In [1]:
```
!pip freeze
```

```
anyio==4.6.2.post1
argon2-cffi==23.1.0
argon2-cffi-bindings==21.2.0
arrow==1.3.0
asttokens==2.4.1
async-lru==2.0.4
attrs==24.2.0
babel==2.16.0
beautifulsoup4==4.12.3
bleach==6.2.0
blinker==1.8.2
certifi==2024.8.30
cffi==1.17.1
charset-normalizer==3.4.0
click==8.1.7
colorama==0.4.6
comm==0.2.2
debugpy==1.8.7
decorator==5.1.1
defusedxml==0.7.1
executing==2.1.0
fastjsonschema==2.20.0
Flask==3.0.3
fqdn==1.5.1
h11==0.14.0
httpcore==1.0.6
httpx==0.27.2
idna==3.10
ipykernel==6.29.5
ipython==8.29.0
isoduration==20.11.0
itsdangerous==2.2.0
jedi==0.19.1
Jinja2==3.1.4
json5==0.9.25
jsonpointer==3.0.0
jsonschema==4.23.0
jsonschema-specifications==2024.10.1
jupyter-events==0.10.0
jupyter-lsp==2.2.5
jupyter_client==8.6.3
jupyter_core==5.7.2
jupyter_server==2.14.2
jupyter_server_terminals==0.5.3
jupyterlab==4.2.5
jupyterlab_pygments==0.3.0
jupyterlab_server==2.27.3
MarkupSafe==3.0.2
matplotlib-inline==0.1.7
mistune==3.0.2
nbclient==0.10.0
nbconvert==7.16.4
nbformat==5.10.4
nest-asyncio==1.6.0
notebook==7.2.2
notebook_shim==0.2.4
numpy==2.1.3
overrides==7.7.0
packaging==24.1
pandas==2.2.3
```

```
pandocfilters==1.5.1
parso==0.8.4
platformdirs==4.3.6
prometheus_client==0.21.0
prompt_toolkit==3.0.48
psutil==6.1.0
pure_eval==0.2.3
pycparser==2.22
Pygments==2.18.0
python-dateutil==2.9.0.post0
python-json-logger==2.0.7
pytz==2024.2
pywin32==308
pywinpty==2.0.14
PyYAML==6.0.2
pyzmq==26.2.0
referencing==0.35.1
requests==2.32.3
rfc3339-validator==0.1.4
rfc3986-validator==0.1.1
rpds-py==0.20.1
Send2Trash==1.8.3
six==1.16.0
sniffio==1.3.1
soupsieve==2.6
stack-data==0.6.3
terminado==0.18.1
tinycss2==1.4.0
tornado==6.4.1
traitlets==5.14.3
types-python-dateutil==2.9.0.20241003
typing_extensions==4.12.2
tzdata==2024.2
uri-template==1.3.0
urllib3==2.2.3
wcwidth==0.2.13
webcolors==24.8.0
webencodings==0.5.1
websocket-client==1.8.0
Werkzeug==3.1.1
```

In [2]:
```
import pandas as pd
```

In [3]:
```
df=pd.read_csv('adv.csv')
```

In [4]:
```
df
```

Out[4]:

	Unnamed:0	TV	radio	newspaper	sales
0	1	230.1	37.8	69.2	22.1
1	2	44.5	39.3	45.1	10.4
2	3	17.2	45.9	69.3	9.3
3	4	151.5	41.3	58.5	18.5
4	5	180.8	10.8	58.4	12.9
5	6	175.2	32.4	87.1	19.2
6	7	65.1	54.3	81.2	13.4
7	8	18.2	48.1	73.4	8.2
8	9	240.0	32.0	71.0	24.4
9	10	174.0	59.0	71.2	19.3

In [5]:
```
df.head()
```

Out[5]:

	Unnamed:0	TV	radio	newspaper	sales
0	1	230.1	37.8	69.2	22.1
1	2	44.5	39.3	45.1	10.4
2	3	17.2	45.9	69.3	9.3
3	4	151.5	41.3	58.5	18.5
4	5	180.8	10.8	58.4	12.9

In [6]:
```
df.tail()
```

Out[6]:

	Unnamed:0	TV	radio	newspaper	sales
5	6	175.2	32.4	87.1	19.2
6	7	65.1	54.3	81.2	13.4
7	8	18.2	48.1	73.4	8.2
8	9	240.0	32.0	71.0	24.4
9	10	174.0	59.0	71.2	19.3

In [7]:
```
df.tail(2)
```

Out[7]:

	Unnamed:0	TV	radio	newspaper	sales
8	9	240.0	32.0	71.0	24.4
9	10	174.0	59.0	71.2	19.3

In [8]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10 entries, 0 to 9
Data columns (total 5 columns):
 #   Column      Non-Null Count  Dtype  
---  ------      --------------  -----  
 0   Unnamed:0   10 non-null     int64  
 1   TV          10 non-null     float64
 2   radio       10 non-null     float64
 3   newspaper   10 non-null     float64
 4   sales       10 non-null     float64
dtypes: float64(4), int64(1)
memory usage: 532.0 bytes
```

In [9]:
```
#DROP UNNAMED COLUMN
```

In [10]:
```
df.drop(['Unnamed:0'],axis=1)
df.head()
```

Out[10]:

	Unnamed:0	TV	radio	newspaper	sales
0	1	230.1	37.8	69.2	22.1
1	2	44.5	39.3	45.1	10.4
2	3	17.2	45.9	69.3	9.3
3	4	151.5	41.3	58.5	18.5
4	5	180.8	10.8	58.4	12.9

In [12]:
```
df.drop(['Unnamed:0'],axis=1)
```

Out[12]:

	TV	radio	newspaper	sales
0	230.1	37.8	69.2	22.1
1	44.5	39.3	45.1	10.4
2	17.2	45.9	69.3	9.3
3	151.5	41.3	58.5	18.5
4	180.8	10.8	58.4	12.9
5	175.2	32.4	87.1	19.2
6	65.1	54.3	81.2	13.4
7	18.2	48.1	73.4	8.2
8	240.0	32.0	71.0	24.4
9	174.0	59.0	71.2	19.3

In [13]:
```
df
```

Out[13]:

	Unnamed:0	TV	radio	newspaper	sales
0	1	230.1	37.8	69.2	22.1
1	2	44.5	39.3	45.1	10.4
2	3	17.2	45.9	69.3	9.3
3	4	151.5	41.3	58.5	18.5
4	5	180.8	10.8	58.4	12.9
5	6	175.2	32.4	87.1	19.2
6	7	65.1	54.3	81.2	13.4
7	8	18.2	48.1	73.4	8.2
8	9	240.0	32.0	71.0	24.4
9	10	174.0	59.0	71.2	19.3

In [14]:
```
df=df.drop(['Unnamed:0'],axis=1)
df
```

Out[14]:

	TV	radio	newspaper	sales
0	230.1	37.8	69.2	22.1
1	44.5	39.3	45.1	10.4
2	17.2	45.9	69.3	9.3
3	151.5	41.3	58.5	18.5
4	180.8	10.8	58.4	12.9
5	175.2	32.4	87.1	19.2
6	65.1	54.3	81.2	13.4
7	18.2	48.1	73.4	8.2
8	240.0	32.0	71.0	24.4
9	174.0	59.0	71.2	19.3

In [15]:
```
#SPLIT DATA IN FEATURES AND TARGET
```

In [16]:
```
tobesplitted=df.loc[:,df.columns != 'sales']
```

In [17]:
```
tobesplitted
```

Out[17]:

	TV	radio	newspaper
0	230.1	37.8	69.2
1	44.5	39.3	45.1
2	17.2	45.9	69.3
3	151.5	41.3	58.5
4	180.8	10.8	58.4
5	175.2	32.4	87.1
6	65.1	54.3	81.2
7	18.2	48.1	73.4
8	240.0	32.0	71.0
9	174.0	59.0	71.2

In [18]:
```
result=df['sales']
```

In [19]:
```
result
```

Out[19]:
```
0    22.1
1    10.4
2     9.3
3    18.5
4    12.9
5    19.2
6    13.4
7     8.2
8    24.4
9    19.3
Name: sales, dtype: float64
```

In [20]:
```
X=df.loc[:,df.columns != 'sales']
```

In [21]:
```
X
```

Out[21]:

	TV	radio	newspaper
0	230.1	37.8	69.2
1	44.5	39.3	45.1
2	17.2	45.9	69.3
3	151.5	41.3	58.5
4	180.8	10.8	58.4
5	175.2	32.4	87.1
6	65.1	54.3	81.2
7	18.2	48.1	73.4
8	240.0	32.0	71.0
9	174.0	59.0	71.2

In [22]:
```
Y=df['sales']
```

In [23]:
```
X
```

Out[23]:

	TV	radio	newspaper
0	230.1	37.8	69.2
1	44.5	39.3	45.1
2	17.2	45.9	69.3
3	151.5	41.3	58.5
4	180.8	10.8	58.4
5	175.2	32.4	87.1
6	65.1	54.3	81.2
7	18.2	48.1	73.4
8	240.0	32.0	71.0
9	174.0	59.0	71.2

In [24]:
```
Y
```

Out[24]:
```
0    22.1
1    10.4
2     9.3
3    18.5
4    12.9
5    19.2
6    13.4
7     8.2
8    24.4
9    19.3
Name: sales, dtype: float64
```

In [25]:
```
!pip install scikit-learn
```

```
Collecting scikit-learn
  Downloading scikit_learn-1.5.2-cp311-cp311-win_amd64.whl.metadata (13 kB)
Requirement already satisfied: numpy>=1.19.5 in d:\advertize\lib\site-packages (f
rom scikit-learn) (2.1.3)
Collecting scipy>=1.6.0 (from scikit-learn)
  Downloading scipy-1.14.1-cp311-cp311-win_amd64.whl.metadata (60 kB)
Collecting joblib>=1.2.0 (from scikit-learn)
  Downloading joblib-1.4.2-py3-none-any.whl.metadata (5.4 kB)
Collecting threadpoolctl>=3.1.0 (from scikit-learn)
  Downloading threadpoolctl-3.5.0-py3-none-any.whl.metadata (13 kB)
Downloading scikit_learn-1.5.2-cp311-cp311-win_amd64.whl (11.0 MB)
   ---------------------------------------- 0.0/11.0 MB ? eta -:--:--
   ---------------------------------------  10.7/11.0 MB 55.8 MB/s eta 0:00:01
   ---------------------------------------- 11.0/11.0 MB 49.1 MB/s eta 0:00:00
Downloading joblib-1.4.2-py3-none-any.whl (301 kB)
Downloading scipy-1.14.1-cp311-cp311-win_amd64.whl (44.8 MB)
   ---------------------------------------- 0.0/44.8 MB ? eta -:--:--
   ------------ -------------------------- 13.9/44.8 MB 66.8 MB/s eta 0:00:01
   -------------- ------------------------ 17.0/44.8 MB 42.9 MB/s eta 0:00:01
   ------------------ --------------------- 21.0/44.8 MB 34.9 MB/s eta 0:00:01
   ---------------------- ----------------- 24.6/44.8 MB 30.0 MB/s eta 0:00:01
   ------------------------- -------------- 28.3/44.8 MB 27.6 MB/s eta 0:00:01
   ---------------------------- ----------- 31.7/44.8 MB 25.5 MB/s eta 0:00:01
   ------------------------------- -------- 35.7/44.8 MB 24.6 MB/s eta 0:00:01
   ---------------------------------- ----- 39.3/44.8 MB 23.6 MB/s eta 0:00:01
   ------------------------------------ --- 42.7/44.8 MB 23.0 MB/s eta 0:00:01
   ---------------------------------------- 44.8/44.8 MB 21.9 MB/s eta 0:00:00
Downloading threadpoolctl-3.5.0-py3-none-any.whl (18 kB)
Installing collected packages: threadpoolctl, scipy, joblib, scikit-learn
Successfully installed joblib-1.4.2 scikit-learn-1.5.2 scipy-1.14.1 threadpoolctl
-3.5.0
```

In [26]:
```
from sklearn.model_selection import train_test_split
```

In [27]:
```
X_train,X_test,y_train,y_test=train_test_split(X,Y,test_size=0.25)
```

In [28]:
```
print(X_train.shape)
```

```
(7, 3)
```

In [29]:
```
print(X_train.head())
```

```
      TV  radio  newspaper
9  174.0   59.0       71.2
6   65.1   54.3       81.2
1   44.5   39.3       45.1
0  230.1   37.8       69.2
8  240.0   32.0       71.0
```

In [32]:
```
print(X_test.head())
print(X_test.shape)
```

```
      TV  radio  newspaper
7   18.2   48.1       73.4
4  180.8   10.8       58.4
5  175.2   32.4       87.1
(3, 3)
```

In [33]:
```
#MODEL CREATION
```

In [34]:
```
from sklearn.linear_model import LinearRegression
```

In [35]:
```
model=LinearRegression()
```

In [36]:
```
#TRAIN THE MODEL
```

In [37]:
```
model.fit(X_train,y_train)
```

Out[37]:
```
LinearRegression
LinearRegression()
```

In [38]:
```
prediction=model.predict(X_test)
```

In [39]:
```
prediction[:10]
```

Out[39]:
```
array([ 9.88481262, 19.96185082, 20.45316035])
```

In [40]:
```
prediction
```

Out[40]:
```
array([ 9.88481262, 19.96185082, 20.45316035])
```

In [41]:
```
print(y_test.head())
```

```
7     8.2
4    12.9
5    19.2
Name: sales, dtype: float64
```

In [42]:
```
#SAVE THE MODEL
```

In [43]:
```
import joblib
```

In [44]:
```
joblib.dump(model,'preetha.pkl')
```

Out[44]:
```
['preetha.pkl']
```

In [45]:
```
#TEST WITH NEW DATA
```

In [46]:
```
import numpy as np
```

In [100…
```
data=[[259.1,167.1,190.3]]
```

In [48]:
```
#convert to numpy array and reshape
```

In [101…
```
data=np.array(data).astype(float)
```

In [102…
```
data.reshape(1,-1)
```

Out[102…
```
array([[259.1, 167.1, 190.3]])
```

In [103…
```
data
```

Out[103…
```
array([[259.1, 167.1, 190.3]])
```

In [94]:
```
#load the saved data
```

In [104…
```
f='preetha.pkl'
```

In [105…
```
file=open(f,'rb')
```

In [106…
```
trained_model=joblib.load(file)
```

In [107…
```
prediction=trained_model.predict(data)
```

```
D:\ADVERTIZE\Lib\site-packages\sklearn\base.py:493: UserWarning: X does not have
valid feature names, but LinearRegression was fitted with feature names
  warnings.warn(
```

In [108…
```
print(prediction)
```

```
[27.97810529]
```

In [109…
```
#Now create a web application to host in flask then transfer the total project i
```

In [ ]:
```

```