In [2]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
from sklearn import linear_model
```

In [3]:

```
df = pd.read_csv("F:\clg\Home.csv")
```

In [4]:

```
df
```

Out[4]:

	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

In [5]:

```
print(df['Bedrooms'].median())
```

```
3.5
```

In [6]:

```
median_bedrooms = int(df.Bedrooms.median())
median_bedrooms
```

Out[6]:

```
3
```

In [7]:

```
df.Bedrooms.fillna(median_bedrooms)
```

Out[7]:

```
0    3.0
1    4.0
2    3.0
3    3.0
4    5.0
Name: Bedrooms, dtype: float64
```

In [8]:

```
df.Bedrooms = df.Bedrooms.fillna(median_bedrooms)
df
```

Out[8]:

	Area	Bedrooms	Age	Price
0	2600	3.0	20	550000
1	3000	4.0	15	565000
2	3200	3.0	18	610000
3	3600	3.0	30	595000
4	4000	5.0	8	760000

In [9]:

```
reg = linear_model.LinearRegression()
```

In [10]:

```
reg.fit(df[['Area', 'Bedrooms', 'Age']], df.Price)
```

Out[10]:

```
LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=F
alse)
```

In [11]:

```
reg.coef_                #1.Calculate the coefficient of each factor.
```

Out[11]:

```
array([   137.25, -26025.  ,  -6825.  ])
```

In [12]:

```
reg.intercept_               #2.Calculate the intercept value.
```

Out[12]:

```
383724.9999999998
```

In [13]:

```
reg.predict([[4000,4,30]])  #3. Predict the price of a property which is in 4000 unit i
n area, 4 BHK and 30 years old.
```

Out[13]:

```
array([623875.])
```

In [14]:

```
reg.predict([[4000,4,3]])   #4. Predict the price of a property which is in 4000 unit i
n area, 4 BHK and 3 years old.
```

Out[14]:

```
array([808150.])
```

Based on point 3 and 4 there predict price is change because price is depend on year so Q.3 there is old property and Q.4 there is new property.always new property rate is higher then old property.

In [ ]: