.338106
41	M4K	East Toronto	The Danforth West,Riverdale	43.679557	-79.352188
42	M4L	East Toronto	The Beaches West,India Bazaar	43.668999	-79.315572
43	M4M	East Toronto	Studio District	43.659526	-79.340923
44	M4N	Central Toronto	Lawrence Park	43.728020	-79.388790
45	M4P	Central Toronto	Davisville North	43.712751	-79.390197
46	M4R	Central Toronto	North Toronto West	43.715383	-79.405678
47	M4S	Central Toronto	Davisville	43.704324	-79.388790
48	M4T	Central Toronto	Moore Park,Summerhill East	43.689574	-79.383160
49	M4V	Central Toronto	Deer Park,Forest Hill SE,Rathnelly,South Hill,...	43.686412	-79.400049

	Postcode	Borough	Neighbourhood	Latitude	Longitude
50	M4W	Downtown Toronto	Rosedale	43.679563	-79.377529
51	M4X	Downtown Toronto	Cabbagetown	43.667967	-79.367675
52	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160
53	M5A	Downtown Toronto	Harbourfront	43.654260	-79.360636
54	M5B	Downtown Toronto	Ryerson,Garden District	43.657162	-79.378937
55	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306
56	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383
57	M5H	Downtown Toronto	Adelaide,King,Richmond	43.650571	-79.384568
58	M5J	Downtown Toronto	Harbourfront East,Toronto Islands,Union Station	43.640816	-79.381752
59	M5K	Downtown Toronto	Design Exchange,Toronto Dominion Centre	43.647177	-79.381576
60	M5L	Downtown Toronto	Commerce Court,Victoria Hotel	43.648198	-79.379817
61	M5M	North York	Bedford Park,Lawrence Manor East	43.733283	-79.419750
62	M5N	Central Toronto	Roselawn	43.711695	-79.416936
63	M5P	Central Toronto	Forest Hill North,Forest Hill West	43.696948	-79.411307
64	M5R	Central Toronto	The Annex,North Midtown,Yorkville	43.672710	-79.405678
65	M5S	Downtown Toronto	Harbord,University of Toronto	43.662696	-79.400049
66	M5T	Downtown Toronto	Chinatown,Grange Park,Kensington Market	43.653206	-79.400049
67	M5V	Downtown Toronto	CN Tower,Bathurst Quay,Island airport,Harbourf...	43.628947	-79.394420
68	M5W	Downtown Toronto	Stn A PO Boxes 25 The Esplanade	43.646435	-79.374846
69	M5X	Downtown Toronto	First Canadian Place,Underground city	43.648429	-79.382280
70	M6A	North York	Lawrence Heights,Lawrence Manor	43.718518	-79.464763
71	M6B	North York	Glencairn	43.709577	-79.445073
72	M6C	York	Humewood-Cedarvale	43.693781	-79.428191
73	M6E	York	Caledonia-Fairbanks	43.689026	-79.453512
74	M6G	Downtown Toronto	Christie	43.669542	-79.422564
75	M6H	West Toronto	Dovercourt Village,Dufferin	43.669005	-79.442259
76	M6J	West Toronto	Little Portugal,Trinity	43.647927	-79.419750
77	M6K	West Toronto	Brockton,Exhibition Place,Parkdale Village	43.636847	-79.428191
78	M6L	North York	Downsview,North Park,Upwood Park	43.713756	-79.490074
79	M6M	York	Del Ray,Keelesdale,Mount Dennis,Silverthorn	43.691116	-79.476013
80	M6N	York	The Junction North	43.673185	-79.487262
81	M6P	West Toronto	High Park,The Junction South	43.661608	-79.464763
82	M6R	West Toronto	Parkdale,Roncesvalles	43.648960	-79.456325
83	M6S	West Toronto	Swansea	43.651571	-79.484450
84	M7A	Downtown Toronto	Queen's Park	43.662301	-79.389494
85	M7R	Mississauga	Canada Post Gateway Processing Centre	43.636966	-79.615819
86	M7Y	East Toronto	Business Reply Mail Processing Centre 969 Eastern	43.662744	-79.321558
87	M8V	Etobicoke	Humber Bay Shores,Mimico South,New Toronto	43.605647	-79.501321
88	M8W	Etobicoke	Alderwood,Long Branch	43.602414	-79.543484
89	M8X	Etobicoke	The Kingsway,Montgomery Road,Old Mill North	43.653654	-79.506944
90	M8Y	Etobicoke	Humber Bay,King's Mill Park,Kingsway Park Sout...	43.636258	-79.498509
91	M8Z	Etobicoke	Kingsway Park South West,Mimico NW,The Queensw...	43.628841	-79.520999
92	M9A	Etobicoke	Islington Avenue	43.667856	-79.532242
93	M9B	Etobicoke	Cloverdale,Islington,Martin Grove,Princess Gar...	43.650943	-79.554724
94	M9C	Etobicoke	Bloordale Gardens,Eringate,Markland Wood,Old B...	43.643515	-79.577201
95	M9L	North York	Humber Summit	43.756303	-79.565963
96	M9M	North York	Emery,Humberlea	43.724766	-79.532242
97	M9N	York	Weston	43.706876	-79.518188
98	M9P	Etobicoke	Westmount	43.696319	-79.532242
99	M9R	Etobicoke	Kingsview Village,Martin Grove Gardens,Richvie...	43.688905	-79.554724
100	M9V	Etobicoke	Albion Gardens,Beaumond Heights,Humbergate,Jam...	43.739416	-79.588437
101	M9W	Etobicoke	Northwest	43.706748	-79.594054

Cluster the neighbourhoods and create visualizations.

Import the necessary libraries first.

In [81]:
```python
import pandas as pd
pd.set_option('display.max_columns', None)
pd.set_option('display.max_rows', None)

import json # library to handle JSON files

#!conda install -c conda-forge geopy --yes # uncomment this line if you haven't completed the Foursquare API lab
from geopy.geocoders import Nominatim # convert an address into latitude and longitude values

import requests # library to handle requests
from pandas.io.json import json_normalize # tranform JSON file into a pandas dataframe

# Matplotlib and associated plotting modules
import matplotlib.cm as cm
import matplotlib.colors as colors

# import k-means from clustering stage
from sklearn.cluster import KMeans

#!conda install -c conda-forge folium=0.5.0 --yes # uncomment this line if you haven't completed the Foursquare API lab
import folium # map rendering library
```

Get the latitude and longitude for Toronto City.

In [82]:
```python
address = 'Toronto, Canada'

geolocator = Nominatim(user_agent="ny_explorer")
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geograpical coordinate of Toronto, Canada are {}, {}.'.format(latitude, longitude))
```

The geograpical coordinate of Toronto, Canada are 43.653963, -79.387207.

Create a map of Toronto City with the neighbourhoods superimposed on it.

```
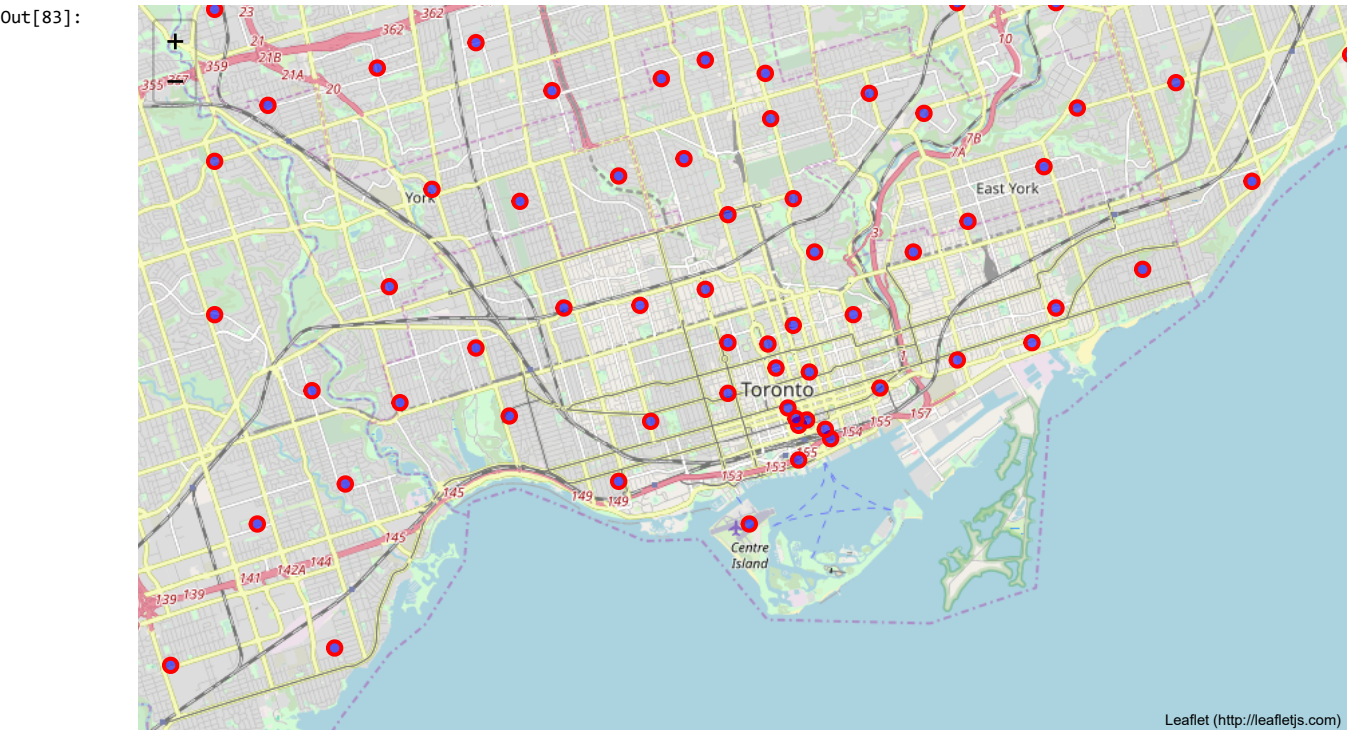In [83]:    # Toronto latitude and longitude.
            latitude= 43.653963
            longitude=-79.387207

            # create map and display it
            toronto_map = folium.Map(location=[latitude, longitude], zoom_start=12)

            # instantiate a feature group for the incidents in the dataframe
            incidents = folium.map.FeatureGroup()

            # loop through the 100 crimes and add each to the incidents feature group
            for lat, lng, in zip(Geomerge_df.Latitude, Geomerge_df.Longitude):
                incidents.add_child(
                    folium.features.CircleMarker(
                        [lat, lng],
                        radius=5, # define how big you want the circle markers to be
                        color='red',
                        fill=True,
                        fill_color='blue',
                        fill_opacity=0.6
                    )
                )

            # add incidents to map
            toronto_map.add_child(incidents)
```

Out[83]:

Leaflet (http://leafletjs.com)

Simplify the above map and segment and cluster only the neighborhoods with "Toronto" in their borough names. These are the neighbourhoods associated with the following boroughs: Central Toronto, Downtown Toronto, East Toronto and West Toronta. So let's slice the original dataframe and create a new dataframe for the Toronto boroughs.

```
In [84]:    toronto_data1 = Geomerge_df[Geomerge_df['Borough'] == 'Central Toronto'].reset_index(drop=True)
            toronto_data2 = Geomerge_df[Geomerge_df['Borough'] == 'Downtown Toronto'].reset_index(drop=True)
            toronto_data3 = Geomerge_df[Geomerge_df['Borough'] == 'East Toronto'].reset_index(drop=True)
            toronto_data4 = Geomerge_df[Geomerge_df['Borough'] == 'West Toronto'].reset_index(drop=True)
            frames=[toronto_data1,toronto_data2,toronto_data3,toronto_data4]
            toronto_data=pd.concat(frames)
            toronto_data.head()
```

Out[84]:

	Postcode	Borough	Neighbourhood	Latitude	Longitude
0	M4N	Central Toronto	Lawrence Park	43.728020	-79.388790
1	M4P	Central Toronto	Davisville North	43.712751	-79.390197
2	M4R	Central Toronto	North Toronto West	43.715383	-79.405678
3	M4S	Central Toronto	Davisville	43.704324	-79.388790
4	M4T	Central Toronto	Moore Park,Summerhill East	43.689574	-79.383160

Let's get the latitude and longitude for Downtown Toronto.

Visualize the Toronto boroughs with their neighbourhoods.

In [85]:

```
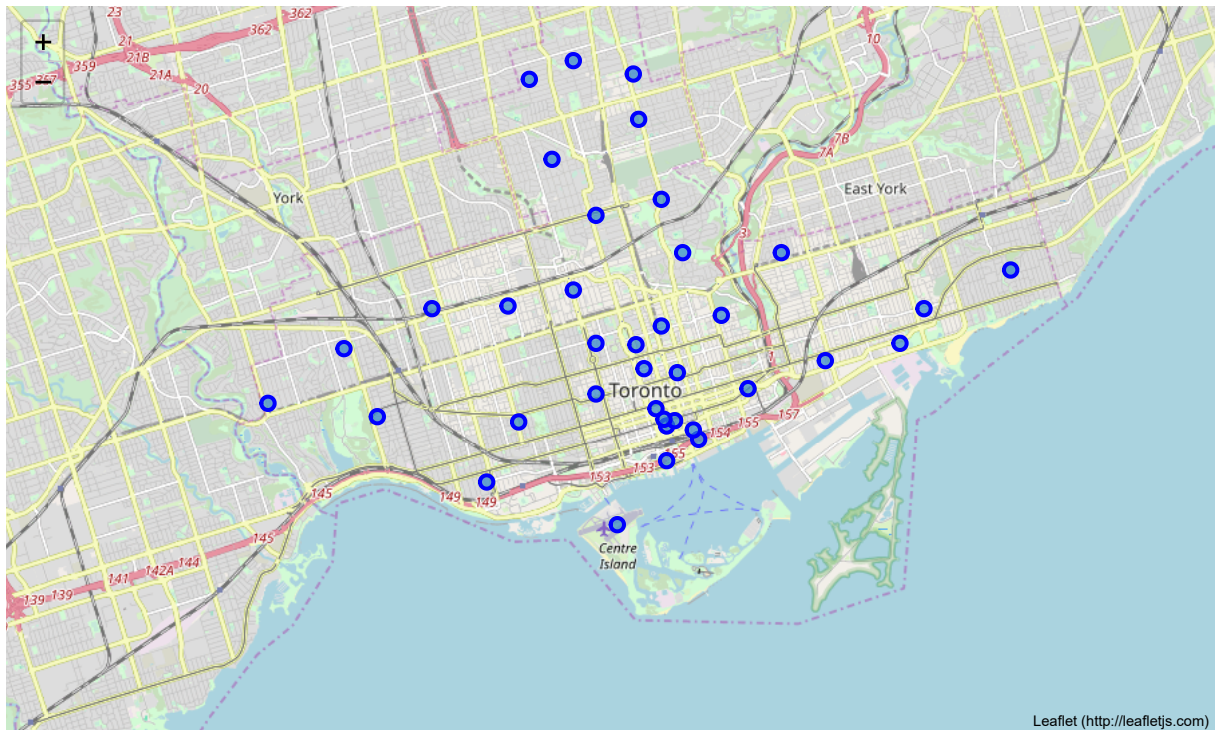# create map of Manhattan using latitude and longitude values
toronto_map = folium.Map(location=[latitude, longitude], zoom_start=12)

# add markers to map
for lat, lng, label in zip(toronto_data['Latitude'], toronto_data['Longitude'], toronto_data['Borough']):
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        color='blue',
        fill=True,
        fill_color='#3186cc',
        fill_opacity=0.7,
        parse_html=False).add_to(toronto_map)

toronto_map
```

Out[85]:

Let's cluster the neighbourhoods in Downtown Toronto.

```
In [86]:    # Onehot coding
            toronto_data_onehot = pd.get_dummies(toronto_data[['Neighbourhood']], prefix="", prefix_sep="")

            # add Borough column back to dataframe
            toronto_data_onehot['Borough'] = toronto_data['Borough']

            # move neighborhood column to the first column
            fixed_columns = [toronto_data_onehot.columns[-1]] + list(toronto_data_onehot.columns[:-1])
            toronto_data_onehot = toronto_data_onehot[fixed_columns]

            toronto_data_onehot.head(10)
```

Out[86]:

	Borough	Adelaide,King,Richmond	Berczy Park	Brockton,Exhibition Place,Parkdale Village	Business Reply Mail Processing Centre 969 Eastern	CN Tower,Bathurst Quay,Island airport,Harbourfront West,King and Spadina,Railway Lands,South Niagara	Cabbagetown	Central Bay Street	Chinatown,Grange Park,Kensington Market	Ch
0	Central Toronto	0	0	0	0	0	0	0	0	
1	Central Toronto	0	0	0	0	0	0	0	0	
2	Central Toronto	0	0	0	0	0	0	0	0	
3	Central Toronto	0	0	0	0	0	0	0	0	
4	Central Toronto	0	0	0	0	0	0	0	0	
5	Central Toronto	0	0	0	0	0	0	0	0	
6	Central Toronto	0	0	0	0	0	0	0	0	
7	Central Toronto	0	0	0	0	0	0	0	0	
8	Central Toronto	0	0	0	0	0	0	0	0	
0	Downtown Toronto	0	0	0	0	0	0	0	0	

Group the rows by boroughs and by taking the mean of the frequency of occurrence of each category.

```
In [87]:    toronto_grouped =toronto_data_onehot.groupby('Borough').mean().reset_index()
            toronto_grouped
```

Out[87]:

	Borough	Adelaide,King,Richmond	Berczy Park	Brockton,Exhibition Place,Parkdale Village	Business Reply Mail Processing Centre 969 Eastern	CN Tower,Bathurst Quay,Island airport,Harbourfront West,King and Spadina,Railway Lands,South Niagara	Cabbagetown	Central Bay Street	Chinatown,Grange Park,Kensington Market
0	Central Toronto	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.000000
1	Downtown Toronto	0.055556	0.055556	0.000000	0.0	0.055556	0.055556	0.055556	0.055556
2	East Toronto	0.000000	0.000000	0.000000	0.2	0.000000	0.000000	0.000000	0.000000
3	West Toronto	0.000000	0.000000	0.166667	0.0	0.000000	0.000000	0.000000	0.000000

Cluster the neighbourhoods in Central Toronto, Downtown Toronto, East Toronto and West Toronto.

```
In [88]:    from sklearn.cluster import KMeans
            # set number of clusters
            kclusters = 4

            toronto_clustering = toronto_grouped.drop('Borough', 1)

            # run k-means clustering
            kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(toronto_clustering)

            # check cluster labels generated for each row in the dataframe
            kmeans.labels_[0:3]
```

Out[88]:   array([0, 3, 1])

Visualize the clusters of the neighbourhoods in Central Toronto, Downtown Toronto, East Toronto and West Toronto on the map.

```
In [89]:    # Add cluster labels
            toronto_grouped.insert(0,"Cluster label",kmeans.labels_)
```

```
In [90]:    toronto_data=toronto_data.join(toronto_grouped.set_index("Borough"),on="Borough")
            toronto_data.head()
```

Out[90]:

	Postcode	Borough	Neighbourhood	Latitude	Longitude	Cluster label	Adelaide,King,Richmond	Berczy Park	Brockton,Exhibition Place,Parkdale Village	Business Reply Mail Processing Centre 969 Eastern	CN Tow (airport,H We Spad L
0	M4N	Central Toronto	Lawrence Park	43.728020	-79.388790	0		0.0	0.0	0.0	0.0
1	M4P	Central Toronto	Davisville North	43.712751	-79.390197	0		0.0	0.0	0.0	0.0
2	M4R	Central Toronto	North Toronto West	43.715383	-79.405678	0		0.0	0.0	0.0	0.0
3	M4S	Central Toronto	Davisville	43.704324	-79.388790	0		0.0	0.0	0.0	0.0
4	M4T	Central Toronto	Moore Park,Summerhill East	43.689574	-79.383160	0		0.0	0.0	0.0	0.0

In [95]:

```
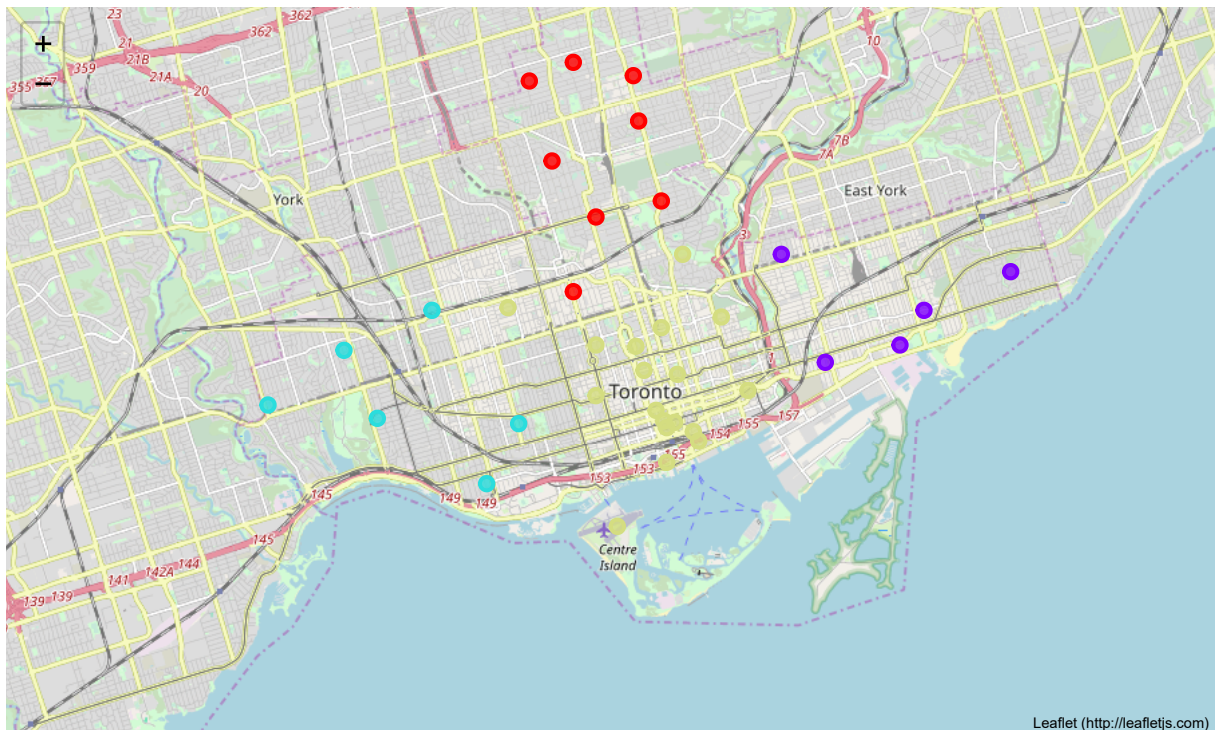# create map
map_clusters = folium.Map(location=[latitude, longitude], zoom_start=11.5)

# set color scheme for the clusters
x = np.arange(kclusters)
ys = [i + x + (i*x)**2 for i in range(kclusters)]
colors_array = cm.rainbow(np.linspace(0, 1, len(ys)))
rainbow = [colors.rgb2hex(i) for i in colors_array]

# add markers to the map
markers_colors = []
for lat, lon, poi, cluster in zip(toronto_data['Latitude'], toronto_data['Longitude'], toronto_data['Borough'],toronto_data[
    label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), parse_html=True)
    folium.CircleMarker(
        [lat, lon],
        radius=5,
        popup=label,
        color=rainbow[cluster-1],
        fill=True,
        fill_color=rainbow[cluster-1],
        fill_opacity=0.8).add_to(map_clusters)

map_clusters
```

Out[95]:

There we have the map of Toronto City with the four boroughs with Toronton in their names (Central Toronto, Downtown Toronto, East Toronto and West Toronto) superimposed on it. Each color represents a different borough, while each dot represents a neighbbourhood. To see which borough and cluster a dot represents, hover the cursor over the dot and click on it.

Thanks for reading!

In [ ]: