Assessment

I am going to provide two .csv files , you are supposed to work on them and have to provide solutions to the following problems

import necessary libraries

In [1]:

```python
import pandas as pd
import numpy as np
```

merge those two csv files (after getting as dataframes, get them as a single dataframe)

In [4]:

```python
data1=pd.read_csv("C:/Users/91984/Downloads/college_1.csv")
data2=pd.read_csv("C:/Users/91984/Downloads/college_2.csv")
```

In [11]:

```
data1
```

Out[11]:

	Name	python	mysql	Previous Geekions	CodeKata Score	Department	Rising
0	A.Dharani	82.0	20.0	24500	24500	Computer Science and Engineering	0
1	V.JEEVITHA	82.0	20.0	21740	21740	Computer Science and Engineering	0
2	HEMAVATHI.R	100.0	100.0	19680	19680	Computer Science and Engineering	0
3	Mugunthan S	100.0	47.0	10610	10610	Computer Science and Engineering	0
4	Sathammai.S	100.0	8.0	8980	8980	Computer Science and Engineering	0
...	...	...	...	...	...	...	...
79	KarthikeyanS	45.0	0.0	0	0	Electronics and Electrical Engineering	0
80	BARATH.P	29.0	0.0	0	0	Electronics and Electrical Engineering	0
81	N.Ajith kumar	82.0	0.0	0	0	Electronics and Electrical Engineering	0
82	mohamed nabi	0.0	0.0	0	0	Electronics and Electrical Engineering	0
83	yaser ahamed.A	0.0	27.0	0	0	Electronics and Electrical Engineering	0

84 rows × 7 columns

In [10]:

```
data2
```

Out[10]:

	Name	python	python_en	mysql	computational_thinking	Previous Geekions	CodeKat Sco
0	ASHOK KUMAR K	100	0	31	9	19400	1940
1	Chandru	0	20	0	0	14150	1449
2	Ganesh Ramkumar R	-1	55	24	6	8790	1079
3	Bodipudi Harini	-1	0	35	-1	10040	1004
4	VIINU V P	-1	100	24	-1	9150	915
5	Poojitha Y	0	78	35	0	7210	897
6	Gowtham Chandrasekaran	-1	0	0	3	7250	895
7	Gowtham R	-1	60	4	6	8650	865
8	shifak N	58	0	0	0	5180	832
9	Balaji	-1	100	20	0	6170	816
10	Rushitha	-1	15	62	9	7470	809
11	Balaraman Muthupandi	-1	9	0	0	7670	805
12	JPradhap	100	0	0	0	7340	803
13	SRIRAMKUMAR B	-1	0	0	0	6400	794
14	Praneetha	-1	23	100	-1	7170	773
15	Rahul Raj K	-1	0	0	0	6280	764
16	Guhan S	-1	72	35	39	7310	763

	Name	python	python_en	mysql	computational_thinking	Previous Geekions	CodeKat Scor
17	Suryarajan S	-1	63	31	0	6710	755
18	Narasimhan Y L	-1	-1	0	0	4800	680
19	Leelakrishna	-1	6	4	0	5300	664
20	PremsagarReddy	-1	20	0	0	5290	629
21	pradeeban	16	20	24	0	6060	609
22	santhoshini	-1	43	0	0	3910	574
23	MUKESH S M	-1	46	0	0	5200	520
24	P.GOVARDHANAN	-1	40	0	-1	5050	505
25	SHAIK DANISH	-1	43	0	0	4050	482
26	sridhar.s	-1	52	0	0	3860	444
27	Vimalesh Fernando	-1	0	0	0	4170	417
28	M.Karthikeyan	-1	0	4	6	4020	402
29	Abishak s	-1	55	35	6	3220	322
30	praveen raj j	24	-1	0	0	2380	238
31	AMARNATH D	-1	52	12	-1	1890	189
32	bala	32	49	0	-1	1720	172
33	XY Z	-1	20	-1	-1	0	
34	Hariharan	-1	-1	-1	0	0	

In [7]:

```
data=pd.concat([data1,data2])
```

In [8]:

```
data
```

Out[8]:

	Name	python	mysql	Previous Geekions	CodeKata Score	Department	Rising	python_en	c
0	A.Dharani	82.0	20.0	24500	24500	Computer Science and Engineering	0	NaN	
1	V.JEEVITHA	82.0	20.0	21740	21740	Computer Science and Engineering	0	NaN	
2	HEMAVATHI.R	100.0	100.0	19680	19680	Computer Science and Engineering	0	NaN	
3	Mugunthan S	100.0	47.0	10610	10610	Computer Science and Engineering	0	NaN	
4	Sathammai.S	100.0	8.0	8980	8980	Computer Science and Engineering	0	NaN	
...	...	...	...	...	...	...	...	...	
30	praveen raj j	24.0	0.0	2380	2380	Computer Science and Engineering	0	-1.0	
31	AMARNATH D	-1.0	12.0	1890	1890	Electronics and Communication Engineering	0	52.0	
32	bala	32.0	0.0	1720	1720	Electronics and Communication Engineering	0	49.0	
33	XY Z	-1.0	-1.0	0	0	Computer Science and Engineering	0	20.0	
34	Hariharan	-1.0	-1.0	0	0	Computer Science and Engineering	0	-1.0	

119 rows × 9 columns

Take each csv file , split that csv file into multiple categories (example csv files are added in the repo)

consider if the codekata score exceeds 15000 points(present week) then make a csv on those observations as Exceeded expectations.csv

if 10000<codekata score<15000 (Reached_expectations.csv)

if 7000<codekata score<10000 (Needs_Improvement.csv)

if codekate score < 7000 (Unsatisfactory.csv)

In [29]:

```
Exceeded=data[data["CodeKata Score"]>15000]

Exceeded.to_csv("Exceeded_expectations.csv")
Exceeded
```

Out[29]:

	Name	python	mysql	Previous Geekions	CodeKata Score	Department	Rising	python_en	comp
0	A.Dharani	82.0	20.0	24500	24500	Computer Science and Engineering	0	NaN	
1	V.JEEVITHA	82.0	20.0	21740	21740	Computer Science and Engineering	0	NaN	
2	HEMAVATHI.R	100.0	100.0	19680	19680	Computer Science and Engineering	0	NaN	
0	ASHOK KUMAR K	100.0	31.0	19400	19400	Electronics and Electrical Engineering	0	0.0	

In [37]:

```
Reached=data[data["CodeKata Score"]>10000]
Reached_expectations=Reached[Reached["CodeKata Score"]<15000]
Reached_expectations.to_csv("Reached_expectations.csv")
```

In [38]:

```
Reached_expectations
```

Out[38]:

	Name	python	mysql	Previous Geekions	CodeKata Score	Department	Rising	python_en	comp
3	Mugunthan S	100.0	47.0	10610	10610	Computer Science and Engineering	0	NaN	
1	Chandru	0.0	0.0	14150	14490	Computer Science and Engineering	340	20.0	
2	Ganesh Ramkumar R	-1.0	24.0	8790	10790	Computer Science and Engineering	2000	55.0	
3	Bodipudi Harini	-1.0	35.0	10040	10040	Electronics and Communication Engineering	0	0.0	

◄ ▐▐▐ ▶

In [39]:

```
Needs=data[data["CodeKata Score"]>7000]
Needs_Improvement=Needs[Needs["CodeKata Score"]<10000]
```

In [41]:

```
Needs_Improvement.to_csv("Needs_Improvement.csv")
```

In [42]:

```
Unsatisfactory=data[data["CodeKata Score"]<7000]
Unsatisfactory
```

Out[42]:

	Name	python	mysql	Previous Geekions	CodeKata Score	Department	Rising	python_en
8	J.SUGANTHI	27.0	50.00	5860	5860	Electronics and Communication Engineering	0	NaN
9	thamizhpaana	29.0	0.00	5500	5500	Computer Science and Engineering	0	NaN
10	Iyappan Samiraj	50.0	24.00	5400	5400	Electronics and Communication Engineering	0	NaN
11	Ponniyamma.R	100.0	11.25	5380	5380	Computer Science and Engineering	0	NaN
12	SABAPATHI	100.0	0.00	5280	5280	Computer Science and Engineering	0	NaN
...	...	...	...	...	...	...	...	...
30	praveen raj j	24.0	0.00	2380	2380	Computer Science and Engineering	0	-1.0
31	AMARNATH D	-1.0	12.00	1890	1890	Electronics and Communication Engineering	0	52.0
32	bala	32.0	0.00	1720	1720	Electronics and Communication Engineering	0	49.0
33	XY Z	-1.0	-1.00	0	0	Computer Science and Engineering	0	20.0
34	Hariharan	-1.0	-1.00	0	0	Computer Science and Engineering	0	-1.0

93 rows × 9 columns

In [43]:

```
Unsatisfactory.to_csv("Unsatisfactory.csv")
```

Average of previous week geekions vs this week geekions (i.e Previous Geekions vs CodeKata Score)

In [46]:

```python
new_data=data[['Previous Geekions', 'CodeKata Score']].agg(['sum','mean']).T.rename(col
umns={'sum':'total','mean':'Avg'})
print(new_data)
```

```
                     total           Avg
Previous Geekions   457240.0   3842.352941
CodeKata Score      482160.0   4051.764706
```

No of students participated

In [ ]:

#Average completion of python course or my_sql or python english or computational thinking

In [48]:

```python
 new_data=data[['mysql','python_en','computational_thinking']].agg(['sum','mean']).T.re
name(columns={'sum':'total','mean':'Avg'})
print(new_data)
```

```
                          total        Avg
mysql                     2796.5   23.500000
python_en                 1038.0   29.657143
computational_thinking      77.0    2.200000
```

rising star of the week (top 3 candidate who performed well in that particular week)

In [51]:

```python
data.nlargest(3,['CodeKata Score','python','mysql'])
```

Out[51]:

	Name	python	mysql	Previous Geekions	CodeKata Score	Department	Rising	python_en	comp
0	A.Dharani	82.0	20.0	24500	24500	Computer Science and Engineering	0	NaN	
1	V.JEEVITHA	82.0	20.0	21740	21740	Computer Science and Engineering	0	NaN	
2	HEMAVATHI.R	100.0	100.0	19680	19680	Computer Science and Engineering	0	NaN	

Shining stars of the week (top 3 candidates who has highest geekions)

In [52]:

```python
data.nlargest(3,['Previous Geekions','CodeKata Score'])
```

Out[52]:

	Name	python	mysql	Previous Geekions	CodeKata Score	Department	Rising	python_en	comp
0	A.Dharani	82.0	20.0	24500	24500	Computer Science and Engineering	0	NaN	
1	V.JEEVITHA	82.0	20.0	21740	21740	Computer Science and Engineering	0	NaN	
2	HEMAVATHI.R	100.0	100.0	19680	19680	Computer Science and Engineering	0	NaN	

Department wise codekata performance (pie chart)

In [57]:

```python
import matplotlib.ticker as ticker
import matplotlib.cm as cm
import matplotlib as mpl
from matplotlib.gridspec import GridSpec
import matplotlib.pyplot as plt

plt.figure(1, figsize=(20,10))
the_grid = GridSpec(2, 2)

cmap = plt.get_cmap('Spectral')
colors = [cmap(i) for i in np.linspace(0, 1, 8)]

dep = data.groupby('Department').agg('count')
dep
```

Out[57]:

	Name	python	mysql	Previous Geekions	CodeKata Score	Rising	python_en	computation
Department								
Computer Science and Engineering	63	63	63	63	63	63	22	
Electronics and Communication Engineering	39	39	39	39	39	39	8	
Electronics and Electrical Engineering	17	17	17	17	17	17	5	

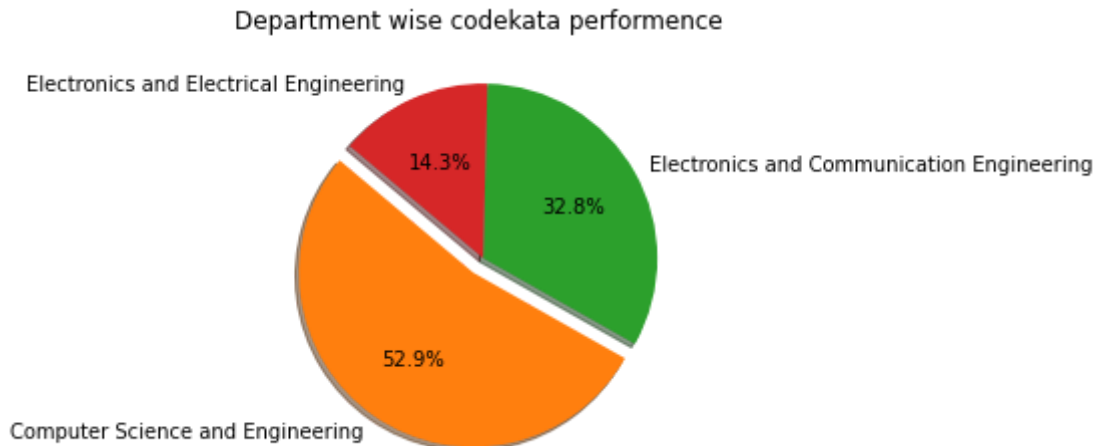```
<Figure size 1440x720 with 0 Axes>
```

In [68]:

```python
colors = [ "#ff7f0e", "#2ca02c", "#d62728"]
explode = (0.1, 0, 0)
plt.pie(dep["CodeKata Score"], labels=dep.index, explode=explode, colors=colors,
autopct='%1.1f%%', shadow=True, startangle=140)
plt.title("Department wise codekata performance ")
plt.show()
```

Department wise codekata performance

Department wise toppers (horizantal bar graph or any visual representations of your choice)

In [108]:

```python
idx = data.groupby(['Department'])['CodeKata Score'].transform(max) == data['CodeKata
Score']
```

In [111]:

```python
new=data[idx]
```
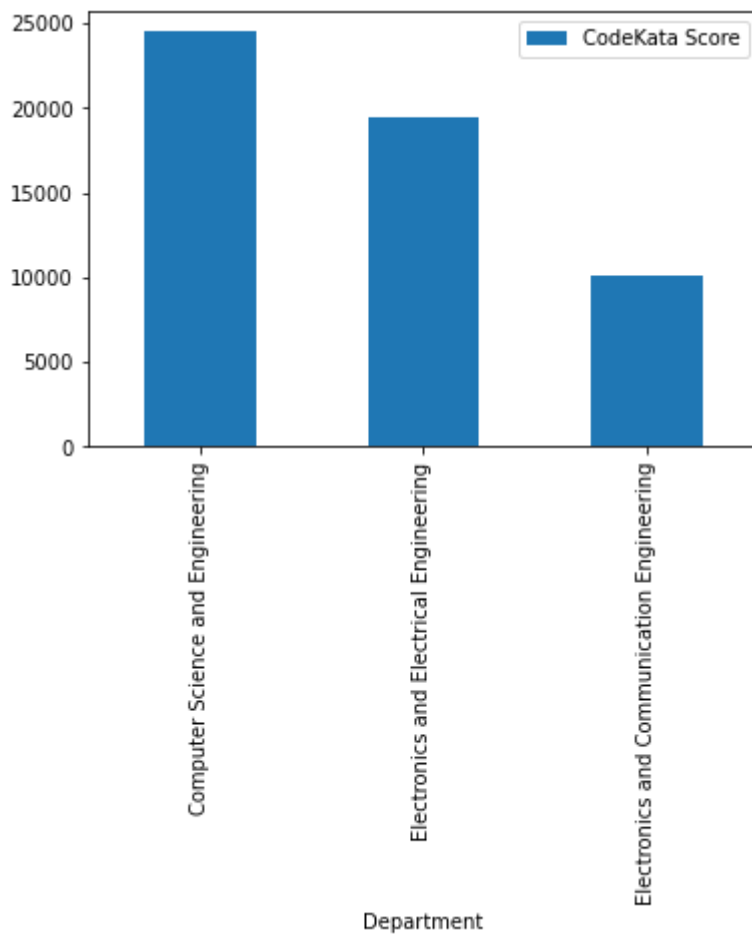
In [112]:

```python
new
```

Out[112]:

	Name	python	mysql	Previous Geekions	CodeKata Score	Department	Rising	python_en	compu
0	A.Dharani	82.0	20.0	24500	24500	Computer Science and Engineering	0	NaN	
0	ASHOK KUMAR K	100.0	31.0	19400	19400	Electronics and Electrical Engineering	0	0.0	
3	Bodipudi Harini	-1.0	35.0	10040	10040	Electronics and Communication Engineering	0	0.0	

In [115]:

```python
new.plot(x="Department", y=["CodeKata Score"], kind="bar")
```

Out[115]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1417c723370>
```


In [ ]:

In [ ]: