

```
In [1]: ▶ import pandas as pd
import numpy as np
from sklearn import linear_model
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings("ignore", category=UserWarning)
```

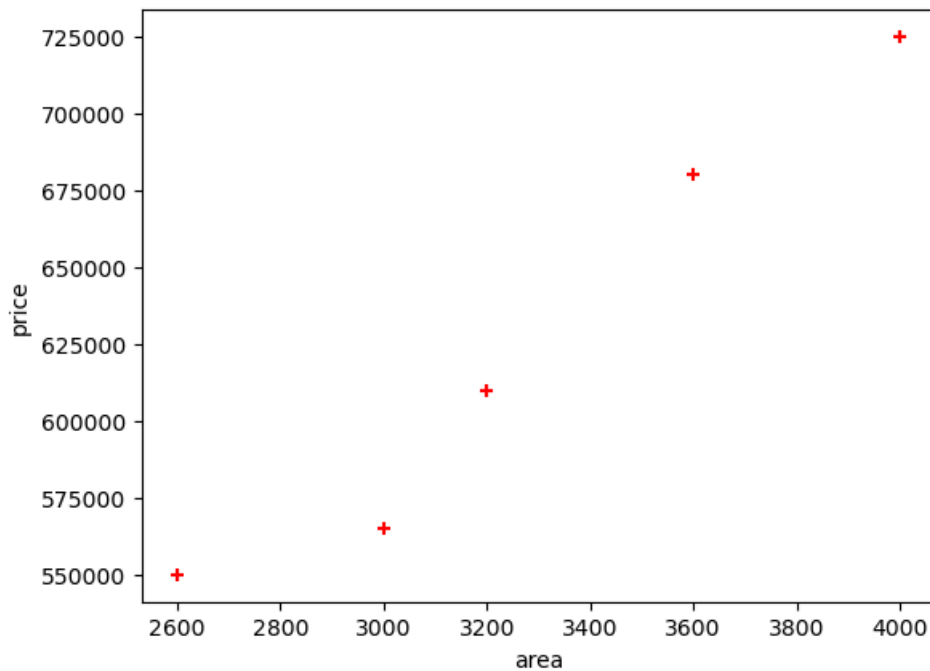
```
In [2]: ▶ df = pd.read_csv('C:\\Users\\Asus\\Downloads\\homeprices.csv')
df
```

Out[2]:

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

```
In [3]: ▶ %matplotlib inline
plt.xlabel('area')
plt.ylabel('price')
plt.scatter(df.area,df.price,color='red',marker='+')
```

Out[3]: <matplotlib.collections.PathCollection at 0x19ac5a6aa40>



```
In [4]: ▶ new_df = df.drop('price',axis='columns')
new_df
```

Out[4]:

	area
0	2600
1	3000
2	3200
3	3600
4	4000

```
In [5]: price = df.price
price
```

```
Out[5]: 0    550000
        1    565000
        2    610000
        3    680000
        4    725000
        Name: price, dtype: int64
```

```
In [6]: # Create linear regression object
reg = linear_model.LinearRegression()
reg.fit(new_df,price) #training the model using available dataset.
```

```
Out[6]: LinearRegression
LinearRegression()
```

## (1) Predict price of a home with area = 3300 sqr ft

```
In [7]: reg.predict([[3300]])
```

```
Out[7]: array([628715.75342466])
```

## $Y = m * X + b$ (m is coefficient and b is intercept)

```
In [8]: reg.coef_
```

```
Out[8]: array([135.78767123])
```

```
In [9]: reg.intercept_
```

```
Out[9]: 180616.43835616432
```

```
In [10]: 3300*135.78767123 + 180616.43835616432
```

```
Out[10]: 628715.7534151643
```

## (2) Predict price of a home with area = 5000 sqr ft

```
In [11]: reg.predict([[5000]])
```

```
Out[11]: array([859554.79452055])
```

```
In [ ]:
```