

```
In [4]: import pandas as pd
import numpy as np
from sklearn import linear_model
```

```
In [5]: df = pd.read_csv('C:/Users/aziz/Desktop/Downloads/Machine Learning Python/2_Linear Regression Multiple Variables.csv')
```

```
In [6]: df
```

Out[6]:

	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 [7]: df.bedrooms.median()
```

Out[7]: 4.0

```
In [8]: df.bedrooms = df.bedrooms.fillna(df.bedrooms.median())
df
```

Out[8]:

	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 [9]: reg = linear_model.LinearRegression()
reg.fit(df.drop('price',axis='columns'),df.price)
```

Out[9]: LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=False)

```
In [10]: reg.coef_
```

Out[10]: array([112.06244194, 23388.88007794, -3231.71790863])

```
In [11]: reg.intercept_
```

```
Out[11]: 221323.00186540408
```

```
In [13]: reg.predict([[3000, 3, 40]])
```

```
Out[13]: array([498408.25158031])
```

```
In [14]: 112.06244194*3000 + 23388.88007794*3 + -3231.71790863*40 + 221323.00186540384
```

```
Out[14]: 498408.25157402386
```