

Peer-graded Assignment: Capstone Project - The Battle of Neighborhoods (Week 1)

For this assignment I will be using Toronto web-scraped data for neighbourhoods info, Toronto open datasets for environmental purposes and Foursquare API information.

First two datasets are uploaded as *df1* and *df2*. API requests will be implemented later.

```
In [19]: import pandas as pd
df1 = pd.read_csv('~Downloads/export-2.csv', index_col=[0])
df1.head()
```

Out[19]:

| | 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 [20]: df1 = df[(df["Borough"] != "Not assigned") & (df["Neighbourhood"] != "Not assigned")]
df1.reset_index(drop=True).head()
```

Out[20]:

| | Postcode | Borough | Neighbourhood |
|---|----------|------------------|------------------|
| 0 | M3A | North York | Parkwoods |
| 1 | M4A | North York | Victoria Village |
| 2 | M5A | Downtown Toronto | Harbourfront |
| 3 | M5A | Downtown Toronto | Regent Park |
| 4 | M6A | North York | Lawrence Heights |

```
In [21]: df1.shape
```

Out[21]: (211, 3)

```
In [22]: df2 = pd.read_csv('~Downloads/environment.csv', index_col=None, de
limiter = ';')
df2.head()
```

Out[22]:

| | Neighbourhood | Neighbourhood Id | Green Rebate Programs | Green Spaces | Pollutant Carcinogenic TEP Score | Pollutant Non-Carcinogenic TEP Score | Po R |
|---|-----------------------------------|------------------|-----------------------|--------------|----------------------------------|--------------------------------------|------|
| 0 | West Humber-Clairville | 1 | 428 | 2,078835532 | 5737,87 | 18658529,73 | |
| 1 | Mount Olive-Silverstone-Jamestown | 2 | 250 | 1,048870056 | 29,76 | 2015 | |
| 2 | Thistletown-Beaumont Heights | 3 | 118 | 0,939107957 | 0 | 0 | |
| 3 | Rexdale-Kipling | 4 | 121 | 0,240663012 | 0 | 37632 | |
| 4 | Elms-Old Rexdale | 5 | 73 | 0,730089694 | 0 | 309 | |

```
In [26]: df=df2.merge(df1)
df.head()
```

Out[26]:

| | Neighbourhood | Neighbourhood Id | Green Rebate Programs | Green Spaces | Pollutant Carcinogenic TEP Score | Pollutant Non-Carcinogenic TEP Score | Po R |
|---|---------------|------------------|-----------------------|--------------|----------------------------------|--------------------------------------|------|
| 0 | Markland Wood | 12 | 182 | 0,224255343 | 423,36 | 28665 | |
| 1 | New Toronto | 18 | 109 | 0,593630152 | 0 | 98637 | |
| 2 | Long Branch | 19 | 92 | 0,323676887 | 0 | 3475,8 | |
| 3 | Alderwood | 20 | 192 | 0,292452196 | 107370 | 3099019 | |
| 4 | Humber Summit | 21 | 84 | 0,686271868 | 7785,9 | 24648447,6 | |

In []: