```
import pandas as pd
import numpy as np
from sklearn import linear_model
df = pd.read_csv('homeprices.csv')
print(df)
```

```
   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
```

```
df.bedrooms.median()
```

```
4.0
```

```
df.bedrooms = df.bedrooms.fillna(df.bedrooms.median())
df
```

```
   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
```

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

```
LinearRegression()
```

```
reg.coef_
```

```
array([  112.06244194, 23388.88007794, -3231.71790863])
```

```
reg.intercept_
```

```
221323.00186540408
```

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

```
array([498408.25158031])
```

```
112.06244194*3000 + 23388.88007794*3 + -3231.71790863*40 +
221323.00186540384
```

```
498408.25157402386
```

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

```
array([578876.03748933])
```