This notebook will be mainly used for the capstone project offered by Coursera.

In [42]:
```
import numpy as np
import pandas as pd
print("Hello Capstone Project Course!")
```

```
Hello Capstone Project Course!
```

In [43]:
```
import urllib.request
url="https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M"
page = urllib.request.urlopen(url)
from bs4 import BeautifulSoup
soup = BeautifulSoup(page, "lxml")
all_tables=soup.find_all("table")
```

In [44]:
```
A=[]
B=[]
C=[]
for row in right_table.findAll('tr'):
    cells=row.findAll('td')
    if len(cells)==3:
        A.append(cells[0].find(text=True).rstrip())

        B.append(cells[1].find(text=True).rstrip())
        C.append(cells[2].find(text=True).rstrip())
```

In [78]:
```
data=pd.DataFrame(A,columns=['Postal Code'])
data['Borough']=B
data['Neighbourhood']=C
df=data.copy()
df['Borough'].replace(to_replace = "Not assigned" , value = np.nan,inplace=True)
df.dropna(axis=0,inplace=True)
df.reset_index(inplace = True)
df_post=df.groupby("Postal Code",as_index=False).agg(lambda x:','.join(set(x)))

df_post.shape
```

Out[78]:
```
(103, 3)
```

In [ ]:

In [ ]: