

```
In [ ]: import sys
        !{sys.executable} -m pip install geocoder

        print('Packages installed.')
```

```
In [ ]: import numpy as np # library to handle data in a vectorized manner
        import pandas as pd # library for data analysis
        import geocoder # import geocoder
        import requests
        from bs4 import BeautifulSoup

        print('Libraries imported.')
```

```
In [ ]: URL = "https://www.wikizeroo.org/index.php?q=aHR0cHM6Ly91bi53aWtpcGVkawEub3JnL3dpa2kvTG1zdF9vZl9wb3N0YWxfY29kZXNfb2Z2fQ2FuYWRh019N"
        r = requests.get(URL)

        soup = BeautifulSoup(r.content, 'html5lib')
        table = soup.find('div', attrs = {'id': 'container'})

        # print(soup.prettify())
        print('Page Scrapped.')
```

```
In [ ]: postalCodes = [];
        boroughs= [];
        neighborhoods = [];
        columnNum = 1;
        passVal = False

        for row in soup.find_all('td'):
            for cell in row:
                if cell.string and cell.string[0].isalpha() and len(cell.string) > 2:
                    passVal = False
                    if columnNum == 1:
                        if passVal == False and cell.string[1].isdigit():
                            postalCodes.append(cell.string);
                            columnNum = 2
                        else:
                            continue
                    elif columnNum == 2 :
                        if cell.string == 'Not assigned':
                            passVal = True
                            del postalCodes[-1]
                            columnNum = 1
                            continue
                        else:
                            boroughs.append(cell.string);
                            columnNum = 3
                    elif columnNum == 3 :
                        if cell.string == 'Not assigned\n':
                            neighborhoods.append(boroughs[-1])
                        else:
                            neighborhoods.append(cell.string);
                            columnNum = 1

        print('Data Collected.')
```

```
In [ ]: # define the dataframe columns
        column_names = ['PostalCode', 'Borough', 'Neighborhood', 'Latitude', 'Longitude']

        # instantiate the dataframe
        neighbors = pd.DataFrame(columns=column_names)

        neighbors
```

```
In [ ]: # initialize your variable to None
        lat_lng_coords = None

        for data in range(0, len(postalCodes)-1):
            code = postalCodes[data]
            borough = boroughs[data]
            neighborhood_name = neighborhoods[data]

            g = geocoder.arcgis('{}, Toronto, Ontario'.format(code))
            lat_lng_coords = g.latlng

            neighbors = neighbors.append({ 'PostalCode': code,
                                           'Borough': borough,
                                           'Neighborhood': neighborhood_name,
                                           'Latitude': lat_lng_coords[0],
                                           'Longitude': lat_lng_coords[1]}, ignore_index=True)
```

```
In [ ]: neighbors
```

```
In [ ]: neighbors.shape
```