

In [1]:

```
import pandas as pd
import numpy as np
```

In [2]:

```
dataset = pd.read_csv('house_data.csv')
```

In [15]:

```
dataset.shape
```

Out[15]:

```
(21613, 11)
```

In [4]:

```
X = dataset.iloc[:,[3,4]].values
```

In [5]:

```
y = dataset.iloc[:, -1].values
```

In [6]:

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=0)
```

In [7]:

```
y
```

Out[7]:

```
array([221900., 538000., 180000., ..., 402101., 400000., 325000.])
```

In [8]:

```
from sklearn.linear_model import LinearRegression
regressor = LinearRegression()
regressor.fit(X_train, y_train)
```

Out[8]:

```
LinearRegression()
```

In [9]:

```
y_pred = regressor.predict(X_test)
```

In [10]:

```
y_pred
```

Out[10]:

```
array([795554.35059871, 660789.55506513, 446053.54464655, ...,  
       621652.28245232, 534707.75526181, 452804.43542934])
```

In [11]:

```
from sklearn import metrics  
print('Mean Absolute Error:', metrics.mean_absolute_error(y_test, y_pred))  
print('Mean Squared Error:', metrics.mean_squared_error(y_test, y_pred))  
print('Root Mean Squared Error:', np.sqrt(metrics.mean_squared_error(y_test, y_pred)))
```

```
Mean Absolute Error: 217846.7647697289  
Mean Squared Error: 109993546353.38124  
Root Mean Squared Error: 331652.74965448614
```

In []: