In [3]:

```
import numpy as np
import pandas as pd
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split

data = pd.read_csv('houseprice.csv')
data
```

Out[3]:

	date	price	bedrooms	bathrooms	sqft_living	sqft_lot	floors	waterfront
0	2014-05-02 00:00:00	3.130000e+05	3.0	1.50	1340	7912	1.5	0
1	2014-05-02 00:00:00	2.384000e+06	5.0	2.50	3650	9050	2.0	0
2	2014-05-02 00:00:00	3.420000e+05	3.0	2.00	1930	11947	1.0	0
3	2014-05-02 00:00:00	4.200000e+05	3.0	2.25	2000	8030	1.0	0
4	2014-05-02 00:00:00	5.500000e+05	4.0	2.50	1940	10500	1.0	0
...	...	...	...	...	...	...	...	...
4595	2014-07-09 00:00:00	3.081667e+05	3.0	1.75	1510	6360	1.0	0
4596	2014-07-09 00:00:00	5.343333e+05	3.0	2.50	1460	7573	2.0	0
4597	2014-07-09 00:00:00	4.169042e+05	3.0	2.50	3010	7014	2.0	0
4598	2014-07-10 00:00:00	2.034000e+05	4.0	2.00	2090	6630	1.0	0
4599	2014-07-10 00:00:00	2.206000e+05	3.0	2.50	1490	8102	2.0	0

4600 rows × 18 columns

In [4]:
```
features = data[['sqft_lot', 'bedrooms', 'bathrooms']]
features
```

Out[4]:

	sqft_lot	bedrooms	bathrooms
0	7912	3.0	1.50
1	9050	5.0	2.50
2	11947	3.0	2.00
3	8030	3.0	2.25
4	10500	4.0	2.50
...	...	...	...
4595	6360	3.0	1.75
4596	7573	3.0	2.50
4597	7014	3.0	2.50
4598	6630	4.0	2.00
4599	8102	3.0	2.50

4600 rows × 3 columns

In [5]:
```
target = data['price']
target
```

Out[5]:
```
0       3.130000e+05
1       2.384000e+06
2       3.420000e+05
3       4.200000e+05
4       5.500000e+05
            ...
4595    3.081667e+05
4596    5.343333e+05
4597    4.169042e+05
4598    2.034000e+05
4599    2.206000e+05
Name: price, Length: 4600, dtype: float64
```

In [6]:
```
X_train, X_test, y_train, y_test = train_test_split(features, target,
model = LinearRegression()
model.fit(X_train, y_train)
model.fit
```

Out[6]:
```
<bound method LinearRegression.fit of LinearRegression()>
```

In [7]:
```
y_pred = model.predict(X_test)
y_pred
```

Out[7]:
```
array([580234.64238013, 283724.54211088, 544466.46208747, ...,
       809743.60935936, 579310.78277027, 449688.13337662])
```

In [8]:
```
1 X_train
```

Out[8]:

	sqft_lot	bedrooms	bathrooms
794	15639	3.0	2.50
4525	6413	2.0	1.00
2406	31941	4.0	2.25
844	6455	5.0	2.50
2072	7679	5.0	2.75
...	...	...	...
501	6380	3.0	1.00
44	9400	2.0	1.00
1658	7330	3.0	1.75
2017	4080	2.0	1.50
1665	8880	3.0	1.50

3450 rows × 3 columns

In [ ]:
```
1
```