10/17/2018 scrap web

In [1]:

```
import requests
r = requests.get("http://psgim.ac.in/2017/01/full-time-faculty/")
print(r)
```

<Response [200]>

In [2]:

```
from bs4 import BeautifulSoup
soup = BeautifulSoup(r.content, 'html.parser')
faculty = soup.findAll('div',{"class":"vc_column-inner"})
facultyName = []
for i in range(len(faculty)):
    if faculty[i].find('h4'):
        facultyName.append(faculty[i].find('h4').text)
facultyName.remove('Ms Muthu Janaki')
print('Number of faculties',len(facultyName))
for i in range(len(facultyName)):
    print(str(i+1)+" "+facultyName[i])
```

```
Number of faculties 31
1 Dr Archana Krishnan
2 DR ARUL RAJAN K
3 DR Balasudarsun N L
4 DR DEEPA R
5 Mr Harish V
6 DR JAGAJEEVAN R
7 DR JOSHUA SELVAKUMAR J
8 Mr Karthikeyan M S
9 Mr KARTHIKEYAN L
10 DR KAVITHA D
11 Dr KRISHNAVENI R
12 DR KRISHNAVENI MUTHIAH R
13 Dr MANJU P GEORGE
14 MR Manikandan S S
15 DR MANSURALI A
16 Ms Rajeswari R
17 DR RAMAN H
18 Dr RAMKUMAR N
19 DR SATHISH M
20 DR SATHYANARAYANAN R S
21 DR SEKKIZHAR J
22 Dr Srigayathri Devi K
23 Dr SRIVIDYA V
24 DR SUJATHA R
25 DR SUDHARANI RAVINDRAN D
26 DR SWAMYNATHAN R
27 DR THILAGAM V
28 DR UMA MAHESWARI B
29 DR UMESH CHANDRASEKHAR
30 Dr VIJAYA T G
31 DR VIVEK N
```

10/17/2018 scrap web

```
In [3]:
i = -1
c = 0
fac = soup.findAll('div',{"class":"wpb_text_column"})
print("Faculty with Ph.D: ")
for i in range(len(fac)):
    if fac[i].find('p'):
        if fac[i].find('p').text.find('Qualification')!=-1:
            i=i+1
            if fac[i].find('p').text.find('Ph.D')!=-1 or fac[i].find('p').text.find('Ph.D.,
                c=c+1
                print(c,facultyName[i])
Faculty with Ph.D:
1 DR KAVITHA D
2 Dr KRISHNAVENI R
3 DR KRISHNAVENI MUTHIAH R
4 DR MANSURALI A
5 Ms Rajeswari R
6 DR RAMAN H
7 Dr RAMKUMAR N
8 DR SATHYANARAYANAN R S
AttributeError
                                           Traceback (most recent call last)
<ipython-input-3-12adf6319f49> in <module>()
      7
                if fac[i].find('p').text.find('Qualification')!=-1:
     8
                    i=i+1
                    if fac[i].find('p').text.find('Ph.D')!=-1 or fac[i].find
---> 9
('p').text.find('Ph.D.,')!=-1 or fac[i].find('p').text.find('PhD')!=-1:
     10
                        C=C+1
     11
                        print(c,facultyName[i])
AttributeError: 'NoneType' object has no attribute 'text'
In [ ]:
```