advertisement_model

September 24, 2023

[3]:
```
import pandas as pd
import matplotlib.pyplot as plt
import pickle
from pandas_profiling import ProfileReport
import numpy as np
```

```
/tmp/ipykernel_3542/4240260412.py:4: DeprecationWarning: `import
pandas_profiling` is going to be deprecated by April 1st. Please use `import
ydata_profiling` instead.
  from pandas_profiling import ProfileReport
```

[2]:
```
import sys
!{sys.executable} -m pip install pandas-profiling
```

```
Requirement already satisfied: pandas-profiling in
/opt/conda/lib/python3.10/site-packages (3.6.6)
Requirement already satisfied: ydata-profiling in
/opt/conda/lib/python3.10/site-packages (from pandas-profiling) (4.5.1)
Requirement already satisfied: PyYAML<6.1,>=5.0.0 in
/opt/conda/lib/python3.10/site-packages (from ydata-profiling->pandas-profiling)
(6.0)
Requirement already satisfied: requests<3,>=2.24.0 in
/opt/conda/lib/python3.10/site-packages (from ydata-profiling->pandas-profiling)
(2.28.1)
Requirement already satisfied: jinja2<3.2,>=2.11.1 in
/opt/conda/lib/python3.10/site-packages (from ydata-profiling->pandas-profiling)
(3.1.2)
Requirement already satisfied: visions[type_image_path]==0.7.5 in
/opt/conda/lib/python3.10/site-packages (from ydata-profiling->pandas-profiling)
(0.7.5)
Requirement already satisfied: tqdm<5,>=4.48.2 in
/opt/conda/lib/python3.10/site-packages (from ydata-profiling->pandas-profiling)
(4.64.1)
Requirement already satisfied: phik<0.13,>=0.11.1 in
/opt/conda/lib/python3.10/site-packages (from ydata-profiling->pandas-profiling)
(0.12.3)
Requirement already satisfied: multimethod<2,>=1.4 in
/opt/conda/lib/python3.10/site-packages (from ydata-profiling->pandas-profiling)
```

```
(1.10)
Requirement already satisfied: wordcloud>=1.9.1 in
/opt/conda/lib/python3.10/site-packages (from ydata-profiling->pandas-profiling)
(1.9.2)
Requirement already satisfied: statsmodels<1,>=0.13.2 in
/opt/conda/lib/python3.10/site-packages (from ydata-profiling->pandas-profiling)
(0.13.2)
Requirement already satisfied: pydantic<2,>=1.8.1 in
/opt/conda/lib/python3.10/site-packages (from ydata-profiling->pandas-profiling)
(1.10.12)
Requirement already satisfied: seaborn<0.13,>=0.10.1 in
/opt/conda/lib/python3.10/site-packages (from ydata-profiling->pandas-profiling)
(0.12.0)
Requirement already satisfied: pandas!=1.4.0,<2.1,>1.1 in
/opt/conda/lib/python3.10/site-packages (from ydata-profiling->pandas-profiling)
(1.5.1)
Requirement already satisfied: imagehash==4.3.1 in
/opt/conda/lib/python3.10/site-packages (from ydata-profiling->pandas-profiling)
(4.3.1)
Requirement already satisfied: typeguard<3,>=2.13.2 in
/opt/conda/lib/python3.10/site-packages (from ydata-profiling->pandas-profiling)
(2.13.3)
Requirement already satisfied: scipy<1.12,>=1.4.1 in
/opt/conda/lib/python3.10/site-packages (from ydata-profiling->pandas-profiling)
(1.9.3)
Requirement already satisfied: numpy<1.24,>=1.16.0 in
/opt/conda/lib/python3.10/site-packages (from ydata-profiling->pandas-profiling)
(1.23.4)
Requirement already satisfied: dacite>=1.8 in /opt/conda/lib/python3.10/site-
packages (from ydata-profiling->pandas-profiling) (1.8.1)
Requirement already satisfied: matplotlib<4,>=3.2 in
/opt/conda/lib/python3.10/site-packages (from ydata-profiling->pandas-profiling)
(3.6.1)
Requirement already satisfied: htmlmin==0.1.12 in
/opt/conda/lib/python3.10/site-packages (from ydata-profiling->pandas-profiling)
(0.1.12)
Requirement already satisfied: pillow in /opt/conda/lib/python3.10/site-packages
(from imagehash==4.3.1->ydata-profiling->pandas-profiling) (9.2.0)
Requirement already satisfied: PyWavelets in /opt/conda/lib/python3.10/site-
packages (from imagehash==4.3.1->ydata-profiling->pandas-profiling) (1.3.0)
Requirement already satisfied: networkx>=2.4 in /opt/conda/lib/python3.10/site-
packages (from visions[type_image_path]==0.7.5->ydata-profiling->pandas-
profiling) (2.8.7)
Requirement already satisfied: tangled-up-in-unicode>=0.0.4 in
/opt/conda/lib/python3.10/site-packages (from
visions[type_image_path]==0.7.5->ydata-profiling->pandas-profiling) (0.2.0)
Requirement already satisfied: attrs>=19.3.0 in /opt/conda/lib/python3.10/site-
packages (from visions[type_image_path]==0.7.5->ydata-profiling->pandas-
```

```
profiling) (22.1.0)
Requirement already satisfied: MarkupSafe>=2.0 in
/opt/conda/lib/python3.10/site-packages (from jinja2<3.2,>=2.11.1->ydata-
profiling->pandas-profiling) (2.1.1)
Requirement already satisfied: python-dateutil>=2.7 in
/opt/conda/lib/python3.10/site-packages (from matplotlib<4,>=3.2->ydata-
profiling->pandas-profiling) (2.8.2)
Requirement already satisfied: cycler>=0.10 in /opt/conda/lib/python3.10/site-
packages (from matplotlib<4,>=3.2->ydata-profiling->pandas-profiling) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in
/opt/conda/lib/python3.10/site-packages (from matplotlib<4,>=3.2->ydata-
profiling->pandas-profiling) (4.38.0)
Requirement already satisfied: contourpy>=1.0.1 in
/opt/conda/lib/python3.10/site-packages (from matplotlib<4,>=3.2->ydata-
profiling->pandas-profiling) (1.0.5)
Requirement already satisfied: kiwisolver>=1.0.1 in
/opt/conda/lib/python3.10/site-packages (from matplotlib<4,>=3.2->ydata-
profiling->pandas-profiling) (1.4.4)
Requirement already satisfied: packaging>=20.0 in
/opt/conda/lib/python3.10/site-packages (from matplotlib<4,>=3.2->ydata-
profiling->pandas-profiling) (21.3)
Requirement already satisfied: pyparsing>=2.2.1 in
/opt/conda/lib/python3.10/site-packages (from matplotlib<4,>=3.2->ydata-
profiling->pandas-profiling) (3.0.9)
Requirement already satisfied: pytz>=2020.1 in /opt/conda/lib/python3.10/site-
packages (from pandas!=1.4.0,<2.1,>1.1->ydata-profiling->pandas-profiling)
(2022.5)
Requirement already satisfied: joblib>=0.14.1 in /opt/conda/lib/python3.10/site-
packages (from phik<0.13,>=0.11.1->ydata-profiling->pandas-profiling) (1.2.0)
Requirement already satisfied: typing-extensions>=4.2.0 in
/opt/conda/lib/python3.10/site-packages (from pydantic<2,>=1.8.1->ydata-
profiling->pandas-profiling) (4.4.0)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in
/opt/conda/lib/python3.10/site-packages (from requests<3,>=2.24.0->ydata-
profiling->pandas-profiling) (1.26.11)
Requirement already satisfied: certifi>=2017.4.17 in
/opt/conda/lib/python3.10/site-packages (from requests<3,>=2.24.0->ydata-
profiling->pandas-profiling) (2022.9.24)
Requirement already satisfied: charset-normalizer<3,>=2 in
/opt/conda/lib/python3.10/site-packages (from requests<3,>=2.24.0->ydata-
profiling->pandas-profiling) (2.1.1)
Requirement already satisfied: idna<4,>=2.5 in /opt/conda/lib/python3.10/site-
packages (from requests<3,>=2.24.0->ydata-profiling->pandas-profiling) (3.4)
Requirement already satisfied: patsy>=0.5.2 in /opt/conda/lib/python3.10/site-
packages (from statsmodels<1,>=0.13.2->ydata-profiling->pandas-profiling)
(0.5.3)
Requirement already satisfied: six in /opt/conda/lib/python3.10/site-packages
(from patsy>=0.5.2->statsmodels<1,>=0.13.2->ydata-profiling->pandas-profiling)
```

```
(1.16.0)
```

[4]:
```
df = pd.read_csv("https://raw.githubusercontent.com/justmarkham/
  ↪scikit-learn-videos/master/data/Advertising.csv")
df
```

[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
..          …    …     …          …     …
195         196   38.2    3.7       13.8    7.6
196         197   94.2    4.9        8.1    9.7
197         198  177.0    9.3        6.4   12.8
198         199  283.6   42.0       66.2   25.5
199         200  232.1    8.6        8.7   13.4

[200 rows x 5 columns]
```

[5]:
```
df
```

[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
..          …    …     …          …     …
195         196   38.2    3.7       13.8    7.6
196         197   94.2    4.9        8.1    9.7
197         198  177.0    9.3        6.4   12.8
198         199  283.6   42.0       66.2   25.5
199         200  232.1    8.6        8.7   13.4

[200 rows x 5 columns]
```

[6]:
```
df.head()
```

[6]:
```
   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
```

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

[7]:

	Unnamed: 0	TV	Radio	Newspaper	Sales
195	196	38.2	3.7	13.8	7.6
196	197	94.2	4.9	8.1	9.7
197	198	177.0	9.3	6.4	12.8
198	199	283.6	42.0	66.2	25.5
199	200	232.1	8.6	8.7	13.4

[8]:
```
df.describe()
```

[8]:

	Unnamed: 0	TV	Radio	Newspaper	Sales
count	200.000000	200.000000	200.000000	200.000000	200.000000
mean	100.500000	147.042500	23.264000	30.554000	14.022500
std	57.879185	85.854236	14.846809	21.778621	5.217457
min	1.000000	0.700000	0.000000	0.300000	1.600000
25%	50.750000	74.375000	9.975000	12.750000	10.375000
50%	100.500000	149.750000	22.900000	25.750000	12.900000
75%	150.250000	218.825000	36.525000	45.100000	17.400000
max	200.000000	296.400000	49.600000	114.000000	27.000000

[9]:
```
x = df["TV"]
x
```

[9]:
```
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
Name: TV, Length: 200, dtype: float64
```

[10]:
```
y = df.Sales
y
```

[10]:
```
0      22.1
1      10.4
2       9.3
3      18.5
4      12.9
       …
195     7.6
```

```
196     9.7
197    12.8
198    25.5
199    13.4
Name: Sales, Length: 200, dtype: float64
```

[11]:
```
print("Shape of x:", x.shape)
print("Shape of y:", y.shape)
```

```
Shape of x: (200,)
Shape of y: (200,)
```

[12]:
```
import numpy as np
from sklearn.linear_model import LinearRegression

# Assuming x is a pandas Series
x_array = x.to_numpy().reshape(-1, 1)

lr = LinearRegression()
lr.fit(x_array, y)
```

[12]:
```
LinearRegression()
```

[13]:
```
lr.intercept_
```

[13]:
```
7.032593549127695
```

[14]:
```
lr.coef_
```

[14]:
```
array([0.04753664])
```

[20]:
```
import numpy as np
from sklearn.linear_model import LinearRegression

# Assuming x is a pandas Series
x_array = np.array(x).reshape(-1, 1)

lr = LinearRegression()
lr.fit(x_array, y)

# Now, make predictions
prediction = lr.predict([[45]])  # Use double brackets to provide a 2D array as␣
 ↪input
print("Predicted value:", prediction)
```

```
Predicted value: [9.17174237]
```