

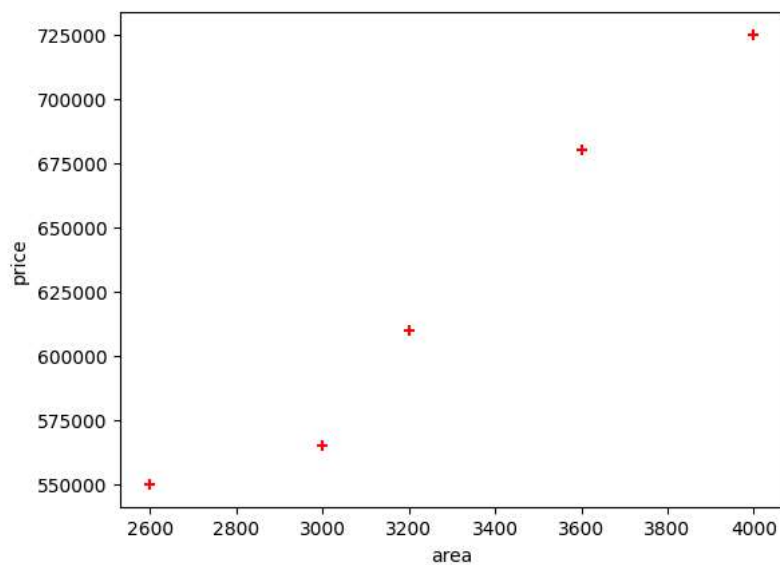
```
import pandas as pd
import numpy as np
from sklearn import linear_model
from sklearn.linear_model import LinearRegression
import matplotlib.pyplot as plt
```

```
df=pd.read_csv('/content/houseprices.csv')
df
```

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

```
plt.xlabel('area')
plt.ylabel('price')
plt.scatter(df.area,df.price,color='red',marker='+')
```

 <matplotlib.collections.PathCollection at 0x7b1ccc9cfee0>



```
new_df = df.drop('price',axis='columns')
new_df
```

	area
0	2600
1	3000
2	3200
3	3600
4	4000

```
price = df.price
price
```

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

```
#Create LinearRegression object
reg = linear_model.LinearRegression()
reg.fit(new_df,price)
```

```
▼ LinearRegression
LinearRegression()
```

```
model=LinearRegression()
```

```
LinearRegression(copy_X = True,fit_intercept=True)
```

```
▼ LinearRegression
LinearRegression()
```

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

```
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: UserWarning: X does not have valid feature names, but LinearRegression was
warnings.warn(
array([628715.75342466])
```

```
reg.coef_
```

```
array([135.78767123])
```

```
reg.intercept_
```

```
180616.43835616432
```

```
##y = mx+b; (m is co-efficient and b is intercept)
135.78*3300+180616.438
```

```
628690.438
```

```
reg.predict([[5500]])
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: UserWarning: X does not have valid feature names, but LinearRegression was
warnings.warn(
array([927448.63013699])
```

```
area_df=pd.read_csv("/content/Areas.csv")
area_df
```

```
area
0 1000
1 2000
2 3500
3 4000
```

```
p=reg.predict(area_df)
p
```

```
array([316404.10958904, 452191.78082192, 655873.28767123, 723767.12328767])
```

```
area_df['price']=p
area_df
area_df.to_csv("Prediction.csv")
```

