

```
! pip install [https://github.com/pandas-profiling/pandas-  
profiling/archive/master.zip](https://github.com/pandas-profiling/  
pandas-profiling/archive/master.zip)
```

```
/bin/bash: -c: line 0: syntax error near unexpected token `('`  
/bin/bash: -c: line 0: `pip install [https://github.com/pandas-  
profiling/pandas-profiling/archive/master.zip](https://github.com/  
pandas-profiling/pandas-profiling/archive/master.zip)'
```

```
pip install pandas-profiling
```

```
Looking in indexes: https://pypi.org/simple, https://us-  
python.pkg.dev/colab-wheels/public/simple/  
Requirement already satisfied: pandas-profiling in  
/usr/local/lib/python3.9/dist-packages (3.6.6)  
Requirement already satisfied: ydata-profiling in  
/usr/local/lib/python3.9/dist-packages (from pandas-profiling) (4.1.2)  
Requirement already satisfied: numpy<1.24,>=1.16.0 in  
/usr/local/lib/python3.9/dist-packages (from ydata-profiling->pandas-  
profiling) (1.22.4)  
Requirement already satisfied: tqdm<4.65,>=4.48.2 in  
/usr/local/lib/python3.9/dist-packages (from ydata-profiling->pandas-  
profiling) (4.64.1)  
Requirement already satisfied: jinja2<3.2,>=2.11.1 in  
/usr/local/lib/python3.9/dist-packages (from ydata-profiling->pandas-  
profiling) (3.1.2)  
Requirement already satisfied: requests<2.29,>=2.24.0 in  
/usr/local/lib/python3.9/dist-packages (from ydata-profiling->pandas-  
profiling) (2.27.1)  
Requirement already satisfied: htmlmin==0.1.12 in  
/usr/local/lib/python3.9/dist-packages (from ydata-profiling->pandas-  
profiling) (0.1.12)  
Requirement already satisfied: phik<0.13,>=0.11.1 in  
/usr/local/lib/python3.9/dist-packages (from ydata-profiling->pandas-  
profiling) (0.12.3)  
Requirement already satisfied: typeguard<2.14,>=2.13.2 in  
/usr/local/lib/python3.9/dist-packages (from ydata-profiling->pandas-  
profiling) (2.13.3)  
Requirement already satisfied: seaborn<0.13,>=0.10.1 in  
/usr/local/lib/python3.9/dist-packages (from ydata-profiling->pandas-  
profiling) (0.12.2)  
Requirement already satisfied: imagehash==4.3.1 in  
/usr/local/lib/python3.9/dist-packages (from ydata-profiling->pandas-  
profiling) (4.3.1)  
Requirement already satisfied: matplotlib<3.7,>=3.2 in  
/usr/local/lib/python3.9/dist-packages (from ydata-profiling->pandas-  
profiling) (3.6.3)  
Requirement already satisfied: statsmodels<0.14,>=0.13.2 in  
/usr/local/lib/python3.9/dist-packages (from ydata-profiling->pandas-  
profiling) (0.13.5)  
Requirement already satisfied: multimethod<1.10,>=1.4 in
```

/usr/local/lib/python3.9/dist-packages (from ydata-profiling->pandas-profiling) (1.9.1)
Requirement already satisfied: visions[type_image_path]==0.7.5 in /usr/local/lib/python3.9/dist-packages (from ydata-profiling->pandas-profiling) (0.7.5)
Requirement already satisfied: pandas!=1.4.0,<1.6,>1.1 in /usr/local/lib/python3.9/dist-packages (from ydata-profiling->pandas-profiling) (1.5.3)
Requirement already satisfied: pydantic<1.11,>=1.8.1 in /usr/local/lib/python3.9/dist-packages (from ydata-profiling->pandas-profiling) (1.10.7)
Requirement already satisfied: PyYAML<6.1,>=5.0.0 in /usr/local/lib/python3.9/dist-packages (from ydata-profiling->pandas-profiling) (5.4.1)
Requirement already satisfied: scipy<1.10,>=1.4.1 in /usr/local/lib/python3.9/dist-packages (from ydata-profiling->pandas-profiling) (1.9.3)
Requirement already satisfied: PyWavelets in /usr/local/lib/python3.9/dist-packages (from imagehash==4.3.1->ydata-profiling->pandas-profiling) (1.4.1)
Requirement already satisfied: pillow in /usr/local/lib/python3.9/dist-packages (from imagehash==4.3.1->ydata-profiling->pandas-profiling) (8.4.0)
Requirement already satisfied: attrs>=19.3.0 in /usr/local/lib/python3.9/dist-packages (from visions[type_image_path]==0.7.5->ydata-profiling->pandas-profiling) (22.2.0)
Requirement already satisfied: networkx>=2.4 in /usr/local/lib/python3.9/dist-packages (from visions[type_image_path]==0.7.5->ydata-profiling->pandas-profiling) (3.1)
Requirement already satisfied: tangled-up-in-unicode>=0.0.4 in /usr/local/lib/python3.9/dist-packages (from visions[type_image_path]==0.7.5->ydata-profiling->pandas-profiling) (0.2.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.9/dist-packages (from jinja2<3.2,>=2.11.1->ydata-profiling->pandas-profiling) (2.1.2)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.9/dist-packages (from matplotlib<3.7,>=3.2->ydata-profiling->pandas-profiling) (23.0)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.9/dist-packages (from matplotlib<3.7,>=3.2->ydata-profiling->pandas-profiling) (4.39.3)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.9/dist-packages (from matplotlib<3.7,>=3.2->ydata-profiling->pandas-profiling) (1.0.7)
Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.9/dist-packages (from matplotlib<3.7,>=3.2->ydata-profiling->pandas-profiling) (1.4.4)

Requirement already satisfied: pyparsing>=2.2.1 in
/usr/local/lib/python3.9/dist-packages (from matplotlib<3.7,>=3.2-
>ydata-profiling->pandas-profiling) (3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in
/usr/local/lib/python3.9/dist-packages (from matplotlib<3.7,>=3.2-
>ydata-profiling->pandas-profiling) (2.8.2)
Requirement already satisfied: cycloper>=0.10 in
/usr/local/lib/python3.9/dist-packages (from matplotlib<3.7,>=3.2-
>ydata-profiling->pandas-profiling) (0.11.0)
Requirement already satisfied: pytz>=2020.1 in
/usr/local/lib/python3.9/dist-packages (from pandas!=1.4.0,<1.6,>1.1-
>ydata-profiling->pandas-profiling) (2022.7.1)
Requirement already satisfied: joblib>=0.14.1 in
/usr/local/lib/python3.9/dist-packages (from phik<0.13,>=0.11.1-
>ydata-profiling->pandas-profiling) (1.2.0)
Requirement already satisfied: typing-extensions>=4.2.0 in
/usr/local/lib/python3.9/dist-packages (from pydantic<1.11,>=1.8.1-
>ydata-profiling->pandas-profiling) (4.5.0)
Requirement already satisfied: charset-normalizer~=2.0.0 in
/usr/local/lib/python3.9/dist-packages (from requests<2.29,>=2.24.0-
>ydata-profiling->pandas-profiling) (2.0.12)
Requirement already satisfied: idna<4,>=2.5 in
/usr/local/lib/python3.9/dist-packages (from requests<2.29,>=2.24.0-
>ydata-profiling->pandas-profiling) (3.4)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.9/dist-packages (from requests<2.29,>=2.24.0-
>ydata-profiling->pandas-profiling) (2022.12.7)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in
/usr/local/lib/python3.9/dist-packages (from requests<2.29,>=2.24.0-
>ydata-profiling->pandas-profiling) (1.26.15)
Requirement already satisfied: patsy>=0.5.2 in
/usr/local/lib/python3.9/dist-packages (from
statsmodels<0.14,>=0.13.2->ydata-profiling->pandas-profiling) (0.5.3)
Requirement already satisfied: six in /usr/local/lib/python3.9/dist-
packages (from patsy>=0.5.2->statsmodels<0.14,>=0.13.2->ydata-
profiling->pandas-profiling) (1.16.0)

```
import pandas_profiling
```

```
import pandas as pd
```

```
import matplotlib.pyplot as plt
```

```
import pickle
```

```
!pip install cartopy
```

```
import cartopy
```

Looking in indexes: <https://pypi.org/simple>, <https://us-python.pkg.dev/colab-wheels/public/simple/>

Requirement already satisfied: cartopy in

```
/usr/local/lib/python3.9/dist-packages (0.21.1)
Requirement already satisfied: pyshp>=2.1 in
/usr/local/lib/python3.9/dist-packages (from cartopy) (2.3.1)
Requirement already satisfied: pyproj>=3.0.0 in
/usr/local/lib/python3.9/dist-packages (from cartopy) (3.5.0)
Requirement already satisfied: matplotlib>=3.1 in
/usr/local/lib/python3.9/dist-packages (from cartopy) (3.6.3)
Requirement already satisfied: numpy>=1.18 in
/usr/local/lib/python3.9/dist-packages (from cartopy) (1.22.4)
Requirement already satisfied: shapely>=1.6.4 in
/usr/local/lib/python3.9/dist-packages (from cartopy) (2.0.1)
Requirement already satisfied: pillow>=6.2.0 in
/usr/local/lib/python3.9/dist-packages (from matplotlib>=3.1->cartopy)
(8.4.0)
Requirement already satisfied: contourpy>=1.0.1 in
/usr/local/lib/python3.9/dist-packages (from matplotlib>=3.1->cartopy)
(1.0.7)
Requirement already satisfied: cycler>=0.10 in
/usr/local/lib/python3.9/dist-packages (from matplotlib>=3.1->cartopy)
(0.11.0)
Requirement already satisfied: pyparsing>=2.2.1 in
/usr/local/lib/python3.9/dist-packages (from matplotlib>=3.1->cartopy)
(3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in
/usr/local/lib/python3.9/dist-packages (from matplotlib>=3.1->cartopy)
(2.8.2)
Requirement already satisfied: kiwisolver>=1.0.1 in
/usr/local/lib/python3.9/dist-packages (from matplotlib>=3.1->cartopy)
(1.4.4)
Requirement already satisfied: fonttools>=4.22.0 in
/usr/local/lib/python3.9/dist-packages (from matplotlib>=3.1->cartopy)
(4.39.3)
Requirement already satisfied: packaging>=20.0 in
/usr/local/lib/python3.9/dist-packages (from matplotlib>=3.1->cartopy)
(23.0)
Requirement already satisfied: certifi in
/usr/local/lib/python3.9/dist-packages (from pyproj>=3.0.0->cartopy)
(2022.12.7)
Requirement already satisfied: six>=1.5 in
/usr/local/lib/python3.9/dist-packages (from python-dateutil>=2.7-
>matplotlib>=3.1->cartopy) (1.16.0)
```

```
import numpy as np
```

```
from google.colab import files
```

```
uploaded = files.upload()
```

```
<IPython.core.display.HTML object>
```

Saving advertising.csv to advertising (1).csv

```
df = pd.read_csv('advertising.csv')
```

```
df.head()
```

	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	12.0
3	151.5	41.3	58.5	16.5
4	180.8	10.8	58.4	17.9

```
df
```

	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	12.0
3	151.5	41.3	58.5	16.5
4	180.8	10.8	58.4	17.9
...
195	38.2	3.7	13.8	7.6
196	94.2	4.9	8.1	14.0
197	177.0	9.3	6.4	14.8
198	283.6	42.0	66.2	25.5
199	232.1	8.6	8.7	18.4

[200 rows x 4 columns]

```
df.head(10)
```

	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	12.0
3	151.5	41.3	58.5	16.5
4	180.8	10.8	58.4	17.9
5	8.7	48.9	75.0	7.2
6	57.5	32.8	23.5	11.8
7	120.2	19.6	11.6	13.2
8	8.6	2.1	1.0	4.8
9	199.8	2.6	21.2	15.6

```
df.tail()
```

	TV	Radio	Newspaper	Sales
195	38.2	3.7	13.8	7.6
196	94.2	4.9	8.1	14.0
197	177.0	9.3	6.4	14.8
198	283.6	42.0	66.2	25.5
199	232.1	8.6	8.7	18.4

```
df.describe()
```

	TV	Radio	Newspaper	Sales
count	200.000000	200.000000	200.000000	200.000000
mean	147.042500	23.264000	30.554000	15.130500
std	85.854236	14.846809	21.778621	5.283892
min	0.700000	0.000000	0.300000	1.600000
25%	74.375000	9.975000	12.750000	11.000000
50%	149.750000	22.900000	25.750000	16.000000
75%	218.825000	36.525000	45.100000	19.050000
max	296.400000	49.600000	114.000000	27.000000

```
from pandas_profiling import ProfileReport
```

```
ProfileReport(df)
```

```
{"model_id": "df6a984b03434b5bbd7eaf31e0e6220c", "version_major": 2, "version_minor": 0}
```

```
{"model_id": "f4d112887fde43fcaea09051b972bfb8", "version_major": 2, "version_minor": 0}
```

```
{"model_id": "9779c8393d2b496bbf44a003ff430145", "version_major": 2, "version_minor": 0}
```

```
<IPython.core.display.HTML object>
```

```
pf = ProfileReport(df)
```

```
pf.to_file('Advertising.html')
```

```
{"model_id": "eab31ce2b1764727b93fe9a75926c2a5", "version_major": 2, "version_minor": 0}
```

```
{"model_id": "baa5d56ae9bb40a3a97676cd14b0d574", "version_major": 2, "version_minor": 0}
```

```
{"model_id": "f2dedaa37e6a4d25b436d89fe58366ba", "version_major": 2, "version_minor": 0}
```

```
{"model_id": "3eeldbf7141448f390ef99706b711f54", "version_major": 2, "version_minor": 0}
```

```
pf.to_notebook_iframe()
```

```
<IPython.core.display.HTML object>
```

```
!pip install pyyaml==5.4.1
```

Looking in indexes: <https://pypi.org/simple>, <https://us-python.pkg.dev/colab-wheels/public/simple/>

Requirement already satisfied: pyyaml==5.4.1 in
/usr/local/lib/python3.9/dist-packages (5.4.1)

df

	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	12.0
3	151.5	41.3	58.5	16.5
4	180.8	10.8	58.4	17.9
...
195	38.2	3.7	13.8	7.6
196	94.2	4.9	8.1	14.0
197	177.0	9.3	6.4	14.8
198	283.6	42.0	66.2	25.5
199	232.1	8.6	8.7	18.4

[200 rows x 4 columns]

Model Building of TV Vs Sales

```
x = df[["TV"]]
```

x

	TV
0	230.1
1	44.5
2	17.2
3	151.5
4	180.8
...	...
195	38.2
196	94.2
197	177.0
198	283.6
199	232.1

[200 rows x 1 columns]

```
y = df.Sales
```

y

0	22.1
1	10.4
2	12.0
3	16.5
4	17.9
...	...
195	7.6

```

196     14.0
197     14.8
198     25.5
199     18.4
Name: Sales, Length: 200, dtype: float64

from sklearn.linear_model import LinearRegression
linear = LinearRegression()
linear.fit(x,y)
LinearRegression()
linear.intercept_
6.974821488229891
linear.coef_
array([0.05546477])
linear.predict
<bound method LinearModel.predict of LinearRegression(>
file = 'linear_reg.save'
pickle.dump(linear, open(file, 'wb'))
saved = pickle.load(open(file, 'rb'))
saved.predict([[45]])

/usr/local/lib/python3.9/dist-packages/sklearn/base.py:439:
UserWarning: X does not have valid feature names, but LinearRegression
was fitted with feature names
  warnings.warn(
array([9.47073616])

```