

Segmenting_and_Clustering_Neighborhoods

June 26, 2020

```
[38]: import numpy as np # library to handle data in a vectorized manner

import pandas as pd # library for data analysis
pd.set_option('display.max_columns', None)
pd.set_option('display.max_rows', None)

import json # library to handle JSON files
```

```
[39]: !conda install -c anaconda beautifulsoup4 --yes
```

```
Collecting package metadata (current_repodata.json): done
Solving environment: done
```

```
# All requested packages already installed.
```

```
[40]: !conda install -c anaconda lxml --yes
```

```
Collecting package metadata (current_repodata.json): done
Solving environment: done
```

```
# All requested packages already installed.
```

```
[41]: import requests # library to handle requests

from bs4 import BeautifulSoup
import xml
```

```
[66]: List_url = "https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M"
source = requests.get(List_url).text
```

```
[67]: s = BeautifulSoup(source, 'xml')
```

```
[68]: table=s.find('table')
```

```
[69]: #Table consists of three columns: PostalCode, Borough, and Neighborhood. Now
      ↪make it as a data frame.
      column_names = ['Postalcode', 'Borough', 'Neighborhood']
      df = pd.DataFrame(columns = column_names)
```

```
[70]: for tr_cell in table.find_all('tr'):
      row_data=[]
      for td_cell in tr_cell.find_all('td'):
          row_data.append(td_cell.text.strip())
      if len(row_data)==3:
          df.loc[len(df)] = row_data
```

```
[71]: df.head()
```

```
[71]:   Postalcode      Borough      Neighborhood
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
```

```
[73]: df=df[df['Borough']!='Not assigned']
```

```
[85]: df
```

```
[85]:   Postalcode      Borough \
0         M3A    North York
1         M4A    North York
2         M5A  Downtown Toronto
3         M6A    North York
4         M7A  Downtown Toronto
5         M9A      Etobicoke
6         M1B    Scarborough
7         M3B    North York
8         M4B      East York
9         M5B  Downtown Toronto
10        M6B    North York
11        M9B      Etobicoke
12        M1C    Scarborough
13        M3C    North York
14        M4C      East York
15        M5C  Downtown Toronto
16        M6C      York
17        M9C      Etobicoke
18        M1E    Scarborough
19        M4E    East Toronto
20        M5E  Downtown Toronto
```

21	M6E	York
22	M1G	Scarborough
23	M4G	East York
24	M5G	Downtown Toronto
25	M6G	Downtown Toronto
26	M1H	Scarborough
27	M2H	North York
28	M3H	North York
29	M4H	East York
30	M5H	Downtown Toronto
31	M6H	West Toronto
32	M1J	Scarborough
33	M2J	North York
34	M3J	North York
35	M4J	East York
36	M5J	Downtown Toronto
37	M6J	West Toronto
38	M1K	Scarborough
39	M2K	North York
40	M3K	North York
41	M4K	East Toronto
42	M5K	Downtown Toronto
43	M6K	West Toronto
44	M1L	Scarborough
45	M2L	North York
46	M3L	North York
47	M4L	East Toronto
48	M5L	Downtown Toronto
49	M6L	North York
50	M9L	North York
51	M1M	Scarborough
52	M2M	North York
53	M3M	North York
54	M4M	East Toronto
55	M5M	North York
56	M6M	York
57	M9M	North York
58	M1N	Scarborough
59	M2N	North York
60	M3N	North York
61	M4N	Central Toronto
62	M5N	Central Toronto
63	M6N	York
64	M9N	York
65	M1P	Scarborough
66	M2P	North York
67	M4P	Central Toronto

68	M5P	Central Toronto
69	M6P	West Toronto
70	M9P	Etobicoke
71	M1R	Scarborough
72	M2R	North York
73	M4R	Central Toronto
74	M5R	Central Toronto
75	M6R	West Toronto
76	M7R	Mississauga
77	M9R	Etobicoke
78	M1S	Scarborough
79	M4S	Central Toronto
80	M5S	Downtown Toronto
81	M6S	West Toronto
82	M1T	Scarborough
83	M4T	Central Toronto
84	M5T	Downtown Toronto
85	M1V	Scarborough
86	M4V	Central Toronto
87	M5V	Downtown Toronto
88	M8V	Etobicoke
89	M9V	Etobicoke
90	M1W	Scarborough
91	M4W	Downtown Toronto
92	M5W	Downtown Toronto
93	M8W	Etobicoke
94	M9W	Etobicoke
95	M1X	Scarborough
96	M4X	Downtown Toronto
97	M5X	Downtown Toronto
98	M8X	Etobicoke
99	M4Y	Downtown Toronto
100	M7Y	East Toronto
101	M8Y	Etobicoke
102	M8Z	Etobicoke

	Neighborhood
0	Parkwoods
1	Victoria Village
2	Regent Park, Harbourfront
3	Lawrence Manor, Lawrence Heights
4	Queen's Park, Ontario Provincial Government
5	Islington Avenue, Humber Valley Village
6	Malvern, Rouge
7	Don Mills
8	Parkview Hill, Woodbine Gardens
9	Garden District, Ryerson

10 Glencairn
 11 West Deane Park, Princess Gardens, Martin Grov...
 12 Rouge Hill, Port Union, Highland Creek
 13 Don Mills
 14 Woodbine Heights
 15 St. James Town
 16 Humewood-Cedarvale
 17 Eringate, Bloordale Gardens, Old Burnhamthorpe...
 18 Guildwood, Morningside, West Hill
 19 The Beaches
 20 Berczy Park
 21 Caledonia-Fairbanks
 22 Woburn
 23 Leaside
 24 Central Bay Street
 25 Christie
 26 Cedarbrae
 27 Hillcrest Village
 28 Bathurst Manor, Wilson Heights, Downsview North
 29 Thorncliffe Park
 30 Richmond, Adelaide, King
 31 Dufferin, Dovercourt Village
 32 Scarborough Village
 33 Fairview, Henry Farm, Oriole
 34 Northwood Park, York University
 35 East Toronto, Broadview North (Old East York)
 36 Harbourfront East, Union Station, Toronto Islands
 37 Little Portugal, Trinity
 38 Kennedy Park, Ionview, East Birchmount Park
 39 Bayview Village
 40 Downsview
 41 The Danforth West, Riverdale
 42 Toronto Dominion Centre, Design Exchange
 43 Brockton, Parkdale Village, Exhibition Place
 44 Golden Mile, Clairlea, Oakridge
 45 York Mills, Silver Hills
 46 Downsview
 47 India Bazaar, The Beaches West
 48 Commerce Court, Victoria Hotel
 49 North Park, Maple Leaf Park, Upwood Park
 50 Humber Summit
 51 Cliffside, Cliffcrest, Scarborough Village West
 52 Willowdale, Newtonbrook
 53 Downsview
 54 Studio District
 55 Bedford Park, Lawrence Manor East
 56 Del Ray, Mount Dennis, Keelsdale and Silverthorn

57 Humberlea, Emery
 58 Birch Cliff, Cliffside West
 59 Willowdale, Willowdale East
 60 Downsview
 61 Lawrence Park
 62 Roselawn
 63 Runnymede, The Junction North
 64 Weston
 65 Dorset Park, Wexford Heights, Scarborough Town...
 66 York Mills West
 67 Davisville North
 68 Forest Hill North & West, Forest Hill Road Park
 69 High Park, The Junction South
 70 Westmount
 71 Wexford, Maryvale
 72 Willowdale, Willowdale West
 73 North Toronto West, Lawrence Park
 74 The Annex, North Midtown, Yorkville
 75 Parkdale, Roncesvalles
 76 Canada Post Gateway Processing Centre
 77 Kingsview Village, St. Phillips, Martin Grove ...
 78 Agincourt
 79 Davisville
 80 University of Toronto, Harbord
 81 Runnymede, Swansea
 82 Clarks Corners, Tam O'Shanter, Sullivan
 83 Moore Park, Summerhill East
 84 Kensington Market, Chinatown, Grange Park
 85 Milliken, Agincourt North, Steeles East, L'Amo...
 86 Summerhill West, Rathnelly, South Hill, Forest...
 87 CN Tower, King and Spadina, Railway Lands, Har...
 88 New Toronto, Mimico South, Humber Bay Shores
 89 South Steeles, Silverstone, Humbergate, Jamest...
 90 Steeles West, L'Amoreaux West
 91 Rosedale
 92 Stn A PO Boxes
 93 Alderwood, Long Branch
 94 Northwest, West Humber - Clairville
 95 Upper Rouge
 96 St. James Town, Cabbagetown
 97 First Canadian Place, Underground city
 98 The Kingsway, Montgomery Road, Old Mill North
 99 Church and Wellesley
 100 Business reply mail Processing Centre, South C...
 101 Old Mill South, King's Mill Park, Sunnylea, Hu...
 102 Mimico NW, The Queensway West, South of Bloor,...

```
[83]: df.shape
```

```
[83]: (103, 3)
```