Import required libraries

In [40]:
```
import numpy as np
import pandas as pd
```

Create Data Frame

In [41]:
```
df = ['Jason','Miller',42,4,2500],['Molly','Jacobson',52,24,94000],['Tina','.',36,31,57],['Jake','Milner',24,'.',62],['A
```

In [42]:
```
df = pd.DataFrame(df, columns=['First_name','Last_name','Age','PreTestScore','PostTestScore'])
```

In [43]:
```
df
```

Out[43]:

	First_name	Last_name	Age	PreTestScore	PostTestScore
0	Jason	Miller	42	4	2500
1	Molly	Jacobson	52	24	94000
2	Tina	.	36	31	57
3	Jake	Milner	24	.	62
4	Amy	Cooze	73	.	70

Save the data frame into a .csv file as project.csv

In [44]:
```
df.to_csv('project.csv')
```

Read the project.csv file and print the data frame

In [45]:
```
df = pd.read_csv('project.csv')
```

In [46]:
```
print(df)
```

```
   Unnamed: 0 First_name Last_name  Age PreTestScore  PostTestScore
0           0       Jason    Miller   42            4           2500
1           1       Molly  Jacobson   52           24          94000
2           2        Tina         .   36           31             57
3           3        Jake    Milner   24            .             62
4           4         Amy     Cooze   73            .             70
```

Read the project.csv file without column heading

In [54]:
```
df1 = pd.read_csv('project.csv', header=None )
```

In [55]:
```
df1.head()
```

Out[55]:

	0	1	2	3	4	5
0	NaN	First_name	Last_name	Age	PreTestScore	PostTestScore
1	0.0	Jason	Miller	42	4	2500
2	1.0	Molly	Jacobson	52	24	94000
3	2.0	Tina	.	36	31	57
4	3.0	Jake	Milner	24	.	62

Load a csv while specifying column names

In [56]:
```
df = pd.read_csv('project.csv', names =['ID', 'First_name', 'Last_name', 'Age', 'PreTestScore', 'PostTestScore'])
print(df)
```

```
    ID  First_name  Last_name  Age  PreTestScore  PostTestScore
0  NaN  First_name  Last_name  Age  PreTestScore  PostTestScore
1  0.0       Jason     Miller   42             4           2500
2  1.0       Molly   Jacobson   52            24          94000
3  2.0        Tina          .   36            31             57
4  3.0        Jake     Milner   24             .             62
5  4.0         Amy      Cooze   73             .             70
```

Load a csv with setting the index column to ID

In [59]:
```
df= pd.read_csv('project.csv', index_col='ID', names =['ID', 'First_name', 'Last_name', 'Age', 'PreTestScore', 'PostTest
print(df)
```

```
     First_name  Last_name  Age  PreTestScore  PostTestScore
ID
NaN  First_name  Last_name  Age  PreTestScore  PostTestScore
0.0       Jason     Miller   42             4           2500
1.0       Molly   Jacobson   52            24          94000
2.0        Tina          .   36            31             57
3.0        Jake     Milner   24             .             62
4.0         Amy      Cooze   73             .             70
```

Read the project.csv file and make two index columns, namely, ‘First Name’ and ‘Last Name'

In [65]:
```
df= pd.read_csv("project.csv", index_col = ['First_name','Last_name'], names=['ID', 'First_name','Last_name', 'Age', 'Pr
print(df)
```

```
                       ID  Age  PreTestScore  PostTestScore
First_name Last_name
First_name Last_name  NaN  Age  PreTestScore  PostTestScore
Jason      Miller     0.0   42             4           2500
Molly      Jacobson   1.0   52            24          94000
Tina       .          2.0   36            31             57
Jake       Milner     3.0   24             .             62
Amy        Cooze      4.0   73             .             70
```

Print the data frame in a Boolean form as True or False. True for Null/ NaN values and false for non-null values

In [66]:
```
df = pd.read_csv('project.csv', na_values=['.'])
pd.isnull(df)
```

Out[66]:

	Unnamed: 0	First_name	Last_name	Age	PreTestScore	PostTestScore
0	False	False	False	False	False	False
1	False	False	False	False	False	False
2	False	False	True	False	False	False
3	False	False	False	False	True	False
4	False	False	False	False	True	False

Read the data frame by skipping the first 3 rows and print the data frame

In [27]:
```
df = pd.read_csv('project.csv', skiprows=3)
```

In [28]:
```
df.head()
```

Out[28]:

	2	Tina	.	36	31	57
0	3	Jake	Milner	24	.	62
1	4	Amy	Cooze	73	.	70