In [ ]:
```
#scrapes rice university website for FALL semester (takes around 30 minutes to
run)
```

In [ ]:
```
from bs4 import BeautifulSoup
from requests import get
import re
from selenium import webdriver
import pandas as pd
import csv

with open('RICE files_fall.csv', mode = 'w', newline = '') as csvfile:
    writer = csv.DictWriter(csvfile, fieldnames = ['num_people_enrolled', 'tot
al_class_size', 'class_number', 'term','professor'])
    writer.writeheader()

    #runs through 2004 through 2019
    for i in range(5,20):

        for term in range(0,1):
            season = '2'
            seasons = 'SPRING'
            if term is 0:
                season = '1'
                seasons = 'FALL'
                if i <= 10:
                    years = '200' + str(i-1)
                else:
                    years = '20' + str(i-1)
            if i <= 9:
                year = '200' + str(i)
            else:
                year = '20' + str(i)
            url = 'https://courses.rice.edu/courses/!SWKSCAT.cat?p_action=QUER
Y&p_term=' + year + season +'0&p_subj=COMP'
            #print(url)
            print(year)
            response = BeautifulSoup(get(url).text, 'lxml')
            body = response.find('table', {'class': 'table table-condensed'})
            sub_body = body.find('tbody')
            each_class = sub_body.find_all('tr')


            for i in each_class:
                block = i.find_all('td')
                class_name = block[1].text
                professor = block[4].text

                sub_urlnum = block[0].text

                #parsing class size/enrollment (opens up the aref link to see
 class details)
                sub_url = 'https://courses.rice.edu/courses/!SWKSCAT.cat?p_act
ion=COURSE&p_term=' + year + season + '0&p_crn=' + sub_urlnum
                sub_response = BeautifulSoup(get(sub_url).text, 'lxml')
                class_body = sub_response.find_all('div', {'class': 'col-lg-6'
})
                section = class_body[1]
                test = section.find_all('div')
                enrollment = test[0].text
```

```
            enrolled = enrollment.split(": ",1)[1]

            capacity1 = test[1].text
            capacity = capacity1.split(": ",1)[1]

            #writes dataframe to csv file
            if term is 0:
                class_info = {
                    'num_people_enrolled' : enrolled,
                    'total_class_size': capacity,
                    'class_number': class_name,
                    'term': seasons+years,
                    'professor':professor
                }
                writer.writerow(class_info)
            if term is 1:
                class_info = {
                    'num_people_enrolled' : enrolled,
                    'total_class_size': capacity,
                    'class_number': class_name,
                    'term': seasons+year,
                    'professor':professor
                }
                writer.writerow(class_info)
```