27.07.2023

```
In [190]: # import libraries
          import numpy as np
          import pandas as pd
          import matplotlib.pyplot as plt
          import seaborn as sns
```

```
In [191]: a=pd.read_csv(r"C:\Users\user\Downloads\3_Fitness-1.csv")
          a
```

Out[191]:

	Row Labels	Sum of Jan	Sum of Feb	Sum of Mar	Sum of Total Sales
0	A	5.62%	7.73%	6.16%	75
1	B	4.21%	17.27%	19.21%	160
2	C	9.83%	11.60%	5.17%	101
3	D	2.81%	21.91%	7.88%	127
4	E	25.28%	10.57%	11.82%	179
5	F	8.15%	16.24%	18.47%	167
6	G	18.54%	8.76%	17.49%	171
7	H	25.56%	5.93%	13.79%	170
8	Grand Total	100.00%	100.00%	100.00%	1150

```
In [192]: a.info()

          <class 'pandas.core.frame.DataFrame'>
          RangeIndex: 9 entries, 0 to 8
          Data columns (total 5 columns):
           #   Column              Non-Null Count  Dtype
          ---  ------              --------------  -----
           0   Row Labels          9 non-null      object
           1   Sum of Jan          9 non-null      object
           2   Sum of Feb          9 non-null      object
           3   Sum of Mar          9 non-null      object
           4   Sum of Total Sales  9 non-null      int64
          dtypes: int64(1), object(4)
          memory usage: 488.0+ bytes
```

```
In [193]: a.columns
```

Out[193]: Index(['Row Labels', 'Sum of Jan', 'Sum of Feb', 'Sum of Mar',
               'Sum of Total Sales'],
              dtype='object')

```
In [194]: a.head(50)
```

Out[194]:

	Row Labels	Sum of Jan	Sum of Feb	Sum of Mar	Sum of Total Sales
0	A	5.62%	7.73%	6.16%	75
1	B	4.21%	17.27%	19.21%	160
2	C	9.83%	11.60%	5.17%	101
3	D	2.81%	21.91%	7.88%	127
4	E	25.28%	10.57%	11.82%	179
5	F	8.15%	16.24%	18.47%	167
6	G	18.54%	8.76%	17.49%	171
7	H	25.56%	5.93%	13.79%	170
8	Grand Total	100.00%	100.00%	100.00%	1150

```
In [195]: a.describe()
```

Out[195]:

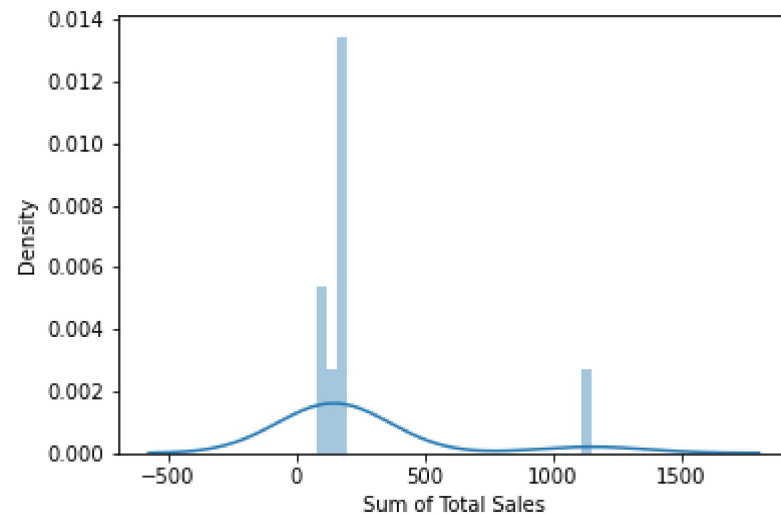|	Sum of Total Sales
count	9.000000
mean	255.555556
std	337.332963
min	75.000000
25%	127.000000
50%	167.000000
75%	171.000000
max	1150.000000

```
In [196]: sns.pairplot(a)
```

Out[196]: <seaborn.axisgrid.PairGrid at 0x18df61bebb0>
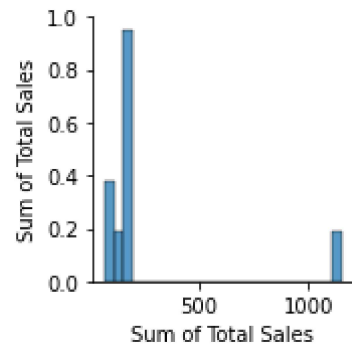
In [197]: `sns.distplot(a['Sum of Total Sales'])`

C:\ProgramData\Anaconda3\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a dep
recated function and will be removed in a future version. Please adapt your code to use either `displot` (a
figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)

Out[197]: `<AxesSubplot:xlabel='Sum of Total Sales', ylabel='Density'>`

In [198]: `x1=a[['Sum of Total Sales']]`

```python
In [199]:  sns.heatmap(x1.corr())

Out[199]:  <AxesSubplot:>
```



```python
In [200]:  from sklearn.model_selection import train_test_split

           x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.3)
```

```python
In [201]:  x=a[['Sum of Total Sales']]
           y=a['Sum of Total Sales']
```

```python
In [202]:  from sklearn.linear_model import LinearRegression
           lr=LinearRegression()
           lr.fit(x_train,y_train)

Out[202]:  LinearRegression()
```

```python
In [203]:  print(lr.intercept_)

           0.0
```
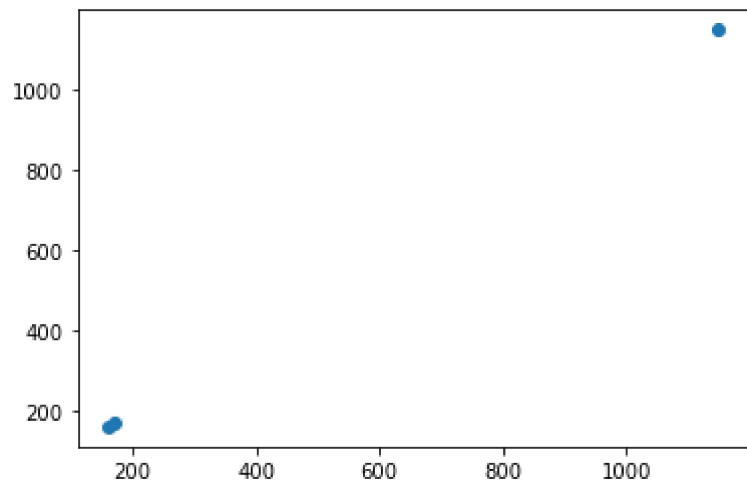
```
In [204]: coeff=pd.DataFrame(lr.coef_,x.columns,columns=['Co-efficient'])
          coeff
```

Out[204]:

	Co-efficient
Sum of Total Sales	1.0

```
In [205]: prediction=lr.predict(x_test)
          plt.scatter(y_test,prediction)
```

Out[205]: <matplotlib.collections.PathCollection at 0x18df6324c10>


```
In [206]: print(lr.score(x_test,y_test))
```

1.0