In [28]:

```
#Case Study On Pandas
#2111CS020505
#OMEGA AIML
#N.Shashank Sagar
```

In [27]:

```
import pandas as pd
a=pd.read_csv("class_marks.csv")
a
```

Out[27]:

	TOTAL	Q1aM4	Q1bM6	Q2aM6	Q2bM4	Q3aM5	Q3bM5	Q4aM3	Q4bM7	Q5M10	Q6aM4
0	37	4	5	6	4	2	1	0	5	8	4
1	32	4	3	4	3	0	0	3	6	9	0
2	33	4	5	5	1	5	5	0	0	8	0
3	24	4	6	6	3	2	2	0	0	0	2
4	36	3	6	4	4	5	4	0	0	10	0
...	...	...	...	...	...	...	...	...	...	...	...
81	32	3	6	3	4	5	3	0	0	0	4
82	27	2	2	5	3	0	0	0	0	7	3
83	37	4	6	6	2	0	0	0	0	9	4
84	28	4	0	5	4	5	4	0	0	6	0
85	29	4	6	0	0	0	0	3	5	7	1

86 rows × 12 columns

In [2]:

```
#1) Read class_marks.csv file and display 15th to 20th Records only
a.loc[15:20]
```

Out[2]:

	TOTAL	Q1aM4	Q1bM6	Q2aM6	Q2bM4	Q3aM5	Q3bM5	Q4aM3	Q4bM7	Q5M10	Q6aM4
15	32	3	0	2	1	5	5	3	7	0	4
16	30	4	0	6	4	5	2	0	0	8	0
17	36	3	4	6	4	0	0	0	0	9	4
18	24	3	3	5	3	0	0	2	1	7	0
19	26	3	0	6	4	0	2	2	1	7	3
20	29	2	6	2	2	5	5	0	0	7	0

In [3]:

```
#2) Create two columns as 2Q,5Q and remove Q5M10 feature from class_marks.csv file.
a["2Q"]=a["Q1aM4"]+a["Q1bM6"]
a["5Q"]=a["Q2aM6"]+a["Q2bM4"]
a.drop("Q5M10",axis=1,inplace=True)
a
```

Out[3]:

	TOTAL	Q1aM4	Q1bM6	Q2aM6	Q2bM4	Q3aM5	Q3bM5	Q4aM3	Q4bM7	Q6aM4	Q6bM6
0	37	4	5	6	4	2	1	0	5	4	6
1	32	4	3	4	3	0	0	3	6	0	0
2	33	4	5	5	1	5	5	0	0	0	0
3	24	4	6	6	3	2	2	0	0	2	0
4	36	3	6	4	4	5	4	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...
81	32	3	6	3	4	5	3	0	0	4	6
82	27	2	2	5	3	0	0	0	0	3	5
83	37	4	6	6	2	0	0	0	0	4	6
84	28	4	0	5	4	5	4	0	0	0	0
85	29	4	6	0	0	0	0	3	5	1	4

86 rows × 13 columns

In [4]:

```
#3) Read class_marks.csv file from Desktop and write as excel file in Documents folder.
b=pd.read_csv(r"C:\Users\SHASHANK SAGAR\Desktop\class_marks.csv")
b
```

Out[4]:

	TOTAL	Q1aM4	Q1bM6	Q2aM6	Q2bM4	Q3aM5	Q3bM5	Q4aM3	Q4bM7	Q5M10	Q6a
0	37	4	5	6	4	2	1	0	5	8	
1	32	4	3	4	3	0	0	3	6	9	
2	33	4	5	5	1	5	5	0	0	8	
3	24	4	6	6	3	2	2	0	0	0	
4	36	3	6	4	4	5	4	0	0	10	
...	...	...	...	...	...	...	...	...	...	...	
81	32	3	6	3	4	5	3	0	0	0	
82	27	2	2	5	3	0	0	0	0	7	
83	37	4	6	6	2	0	0	0	0	9	
84	28	4	0	5	4	5	4	0	0	6	
85	29	4	6	0	0	0	0	3	5	7	

86 rows × 12 columns

In [5]:

```
b.to_excel(r"C:\Users\SHASHANK SAGAR\Documents\New_class-marks.xlsx")
```

In [6]:

```
#5. Read class_marks.csv file except 5,3,8 records.
c=pd.read_csv(r"C:\Users\SHASHANK SAGAR\Desktop\class_marks.csv",skiprows=[5,3,8])
c
```

Out[6]:

	TOTAL	Q1aM4	Q1bM6	Q2aM6	Q2bM4	Q3aM5	Q3bM5	Q4aM3	Q4bM7	Q5M10	Q6aM4
0	37	4	5	6	4	2	1	0	5	8	4
1	32	4	3	4	3	0	0	3	6	9	0
2	24	4	6	6	3	2	2	0	0	0	2
3	20	4	6	6	4	0	0	0	0	0	0
4	25	3	4	0	2	5	5	0	0	6	0
...	...	...	...	...	...	...	...	...	...	...	...
78	32	3	6	3	4	5	3	0	0	0	4
79	27	2	2	5	3	0	0	0	0	7	3
80	37	4	6	6	2	0	0	0	0	9	4
81	28	4	0	5	4	5	4	0	0	6	0
82	29	4	6	0	0	0	0	3	5	7	1

83 rows × 12 columns

In [7]:

```
#6. Read class_marks.csv file with only two columns Q4aM3,Q4bM7.
d=pd.read_csv(r"C:\Users\SHASHANK SAGAR\Desktop\class_marks.csv",usecols=["Q4aM3","Q4bM7"])
d
```

Out[7]:

	Q4aM3	Q4bM7
0	0	5
1	3	6
2	0	0
3	0	0
4	0	0
...	...	...
81	0	0
82	0	0
83	0	0
84	0	0
85	3	5

86 rows × 2 columns

In [8]:

```
#7. Read class_marks.csv file except first six records.
e=pd.read_csv("C:\\Users\\SHASHANK SAGAR\Desktop\\class_marks.csv",skiprows=[1,2,3,4,5,6])
e
```

Out[8]:

	TOTAL	Q1aM4	Q1bM6	Q2aM6	Q2bM4	Q3aM5	Q3bM5	Q4aM3	Q4bM7	Q5M10	Q6aM4
0	25	3	4	0	2	5	5	0	0	6	0
1	34	4	6	6	4	0	0	2	0	0	0
2	27	3	5	5	0	0	0	0	0	9	3
3	35	2	4	5	4	5	5	0	0	0	4
4	37	3	5	6	4	0	0	3	6	0	4
...	...	...	...	...	...	...	...	...	...	...	...
75	32	3	6	3	4	5	3	0	0	0	4
76	27	2	2	5	3	0	0	0	0	7	3
77	37	4	6	6	2	0	0	0	0	9	4
78	28	4	0	5	4	5	4	0	0	6	0
79	29	4	6	0	0	0	0	3	5	7	1

80 rows × 12 columns

In [9]:

```
#8. Read class_marks.csv file without header.
f=pd.read_csv("C:\\Users\\SHASHANK SAGAR\Desktop\\class_marks.csv",header=None)
f
```

Out[9]:

	0	1	2	3	4	5	6	7	8	9	10
0	TOTAL	Q1aM4	Q1bM6	Q2aM6	Q2bM4	Q3aM5	Q3bM5	Q4aM3	Q4bM7	Q5M10	Q6aM4
1	37	4	5	6	4	2	1	0	5	8	4
2	32	4	3	4	3	0	0	3	6	9	0
3	33	4	5	5	1	5	5	0	0	8	0
4	24	4	6	6	3	2	2	0	0	0	2
...	...	...	...	...	...	...	...	...	...	...	...
82	32	3	6	3	4	5	3	0	0	0	4
83	27	2	2	5	3	0	0	0	0	7	3
84	37	4	6	6	2	0	0	0	0	9	4
85	28	4	0	5	4	5	4	0	0	6	0
86	29	4	6	0	0	0	0	3	5	7	1

87 rows × 12 columns

In [10]:

```
#9. Read class_marks.csv file and display question 5 unique values with each count.
g=pd.read_csv("C:\\Users\\SHASHANK SAGAR\Desktop\\class_marks.csv")
g["Q5"]=g["Q5M10"]
g.Q5.value_counts()
```

Out[10]:

```
0     30
10    15
7     12
8     10
9      9
6      5
5      3
4      2
Name: Q5, dtype: int64
```

In [11]:

```
# 11. Read class_marks.csv file and give full marks to students who are attempted 4a due to
a=pd.read_csv("C:\\Users\\SHASHANK SAGAR\Desktop\\class_marks.csv")
a.rename(columns={"Q4aM3":"Q4a"},inplace=True)
a.loc[(a.Q4a>=0)]=3
a
```

Out[11]:

	TOTAL	Q1aM4	Q1bM6	Q2aM6	Q2bM4	Q3aM5	Q3bM5	Q4a	Q4bM7	Q5M10	Q6aM4	C
0	3	3	3	3	3	3	3	3	3	3	3	
1	3	3	3	3	3	3	3	3	3	3	3	
2	3	3	3	3	3	3	3	3	3	3	3	
3	3	3	3	3	3	3	3	3	3	3	3	
4	3	3	3	3	3	3	3	3	3	3	3	
...	...	...	...	...	...	...	...	...	...	...	...	
81	3	3	3	3	3	3	3	3	3	3	3	
82	3	3	3	3	3	3	3	3	3	3	3	
83	3	3	3	3	3	3	3	3	3	3	3	
84	3	3	3	3	3	3	3	3	3	3	3	
85	3	3	3	3	3	3	3	3	3	3	3	

86 rows × 12 columns

In [12]:

```
# 12. Read class_marks.csv file except first Ten records.
a=pd.read_csv("C:\\Users\\SHASHANK SAGAR\Desktop\\class_marks.csv",skiprows=10)
a
```

Out[12]:

	35	2	4	5	4.1	5.1	5.2	0	0.1	0.2	4.2	6
0	37	3	5	6	4	0	0	3	6	0	4	6
1	8	2	2	0	3	1	0	0	0	0	0	0
2	34	4	4	5	3	2	2	2	1	9	2	2
3	32	3	3	6	4	3	5	0	0	9	0	0
4	30	4	6	6	2	4	5	3	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...
71	32	3	6	3	4	5	3	0	0	0	4	6
72	27	2	2	5	3	0	0	0	0	7	3	5
73	37	4	6	6	2	0	0	0	0	9	4	6
74	28	4	0	5	4	5	4	0	0	6	0	0
75	29	4	6	0	0	0	0	3	5	7	1	4

76 rows × 12 columns

In [13]:

```
# 14. Read class_marks.csv file and Display subject failures count. a=read_csv("class_marks
a=pd.read_csv("C:\\Users\\SHASHANK SAGAR\Desktop\\class_marks.csv")
a.loc[(a.TOTAL<=16)].TOTAL.count()
```

Out[13]:

4

In [14]:

```
# 15. Read class_marks.csv file except first Ten records.
a=pd.read_csv("C:\\Users\\SHASHANK SAGAR\Desktop\\class_marks.csv",skiprows=10)
a
```

Out[14]:

	35	2	4	5	4.1	5.1	5.2	0	0.1	0.2	4.2	6
0	37	3	5	6	4	0	0	3	6	0	4	6
1	8	2	2	0	3	1	0	0	0	0	0	0
2	34	4	4	5	3	2	2	2	1	9	2	2
3	32	3	3	6	4	3	5	0	0	9	0	0
4	30	4	6	6	2	4	5	3	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...
71	32	3	6	3	4	5	3	0	0	0	4	6
72	27	2	2	5	3	0	0	0	0	7	3	5
73	37	4	6	6	2	0	0	0	0	9	4	6
74	28	4	0	5	4	5	4	0	0	6	0	0
75	29	4	6	0	0	0	0	3	5	7	1	4

76 rows × 12 columns

In [15]:

```
## 16. Marks are in present in decimals, convert any column marks into numbers in class_mar
df=pd.read_csv("C:\\Users\\SHASHANK SAGAR\Desktop\\class_marks.csv")
df.fillna(0,inplace=True)
df["Q1aM4"].astype(int)
```

Out[15]:

```
0     4
1     4
2     4
3     4
4     3
     ..
81    3
82    2
83    4
84    4
85    4
Name: Q1aM4, Length: 86, dtype: int32
```

In [16]:

```
#17. Display who scored 40 marks in asending order from class_marks.csv file.
a=pd.read_csv("class_marks.csv")
a.sort_values("TOTAL")
```

Out[16]:

	TOTAL	Q1aM4	Q1bM6	Q2aM6	Q2bM4	Q3aM5	Q3bM5	Q4aM3	Q4bM7	Q5M10	Q6aM4
69	3	1	0	1	0	0	0	1	0	0	0
11	8	2	2	0	3	1	0	0	0	0	0
23	9	4	3	0	0	0	0	0	0	0	1
22	14	4	4	5	2	0	0	0	0	0	0
57	17	3	0	0	4	0	0	3	7	0	4
...	...	...	...	...	...	...	...	...	...	...	...
73	40	4	6	0	0	5	5	3	0	10	4
53	40	4	6	6	4	5	5	0	0	10	0
51	40	0	0	6	4	0	0	3	7	10	0
33	40	0	0	6	4	5	5	3	7	0	4
65	40	4	6	6	4	5	5	0	0	10	0

86 rows × 12 columns

In [17]:

```
## 18. Change question name as Q5 instead of Q5M10 in class_marks.csv file.
a=pd.read_csv("C:\\Users\\SHASHANK SAGAR\Desktop\\class_marks.csv")
a.rename(columns={'Q5M10':'Q5'},inplace=True)
a
```

Out[17]:

	TOTAL	Q1aM4	Q1bM6	Q2aM6	Q2bM4	Q3aM5	Q3bM5	Q4aM3	Q4bM7	Q5	Q6aM4	Q6
0	37	4	5	6	4	2	1	0	5	8	4	
1	32	4	3	4	3	0	0	3	6	9	0	
2	33	4	5	5	1	5	5	0	0	8	0	
3	24	4	6	6	3	2	2	0	0	0	2	
4	36	3	6	4	4	5	4	0	0	10	0	
...	...	...	...	...	...	...	...	...	...	...	...	...
81	32	3	6	3	4	5	3	0	0	0	4	
82	27	2	2	5	3	0	0	0	0	7	3	
83	37	4	6	6	2	0	0	0	0	9	4	
84	28	4	0	5	4	5	4	0	0	6	0	
85	29	4	6	0	0	0	0	3	5	7	1	

86 rows × 12 columns

In [18]:

```
# 19. Read class_marks.csv file except 5th,3rd,8th records.
a=pd.read_csv("class_marks.csv",skiprows=[3,5,8])
a
```

Out[18]:

	TOTAL	Q1aM4	Q1bM6	Q2aM6	Q2bM4	Q3aM5	Q3bM5	Q4aM3	Q4bM7	Q5M10	Q6aM4
0	37	4	5	6	4	2	1	0	5	8	4
1	32	4	3	4	3	0	0	3	6	9	0
2	24	4	6	6	3	2	2	0	0	0	2
3	20	4	6	6	4	0	0	0	0	0	0
4	25	3	4	0	2	5	5	0	0	6	0
...	...	...	...	...	...	...	...	...	...	...	...
78	32	3	6	3	4	5	3	0	0	0	4
79	27	2	2	5	3	0	0	0	0	7	3
80	37	4	6	6	2	0	0	0	0	9	4
81	28	4	0	5	4	5	4	0	0	6	0
82	29	4	6	0	0	0	0	3	5	7	1

83 rows × 12 columns

In [19]:

```
# 20. Total column name should convert into capital letters in class_marks.csv file.
a=pd.read_csv("class_marks.csv")
a.rename(columns={'TOTAL':'Total'},inplace=True)
a
```

Out[19]:

	Total	Q1aM4	Q1bM6	Q2aM6	Q2bM4	Q3aM5	Q3bM5	Q4aM3	Q4bM7	Q5M10	Q6aM4
0	37	4	5	6	4	2	1	0	5	8	4
1	32	4	3	4	3	0	0	3	6	9	0
2	33	4	5	5	1	5	5	0	0	8	0
3	24	4	6	6	3	2	2	0	0	0	2
4	36	3	6	4	4	5	4	0	0	10	0
...	...	...	...	...	...	...	...	...	...	...	...
81	32	3	6	3	4	5	3	0	0	0	4
82	27	2	2	5	3	0	0	0	0	7	3
83	37	4	6	6	2	0	0	0	0	9	4
84	28	4	0	5	4	5	4	0	0	6	0
85	29	4	6	0	0	0	0	3	5	7	1

86 rows × 12 columns

In [20]:

```
# 21. Read class_marks.csv file and display how many members has not attempt the question 2
a=pd.read_csv("class_marks.csv")
a.rename(columns={"Q2aM6":"Q2a"},inplace=True)
a.loc[:,["Q2a","Q2bM4"]].isnull().sum()
```

Out[20]:

```
Q2a      0
Q2bM4    0
dtype: int64
```

In [21]:

```
# 22. Calculate total marks based on marks scored for each question of each student record
a=pd.read_csv("class_marks.csv")
a.fillna(0)
a["Q1"]=a["Q1aM4"]+a["Q1bM6"]
a["Q2"]=a["Q2aM6"]+a["Q2bM4"]
a["Q3"]=a["Q3aM5"]+a["Q3bM5"]
a["Q4"]=a["Q4aM3"]+a["Q4bM7"]
a["Q5"]=a["Q5M10"]
a["Q6"]=a["Q6aM4"]+a["Q6bM6"]
a["TOTAL"]=None
a.sort_values("Q1")
```

Out[21]:

	TOTAL	Q1aM4	Q1bM6	Q2aM6	Q2bM4	Q3aM5	Q3bM5	Q4aM3	Q4bM7	Q5M10	Q6aM4
33	None	0	0	6	4	5	5	3	7	0	4
51	None	0	0	6	4	0	0	3	7	10	0
69	None	1	0	1	0	0	0	1	0	0	0
72	None	2	0	4	4	0	5	3	7	0	0
36	None	2	0	6	4	5	5	0	0	10	4
...	...	...	...	...	...	...	...	...	...	...	...
7	None	4	6	6	4	0	0	2	0	0	0
5	None	4	6	6	4	0	0	0	0	0	0
3	None	4	6	6	3	2	2	0	0	0	2
70	None	4	6	6	4	5	5	0	0	5	0
85	None	4	6	0	0	0	0	3	5	7	1

86 rows × 18 columns

In [22]:

```
# 23. Read class_marks.csv file except 10 to 15 records.
a=pd.read_csv(r"C:\Users\SHASHANK SAGAR\Desktop\class_marks.csv",skiprows=[10,11,12,13,14,1
a
```

Out[22]:

	TOTAL	Q1aM4	Q1bM6	Q2aM6	Q2bM4	Q3aM5	Q3bM5	Q4aM3	Q4bM7	Q5M10	Q6aM4
0	37	4	5	6	4	2	1	0	5	8	4
1	32	4	3	4	3	0	0	3	6	9	0
2	33	4	5	5	1	5	5	0	0	8	0
3	24	4	6	6	3	2	2	0	0	0	2
4	36	3	6	4	4	5	4	0	0	10	0
...	...	...	...	...	...	...	...	...	...	...	...
75	32	3	6	3	4	5	3	0	0	0	4
76	27	2	2	5	3	0	0	0	0	7	3
77	37	4	6	6	2	0	0	0	0	9	4
78	28	4	0	5	4	5	4	0	0	6	0
79	29	4	6	0	0	0	0	3	5	7	1

80 rows × 12 columns

In [23]:

```
# 24. Read class_marks.csv file and display last 15 records.
a=pd.read_csv("class_marks.csv")
a.tail(15)
```

Out[23]:

	TOTAL	Q1aM4	Q1bM6	Q2aM6	Q2bM4	Q3aM5	Q3bM5	Q4aM3	Q4bM7	Q5M10	Q6aM4
71	36	4	5	6	4	5	5	0	0	7	0
72	25	2	0	4	4	0	5	3	7	0	0
73	40	4	6	0	0	5	5	3	0	10	4
74	25	1	5	6	4	0	0	0	0	9	0
75	21	4	0	6	1	1	1	0	0	8	0
76	17	2	3	4	2	4	2	0	0	0	0
77	22	4	5	0	0	3	2	2	0	0	4
78	33	2	3	6	4	5	5	0	0	8	0
79	27	2	6	0	3	2	5	0	0	9	0
80	31	4	6	6	2	2	5	0	0	6	0
81	32	3	6	3	4	5	3	0	0	0	4
82	27	2	2	5	3	0	0	0	0	7	3
83	37	4	6	6	2	0	0	0	0	9	4
84	28	4	0	5	4	5	4	0	0	6	0
85	29	4	6	0	0	0	0	3	5	7	1

In [24]:

```
# 25. Display students count who attempted Q4
a=pd.read_csv("class_marks.csv")
a["Q4"]=a["Q4aM3"]+a["Q4bM7"]
a.fillna(0)
a.loc[(a.Q4>0)].Q4.count()
```

Out[24]:

```
36
```