# Data Science Capstone Week 3 Assignment - PART 2

Intent: Create a Dataframe containing Toronto Postalcode Data (wrangled and formed in PART 1 of this project) with longitude and latitude information of that Postalcode. Latitude and Longitude information is procured from http://cocl.us/Geospatial_data (http://cocl.us/Geospatial_data)

## Read in Dataframe formed in PART I of this project

In [3]:

```
import pandas as pd
import numpy as np
```

In [14]:

```
part1DF = pd.read_csv('/resources/data/DataScienceCAPSTONE/Week3/regionsDF.csv')
part1DF.head()
```

Out[14]:

| | Unnamed: 0 | Postcode | Borough | Neighbourhood |
|---|---|---|---|---|
| 0 | 0 | M3A | North York | Parkwoods |
| 1 | 1 | M4A | North York | Victoria Village |
| 2 | 2 | M9A | Etobicoke | Islington Avenue |
| 3 | 3 | M3B | North York | Don Mills North |
| 4 | 4 | M6B | North York | Glencairn |

### Read in Longitude and Latitude Coordinates procured from http://cocl.us/Geospatial_data and placed in local IBM repository

In [22]:

```
longLatDF = pd.read_csv('/resources/data/Geospatial_Coordinates.csv')
longLatDF.columns=['Postcode', 'Latitude', 'Longitude']
longLatDF.head()
```

Out[22]:

| | 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 |

# Standardize Column Headers and then MERGE the two dataframes into final version

In [26]:

```
neighInfoDF = part1DF[['Postcode','Borough','Neighbourhood']]
neighInfoDF.head()
```

Out[26]:

| | Postcode | Borough | Neighbourhood |
|---|---|---|---|
| 0 | M3A | North York | Parkwoods |
| 1 | M4A | North York | Victoria Village |
| 2 | M9A | Etobicoke | Islington Avenue |
| 3 | M3B | North York | Don Mills North |
| 4 | M6B | North York | Glencairn |

In [27]:

```
#result = pd.merge(left, right, on='key')

completeNeighDF = pd.merge(neighInfoDF, longLatDF, on='Postcode')
completeNeighDF.head(15)
```

Out[27]:

| | Postcode | Borough | Neighbourhood | Latitude | Longitude |
|---|---|---|---|---|---|
| 0 | M3A | North York | Parkwoods | 43.753259 | -79.329656 |
| 1 | M4A | North York | Victoria Village | 43.725882 | -79.315572 |
| 2 | M9A | Etobicoke | Islington Avenue | 43.667856 | -79.532242 |
| 3 | M3B | North York | Don Mills North | 43.745906 | -79.352188 |
| 4 | M6B | North York | Glencairn | 43.709577 | -79.445073 |
| 5 | M4C | East York | Woodbine Heights | 43.695344 | -79.318389 |
| 6 | M5C | Downtown Toronto | St. James Town | 43.651494 | -79.375418 |
| 7 | M6C | York | Humewood-Cedarvale | 43.693781 | -79.428191 |
| 8 | M4E | East Toronto | The Beaches | 43.676357 | -79.293031 |
| 9 | M5E | Downtown Toronto | Berczy Park | 43.644771 | -79.373306 |
| 10 | M6E | York | Caledonia-Fairbanks | 43.689026 | -79.453512 |
| 11 | M1G | Scarborough | Woburn | 43.770992 | -79.216917 |
| 12 | M4G | East York | Leaside | 43.709060 | -79.363452 |
| 13 | M5G | Downtown Toronto | Central Bay Street | 43.657952 | -79.387383 |
| 14 | M6G | Downtown Toronto | Christie | 43.669542 | -79.422564 |

In [29]:

```
completeNeighDF.shape
```

Out[29]:

```
(103, 5)
```

In [30]:

```
completeNeighDF.to_csv('/resources/data/DataScienceCAPSTONE/Week3/completePart2DF.csv')
print("csv file written to '/resources/data/DataScienceCAPSTONE/Week3/completePart2DF.csv' ")
```

```
csv file written to '/resources/data/DataScienceCAPSTONE/Week3/completePart2DF.csv'
```