area price prediction

June 4, 2022

[2]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn import linear_model
```

[5]:
```
df=pd.read_csv('area.csv')
df
```

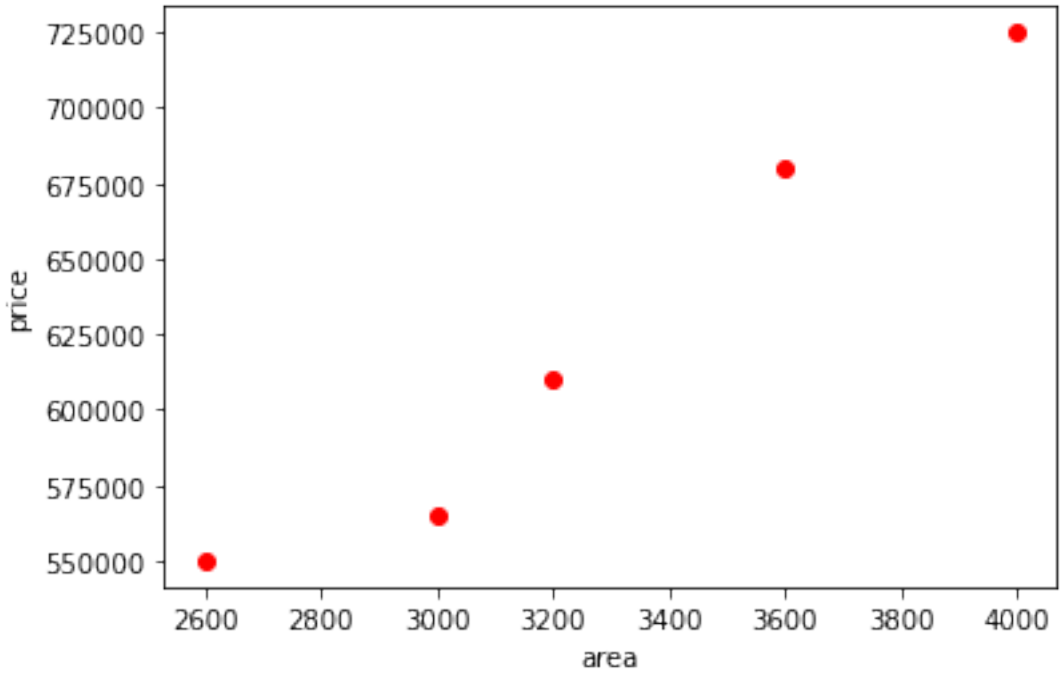
[5]:

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

[3]:
```
%matplotlib inline
```

[6]:
```
plt.xlabel('area')
plt.ylabel('price')
plt.scatter(df.area,df.price,color="red")
```

[6]:
```
<matplotlib.collections.PathCollection at 0x2755f387d30>
```



[7]:
```
new_df = df.drop('price',axis='columns')
new_df
```

[7]:
```
   area
0  2600
1  3000
2  3200
3  3600
4  4000
```

[8]:
```
price = df.price
price
```

[8]:
```
0    550000
1    565000
2    610000
3    680000
4    725000
Name: price, dtype: int64
```

[9]:
```
reg=linear_model.LinearRegression()
reg.fit(new_df,price)
```

[9]:
```
LinearRegression()
```

[10]:
```
reg.predict([[3300]])
```

[10]:
```
array([628715.75342466])
```

[11]:
```
d=pd.read_csv('areaprediction.csv')
```

[12]:
```
d.head()
```

[12]:
```
   area
0  1000
1  1500
2  2300
3  3540
4  4120
```

[13]:
```
p=reg.predict(d)
p
```

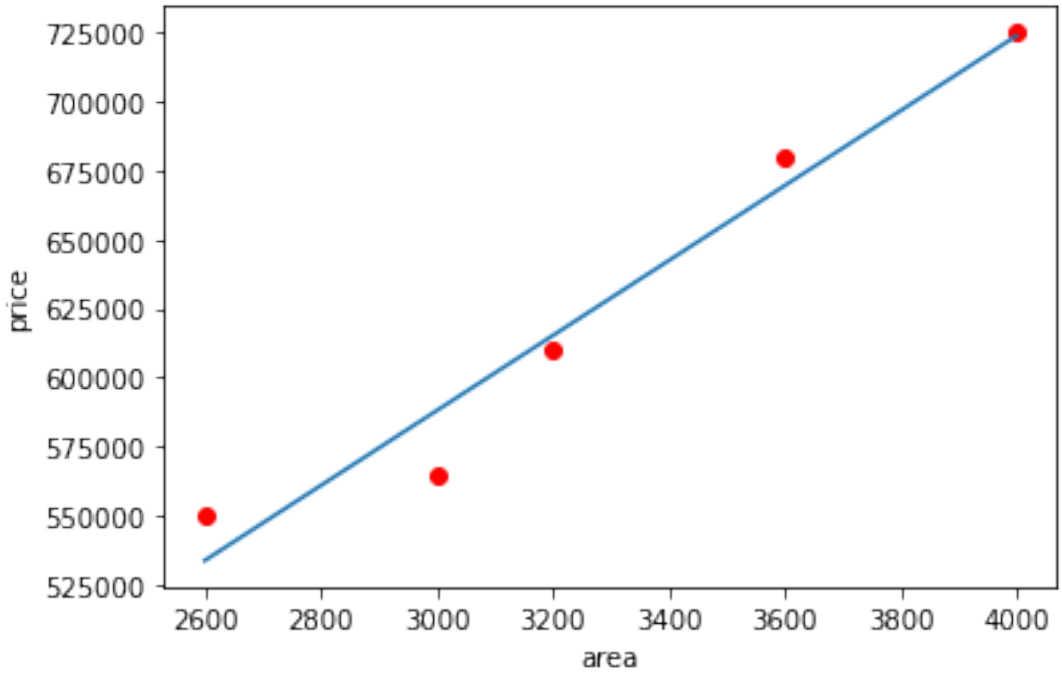
[13]:
```
array([ 316404.10958904,  384297.94520548,  492928.08219178,
        661304.79452055,  740061.64383562,  799808.21917808,
        926090.75342466,  650441.78082192,  825607.87671233,
        492928.08219178, 1402705.47945205, 1348390.4109589 ,
       1144708.90410959])
```

[14]:
```
d['price predicted']=p
```

[15]:
```
d.to_csv('prediction.csv')
```

[17]:
```
plt.xlabel('area')
plt.ylabel('price')
plt.scatter(df.area,df.price,color="red")
plt.plot(df.area,reg.predict(df[['area']]))
```

[17]:
```
[<matplotlib.lines.Line2D at 0x2755fb32640>]
```


725000
700000
675000
650000
625000
600000
575000
550000
525000
price
2600
2800
3000
3200
3400
3600
3800
4000
area

[ ]:

[ ]: