ml-1

October 8, 2023

[3]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn import linear_model
```

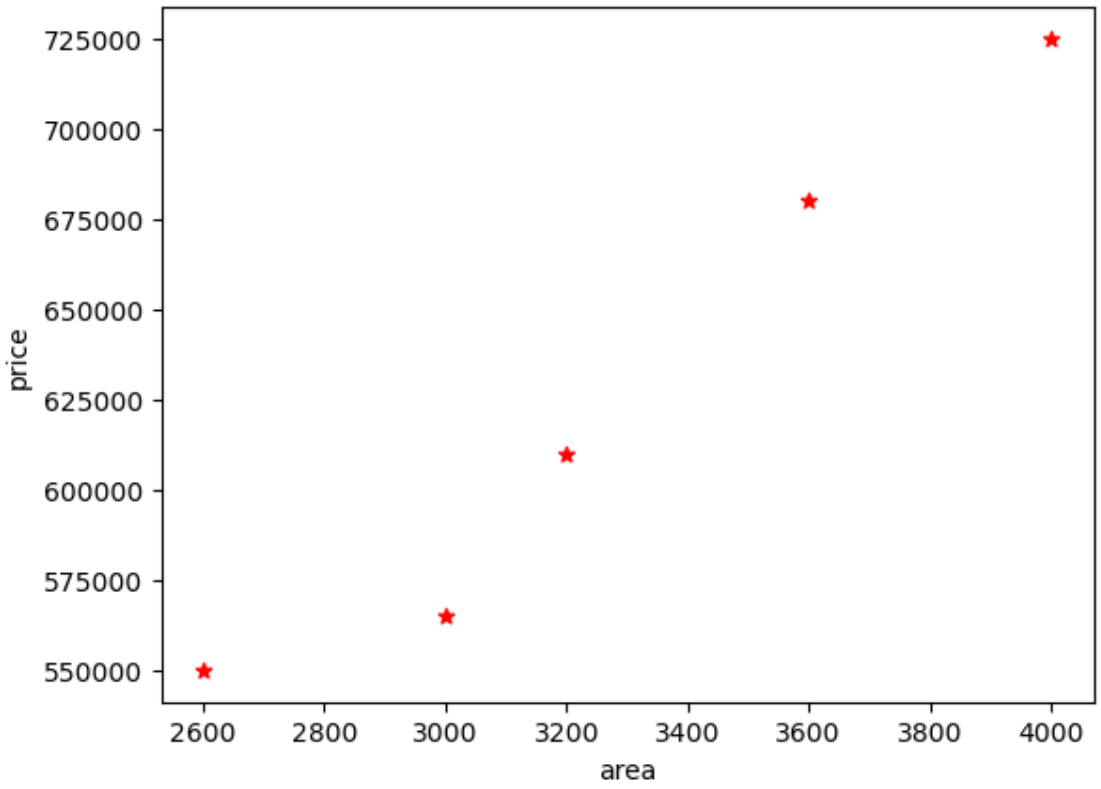
[5]:
```
df = pd.read_csv(r"C:\Users\nandini sharma\Downloads\homeprices.csv")
df
```

[5]:

	area	price
0	2600	550000
1	3000	565000
2	3200	610000
3	3600	680000
4	4000	725000

[9]:
```
plt.xlabel('area')
plt.ylabel('price')
plt.scatter(df.area,df.price,color='red', marker='*')
```

[9]:
```
<matplotlib.collections.PathCollection at 0x198539951e0>
```



[10]:
```
new_df= df.drop('price',axis='columns')
new_df
```

[10]:
```
   area
0  2600
1  3000
2  3200
3  3600
4  4000
```

[11]:
```
price =df.price
price
```

[11]:
```
0    550000
1    565000
2    610000
3    680000
4    725000
Name: price, dtype: int64
```

[14]:
```
reg = linear_model.LinearRegression()
reg.fit(new_df,price)
```

[14]:
```
LinearRegression()
```

[16]:
```
reg.predict([[3300]])
```

```
C:\Users\nandini sharma\anaconda3\lib\site-packages\sklearn\base.py:420:
UserWarning: X does not have valid feature names, but LinearRegression was
fitted with feature names
  warnings.warn(
```

[16]:
```
array([628715.75342466])
```

[17]:
```
reg.coef_
```

[17]:
```
array([135.78767123])
```

[18]:
```
reg.intercept_
```

[18]:
```
180616.43835616432
```

[19]:
```
3300*135.78767123+180616.43835616432
```

[19]:
```
628715.7534151643
```

[20]:
```
reg.predict([[5000]])
```

```
C:\Users\nandini sharma\anaconda3\lib\site-packages\sklearn\base.py:420:
UserWarning: X does not have valid feature names, but LinearRegression was
fitted with feature names
  warnings.warn(
```

[20]:
```
array([859554.79452055])
```

[23]:
```
area_df =pd.read_csv(r"C:\Users\nandini sharma\Downloads\areas.csv")
area_df.head(3)
```

[23]:

	area
0	1000
1	1500
2	2300

[24]:
```
p=reg.predict(area_df)
p
```

[24]:
```
array([ 316404.10958904,  384297.94520548,  492928.08219178,
        661304.79452055,  740061.64383562,  799808.21917808,
        926090.75342466,  650441.78082192,  825607.87671233,
        492928.08219178, 1402705.47945205, 1348390.4109589 ,
       1144708.90410959])
```

[25]:
```
area_df['prices']=p
area_df
```

[25]:

	area	prices
0	1000	3.164041e+05
1	1500	3.842979e+05
2	2300	4.929281e+05
3	3540	6.613048e+05
4	4120	7.400616e+05
5	4560	7.998082e+05
6	5490	9.260908e+05
7	3460	6.504418e+05
8	4750	8.256079e+05
9	2300	4.929281e+05
10	9000	1.402705e+06
11	8600	1.348390e+06
12	7100	1.144709e+06

[ ]: