```python
# import library
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

```python
#get the data set
data=pd.read_csv(r"C:\Users\arati\DATAS SCIENCE NIT\SEPTEMBER\26th- mlr\MLR\Hous
```

```python
data
```

	id	date	price	bedrooms	bathrooms	sqft_living	sqf
0	7129300520	20141013T000000	221900.0	3	1.00	1180	5
1	6414100192	20141209T000000	538000.0	3	2.25	2570	7
2	5631500400	20150225T000000	180000.0	2	1.00	770	10
3	2487200875	20141209T000000	604000.0	4	3.00	1960	5
4	1954400510	20150218T000000	510000.0	3	2.00	1680	8
...	...	...	...	...	...	...	
21608	263000018	20140521T000000	360000.0	3	2.50	1530	1
21609	6600060120	20150223T000000	400000.0	4	2.50	2310	5
21610	1523300141	20140623T000000	402101.0	2	0.75	1020	1
21611	291310100	20150116T000000	400000.0	3	2.50	1600	2
21612	1523300157	20141015T000000	325000.0	2	0.75	1020	1

21613 rows × 21 columns

```python
# devide dependent and independent
# we take for prediction bedrooms,sqft,floors
x=data.iloc[:,[3,5,7,10]].values
y=data.iloc[:,2:3].values
```

```python
x
```

```
array([[3.00e+00, 1.18e+03, 1.00e+00, 3.00e+00],
       [3.00e+00, 2.57e+03, 2.00e+00, 3.00e+00],
       [2.00e+00, 7.70e+02, 1.00e+00, 3.00e+00],
       ...,
       [2.00e+00, 1.02e+03, 2.00e+00, 3.00e+00],
       [3.00e+00, 1.60e+03, 2.00e+00, 3.00e+00],
       [2.00e+00, 1.02e+03, 2.00e+00, 3.00e+00]])
```

```python
y
```

```
Out[7]:  array([[221900.],
                [538000.],
                [180000.],
                ...,
                [402101.],
                [400000.],
                [325000.]])
```

```python
In [10]:  # split the data training and testing
          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)
```

```python
In [13]:  # scalling
          from sklearn.preprocessing import StandardScaler
          sc =  StandardScaler()
          x_train = sc.fit_transform(x_train)
          x_test = sc.transform(x_test)
```

```python
In [14]:  # pass the x_train,y_train to regression model
          from sklearn.linear_model import LinearRegression
          regressor=LinearRegression()
          regressor.fit(x_train,y_train)
```

```
Out[14]:  ▼   LinearRegression  ⓘ ⍰

          LinearRegression()
```

```python
In [15]:  #  prediction on test value
          y_pred=regressor.predict(x_test)
          print(y_pred)
```

```
[[ 429317.02923213]
 [1369097.33064876]
 [ 386530.84117968]
 ...
 [ 325154.97944383]
 [ 183933.6924121 ]
 [ 390778.37876334]]
```

```python
In [16]:  #slope
          m=regressor.coef_
          print(m)
```

```
[[-55678.23017164 292127.93529352  12466.04871852  36158.71128803]]
```

```python
In [17]:  # intercept
          c=regressor.intercept_
          print(c)
```

```
[542799.8377675]
```

```python
In [18]:  # Save the trained model to disk
          import pickle
          filename = 'house price.pkl'
          with open(filename, 'wb') as file:
              pickle.dump(regressor, file)
          print("Model has been pickled and saved as house price.pkl")
```

```
Model has been pickled and saved as house price.pkl
```

```
In [19]:  import os
          print(os.getcwd())
```

C:\Users\arati

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```