4.

Consider two excel files having attendance of a workshop's participants for two days. Each file has three fields 'Name', 'Time of joining', duration (in minutes) where names are unique within a file. Note that duration may take one of three values (30, 40, 50) only. Import the data into two dataframes and do the following: .
a. Perform merging of the two dataframes to find the names of students who had attended the workshop on both days.
b. Find names of all students who have attended workshop on either of the days. .
c. Merge two data frames row-wise and find the total number of records in the data frame. .
d. Merge two data frames and use two columns names and duration as multi-row indexes. Generate descriptive statistics for this multi-index. .

In [17]:

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

In [18]:

```
ls
```

```
Prac2 .pdf          Prac3 .pdf          Prac4.ipynb         delete.ipynb
Prac2.ipynb         Prac3.ipynb         Prac_4_Sheet.csv
```

In [19]:

```
df=pd.read_csv("Prac_4_Sheet.csv")
```

In [20]:

```
df.T
```

Out[20]:

	0	1	2	3	4	5	6	7	8	9
Name	Braelyn McIntosh	Kristian Sanders	Emma Strong	Axl Jaramillo	Guadalupe Rush	Kaiser Higgins	Leighton Tang	Rogelio Marin	Celia Pena	Mai McDo
Time of Joining	10:30	11:30	12:30	13:30	14:30	15:30	16:30	17:30	9:00	1(
Duration	50	40	30	50	40	30	50	40	30	
Day	1	1	1	1	1	1	1	1	1	

4 rows × 49 columns

In [21]:

```
df1=df[df['Day']==1].drop('Day',axis=1)
df2=df[df['Day']==2].drop('Day',axis=1)
```

Day 1 Participants

In [22]:

```
df1.T
```

Out[22]:

	0	1	2	3	4	5	6	7	8	9
Name	Braelyn McIntosh	Kristian Sanders	Emma Strong	Axl Jaramillo	Guadalupe Rush	Kaiser Higgins	Leighton Tang	Rogelio Marin	Celia Pena	Mar McDo
Time of Joining	10:30	11:30	12:30	13:30	14:30	15:30	16:30	17:30	9:00	1(
Duration	50	40	30	50	40	30	50	40	30	

3 rows × 28 columns

Day 2 Participants

In [23]:

```
df2.T
```

Out[23]:

	28	29	30	31	32	33	34	35	36	37	...	39
Name	Jocelyn Sawyer	Jefferson Wilson	Luna McCann	Heath Mathis	Anne Wells	Max Snyder	Callie Moses	Niklaus Frost	Paula Park	Enzo Farley	...	[Jol
Time of Joining	16:30	17:30	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	...	
Duration	30	50	40	30	50	40	30	50	40	30	...	

3 rows × 21 columns

A-> Perform merging of the two dataframes to find the names of students who had attended the workshop on both days.

In [24]:

```
pd.merge(df1,df2,'inner',"Name")
```

Out[24]:

	Name	Time of Joining_x	Duration_x	Time of Joining_y	Duration_y
0	Emma Strong	12:30	30	12:15	30
1	Axl Jaramillo	13:30	50	13:15	30
2	Guadalupe Rush	14:30	40	14:15	50
3	Kaiser Higgins	15:30	30	15:15	40
4	Leighton Tang	16:30	50	16:15	30
5	Rogelio Marin	17:30	40	17:15	40
6	Celia Pena	9:00	30	13:00	30
7	Marcus McDonald	10:00	50	14:00	50
8	Daisy Villegas	11:00	40	15:00	40

B-> Find names of all students who have attended workshop on either of the days.

In [27]:

```
pd.merge(df1,df2,'outer').Name
```

Out[27]:

```
0       Braelyn McIntosh
1       Kristian Sanders
2            Emma Strong
3          Axl Jaramillo
4         Guadalupe Rush
5         Kaiser Higgins
6          Leighton Tang
7          Rogelio Marin
8             Celia Pena
9        Marcus McDonald
10        Daisy Villegas
11         Kieran Golden
12     Giuliana Reynolds
13         Vincent Chung
14         Rivka McBride
15        Denver Compton
16         Elina Hawkins
17        Victor Santana
18             Myra Holt
19             Niko Pena
20          Rachel Moran
21           Tate Harper
22              Ana Ward
23     Jameson Chambers
24        Makayla Bauer
25     Kieran Blackwell
26       Saoirse McCall
27         Kiaan Gilbert
28       Jocelyn Sawyer
29     Jefferson Wilson
30           Luna McCann
31          Heath Mathis
32            Anne Wells
33            Max Snyder
34          Callie Moses
35         Niklaus Frost
36             Paula Park
37           Enzo Farley
38       Wrenley Cochran
39         Danny Johnson
40           Emma Strong
41         Axl Jaramillo
42        Guadalupe Rush
43        Kaiser Higgins
44         Leighton Tang
45         Rogelio Marin
46            Celia Pena
47       Marcus McDonald
48        Daisy Villegas
Name: Name, dtype: object
```

C-> Merge two data frames row-wise and find the total number of records in the data frame

In [32]:

```
len(pd.concat([df2,df1],axis=0,ignore_index=1))
```

Out[32]:

```
49
```

D-> Merge two data frames and use two columns names and duration as multi-row indexes. Generate descriptive statistics for this multi-index

In [38]:

```
Df3=pd.concat([df1,df2],axis=0).set_index(['Name','Duration'])
```

In [41]:

```
Df3.T
```

Out[41]:

Name	Braelyn McIntosh	Kristian Sanders	Emma Strong	Axl Jaramillo	Guadalupe Rush	Kaiser Higgins	Leighton Tang	Rogelio Marin	Celia Pena	M M
Duration	**50**	**40**	**30**	**50**	**40**	**30**	**50**	**40**	**30**	**5**
Time of Joining	10:30	11:30	12:30	13:30	14:30	15:30	16:30	17:30	9:00	

1 rows × 49 columns

In [45]:

```
Df3.describe(datetime_is_numeric=1)
```

Out[45]:

	Time of Joining
count	49
unique	27
top	14:00
freq	4

In [ ]:

```

```