

```
In [23]: # Importing all the necessary libraries
import pandas as pd
import numpy as np
from sklearn import linear_model
import matplotlib.pyplot as plt
```

```
In [33]: # Reading the file with Pandas
df = pd.read_csv('areas.csv')
```

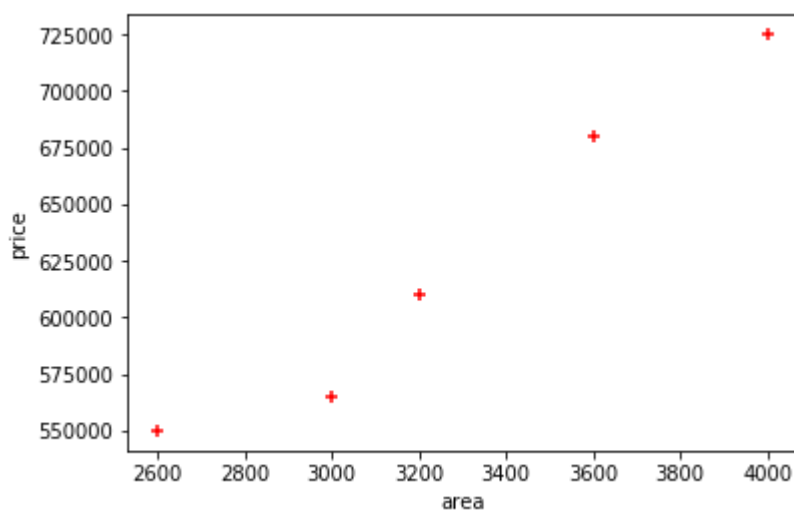
```
In [34]: # Look at the file which is read
df
```

Out[34]:

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

```
In [35]: # matplotlib lib is used to plot the points and xlabel and ylabel is used to define
%matplotlib inline
plt.xlabel('area')
plt.ylabel('price')
plt.scatter(df.area,df.price,color='red',marker='+')
```

Out[35]: <matplotlib.collections.PathCollection at 0x28e9c5d5470>



```
In [36]: # Create linear regression object.  
reg = linear_model.LinearRegression()  
# Fit the data.in other words , training the model based on the available model.  
# 1st argument is has to be a 2D array.  
reg.fit(df.drop('price',axis='columns'),df.price)
```

```
Out[36]: LinearRegression(copy_X=True, fit_intercept=True, n_jobs=1, normalize=False)
```

```
In [37]: # Find price of home with 3000 sqr ft area  
reg.predict([[3300]])
```

```
Out[37]: array([628715.75342466])
```

```
In [38]: # coefficient of the linear regression line  
reg.coef_
```

```
Out[38]: array([135.78767123])
```

```
In [39]: # intercept of the linear regression line  
reg.intercept_
```

```
Out[39]: 180616.43835616432
```