VAC ASSIGNMENT-1 Atmika Parey (RA1911042020042) CSBS

AIM-To pridicting housing price using Linear Regression using scikit-learn

```
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
import seaborn as sns
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn import metrics
from sklearn.metrics import confusion_matrix, accuracy_score
```

```
usa_data = pd.read_csv('/content/USA_Housing.csv')
usa_data.head()
```

	Avg. Area Income	Avg. Area House Age	Avg. Area Number of Rooms	Avg. Area Number of Bedrooms	Area Population	Price	Address
0	79545.458574	5.682861	7.009188	4.09	23086.800503	1.059034e+06	208 Michael Ferry Apt. 674\nLaurabury, NE 3701...
1	79248.642455	6.002900	6.730821	3.09	40173.072174	1.505891e+06	188 Johnson Views Suite 079\nLake Kathleen, CA...
2	61287.067179	5.865890	8.512727	5.13	36882.159400	1.058988e+06	9127 Elizabeth Stravenue\nDanieltown, WI 06482...
3	63345.240046	7.188236	5.586729	3.26	34310.242831	1.260617e+06	USS Barnett\nFPO AP 44820
4	59982.197226	5.040555	7.839388	4.23	26354.109472	6.309435e+05	USNS Raymond\nFPO AE

```
usa_data.isnull().sum()
```

```
Avg. Area Income              0
Avg. Area House Age           0
```

```
Avg. Area Number of Rooms       0
Avg. Area Number of Bedrooms    0
Area Population                 0
Price                           0
Address                         0
dtype: int64
```

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

```
LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=False)
```

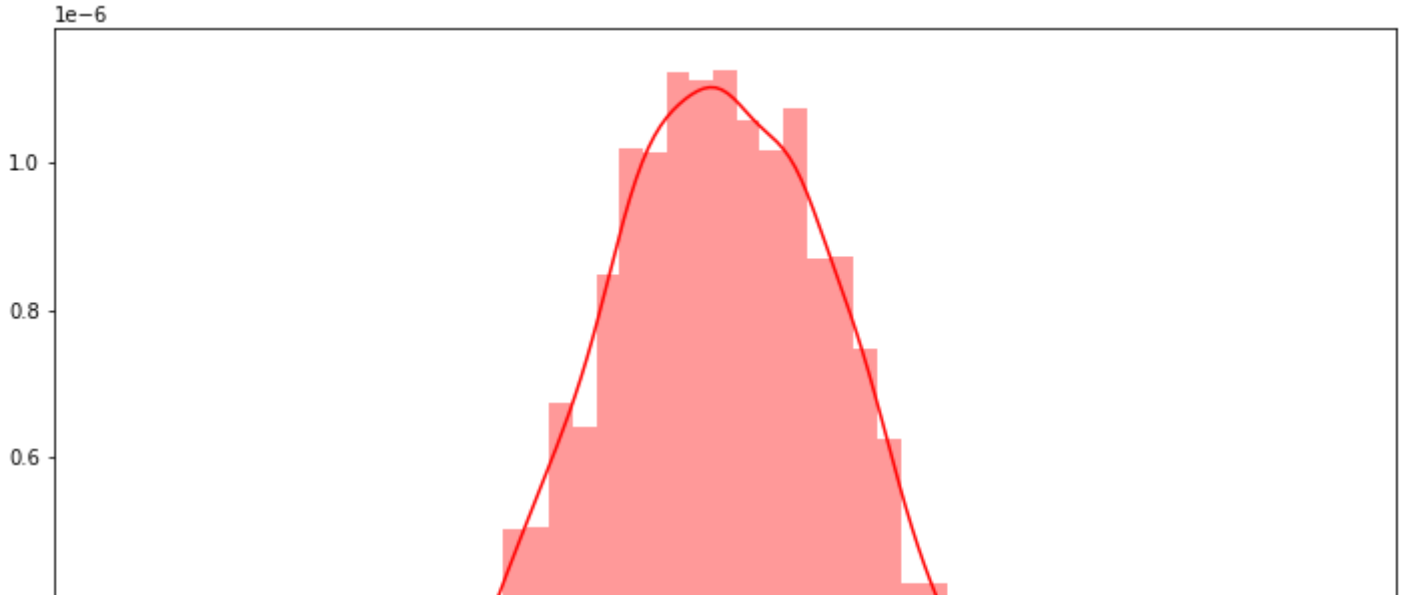
```
x = usa_data[['Avg. Area Income', 'Avg. Area House Age', 'Avg. Area Number of Rooms',
              'Avg. Area Number of Bedrooms', 'Area Population']]

y = usa_data['Price']


x_train, x_test, y_train, y_test = train_test_split(x, y, test_size = 0.2, random_state = 14)


fig = plt.figure(figsize = (12, 8))
sns.distplot(usa_data['Price'], color = 'red')
```

```
sr/local/lib/python3.7/dist-packages/seaborn/distributions.py:2619: FutureWarning: `distplot` is a depre
warnings.warn(msg, FutureWarning)
atplotlib.axes._subplots.AxesSubplot at 0x7f10678e4950>
```



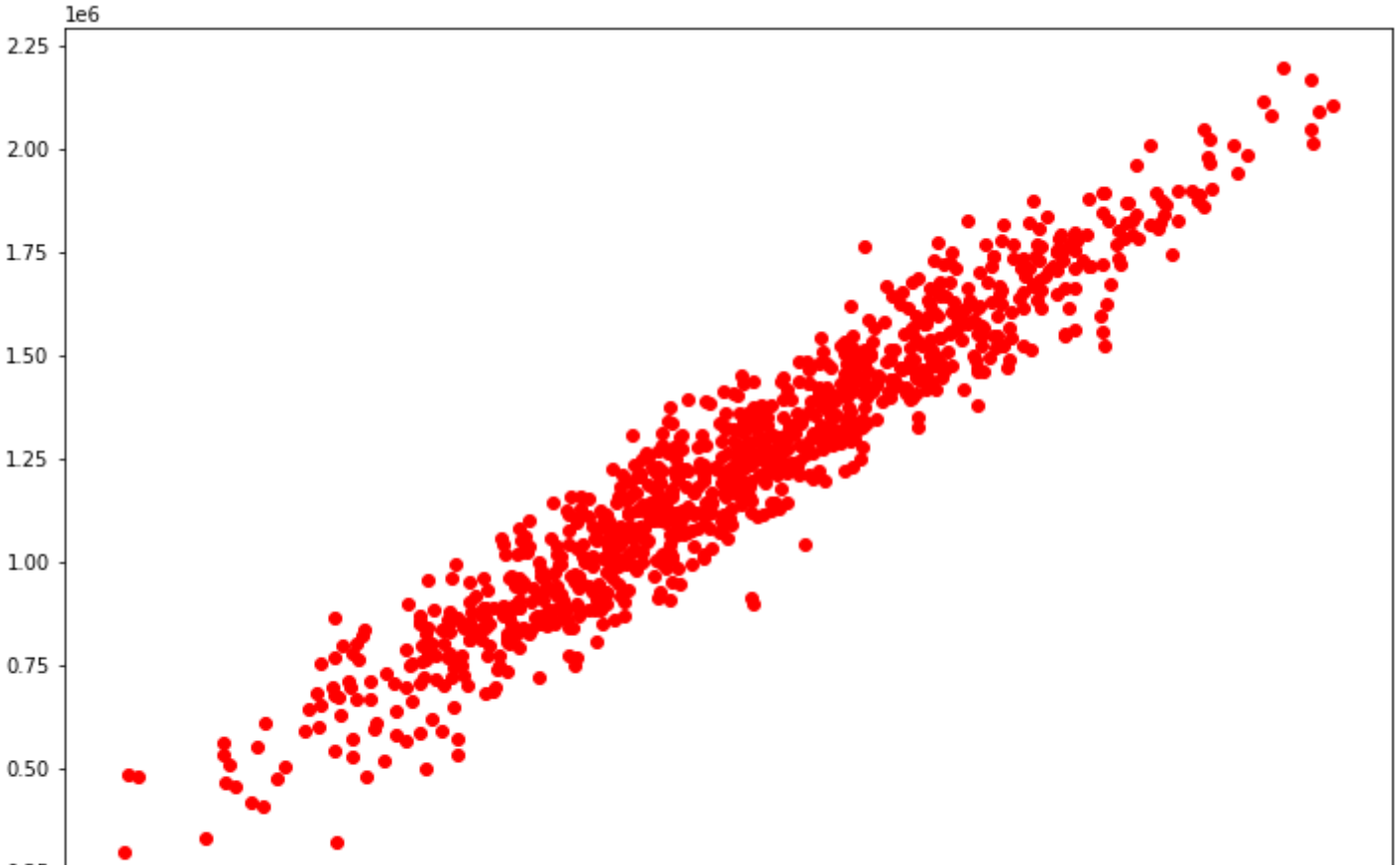

```
linear_model = LinearRegression()
```

```
linear_model.fit(x_train, y_train)
```

```
LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=False)
```

```
y_pred = linear_model.predict(x_test)
```

```
fig = plt.figure(figsize = (12, 8))
plt.scatter(y_test, y_pred, color = 'red')
plt.show()
```



RESULT:- Pridicting housing price using Linear Regression using scikit-learn was implemented.

✓ 0s completed at 12:00 PM