Ex3 - Apply Functions

Step 1. Import the necessary libraries

In [1]:

Step 2. Read from student-mat.csv and assign it to a variable called df.

In [2]:

Out[2]:

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	...	famrel
0	GP	F	18	U	GT3	A	4	4	at_home	teacher	...	4
1	GP	F	17	U	GT3	T	1	1	at_home	other	...	5
2	GP	F	15	U	LE3	T	1	1	at_home	other	...	4
3	GP	F	15	U	GT3	T	4	2	health	services	...	3
4	GP	F	16	U	GT3	T	3	3	other	other	...	4

5 rows × 33 columns

Step 3. Create a function called majority that returns a boolean value to a new column called legal_drinker (Consider majority as older than 17 years old)

In [3]:

In [4]:

Out[4]:

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	...	freetin
0	GP	F	18	U	GT3	A	4	4	at_home	teacher	...	
1	GP	F	17	U	GT3	T	1	1	at_home	other	...	
2	GP	F	15	U	LE3	T	1	1	at_home	other	...	
3	GP	F	15	U	GT3	T	4	2	health	services	...	
4	GP	F	16	U	GT3	T	3	3	other	other	...	

5 rows × 34 columns

Ex 4 - Plotting

Step 1. Import the necessary libraries

In [5]:

Step 2. Create the DataFrame it should look like below.

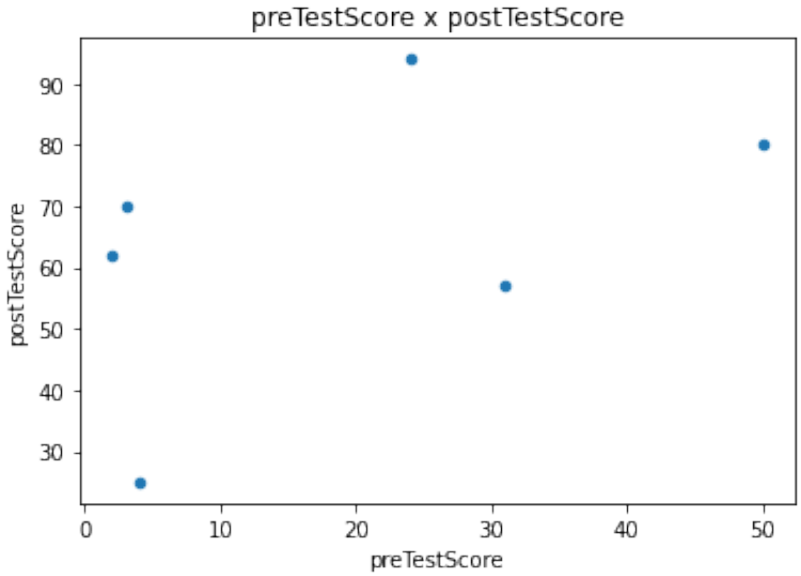
In [6]:

Out[6]:

	first_name	last_name	female	age	preTestScore	postTestScore
0	Jason	Miller	0	42	4	25
1	Molly	Jacobson	1	52	24	94
2	Tina	Ali	1	36	31	57
3	Jake	Milner	0	42	2	62
4	Amy	Cooze	1	42	3	70
5	John	Cooper	0	36	50	80

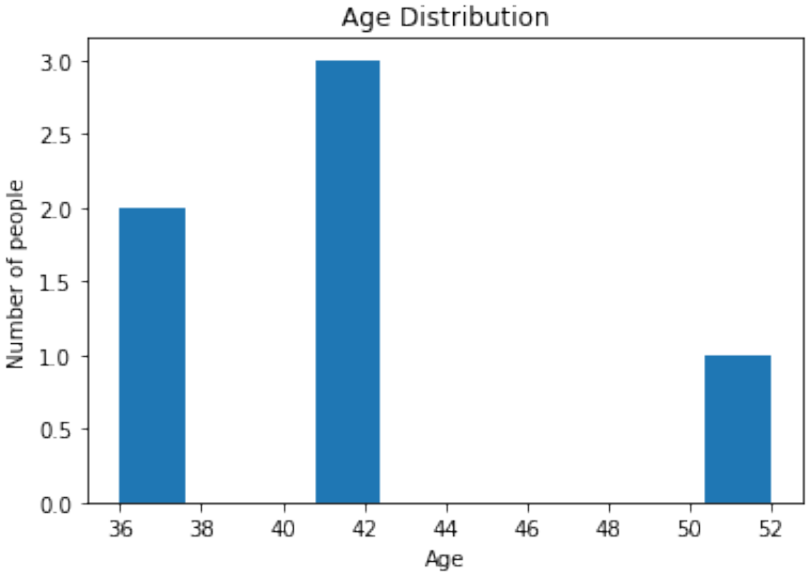
Step 3. Create a Scatterplot with "preTestScore" as the x-axis and "postTestScore" as the y-axis.

In [7]:



Step 4. Create a Histogram using column "Age" to find out the age distribution of participants in dataset.

In [8]:


Age Distribution
Number of people
Age
0.0
0.5
1.0
1.5
2.0
2.5
3.0
36
38
40
42
44
46
48
50
52