In [1]:
```
import pandas as pd
from sklearn import linear_model
import  matplotlib.pyplot as plt
```

In [2]:
```
url="https://raw.githubusercontent.com/apratim777/apratim777/master/homeprices1.csv"
df= pd.read_csv(url)
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
```

In [3]:
```
#filling the missing values by median
df.bedrooms.fillna(df.bedrooms.median(),inplace=True)
print(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
```

In [4]:
```
x=df.drop('price',axis='columns')
y=df.price
print(x.shape)
print(y.shape)
```

```
(6, 3)
(6,)
```

In [5]:
```
reg=linear_model.LinearRegression()
reg.fit(x,y)
acc=reg.score (x,y)
print(acc)
```

```
0.9550196399325819
```

In [6]:
```
pre=reg.predict([[3000,3,40]])
print(pre.round(2)[0])
```

```
498408.25

d:\py\lib\site-packages\sklearn\base.py:450: UserWarning: X does not have valid feature names, but LinearRegression was fitted with feature names
  warnings.warn(
```