

```
In [1]: import pandas as pd
import numpy as np
from sklearn import linear_model
```

```
In [2]: df=pd.read_csv("homeprices2.csv")
df
```

Out[2]:

	area	bedrooms	age	price
0	2600	3.0	20	550000
1	3000	4.0	15	565000
2	3200	NaN	18	610000
3	3600	3.0	30	595000
4	4000	5.0	8	760000
5	4100	6.0	8	810000

```
In [14]: import math
median_bedrooms=math.floor(df.bedrooms.median())
median_bedrooms
```

Out[14]: 4

```
In [15]: df.bedrooms=df.bedrooms.fillna(median_bedrooms)
df
```

Out[15]:

	area	bedrooms	age	price
0	2600	3.0	20	550000
1	3000	4.0	15	565000
2	3200	4.0	18	610000
3	3600	3.0	30	595000
4	4000	5.0	8	760000
5	4100	6.0	8	810000

```
In [16]: reg=linear_model.LinearRegression()
reg.fit(df[['area', 'bedrooms', 'age']],df.price)
```

Out[16]: LinearRegression()

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

Out[17]: array([112.06244194, 23388.88007794, -3231.71790863])

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

Out[18]: 221323.00186540408

```
In [8]: 112.06244194*3000+23388.88007794*3+-3231.71790863*15+221323.00186540408
```

Out[8]: 579201.199289774

```
In [19]: reg.predict([[3000,3,15]])
```

Out[19]: array([579201.19929613])

```
In [20]: reg.predict([[2500,4,5]])
```

Out[20]: array([578876.03748933])