```
# Week 10
# Richa Patel
# Webscraping
http://archive.ics.uci.edu/ml/datasets/auto+mpg
```

In [54]:
```python
import requests
import re
import bs4
from urllib.request import urlopen as uReq
from bs4 import BeautifulSoup as soup


#now open request

openRequest = urllib.request.urlopen('http://archive.ics.uci.edu/ml/datasets/auto+mpg')

myPage = soup(openRequest)

#create empty list

listAll = []

#for loop for itrate

for line in myPage.findAll('a', attrs={'href': re.compile("^http://")}):
    listAll.append(line.get('href'))

print("List of all URl :", listAll)
```

List of all URl : ['http://cml.ics.uci.edu', 'http://rexa.info/paper/9eac95c36cbec1ef4954d3491d
f1ea0a1d09c6f0', 'http://rexa.info/paper/3caf773de7b1ad6c9236cbd03763058bc8846e9d', 'http://rex
a.info/paper/8e674b6eff0f726ba6fff46ef6eaff968dc89f39', 'http://rexa.info/paper/5e9e56b4525a16e
039d75d04d32477b118e36b0d', 'http://rexa.info/paper/0cedee7ef9325127fb946dcdddf8795644e28335',
'http://rexa.info/paper/39037b34a030869c4d11077a2cb171ac3ba3ff0d', 'http://rexa.info/paper/ccde
1df96c7ca5add43e1578b912d95bb86da659', 'http://rexa.info/paper/6d06e4392537787017a43aa34b5412cf
a28e31e8', 'http://rexa.info/paper/ae50055cb7507aa6cad70a2adf7d0f55a25268c7', 'http://rexa.info
/paper/5d88347a087b01681bec3121a13b7b29465cdc71', 'http://rexa.info/paper/d688186b7174c0a53391d
92f70d64a41209d3358', 'http://rexa.info/paper/e7628956b46a78e64774af25d99c0afd0cf90001', 'http:
//rexa.info/paper/48d6beec2a36a87d9d88b6de85dd85a75e5ed24d', 'http://rexa.info', 'http://cml.ic
s.uci.edu']

In [ ]:

In [ ]: