In [1]:
```
import pandas as pd
import numpy as np
from sklearn import linear_model
import warnings
warnings.filterwarnings("ignore", category=UserWarning)
```

In [2]:
```
df = pd.read_csv('C:\\Users\\Asus\\Downloads\\multiple.homeprices.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

Data Preprocessing: Fill NA values with median value of a column

In [3]:
```
df.bedrooms.median()
```

Out[3]:
```
4.0
```

In [4]:
```
df.bedrooms = df.bedrooms.fillna(df.bedrooms.median())
df
```

Out[4]:

	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

y = mx1 + mx2+ mx3+ b

In [5]:
```
reg = linear_model.LinearRegression()
reg.fit(df.drop('price',axis='columns'),df.price) #training the model using available dataset.
```

Out[5]:
```
LinearRegression()
```
In a Jupyter environment, please rerun this cell to show the HTML representation or trust the notebook.
On GitHub, the HTML representation is unable to render, please try loading this page with nbviewer.org.

In [6]:
```
reg.coef_
```

Out[6]:
```
array([  112.06244194, 23388.88007794, -3231.71790863])
```

In [7]:
```
reg.intercept_
```

Out[7]:
```
221323.00186540408
```

Find price of home with 3000 sqr ft area, 3 bedrooms, 40 year old

In [8]:
```
reg.predict([[3000, 3, 40]])
```

Out[8]:
```
array([498408.25158031])
```

In [9]:
```
112.06244194*3000 + 23388.88007794*3 + -3231.71790863*40 + 221323.00186540384
```

Out[9]:
```
498408.25157402386
```

Find price of home with 2500 sqr ft area, 4 bedrooms, 5 year old

In [10]:
```
reg.predict([[2500, 4, 5]])
```

Out[10]:
```
array([578876.03748933])
```