

```
import pandas as pd
import numpy as np
df1=pd.read_csv("/content/CSE.csv")
df2=pd.read_csv("/content/IT.csv")
df3=pd.read_csv("/content/STUDENT.CSV")
df1.rename(columns={'Professional Elective':'PE'},inplace=True)
```

1.Combine the CSE & IT data and display the data

```
a=df1.append([df2],ignore_index=True)
print(a)
```

Output:

	Regd.No	CN	DAA	AFL	OE	PE
0	Y20CS001	10.0	15.0	11.0	16.0	12.0
1	Y20CS002	16.0	14.0	15.0	10.0	13.0
2	Y20CS003	15.0	12.0	32.0	12.0	NaN
3	Y20CS004	12.0	NaN	12.0	NaN	17.0
4	Y20CS005	14.0	16.0	13.0	25.0	6.0
5	Y20CS006	9.0	17.0	9.0	14.0	23.0
6	Y20CS007	13.0	3.0	NaN	17.0	16.0
7	Y20CS008	20.0	12.0	15.0	16.0	11.0
8	Y20CS009	17.0	14.0	12.0	10.0	12.0
9	Y20CS010	5.0	16.0	14.0	9.0	15.0
10	Y20CS011	NaN	13.0	9.0	16.0	14.0
11	Y20CS012	12.0	15.0	13.0	13.0	NaN
12	Y20CS013	13.0	24.0	10.0	17.0	10.0
13	Y20CS014	9.0	12.0	15.0	14.0	16.0
14	Y20CS015	17.0	14.0	8.0	13.0	13.0
15	Y20IT001	12.0	14.0	15.0	NaN	12.0
16	Y20IT002	13.0	15.0	13.0	NaN	15.0
17	Y20IT003	14.0	16.0	12.0	NaN	14.0
18	Y20IT004	NaN	17.0	14.0	NaN	16.0
19	Y20IT005	18.0	13.0	16.0	NaN	18.0
20	Y20IT006	15.0	12.0	9.0	NaN	17.0
21	Y20IT007	22.0	10.0	10.0	NaN	3.0
22	Y20IT008	12.0	11.0	11.0	NaN	125.0
23	Y20IT009	16.0	16.0	13.0	NaN	NaN
24	Y20IT010	14.0	NaN	12.0	NaN	8.0
25	Y20IT011	17.0	14.0	14.0	NaN	14.0
26	Y20IT012	13.0	17.0	15.0	NaN	17.0
27	Y20IT013	11.0	13.0	NaN	NaN	11.0
28	Y20IT014	10.0	12.0	12.0	NaN	13.0
29	Y20IT015	12.0	15.0	16.0	NaN	16.0

2.Display all CSE students marks along with personal information.

```
df_cse=pd.merge(df1,df3,on="Regd.No")
print(df_cse)
```

Output:

	Regd.No	CN	DAA	AFL	OE	PE \
0	Y20CS001	10.0	15.0	11.0	16.0	12.0

1 Y20CS002 16.0 14.0 15.0 10.0 13.0
 2 Y20CS003 15.0 12.0 32.0 12.0 NaN
 3 Y20CS004 12.0 NaN 12.0 NaN 17.0
 4 Y20CS005 14.0 16.0 13.0 25.0 6.0
 5 Y20CS006 9.0 17.0 9.0 14.0 23.0
 6 Y20CS007 13.0 3.0 NaN 17.0 16.0
 7 Y20CS008 20.0 12.0 15.0 16.0 11.0
 8 Y20CS009 17.0 14.0 12.0 10.0 12.0
 9 Y20CS010 5.0 16.0 14.0 9.0 15.0
 10 Y20CS011 NaN 13.0 9.0 16.0 14.0
 11 Y20CS012 12.0 15.0 13.0 13.0 NaN
 12 Y20CS013 13.0 24.0 10.0 17.0 10.0
 13 Y20CS014 9.0 12.0 15.0 14.0 16.0
 14 Y20CS015 17.0 14.0 8.0 13.0 13.0

	Name	Sex	Course	Branch	Address \
0	ADAPA HEMANTH VENKATA SAI PAVAN KUMAR	M	B.Tech	CSE	GUNTUR
1	ALAPARTHI VIVEK MADHAV	F	B.Tech	CSE	GUNTUR
2	ALIFA SHAIK	F	B.Tech	CSE	GUNTUR
3	ALLA NEEHARIKA	M	B.Tech	CSE	TENALI
4	AVYAKTHA	F	B.Tech	CSE	VINUKONDA
5	AMBATI MEGHANA	M	B.Tech	CSE	NARASARAOPET
6	ANCHA PRABANDHA	M	B.Tech	CSE	GUNTUR
7	APPANA HEMA SRI	F	B.Tech	CSE	GUNTUR
8	ARIKATLA VIJAYA LAKSHMI	F	B.Tech	CSE	VINUKONDA
9	AVULA CHAYA PRIYANKA	M	B.Tech	CSE	VIJAYAWADA
10	AVULAPATI ANIL KUMAR	F	B.Tech	CSE	TENALI
11	BALAGA LAVANYA	M	B.Tech	CSE	GUNTUR
12	BANDLA BHAVITHA	F	B.Tech	CSE	NARASARAOPET
13	BEENA VASANTH	F	B.Tech	CSE	VIJAYAWADA
14	BELLAM ABHINAY	M	B.Tech	CSE	VINUKONDA

EAMCET RANK

0 2000
 1 1900
 2 3126
 3 2500
 4 8000
 5 4012
 6 5001
 7 1201
 8 17000
 9 2600
 10 3201
 11 1704
 12 3456
 13 1965
 14 45000

3.Print all students Regd.No, Name and professional elective.

```
df=pd.merge(a,df3,on="Regd.No")
```

```
df[['Regd.No','Name','PE']]
```

Output:

	Regd.No	Name	PE
0	Y20CS001	ADAPA HEMANTH VENKATA SAI PAVAN KUMAR	12.0
1	Y20CS002	ALAPARTHI VIVEK MADHAV	13.0
2	Y20CS003	ALIFA SHAIK	NaN
3	Y20CS004	ALLA NEEHARIKA	17.0
4	Y20CS005	AVYAKTHA	6.0
5	Y20CS006	AMBATI MEGHANA	23.0
6	Y20CS007	ANCHA PRABANDHA	16.0
7	Y20CS008	APPANA HEMA SRI	11.0
8	Y20CS009	ARIKATLA VIJAYA LAKSHMI	12.0
9	Y20CS010	AVULA CHAYA PRIYANKA	15.0
10	Y20CS011	AVULAPATI ANIL KUMAR	14.0
11	Y20CS012	BALAGA LAVANYA	NaN
12	Y20CS013	BANDLA BHAVITHA	10.0
13	Y20CS014	BEENA VASANTH	16.0
14	Y20CS015	BELLAM ABHINAY	13.0
15	Y20IT001	ADAPALA YAVANIKA	12.0
16	Y20IT002	AMBATIPUDI SWETHA	15.0
17	Y20IT003	ARE SRILATHA	14.0
18	Y20IT004	ARVAPALLI SUBHASH	16.0
19	Y20IT005	ATLA RAGHAVENDRA	18.0
20	Y20IT006	AVULA GOWTHAM SAI	17.0
21	Y20IT007	BALABHADRA LAKSHMI VENKATA HANUMA ABHIGNA	3.0
22	Y20IT008	BANAVATHU LAKSHMIDURGA NAIK	125.0
23	Y20IT009	BATCHU VENKATA VIBHAS GUPTA	NaN
24	Y20IT010	BATCHU YASASWINI	8.0
25	Y20IT011	BATHINI BALAKRISHNA	14.0

	Regd.No	Name	PE
26	Y20IT012	BATTINENI HIMAVANTH	17.0
27	Y20IT013	BEERAKA REVANTH NAGESH	11.0
28	Y20IT014	BEZAWADA LOKESH	13.0
29	Y20IT015	BITRA TEJA SRI	16.0

4. Identify the students whose DAA marks are >18.

```
a[a['DAA']>18][['Regd.No']]
```

Output:

```
Regd.No
12  Y20CS013
```

5. Display the names and EAMCET ranks of the students who got minimum 12 marks in all courses.

```
df[(df["CN"]>=12)&(df['OE']>=12)&(df["AFL"]>=12)&(df['PE']>=12)&(df['DAA']>=12)][['Name','EAMCET RANK']]
```

Output:

```
Name  EAMCET RANK
```

6. Calculate mean value of all the subject's marks.

```
a[['CN','DAA','AFL','OE','PE']].mean()
```

Output:

```
CN 13.607143
DAA 14.000000
AFL 13.214286
OE 14.428571
PE 17.666667
dtype: float64
```

7. Display the names common in both CSE & IT along with Regd.No.

```
g=df3.groupby(df3["Branch"])
cse=g.get_group("CSE")
it=g.get_group("IT")
df_n=pd.merge(cse,it,on="Name")
print(df_n)
```

Empty DataFrame

Columns: [Regd.No_x, Name, Sex_x, Course_x, Branch_x, Address_x, EAMCET RANK_x, Regd.No_y, Sex_y, Course_y, Branch_y, Address_y, EAMCET RANK_y]

Index: []

8.Fill the missing values of the data with average marks of the subject of specific group.

```
cse_mean=df1[['CN','DAA','AFL','OE','PE']].mean()
```

```
df1.fillna(cse_mean)
```

```
it_mean=df2[['CN','DAA','AFL','PE']].mean()
```

```
df2.fillna(it_mean)
```

Output:

	Regd.No	AFL	CN	DAA	PE
0	Y20IT001	15.0	12.000000	14.000000	12.000000
1	Y20IT002	13.0	13.000000	15.000000	15.000000
2	Y20IT003	12.0	14.000000	16.000000	14.000000
3	Y20IT004	14.0	14.214286	17.000000	16.000000
4	Y20IT005	16.0	18.000000	13.000000	18.000000
5	Y20IT006	9.0	15.000000	12.000000	17.000000
6	Y20IT007	10.0	22.000000	10.000000	3.000000
7	Y20IT008	11.0	12.000000	11.000000	125.000000
8	Y20IT009	13.0	16.000000	16.000000	21.357143
9	Y20IT010	12.0	14.000000	13.928571	8.000000
10	Y20IT011	14.0	17.000000	14.000000	14.000000
11	Y20IT012	15.0	13.000000	17.000000	17.000000
12	Y20IT013	13.0	11.000000	13.000000	11.000000
13	Y20IT014	12.0	10.000000	12.000000	13.000000
14	Y20IT015	16.0	12.000000	15.000000	16.000000

9.Divide the students into 5 groups based on average marks.

```
a=df1.append(df2)
```

```
df=pd.merge(a,df3,on="Regd.No")
```

```
means=df[['CN','DAA','AFL','PE','OE']].mean()
```

```
d=df.fillna(means)
```

```
d['means']=(d['CN']+d['DAA']+d['AFL']+d['OE']+d['PE'])/5
```

```
d['group']=pd.qcut(d['means'],5,labels=False)
```

d[['Regd.No','Name','group']]

Output:

	Regd.No	Name	group
0	Y20CS001	ADAPA HEMANTH VENKATA SAI PAVAN KUMAR	1
1	Y20CS002	ALAPARTHI VIVEK MADHAV	2
2	Y20CS003	ALIFA SHAIK	4
3	Y20CS004	ALLA NEEHARIKA	2
4	Y20CS005	AVYAKTHA	3
5	Y20CS006	AMBATI MEGHANA	3
6	Y20CS007	ANCHA PRABANDHA	0
7	Y20CS008	APPANA HEMA SRI	3
8	Y20CS009	ARIKATLA VIJAYA LAKSHMI	1
9	Y20CS010	AVULA CHAYA PRIYANKA	0
10	Y20CS011	AVULAPATI ANIL KUMAR	1
11	Y20CS012	BALAGA LAVANYA	2
12	Y20CS013	BANDLA BHAVITHA	3
13	Y20CS014	BEENA VASANTH	1
14	Y20CS015	BELLAM ABHINAY	1
15	Y20IT001	ADAPALA YAVANIKA	1
16	Y20IT002	AMBATIPUDI SWETHA	2
17	Y20IT003	ARE SRILATHA	2
18	Y20IT004	ARVAPALLI SUBHASH	4
19	Y20IT005	ATLA RAGHAVENDRA	4
20	Y20IT006	AVULA GOWTHAM SAI	1
21	Y20IT007	BALABHADRA LAKSHMI VENKATA HANUMA ABHIGNA	0
22	Y20IT008	BANAVATHU LAKSHMIDURGA NAIK	4
23	Y20IT009	BATCHU VENKATA VIBHAS GUPTA	4
24	Y20IT010	BATCHU YASASWINI	0

	Regd.No	Name	group
25	Y20IT011	BATHINI BALAKRISHNA	3
26	Y20IT012	BATTINENI HIMAVANTH	4
27	Y20IT013	BEERAKA REVANTH NAGESH	0
28	Y20IT014	BEZAWADA LOKESH	0
29	Y20IT015	BITRA TEJA SRI	3

10.Create equal sized groups of students based on EAMCET Rank.

```
d['group2']=pd.qcut(d['EAMCET RANK'],6,labels=False)
```

```
d[['Regd.No','Name','group2']]
```

Output:

	Regd.No	Name	group2
0	Y20CS001	ADAPA HEMANTH VENKATA SAI PAVAN KUMAR	0
1	Y20CS002	ALAPARTHI VIVEK MADHAV	0
2	Y20CS003	ALIFA SHAIK	1
3	Y20CS004	ALLA NEEHARIKA	1
4	Y20CS005	AVYAKTHA	2
5	Y20CS006	AMBATI MEGHANA	2
6	Y20CS007	ANCHA PRABANDHA	2
7	Y20CS008	APPANA HEMA SRI	0
8	Y20CS009	ARIKATLA VIJAYA LAKSHMI	4
9	Y20CS010	AVULA CHAYA PRIYANKA	1
10	Y20CS011	AVULAPATI ANIL KUMAR	1
11	Y20CS012	BALAGA LAVANYA	0
12	Y20CS013	BANDLA BHAVITHA	1
13	Y20CS014	BEENA VASANTH	0
14	Y20CS015	BELLAM ABHINAY	5
15	Y20IT001	ADAPALA YAVANIKA	3
16	Y20IT002	AMBATIPUDI SWETHA	3
17	Y20IT003	ARE SRILATHA	3

	Regd.No	Name	group2
18	Y20IT004	ARVAPALLI SUBHASH	5
19	Y20IT005	ATLA RAGHAVENDRA	3
20	Y20IT006	AVULA GOWTHAM SAI	2
21	Y20IT007	BALABHADRA LAKSHMI VENKATA HANUMA ABHIGNA	3
22	Y20IT008	BANAVATHU LAKSHMIDURGA NAIK	5
23	Y20IT009	BATCHU VENKATA VIBHAS GUPTA	4
24	Y20IT010	BATCHU YASASWINI	5
25	Y20IT011	BATHINI BALAKRISHNA	4
26	Y20IT012	BATTINENI HIMAVANTH	5
27	Y20IT013	BEERAKA REVANTH NAGESH	2
28	Y20IT014	BEZAWADA LOKESH	4
29	Y20IT015	BITRA TEJA SRI	4

11.Display the electives and the Regd.No of students who opted the elective along with the subject name.

```
df[(df['OE'].isnull()==False)][['Regd.No','OE']]
```

```
df[(df['PE'].isnull()==False)][['Regd.No','PE']]
```

Output:

	Regd.No	PE
0	Y20CS001	12.0
1	Y20CS002	13.0
3	Y20CS004	17.0
4	Y20CS005	6.0
5	Y20CS006	23.0
6	Y20CS007	16.0
7	Y20CS008	11.0
8	Y20CS009	12.0
9	Y20CS010	15.0
10	Y20CS011	14.0

	Regd.No	PE
12	Y20CS013	10.0
13	Y20CS014	16.0
14	Y20CS015	13.0
15	Y20IT001	12.0
16	Y20IT002	15.0
17	Y20IT003	14.0
18	Y20IT004	16.0
19	Y20IT005	18.0
20	Y20IT006	17.0
21	Y20IT007	3.0
22	Y20IT008	125.0
24	Y20IT010	8.0
25	Y20IT011	14.0
26	Y20IT012	17.0
27	Y20IT013	11.0
28	Y20IT014	13.0
29	Y20IT015	16.0

12.Compare the performance of the students from various cities.

```
r=df.groupby(df['Address']).mean()
print(r)
```

Output:

	CN	DAA	AFL	OE	PE	EAMCET RANK
Address						
CHILAKALURIPET	15.666667	13.666667	15.00	NaN	14.666667	14002.333333
GUNTUR	14.000000	13.222222	15.75	14.0	14.000000	8152.222222
NARASARAOPET	12.666667	18.500000	11.50	15.5	16.333333	11676.500000
TENALI	12.750000	12.333333	10.80	16.0	13.800000	10998.200000
VIJAYAWADA	11.750000	12.750000	13.00	11.5	11.250000	6141.250000
VINUKONDA	14.400000	14.000000	12.00	16.0	34.400000	29000.000000

13.Find the correlation between the marks of DS & DAA

```
df=pd.concat([df1,df2])
df['CN'].corr(df['DAA'])
```

Output: -0.2031529941702545