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

EXERCISE 1

Exercise 1: Create 2 series serA and serB.

- The labels for both series are 'name', 'day', 'time'.
- The values in serA are 'A', 'Mon', 1.
- The values in serB are 'B', 'Fri', 5.
- Give names to the series and print them.
- Print the value for 'time' in serA.
- Print the value for 'day' in serB.

```
# Your solution here
listA = ['A', 'Mon', 1]
listB = ['B', 'Fri', 5]
labels = ['name', 'day', 'time']
serA = pd.Series(data=listA, index=labels)
serB = pd.Series(data=listB, index=labels)

print(serA['time'])
print(serB['day'])

1
Fri
```

EXERCISE 2

Exercise 2: Create a data frame called 'df_cs' from the CSV file 'ComputerSales.csv'.

- Print the first 5 rows.
- Print the first 10 rows.
- Print the total number of rows.
- Print the maximum value of the column 'Profit'
- Print the number of unique values in the column 'Age'
- Print all the rows for which 'Product Type' is 'Tablet'
- Add a column called 'Diff' with values equal to 'Sale Price' - 'Profit'

```
df_cs = pd.read_csv('ComputerSales.csv')
print(df_cs.head()) # first 5 rows

   Sale ID      Contact Sex  Age State Product ID Product Type  Sale
Price  \
0        1  Paul Thomas   M   43    OH  M01-F0024      Desktop
479.99
1        2  Margo Simms   F   37    WV  GT13-0024      Desktop
1249.99
2        3    Sam Stine   M   26    PA      I3670      Desktop
```

```
649.99
3        4    Moe Eggert   M   35     PA       I3593        Laptop
399.99
4        5   Jessica Elk   F   55     PA      15M-ED        Laptop
699.99

   Profit      Lead      Month  Year
0  143.39  Website    January  2018
1  230.89  Flyer 4    January  2018
2  118.64  Website   February  2018
3   72.09  Website      March  2018
4   98.09  Flyer 4      March  2018
```

```python
print(df_cs.head(10)) # first 10 rows
```

```
   Sale ID           Contact Sex  Age State Product ID Product Type  \
0        1      Paul Thomas   M  43     OH  M01-F0024      Desktop
1        2      Margo Simms   F  37     WV  GT13-0024      Desktop
2        3        Sam Stine   M  26     PA      I3670      Desktop
3        4       Moe Eggert   M  35     PA      I3593       Laptop
4        5      Jessica Elk   F  55     PA     15M-ED       Laptop
5        6  Sally Struthers   F  45     PA  GT13-0024      Desktop
6        7   Michelle Samms   F  46     OH    GA401IV       Laptop
7        8     Mick Roberts   M  23     OH    MY2J2LL       Tablet
8        9      Ed Klondike   M  52     OH     81TC00       Laptop
9       10       Phil Jones   M  56     WV  M01-F0024      Desktop

   Sale Price  Profit     Lead     Month  Year
0      479.99  143.39  Website   January  2018
1     1249.99  230.89  Flyer 4   January  2018
2      649.99  118.64  Website  February  2018
3      399.99   72.09  Website     March  2018
4      699.99   98.09  Flyer 4     March  2018
5     1249.99  230.89  Flyer 2     April  2018
6     1349.99  180.34    Email       May  2018
7      999.99  146.69  Website      July  2018
8      649.99  122.34    Email      July  2018
9      479.99  143.39  Flyer 2    August  2018
```

```python
print(f"Number of rows: {df_cs.shape[0]}") # number of rows
```

```
Number of rows: 102
```

```python
print(f"Max value: {max(df_cs['Profit'])}") # max value from the 'Profit' col
```

```
Max value: 1043.39
```

```python
print(f"Number of rows: {df_cs.shape[0]}") # number of rows
```

```
Number of rows: 102
```

```
print(f"Unique values: {df_cs['Age'].nunique()}") # number of unique
values from 'Age' col
```

```
Unique values: 14
```

```
print(df_cs[df_cs['Product Type'] == 'Tablet']) # Product Type is
'Tablet'
```

```
     Sale ID         Contact Sex  Age State     Product ID Product
Type  \
7          8    Mick Roberts   M   23    OH        MY2J2LL
Tablet
14        15      Andy Sands   M   56    OH        MY2J2LL
Tablet
17        18  Michelle Samms   F   46    NY        MY2J2LL
Tablet
34        35  Michelle Samms   F   46    NY        MY2J2LL
Tablet
46        47    Mick Roberts   M   23    OH        MY2J2LL
Tablet
53        54      Andy Sands   M   56    OH  AN515-55-53AG
Tablet
56        57  Michelle Samms   F   46    NY        MY2J2LL
Tablet
72        73    Edna Sanders   F   46    OH  AN515-55-53AG
Tablet
75        76      Jason Case   M   57    PA        MY2J2LL
Tablet
82        83     Jessica Elk   F   55    PA  AN515-55-53AG
Tablet
85        86    Mick Roberts   M   23    OH        MY2J2LL
Tablet
94        95    Doug Johnson   M   51    PA        MY2J2LL
Tablet
101      102     Jessica Elk   F   55    PA  AN515-55-53AG
Tablet

     Sale Price  Profit     Lead     Month  Year
7        999.99  146.69  Website      July  2018
14       999.99  146.69  Flyer 1  December  2018
17       999.99  146.69  Website     March  2019
34       999.99  146.69  Website     April  2020
46       979.99  126.69  Website  December  2018
53       689.99  156.69  Flyer 1  December  2018
56       999.99  146.69  Website    August  2019
72       689.99  156.69  Flyer 1  December  2020
75       999.99  146.69  Website    August  2020
82       669.99  136.69  Flyer 1  December  2018
85       989.99  136.69  Website    August  2018
```

```
94       959.99  106.69  Website   December  2019
101      669.99  136.69  Flyer 1   December  2019

df_cs['Diff'] = df_cs['Sale Price'] - df_cs['Profit']
print(df_cs[['Sale Price', 'Profit', 'Diff']])

     Sale Price  Profit     Diff
0        479.99  143.39   336.60
1       1249.99  230.89  1019.10
2        649.99  118.64   531.35
3        399.99   72.09   327.90
4        699.99   98.09   601.90
..          ...     ...      ...
97       609.99  140.34   469.65
98       889.99  110.89   779.10
99       989.99  111.34   878.65
100      589.99  138.64   451.35
101      669.99  136.69   533.30

[102 rows x 3 columns]
```

EXERCISE 3

Exercise 3: Create a data frame called 'df_deaths' from the CSV file 'Deaths.xlsx'.

- Print the first 2 rows.
- Print the 10th row
- Print the min value of X and Y
- Print the max value of X and Y
- Print the average value of X and Y
- Sort data by X value
- Print the statistics about the data
- Plot the data

```
df_deaths = pd.read_csv('Deaths.csv')
print(df_deaths.head(2)) # first 2 rows

           X         Y
0  13.588010  11.09560
1   9.878124  12.55918

print(df_deaths.iloc[9]) # 10th row

X    11.126390
Y     9.643859
Name: 9, dtype: float64

print(f"Min X: {df_deaths.X.min()} \nMin Y: {df_deaths.Y.min()}") #
min
```

```
Min X: 8.280715
Min Y: 6.090047
```

```
print(f"Max X: {df_deaths.X.max()} \nMax Y: {df_deaths.Y.max()}") #
max
```

```
Max X: 17.93893
Max Y: 16.97276
```

```
print(f"Mean X: {df_deaths.X.mean()} \nMean Y: {df_deaths.Y.mean()}")
# avg
```

```
Mean X: 13.03311612283737
Mean Y: 11.697207534602077
```

```
df_deaths = df_deaths.sort_values(by=['X']) # sort by X
df_deaths
```

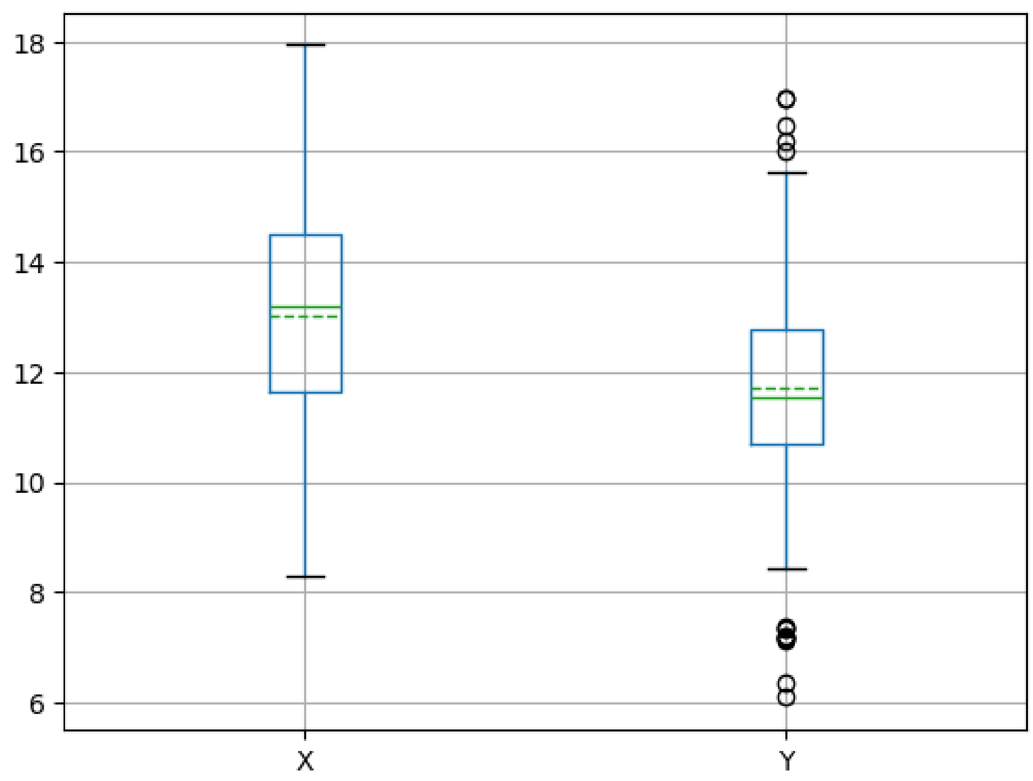
```
             X          Y
434   8.280715  11.568290
108   8.311067   7.202524
312   8.325407   7.166975
201   8.342558   7.136541
117   8.436085   7.393596
..         ...        ...
300  16.839400  11.601700
575  17.271660  11.633800
78   17.515010  11.228880
221  17.595100   7.335869
63   17.938930   7.189272

[578 rows x 2 columns]
```

```
print(df_deaths.describe()) # statistics
```

```
                X           Y
count  578.000000  578.000000
mean    13.033116   11.697208
std      1.953228    1.649661
min      8.280715    6.090047
25%     11.642648   10.680608
50%     13.206795   11.520610
75%     14.515878   12.758265
max     17.938930   16.972760
```

```
df_deaths.boxplot(meanline=True, showmeans=True)
```

```
<Axes: >
```


18
16
14
12
10
8
6
X
Y