

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sb
```

```
In [3]: dataset = pd.read_csv("college_1.csv")
```

```
In [4]: dataset.head()
```

Out[4]:

	Name	python	mysql	Previous Geekions	CodeKata Score	Department	Rising
0	A.Dharani	82.0	20.0	24500	24500	Computer Science and Engineering	0
1	VJEEVITHA	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

```
In [5]: dataset.tail()
```

Out[5]:

	Name	python	mysql	Previous Geekions	CodeKata Score	Department	Rising
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

```
In [6]: dataset = pd.read_csv("college_2.csv")
```

```
In [7]: dataset.head()
```

Out[7]:

	Name	python	python_en	mysql	computational_thinking	Previous Geekions	CodeKata Score	Department	Rising
0	ASHOK KUMAR K	100	0	31	9	19400	19400	Electronics and Electrical Engineering	(
1	Chandru	0	20	0	0	14150	14490	Computer Science and Engineering	34(

2	Ganesh Ramkumar R	-1	55	24	6	8790	10790	Computer Science and Engineering	200
3	Bodipudi Harini	-1	0	35	-1	10040	10040	Electronics and Communication Engineering	(
4	VIINU V P	-1	100	24	-1	9150	9150	Electronics and Electrical Engineering	(

In [8]: `dataset.tail()`

Out[8]:

	Name	python	python_en	mysql	computational_thinking	Previous Geekions	CodeKata Score	Department	Ris
30	praveen raj j	24	-1	0	0	2380	2380	Computer Science and Engineering	
31	AMARNATH D	-1	52	12	-1	1890	1890	Electronics and Communication Engineering	
32	bala	32	49	0	-1	1720	1720	Electronics and Communication Engineering	
33	XY Z	-1	20	-1	-1	0	0	Computer Science and Engineering	
34	Hariharan	-1	-1	-1	0	0	0	Computer Science and Engineering	

In [9]:

```

import pandas as pd
import glob

path = r'C:\Users\swara\AppData\Local\Programs\Python\Python39\Scripts' # use your path
all_files = glob.glob(path + "/*.csv")

li = []

for filename in all_files:
    df = pd.read_csv(filename, index_col=None, header=0)
    li.append(df)

frame = pd.concat(li, axis=0, ignore_index=True)
frame

```

Out[9]:

	Name	python	mysql	Previous Geekions	CodeKata Score	Department	Rising	python_en	computational_think
0	A.Dharani	82.0	20.0	24500.0	24500.0	Computer Science and Engineering	0.0	NaN	N
1	V.JEEVITHA	82.0	20.0	21740.0	21740.0	Computer Science and Engineering	0.0	NaN	N
2	HEMAVATHI.R	100.0	100.0	19680.0	19680.0	Computer Science and Engineering	0.0	NaN	N
3	Mugunthan S	100.0	47.0	10610.0	10610.0	Computer	0.0	NaN	N

						Science and Engineering			
						Computer Science and Engineering			
4	Sathammai.S	100.0	8.0	8980.0	8980.0		0.0	NaN	N
...	...	...	...	...	...	...	...	...	
630	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
631	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
632	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
633	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
634	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N

635 rows × 15 columns

In [10]:

```
# read DataFrame
data = pd.read_csv("college_1.csv")

# no of csv files with row size
k = 2
size = 5

for i in range(k):
    df = data[size*i:size*(i+1)]
    df.to_csv(f'college_{i+1}.csv', index=False)

df_1 = pd.read_csv("college_1.csv")
print(df_1)

df_2 = pd.read_csv("college_2.csv")
print(df_2)
```

	Name	python	mysql	Previous	Geekions	CodeKata	Score \
0	A.Dharani	82.0	20.0		24500		24500
1	V.JEEVITHA	82.0	20.0		21740		21740
2	HEMAVATHI.R	100.0	100.0		19680		19680
3	Mugunthan S	100.0	47.0		10610		10610
4	Sathammai.S	100.0	8.0		8980		8980

	Department	Rising
0	Computer Science and Engineering	0
1	Computer Science and Engineering	0
2	Computer Science and Engineering	0
3	Computer Science and Engineering	0
4	Computer Science and Engineering	0

Empty DataFrame  
Columns: [Name, python, mysql, Previous Geekions, CodeKata Score, Department, Rising]  
Index: []

In [15]:

```
g = df.groupby('CodeKata Score')
g
```

Out[15]:

<pandas.core.groupby.generic.DataFrameGroupBy object at 0x0000024133A47A00>

In [18]:

```
print(
    df['CodeKata Score']
)
```

Series([], Name: CodeKata Score, dtype: int64)

```
In [19]: df = pd.read_csv("college_1.csv")
df
```

```
Out[19]:
```

	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 [24]: df['CodeKata Score']
```

```
Out[24]:
```

0	24500
1	21740
2	19680
3	10610
4	8980
...	...
79	0
80	0
81	0
82	0
83	0

Name: CodeKata Score, Length: 84, dtype: int64

```
In [30]: df = pd.read_csv("college_1.csv")
df.loc[df.python < 50]
```

```
Out[30]:
```

	Name	python	mysql	Previous Geekions	CodeKata Score	Department	Rising
8	J.SUGANTHI	27.0	50.00	5860	5860	Electronics and Communication Engineering	0
9	thamizhpaana	29.0	0.00	5500	5500	Computer Science and Engineering	0
21	kamalishwari	32.0	20.00	3665	3665	Electronics and Communication	0

						Engineering	
32	SarumathiR	16.0	35.00	2240	2240	Electronics and Communication Engineering	0
47	MadhumathiB	16.0	35.00	1130	1130	Electronics and Communication Engineering	0
49	bavithra	0.0	100.00	1020	1020	Electronics and Communication Engineering	0
50	Deivani.S	0.0	86.25	980	980	Electronics and Communication Engineering	0
51	Nivetha.S	0.0	50.00	950	950	Electronics and Communication Engineering	0
52	GUNAMATHI.S	0.0	77.00	850	850	Electronics and Communication Engineering	0
54	u.abarna	3.0	39.00	770	770	Electronics and Communication Engineering	0
55	kaviyarasi.k	16.0	28.75	760	760	Electronics and Communication Engineering	0
57	M.PRIYA	43.0	24.00	690	690	Electronics and Communication Engineering	0
58	M.Nisha@mary	11.0	20.00	580	580	Electronics and Communication Engineering	0
63	E.Nalini	3.0	0.00	270	270	Electronics and Communication Engineering	0
68	sivaranjini	0.0	3.75	160	160	Electronics and Communication Engineering	0
73	rajesh kumar.r	30.0	0.00	100	100	Electronics and Communication Engineering	0
79	KarthikeyanS	45.0	0.00	0	0	Electronics and Electrical Engineering	0
80	BARATH.P	29.0	0.00	0	0	Electronics and Electrical Engineering	0
82	mohamed nabi	0.0	0.00	0	0	Electronics and Electrical Engineering	0
83	yaser ahamed.A	0.0	27.00	0	0	Electronics and Electrical Engineering	0

```
In [37]: df = pd.read_csv("college_1.csv")
df['CodeKata Score'].loc[lambda x: x > 15000]
```

```
Out[37]: 0    24500
1    21740
2    19680
Name: CodeKata Score, dtype: int64
```

```
In [41]: df = pd.read_csv("college_1.csv")
df['CodeKata Score'].loc[lambda x: x > 15000]
print("Reached_expectations")
```

Reached\_expectations

```
In [42]: df = pd.read_csv("college_1.csv")
df['CodeKata Score'].loc[lambda x: x > 10000]
print("Needs_Improvement")
```

Needs\_Improvement

```
In [43]: df = pd.read_csv("college_1.csv")
df['CodeKata Score'].loc[lambda x: x > 7000]
print("Unsatisfactory")
```

Unsatisfactory

```
In [44]: df = pd.read_csv("college_1.csv")
df.mean(axis='index')
```

C:\Users\swara\AppData\Local\Temp\ipykernel\_22280\1588766957.py:2: FutureWarning: Dropping of nuisance columns in DataFrame reductions (with 'numeric\_only=None') is deprecated; in a future version this will raise TypeError. Select only valid columns before calling the reduction.

```
df.mean(axis='index')
```

```
Out[44]: python          73.392857
mysql          27.601190
Previous Geekions  2904.523810
CodeKata Score    2906.071429
Rising          1.547619
dtype: float64
```

```
In [46]: df = pd.read_csv("college_1.csv")
df
```

```
Out[46]:
```

	Name	python	mysql	Previous Geekions	CodeKata Score	Department	Rising
0	A.Dharani	82.0	20.0	24500	24500	Computer Science and Engineering	0
1	VJEEVITHA	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 [47]: df = pd.read_csv("college_2.csv")
df
```

```
Out[47]:
```

	Name	python	python_en	mysql	computational_thinking	Previous	CodeKata	Departmen
--	------	--------	-----------	-------	------------------------	----------	----------	-----------

					Geekions		Score		
0	ASHOK KUMAR K	100	0	31	9	19400	19400	Electronics and Electrical Engineering	
1	Chandru	0	20	0	0	14150	14490	Computer Science and Engineering	
2	Ganesh Ramkumar R	-1	55	24	6	8790	10790	Computer Science and Engineering	
3	Bodipudi Harini	-1	0	35	-1	10040	10040	Electronics and Communication Engineering	
4	VIINU V P	-1	100	24	-1	9150	9150	Electronics and Electrical Engineering	
5	Poojitha Y	0	78	35	0	7210	8970	Computer Science and Engineering	
6	Gowtham Chandrasekaran	-1	0	0	3	7250	8950	Computer Science and Engineering	
7	Gowtham R	-1	60	4	6	8650	8650	Electronics and Communication Engineering	
8	shifak N	58	0	0	0	5180	8320	Electronics and Electrical Engineering	
9	Balaji	-1	100	20	0	6170	8160	Electronics and Electrical Engineering	
10	Rushitha	-1	15	62	9	7470	8090	Computer Science and Engineering	
11	Balaraman Muthupandi	-1	9	0	0	7670	8050	Computer Science and Engineering	
12	JPradhap	100	0	0	0	7340	8030	Electronics and Electrical Engineering	
13	SRIRAMKUMAR B	-1	0	0	0	6400	7940	Computer Science and Engineering	
14	Praneetha	-1	23	100	-1	7170	7730	Electronics and Communication Engineering	
15	Rahul Raj K	-1	0	0	0	6280	7640	Computer Science and Engineering	
16	Guhan S	-1	72	35	39	7310	7630	Electronics and Communication Engineering	
17	Suryarajan S	-1	63	31	0	6710	7550	Computer Science and Engineering	

18	Narasimhan Y L	-1	-1	0	0	4800	6800	Computer Science and Engineering
19	Leelakrishna	-1	6	4	0	5300	6640	Computer Science and Engineering
20	PremasagarReddy	-1	20	0	0	5290	6290	Computer Science and Engineering
21	pradeeban	16	20	24	0	6060	6090	Computer Science and Engineering
22	santhoshini	-1	43	0	0	3910	5740	Computer Science and Engineering
23	MUKESH S M	-1	46	0	0	5200	5200	Computer Science and Engineering
24	P.GOVARDHANAN	-1	40	0	-1	5050	5050	Electronics and Communication Engineering
25	SHAIK DANISH	-1	43	0	0	4050	4820	Computer Science and Engineering
26	sridhar.s	-1	52	0	0	3860	4440	Computer Science and Engineering
27	Vimalesh Fernando	-1	0	0	0	4170	4170	Computer Science and Engineering
28	M.Karthikeyan	-1	0	4	6	4020	4020	Computer Science and Engineering
29	Abishak s	-1	55	35	6	3220	3220	Electronics and Communication Engineering
30	praveen raj j	24	-1	0	0	2380	2380	Computer Science and Engineering
31	AMARNATH D	-1	52	12	-1	1890	1890	Electronics and Communication Engineering
32	bala	32	49	0	-1	1720	1720	Electronics and Communication Engineering
33	XY Z	-1	20	-1	-1	0	0	Computer Science and Engineering
34	Hariharan	-1	-1	-1	0	0	0	Computer Science and Engineering

In [48]:

```
df = pd.read_csv("college_2.csv")
df.mean(axis='index')
```



pping of nuisance columns in DataFrame reductions (with 'numeric\_only=None') is deprecated; in a future version this will raise TypeError. Select only valid columns before calling the reduction.

```
df.mean(axis='index')
```

```
Out[48]: python          8.657143
python_en      29.657143
mysql          13.657143
computational_thinking  2.200000
Previous Geekions  6093.142857
CodeKata Score  6801.428571
Rising         708.285714
dtype: float64
```

```
In [52]: col_max = df.max(axis=0)
col_max
```

```
Out[52]: Name          sridhar.s
python          100
python_en       100
mysql           100
computational_thinking  39
Previous Geekions  19400
CodeKata Score   19400
Department      Electronics and Electrical Engineering
Rising          3140
dtype: object
```

```
In [53]: df = pd.read_csv("college_1.csv")
df
```

```
Out[53]:
```

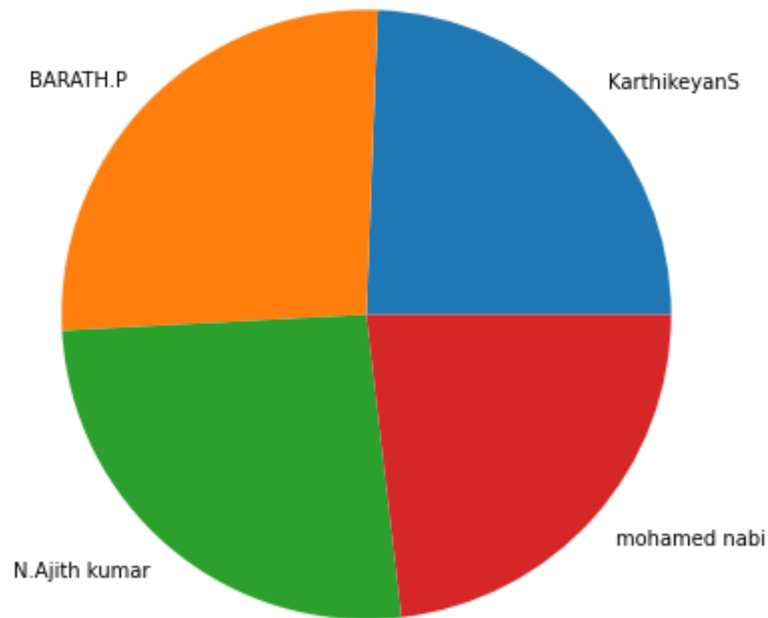
	Name	python	mysql	Previous Geekions	CodeKata Score	Department	Rising
0	A.Dharani	82.0	20.0	24500	24500	Computer Science and Engineering	0
1	VJEEVITHA	82.0	20.0	21740	21740	Computer Science and Engineering	0
2	HEMAVATHI.R	100.0	100.0	19680	19680	Computer Science and Engineering	0
3	Mugunthan S	100.0	47.0	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 [56]: name = ['KarthikeyanS', 'BARATH.P', 'N.Ajith kumar', 'mohamed nabi']
CodeKata_Score = ['24000', '26000', '25567', '22780']
```

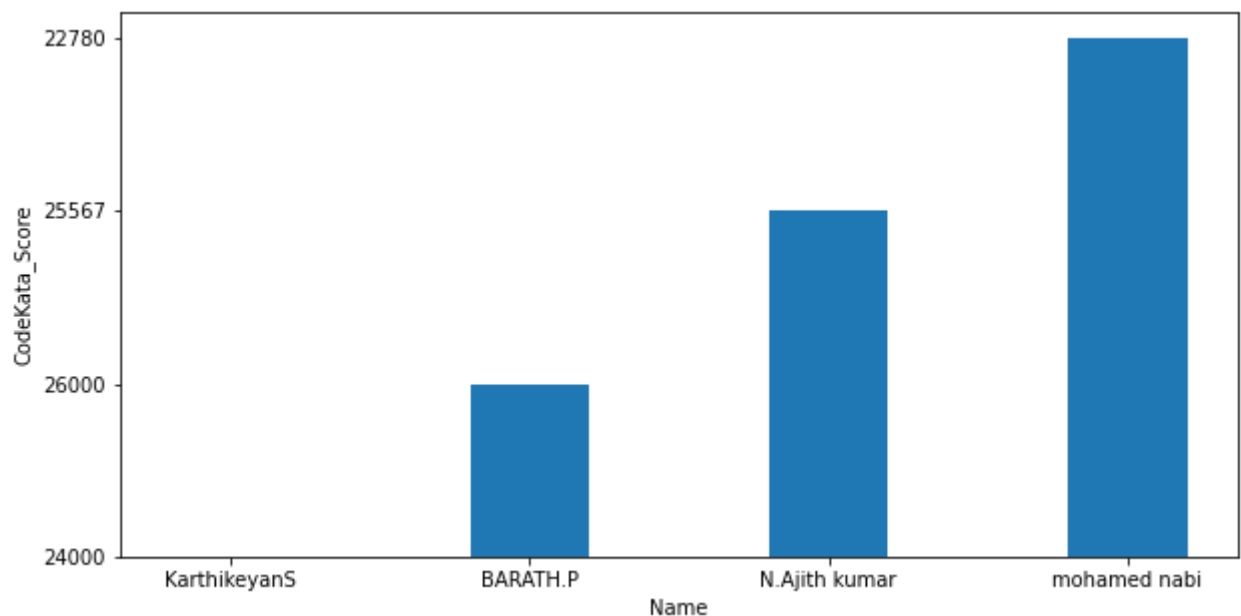
```
fig = plt.figure(figsize=(10, 7))
plt.pie(CodeKata_Score, labels = name )
```

```
Out[56]: ([<matplotlib.patches.Wedge at 0x24135e39ca0>,
<matplotlib.patches.Wedge at 0x24135e51160>,
<matplotlib.patches.Wedge at 0x24135e513a0>,
<matplotlib.patches.Wedge at 0x24135e51880>],
[Text(0.7922587259395201, 0.7631029492622136, 'KarthikeyanS'),
Text(-0.7837514913824506, 0.7718378066380169, 'BARATH.P'),
Text(-0.7097173067582517, -0.8404173632712582, 'N.Ajith kumar'),
Text(0.8213889301232925, -0.7316558107955634, 'mohamed nabi')])
```



```
In [58]: name = ['KarthikeyanS', 'BARATH.P', 'N.Ajith kumar', 'mohamed nabi']
CodeKata_Score = ['24000', '26000', '25567', '22780']

fig = plt.figure(figsize = (10,5))
plt.bar(name,CodeKata_Score, width = 0.4 )
plt.xlabel("Name")
plt.ylabel("CodeKata_Score")
plt.show()
```



In [ ]:

