7/22/23, 5:21 PM fitness\_eda\_dv1

#### 1.IMPORTING LIBRARIES

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

### 2.Importing dataset

In [2]: data=pd.read\_csv(r"C:\Users\user\Downloads\3\_Fitness-1 - 3\_Fitness-1.csv")
 data

Out[2]:		Row Labels	Sum of Jan	Sum of Feb	Sum of Mar	Sum of Total Sales
	0	А	5.62%	7.73%	6.16%	75
	1	В	4.21%	17.27%	19.21%	160
	2	С	9.83%	11.60%	5.17%	101
	3	D	2.81%	21.91%	7.88%	127
	4	Е	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	Н	25.56%	5.93%	13.79%	170
	8	Grand Total	100.00%	100.00%	100.00%	1150

#### 3.head

In [3]: data.head(8)

Out[3]:		Row Labels	Sum of Jan	Sum of Feb	Sum of Mar	Sum of Total Sales
	0	А	5.62%	7.73%	6.16%	75
	1	В	4.21%	17.27%	19.21%	160
	2	С	9.83%	11.60%	5.17%	101
	3	D	2.81%	21.91%	7.88%	127
	4	Е	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	Н	25.56%	5.93%	13.79%	170

#### 4.tail

In [4]:	data.tail(7)		

Out[4]: _	ı	Row Labels	Sum of Jan	Sum of Feb	Sum of Mar	Sum of Total Sales
	2	С	9.83%	11.60%	5.17%	101
3	3	D	2.81%	2.81% 21.91%	7.88%	127
4	4	Е	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	Н	25.56%	5.93%	13.79%	170
:	8	Grand Total	100.00%	100.00%	100.00%	1150

## 5.describe()

```
In [5]: data.describe()
```

Out[5]:		Sum of Total Sales		
	count	9.000000		
	mean	255.555556		
	std	337.332963		
	min	75.000000		
	25%	127.000000		
	50%	167.000000		
	<b>75</b> %	171.000000		
	max	1150.000000		

# 6.shape()

```
In [6]: np.shape(data)
```

Out[6]: (9, 5)

# **7.**size()

```
In [7]: np.size(data)
```

```
Out[7]: 45
```

#### 8.isna()

```
In [8]: data.isna()
Out[8]: Row Labels Sum of Jan Sum of Feb Sum of Mar Sum of Total Sales
```

3]:		Row Labels	Sum of Jan	Sum of Feb	Sum of Mar	Sum of Total Sales
	0	False	False	False	False	False
	1	False	False	False	False	False
	2	False	False	False	False	False
	3	False	False	False	False	False
	4	False	False	False	False	False
	5	False	False	False	False	False
	6	False	False	False	False	False
	7	False	False	False	False	False
	8	False	False	False	False	False

### 9.dropna

```
In [9]: data.dropna()
```

Out[9]:		Row Labels	Sum of Jan	Sum of Feb	Sum of Mar	Sum of Total Sales	
	0	Α	5.62%	7.73%	6.16%	75	
	1	В	4.21%	17.27%	19.21%	160	
	2	С	9.83%	11.60%	5.17%	101	
	3	D	2.81%	21.91%	7.88%	127	
	4	Е	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	Н	25.56%	5.93%	13.79%	170	
	8	Grand Total	100.00%	100.00%	100.00%	1150	

# 10.selecting specific column

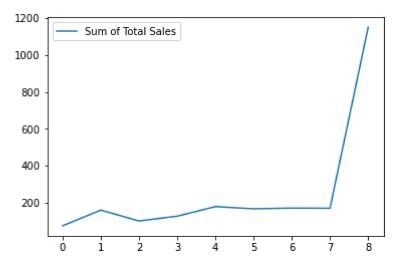
```
In [12]: da=data[["Sum of Jan","Sum of Mar"]] da
```

Out[12]:		Sum of Jan	Sum of Mar
	0	5.62%	6.16%
	1	4.21%	19.21%
	2	9.83%	5.17%
	3	2.81%	7.88%
	4	25.28%	11.82%
	5	8.15%	18.47%
	6	18.54%	17.49%
	7	25.56%	13.79%
	8	100.00%	100.00%

## 11.line plot

```
In [13]: data.plot.line()
```

Out[13]: <AxesSubplot:>

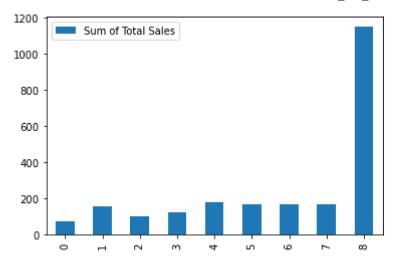


# 12.bar plot

```
In [14]:
    data.plot.bar()
```

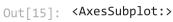
Out[14]: <AxesSubplot:>

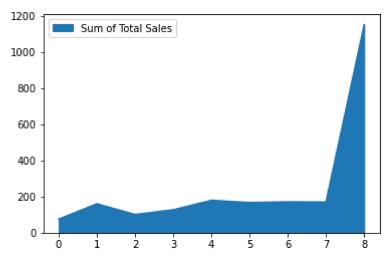
7/22/23, 5:21 PM fitness\_eda\_dv1



## 13.area plot

```
In [15]: data.plot.area()
```



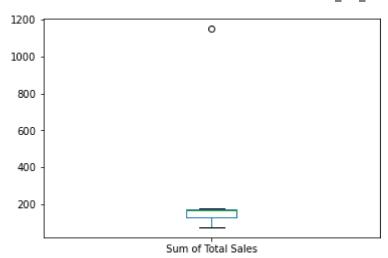


## 14.box plot

```
In [16]: data.plot.box()
```

Out[16]: <AxesSubplot:>

7/22/23, 5:21 PM fitness\_eda\_dv1



## 15.scatter plot

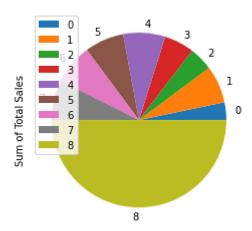
```
In [17]: data.plot.scatter("Sum of Jan", "Sum of Mar")

Out[17]: <AxesSubplot:xlabel='Sum of Jan', ylabel='Sum of Mar'>

100.00%
13.79%
17.49%
5 11.82%
7.88%
5.17%
19.21%
6.16%
5.62% 4.21% 9.83% 2.81% 25.28% 8.15% 18.54% 25.56% 100.00%
Sum of Jan
```

### 16.pie plot

```
In [36]: data.plot.pie(x="Row Labels",y="Sum of Total Sales")
Out[36]: <AxesSubplot:ylabel='Sum of Total Sales'>
```



## 17.histogram

```
In [22]: data.plot.hist()
```

Out[22]: <AxesSubplot:ylabel='Frequency'>

