

Segmenting and Clustering Neighbourhoods in Toronto

- Part 2

In [1]:

```
pip install geocoder
```

```
Requirement already satisfied: geocoder in ./anaconda3/lib/python3.8/site-packages (1.38.1)
```

```
Requirement already satisfied: six in ./anaconda3/lib/python3.8/site-packages (from geocoder) (1.15.0)
```

```
Requirement already satisfied: requests in ./anaconda3/lib/python3.8/site-packages (from geocoder) (2.24.0)
```

```
Requirement already satisfied: ratelim in ./anaconda3/lib/python3.8/site-packages (from geocoder) (0.1.6)
```

```
Requirement already satisfied: click in ./anaconda3/lib/python3.8/site-packages (from geocoder) (7.1.2)
```

```
Requirement already satisfied: future in ./anaconda3/lib/python3.8/site-packages (from geocoder) (0.18.2)
```

```
Requirement already satisfied: idna<3,>=2.5 in ./anaconda3/lib/python3.8/site-packages (from requests->geocoder) (2.10)
```

```
Requirement already satisfied: certifi>=2017.4.17 in ./anaconda3/lib/python3.8/site-packages (from requests->geocoder) (2020.6.20)
```

```
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in ./anaconda3/lib/python3.8/site-packages (from requests->geocoder) (1.25.9)
```

```
Requirement already satisfied: chardet<4,>=3.0.2 in ./anaconda3/lib/python3.8/site-packages (from requests->geocoder) (3.0.4)
```

```
Requirement already satisfied: decorator in ./anaconda3/lib/python3.8/site-packages (from ratelim->geocoder) (4.4.2)
```

```
Note: you may need to restart the kernel to use updated packages.
```

In [2]:

```
import pandas as pd
```

```
import numpy as np
```

```
import geocoder
```

```
print("Imported!")
```

```
Imported!
```

In [3]:

```
df = pd.read_csv('toronto.csv')
```

```
df.head()
```

Out[3]:

| | Postcode | Borough | Neighborhood |
|----------|-----------------|-----------------|---|
| 0 | M1 A n | Not assigned | Not assigned |
| 1 | M1 B | Scarborough | Malvern, Rouge |
| 2 | M1 C | Scarborough | Rouge Hill, Port Union , Highland Creek |
| 3 | M1 E | Scarborough | Guildwood, Morningside, West Hill |
| 4 | M1 G n | Scarborough | Woburn |

In [4]:

```
print(df.shape)
df.describe()

(180, 3)
```

Out[4]:

| | Po sta lco de | B or ou gh | Neig hbor hood |
|--|----------------------------------|---------------------------------------|-------------------------------|
| c o u n t | 180 | 180 | 180 |
| u n i q u e | 180 | 11 | 100 |
| t o p | M 2L \n | N ot as si gn ed \n | Not assig ned\n |
| f r e q | 1 | 77 | 77 |

In [5]:

```
def get_latilong(postal_code):
    lati_long_coords = None
    while(lati_long_coords is None):
        g = geocoder.arcgis('{}, Toronto, Ontario'.format(postal_code))
        lati_long_coords = g.latlng
    return lati_long_coords

get_latilong('M4G')
```

Out[5]:

```
[43.7090200000000066, -79.363489999999996]
```

In [6]:

```
# Retrieving Postal Code Co-ordinates
postal_codes = df['Postalcode']
coords = [ get_latilong(postal_code) for postal_code in postal_codes.tolist()
]

Status code Unknown from https://geocode.arcgis.com/arcgis/rest/services/World/GeocodeServer/find: ERROR - HTTPSConnectionPool(host='geocode.arcgis.com',
port=443): Read timed out. (read timeout=5.0)
```

In [7]:

```
# Adding Columns Latitude & Longitude
df_coords = pd.DataFrame(coords, columns=['Latitude', 'Longitude'])
df['Latitude'] = df_coords['Latitude']
df['Longitude'] = df_coords['Longitude']
```

In [8]:

```
df[df.Postalcode == 'M5G']
```

Out[8]:

| P o s t a l c o d e | B o r o u g h | N e i g h b o r h o o d | L a t i t u d e | L o n g i t u d e |
|--|---------------------------------|--|--------------------------------------|---|
|--|---------------------------------|--|--------------------------------------|---|

In [9]:

```
df.head(15)
```

Out[9]:

| | P o s t a l c o d e | B o r o u g h | N e i g h b o r h o o d | L a t i t u d e | L o n g i t u d e |
|---|--|---|---|--------------------------------------|---|
| 0 | M 1 A \ n | N o t a s s i g n e d \ n | N o t a s s i g n e d \ n | 4 3 . 6 4 8 6 9 | - 7 9 . 3 8 5 4 4 |
| 1 | M 1 B \ n | S c a r b o r o u g h \ n | M a l v e r n , R o u g e | 4 3 . 8 1 1 3 9 | - 7 9 . 1 9 6 6 2 |

| | P o s t a l c o d e | B o r o u g h | N e i g h b o r h o o d | L a t i t u d e | L o n g i t u d e |
|---|--|---|---|--------------------------------------|---|
| 2 | M l C \ n | S c a r b o r o u g h \ n | R o u g e H i l l , P o r t U n i o n , H i g h l a n d C r e | 4 3 . 7 8 5 7 4 | - 7 9 . 1 5 8 7 5 |

| | P o s t a l c o d e | B o r o u g h | N e i g h b o r h o o d | L a t i t u d e | L o n g i t u d e |
|---|--|---|---|--------------------------------------|---|
| | | | e k | | |
| 3 | M l E \ n | S c a r b o r o u g h \ n | G u i l d w o o d , M o r n i n g s i d e , W e s t H | 4 3 . 7 6 5 7 5 | - 7 9 . 1 7 4 7 0 |

| | P o s t a l c o d e | B o r o u g h | N e i g h b o r h o o d | L a t i t u d e | L o n g i t u d e |
|---|--|---|--|--------------------------------------|---|
| | | | i l l | | |
| 4 | M l G \ n | S c a r b o r o u g h \ n | W o b u r n | 4 3 . 7 6 8 1 2 | - 7 9 . 2 1 7 6 1 |
| 5 | M l H \ n | S c a r b o r o u g h | C e d a r b r a e | 4 3 . 7 6 9 4 4 | - 7 9 . 2 3 8 9 2 |

| | P o s t a l c o d e | B o r o u g h | N e i g h b o r h o o d | L a t i t u d e | L o n g i t u d e |
|---|--|---|--|--------------------------------------|---|
| | | \ n | | | |
| 6 | M l J \ n | S c a r b o r o u g h \ n | S c a r b o r o u g h V i l l a g e | 4 3 . 7 4 4 4 6 | - 7 9 . 2 3 1 1 7 |
| 7 | M l K \ n | S c a r b o r | K e n n e d y | 4 3 . 7 2 5 | - 7 9 . 2 6 4 |

| | P o s t a l c o d e | B o r o u g h | N e i g h b o r h o o d | L a t i t u d e | L o n g i t u d e |
|--|--|---------------------------------|---|--------------------------------------|---|
| | | o u g h \ n | P a r k , I o n v i e w , E a s t B i r c h m o u n t P a r k | 8 2 | 6 1 |

| | P o s t a l c o d e | B o r o u g h | N e i g h b o r h o o d | L a t i t u d e | L o n g i t u d e |
|---|--|---|--|--------------------------------------|---|
| 8 | M l L \ n | S c a r b o r o u g h \ n | G o l d e n M i l e , C l a i r l e a , O a k r i d g e | 4 3 . 7 1 2 8 9 | - 7 9 . 2 8 5 0 6 |

| | P o s t a l c o d e | B o r o u g h | N e i g h b o r h o o d | L a t i t u d e | L o n g i t u d e |
|---|--|---|---|--------------------------------------|---|
| 9 | M l M \ n | S c a r b o r o u g h \ n | C l i f f s i d e , C l i f f c r e s t , S c a r b o r o u g | 4 3 . 7 2 3 6 0 | - 7 9 . 2 3 4 9 6 |

| | P o s t a l c o d e | B o r o u g h | N e i g h b o r h o o d | L a t i t u d e | L o n g i t u d e |
|--------|--|---|---|--------------------------------------|---|
| | | | h V i l l a g e W e s t | | |
| 1 0 | M l N \ n | S c a r b o r o u g h \ n | B i r c h C l i f f , C l i f f s | 4 3 . 6 9 5 1 0 | - 7 9 . 2 6 4 6 6 |

| | P o s t a l c o d e | B o r o u g h | N e i g h b o r h o o d | L a t i t u d e | L o n g i t u d e |
|--------|--|---|--|--------------------------------------|---|
| | | | i d e W e s t | | |
| 1 1 | M l P \ n | S c a r b o r o u g h \ n | D o r s e t P a r k , W e x f o r d H e i g | 4 3 : 7 5 9 9 8 | - 7 9 : 2 6 9 4 0 |

| | P o s t a l c o d e | B o r o u g h | N e i g h b o r h o o d | L a t i t u d e | L o n g i t u d e |
|--------|--|---------------------------------|---|--------------------------------------|---|
| | | | h t s , S c a r b o r o u g h T o w n . . . | | |
| 1 2 | M 1 R \ n | S c a r b o r | W e x f o r d | 4 3 . 7 5 0 | - 7 9 . 3 0 0 |

| | P o s t a l c o d e | B o r o u g h | N e i g h b o r h o o d | L a t i t u d e | L o n g i t u d e |
|--------|--|---|--|--------------------------------------|---|
| | | o u g h \ n | , M a r y v a l e | 7 5 | 5 4 |
| 1 3 | M l S \ n | S c a r b o r o u g h \ n | A g i n c o u r t | 4 3 . 7 9 4 5 2 | - 7 9 . 2 6 7 0 8 |
| 1 4 | M l T | S c a r b | C l a r k | 4 3 . 7 8 | - 7 9 . 2 |

| | P o s t a l c o d e | B o r o u g h | N e i g h b o r h o o d | L a t i t u d e | L o n g i t u d e |
|--|--|----------------------------------|--|--------------------------------------|---|
| | \n | o r o u g h \n | s C o r n e r s , T a m O , S h a n t e r , S u l l i v a n | 4 9 1 | 9 7 2 2 |

In [10]:

```
df.to_csv('toronto_part2.csv',index=False)
```