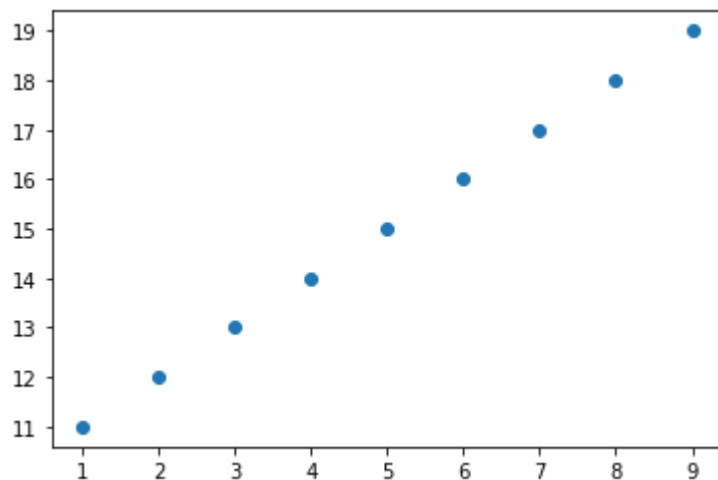
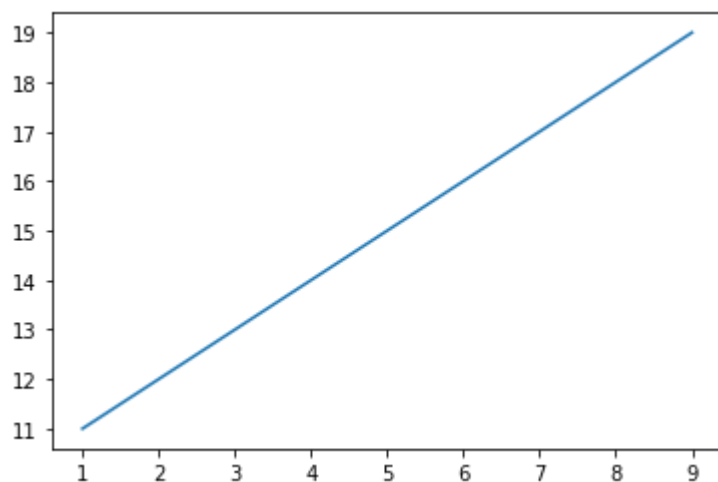


```
plt.show()
```



```
plt.plot(x,y)
```

```
[<matplotlib.lines.Line2D at 0x7f05af155290>]
```



## ▼ Pre-processing

```
import pandas as pd
```

```
df=pd.read_csv('/content/tvmarketing.csv')
```

```
df.head()
```



## Finding Missing values

```
df.isnull().sum()
```

```
TV      0
Sales   0
dtype: int64
```

```
df2=pd.read_csv('/content/train.csv')
```

```
df2.head()
```

	<b>Id</b>	<b>MSSubClass</b>	<b>MSZoning</b>	<b>LotFrontage</b>	<b>LotArea</b>	<b>Street</b>	<b>Alley</b>	<b>LotShape</b>	<b>Lan</b>
<b>0</b>	1	60	RL	65.0	8450	Pave	NaN	Reg	
<b>1</b>	2	20	RL	80.0	9600	Pave	NaN	Reg	
<b>2</b>	3	60	RL	68.0	11250	Pave	NaN	IR1	
<b>3</b>	4	70	RL	60.0	9550	Pave	NaN	IR1	
<b>4</b>	5	60	RL	84.0	14260	Pave	NaN	IR1	

5 rows x 81 columns



```
df2.isnull()
```

	<b>Id</b>	<b>MSSubClass</b>	<b>MSZoning</b>	<b>LotFrontage</b>	<b>LotArea</b>	<b>Street</b>	<b>Alley</b>	<b>LotShape</b>
	0	False	False	False	False	False	True	False
	1	False	False	False	False	False	True	False

df2.fillna(0)

	<b>Id</b>	<b>MSSubClass</b>	<b>MSZoning</b>	<b>LotFrontage</b>	<b>LotArea</b>	<b>Street</b>	<b>Alley</b>	<b>LotShape</b>
<b>0</b>	1	60	RL	65.0	8450	Pave	0	Reg
<b>1</b>	2	20	RL	80.0	9600	Pave	0	Reg
<b>2</b>	3	60	RL	68.0	11250	Pave	0	IR1
<b>3</b>	4	70	RL	60.0	9550	Pave	0	IR1
<b>4</b>	5	60	RL	84.0	14260	Pave	0	IR1
...	...	...	...	...	...	...	...	...
<b>1455</b>	1456	60	RL	62.0	7917	Pave	0	Reg
<b>1456</b>	1457	20	RL	85.0	13175	Pave	0	Reg
<b>1457</b>	1458	70	RL	66.0	9042	Pave	0	Reg
<b>1458</b>	1459	20	RL	68.0	9717	Pave	0	Reg
<b>1459</b>	1460	20	RL	75.0	9937	Pave	0	Reg

1460 rows × 81 columns



```
df3=df2.dropna( )
df3
#it dropped all rows as all rows as all row had some null values
```

	<b>Id</b>	<b>MSSubClass</b>	<b>MSZoning</b>	<b>LotFrontage</b>	<b>LotArea</b>	<b>Street</b>	<b>Alley</b>	<b>LotShape</b>	<b>Land</b>
--	-----------	-------------------	-----------------	--------------------	----------------	---------------	--------------	-----------------	-------------

0 rows × 81 columns



Dropping rows or columns

```
df2.drop([1],axis="rows")
```

	<b>Id</b>	<b>MSSubClass</b>	<b>MSZoning</b>	<b>LotFrontage</b>	<b>LotArea</b>	<b>Street</b>	<b>Alley</b>	<b>LotShape</b>
<b>0</b>	1	60	RL	65.0	8450	Pave	NaN	Reg
<b>2</b>	3	60	RL	68.0	11250	Pave	NaN	IR1
<b>3</b>	4	70	RL	60.0	9550	Pave	NaN	IR1
<b>4</b>	5	60	RL	84.0	14260	Pave	NaN	IR1
<b>5</b>	6	50	RL	85.0	14115	Pave	NaN	IR1
...	...	...	...	...	...	...	...	...
<b>1455</b>	1456	60	RL	62.0	7917	Pave	NaN	Reg
<b>1456</b>	1457	20	RL	85.0	13175	Pave	NaN	Reg
<b>1457</b>	1458	70	RL	66.0	9042	Pave	NaN	Reg
<b>1458</b>	1459	20	RL	68.0	9717	Pave	NaN	Reg
<b>1459</b>	1460	20	RL	75.0	9937	Pave	NaN	Reg

1459 rows x 81 columns

```
df2.drop(["LotArea"],axis="columns")
```

	<b>Id</b>	<b>MSSubClass</b>	<b>MSZoning</b>	<b>LotFrontage</b>	<b>Street</b>	<b>Alley</b>	<b>LotShape</b>	<b>LandCon</b>
<b>0</b>	1	60	RL	65.0	Pave	NaN	Reg	
<b>1</b>	2	20	RL	80.0	Pave	NaN	Reg	
<b>2</b>	3	60	RL	68.0	Pave	NaN	IR1	
<b>3</b>	4	70	RL	60.0	Pave	NaN	IR1	
<b>4</b>	5	60	RL	84.0	Pave	NaN	IR1	
...	...	...	...	...	...	...	...	
<b>1455</b>	1456	60	RL	62.0	Pave	NaN	Reg	
<b>1456</b>	1457	20	RL	85.0	Pave	NaN	Reg	
<b>1457</b>	1458	70	RL	66.0	Pave	NaN	Reg	
<b>1458</b>	1459	20	RL	68.0	Pave	NaN	Reg	
<b>1459</b>	1460	20	RL	75.0	Pave	NaN	Reg	

1460 rows x 80 columns



Encoding

```
from sklearn.preprocessing import LabelEncoder
```

```
le=LabelEncoder()  
  
df2.Street=le.fit_transform(df2.Street)  
df2.head()
```

	<b>Id</b>	<b>MSSubClass</b>	<b>MSZoning</b>	<b>LotFrontage</b>	<b>LotArea</b>	<b>Street</b>	<b>Alley</b>	<b>LotShape</b>	<b>Lan</b>
<b>0</b>	1	60	RL	65.0	8450	1	NaN	Reg	
<b>1</b>	2	20	RL	80.0	9600	1	NaN	Reg	
<b>2</b>	3	60	RL	68.0	11250	1	NaN	IR1	
<b>3</b>	4	70	RL	60.0	9550	1	NaN	IR1	
<b>4</b>	5	60	RL	84.0	14260	1	NaN	IR1	

5 rows × 81 columns



```
df2.LotShape=le.fit_transform(df2.LotShape)  
df2.head()
```

	<b>Id</b>	<b>MSSubClass</b>	<b>MSZoning</b>	<b>LotFrontage</b>	<b>LotArea</b>	<b>Street</b>	<b>Alley</b>	<b>LotShape</b>	<b>Lan</b>
<b>0</b>	1	60	RL	65.0	8450	1	NaN	3	
<b>1</b>	2	20	RL	80.0	9600	1	NaN	3	
<b>2</b>	3	60	RL	68.0	11250	1	NaN	0	
<b>3</b>	4	70	RL	60.0	9550	1	NaN	0	
<b>4</b>	5	60	RL	84.0	14260	1	NaN	0	

5 rows × 81 columns





[Colab paid products](#) - [Cancel contracts here](#)