

```
In [1]: import pandas as pd
import numpy as np
```

```
In [2]: from matplotlib import pyplot as plt
from sklearn.linear_model import LinearRegression
```

```
In [3]: from sklearn.model_selection import train_test_split
```

```
In [4]: data=pd.read_csv("Final_USHousing_DataSet.csv")
data
```

Out[4]:

	Unnamed: 0	DATE	Num_Households	Const_complt	un_constr	Cnstr_not_Strtd	UNRATE	Price_fact	Year	Month
0	0	2001-07-01	305.0	76.0	186.0	43.0	4.6	113.491	2001	7
1	1	2001-08-01	308.0	77.0	187.0	44.0	4.9	114.167	2001	8
2	2	2001-09-01	310.0	79.0	191.0	40.0	5.0	114.812	2001	9
3	3	2001-10-01	308.0	78.0	189.0	41.0	5.3	115.310	2001	10
4	4	2001-11-01	308.0	77.0	191.0	40.0	5.5	115.857	2001	11
...
262	262	2023-05-01	426.0	66.0	268.0	92.0	3.7	302.566	2023	5
263	263	2023-06-01	429.0	70.0	266.0	93.0	3.6	304.593	2023	6
264	264	2023-07-01	429.0	73.0	260.0	96.0	3.5	306.767	2023	7
265	265	2023-08-01	430.0	75.0	258.0	97.0	3.8	309.155	2023	8
266	266	2023-09-01	433.0	75.0	257.0	101.0	3.8	311.175	2023	9

267 rows × 10 columns

```
In [5]: data_df_new=data.drop(columns=['Unnamed: 0', 'DATE', 'Const_complt', 'un_constr', 'Cnstr_not_Strtd', 'Year', 'Month'],axis=1)
data_df_new
```

```
Out[5]:
```

	Num_Households	UNRATE	Price_fact
0	305.0	4.6	113.491
1	308.0	4.9	114.167
2	310.0	5.0	114.812
3	308.0	5.3	115.310
4	308.0	5.5	115.857
...
262	426.0	3.7	302.566
263	429.0	3.6	304.593
264	429.0	3.5	306.767
265	430.0	3.8	309.155
266	433.0	3.8	311.175

267 rows × 3 columns

```
In [6]: X=data_df_new.drop(columns=['Price_fact'],axis=1)
Y=data_df_new['Price_fact']
```

```
In [7]: x_train,x_test,y_train,y_test = train_test_split(X,Y,test_size=0.2)
```

```
In [8]: lr_model=LinearRegression().fit(x_train,y_train)
```

```
In [9]: lr_model.fit(x_train,y_train)
```

```
Out[9]: 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 [10]: x_test.head()
```

```
Out[10]:
```

	Num_Households	UNRATE
214	337.0	3.7
80	470.0	5.1
166	211.0	5.6
112	195.0	9.8
126	149.0	8.3

```
In [11]: Y_pred=lr_model.predict(x_test)
```

```
In [12]: Y_pred
```

```
Out[12]: array([209.69854528, 199.53897303, 179.25985575, 126.23614741,  
142.363793 , 117.0390151 , 170.38096757, 183.40399223,  
183.04683795, 202.4027978 , 210.17976524, 210.95536195,  
211.22698835, 192.77180794, 200.73450156, 183.28763411,  
143.3339405 , 172.83368512, 72.45151278, 151.97936773,  
209.93126152, 184.64576613, 183.07804054, 149.97741225,  
148.3010361 , 164.2965361 , 218.86410264, 201.70427671,  
186.65580153, 198.98838501, 188.07596639, 166.91956471,  
181.36312659, 154.61047626, 169.72906419, 208.29379553,  
143.22528993, 184.60722832, 211.82456644, 196.78417117,  
146.0891147 , 132.12677279, 207.47891632, 207.38530853,  
190.66082955, 193.22182531, 192.93478379, 208.56542194,  
207.01273912, 211.00968723, 203.58291121, 209.44196164,  
196.93906711, 196.93906711])
```

```
In [ ]:
```