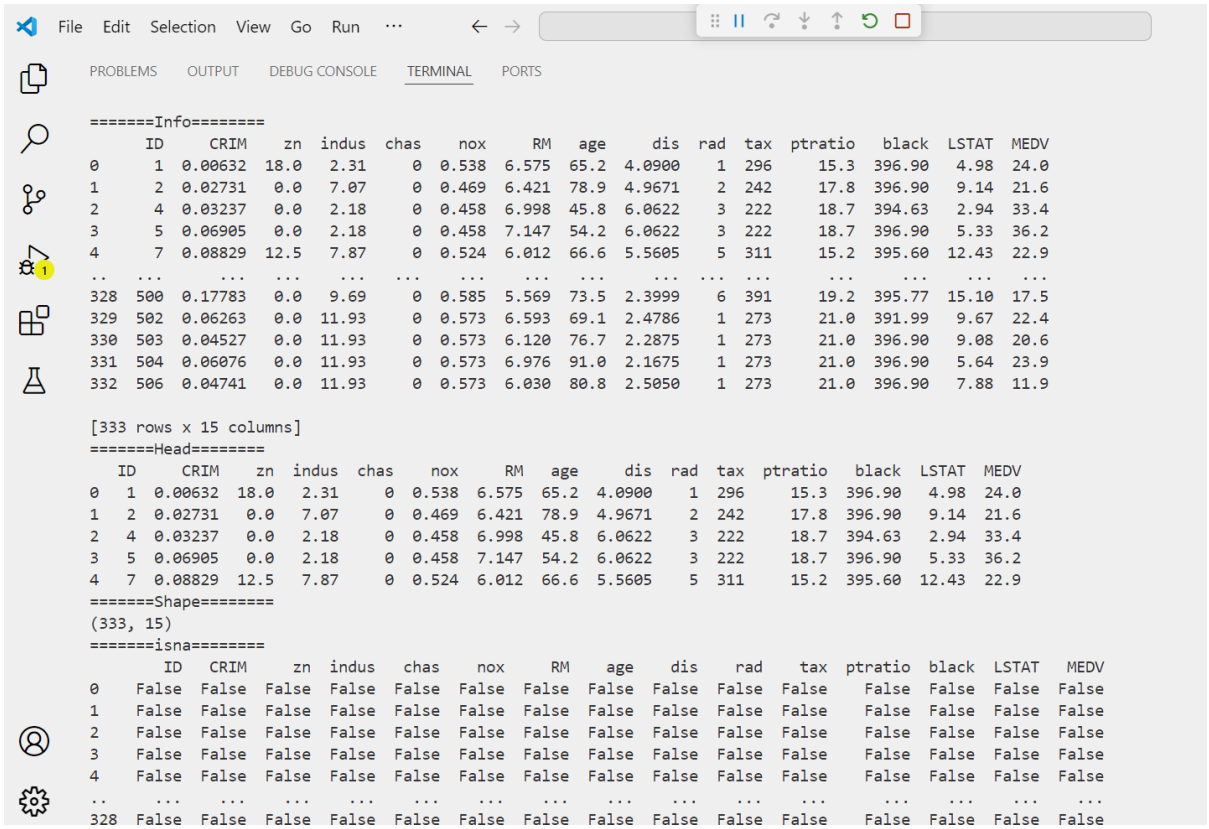
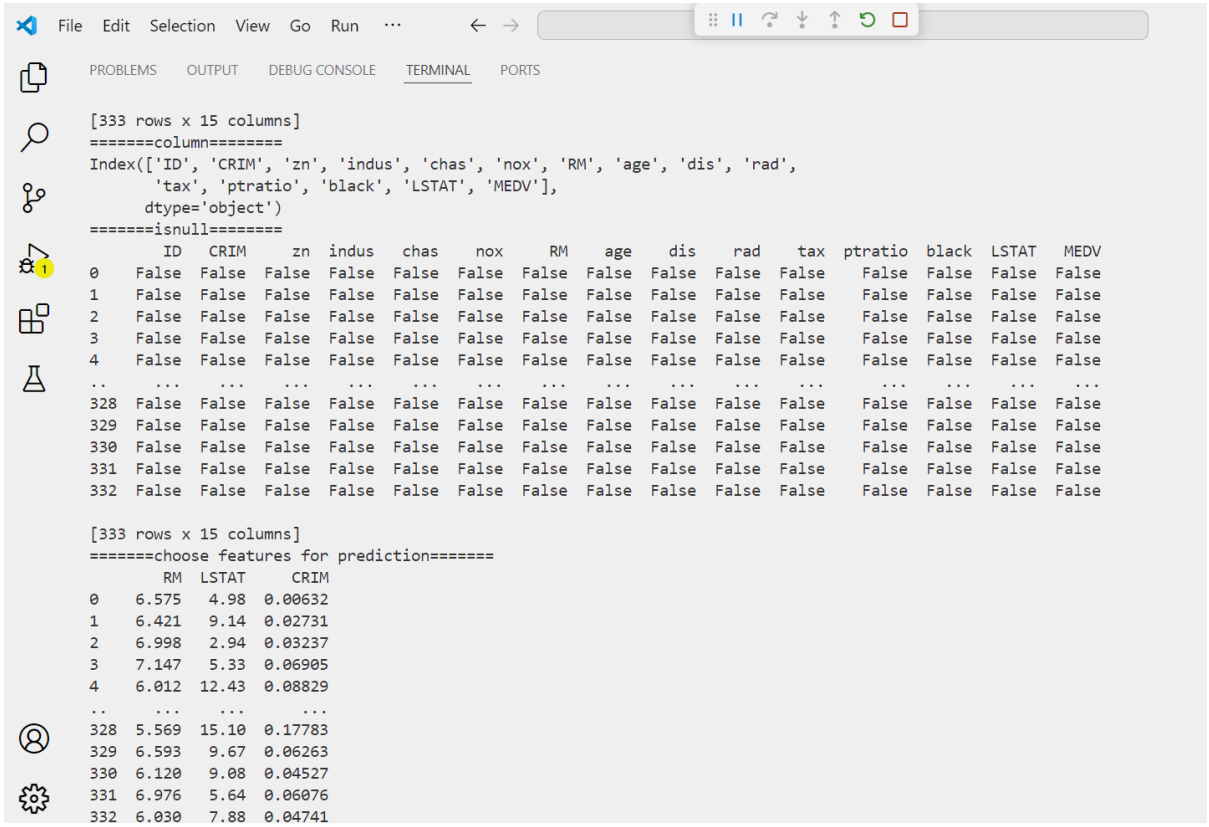
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error,r2_score


df=pd.read_csv(r'train.csv')
print("=======Info========")
print(df)
print("=======Head========")
print(df.head(5))
print("=======Shape========")
print(df.shape)
print("=======isna========")
print(df.isna())
print("=======column========")
print(df.columns)
print("=======isnull========")
print(df.isnull())

print("=======choose features for prediction=======")
X=df[['RM','LSTAT','CRIM']]
Y=df['MEDV']
print(X)
print(Y)

print("==========Split the data into training and testing sets=========")
X_train,X_test,Y_train,Y_test=train_test_split(X,Y,test_size=0.2,random_state=42)
```

```
print("=======x X_train========")
print(X_train.shape)
print("=======Y_train========")
print(Y_train.shape)
print("=======X_test========")
print(X_test.shape)
print("=======Y_test========")
print(Y_test.shape)

print("=========Create a linear regression model=========")
model=LinearRegression()
model.fit(X_train,Y_train)
print("=========Predicted values=========")
Y_pred=model.predict(X_test)
print(Y_pred)
print("=========Mean score=========")
mse=mean_squared_error(Y_test,Y_pred)
print("Mean squared error: ",mse)
print("=========r2_score=========")
r2=r2_score(Y_test,Y_pred)
print("R-squared: ",r2)

print("=========Scatter plot=========")
plt.scatter(Y_test,Y_pred)
plt.plot([min(Y_test),max(Y_test)],[min(Y_pred),max(Y_pred)])
plt.xlabel("Actual prices")
plt.ylabel("Prediced prices")
plt.title("Actual prices vs Prediced prices")
plt.show()
```

OUTPUT –




```
=======Info========
      ID     CRIM    zn  indus  chas    nox     RM   age     dis  rad  tax  ptratio   black  LSTAT  MEDV
0      1  0.00632  18.0   2.31     0  0.538  6.575  65.2  4.0900    1  296     15.3  396.90   4.98  24.0
1      2  0.02731   0.0   7.07     0  0.469  6.421  78.9  4.9671    2  242     17.8  396.90   9.14  21.6
2      4  0.03237   0.0   2.18     0  0.458  6.998  45.8  6.0622    3  222     18.7  394.63   2.94  33.4
3      5  0.06905   0.0   2.18     0  0.458  7.147  54.2  6.0622    3  222     18.7  396.90   5.33  36.2
4      7  0.08829  12.5   7.87     0  0.524  6.012  66.6  5.5605    5  311     15.2  395.60  12.43  22.9
..   ...      ...   ...    ...   ...    ...    ...   ...     ...  ...  ...      ...     ...    ...   ...
328  500  0.17783   0.0   9.69     0  0.585  5.569  73.5  2.3999    6  391     19.2  395.77  15.10  17.5
329  502  0.06263   0.0  11.93     0  0.573  6.593  69.1  2.4786    1  273     21.0  391.99   9.67  22.4
330  503  0.04527   0.0  11.93     0  0.573  6.120  76.7  2.2875    1  273     21.0  396.90   9.08  20.6
331  504  0.06076   0.0  11.93     0  0.573  6.976  91.0  2.1675    1  273     21.0  396.90   5.64  23.9
332  506  0.04741   0.0  11.93     0  0.573  6.030  80.8  2.5050    1  273     21.0  396.90   7.88  11.9

[333 rows x 15 columns]
=======Head========
   ID     CRIM    zn  indus  chas    nox     RM   age     dis  rad  tax  ptratio   black  LSTAT  MEDV
0   1  0.00632  18.0   2.31     0  0.538  6.575  65.2  4.0900    1  296     15.3  396.90   4.98  24.0
1   2  0.02731   0.0   7.07     0  0.469  6.421  78.9  4.9671    2  242     17.8  396.90   9.14  21.6
2   4  0.03237   0.0   2.18     0  0.458  6.998  45.8  6.0622    3  222     18.7  394.63   2.94  33.4
3   5  0.06905   0.0   2.18     0  0.458  7.147  54.2  6.0622    3  222     18.7  396.90   5.33  36.2
4   7  0.08829  12.5   7.87     0  0.524  6.012  66.6  5.5605    5  311     15.2  395.60  12.43  22.9
=======Shape========
(333, 15)
=======isna========
       ID   CRIM     zn  indus   chas    nox     RM    age    dis    rad    tax  ptratio  black  LSTAT   MEDV
0   False  False  False  False  False  False  False  False  False  False  False    False  False  False  False
1   False  False  False  False  False  False  False  False  False  False  False    False  False  False  False
2   False  False  False  False  False  False  False  False  False  False  False    False  False  False  False
3   False  False  False  False  False  False  False  False  False  False  False    False  False  False  False
4   False  False  False  False  False  False  False  False  False  False  False    False  False  False  False
..    ...    ...    ...    ...    ...    ...    ...    ...    ...    ...    ...      ...    ...    ...    ...
328 False  False  False  False  False  False  False  False  False  False  False    False  False  False  False
```




```
[333 rows x 15 columns]
=======column========
Index(['ID', 'CRIM', 'zn', 'indus', 'chas', 'nox', 'RM', 'age', 'dis', 'rad',
       'tax', 'ptratio', 'black', 'LSTAT', 'MEDV'],
      dtype='object')
=======isnull========
       ID   CRIM     zn  indus   chas    nox     RM    age    dis    rad    tax  ptratio  black  LSTAT   MEDV
0   False  False  False  False  False  False  False  False  False  False  False    False  False  False  False
1   False  False  False  False  False  False  False  False  False  False  False    False  False  False  False
2   False  False  False  False  False  False  False  False  False  False  False    False  False  False  False
3   False  False  False  False  False  False  False  False  False  False  False    False  False  False  False
4   False  False  False  False  False  False  False  False  False  False  False    False  False  False  False
..    ...    ...    ...    ...    ...    ...    ...    ...    ...    ...    ...      ...    ...    ...    ...
328 False  False  False  False  False  False  False  False  False  False  False    False  False  False  False
329 False  False  False  False  False  False  False  False  False  False  False    False  False  False  False
330 False  False  False  False  False  False  False  False  False  False  False    False  False  False  False
331 False  False  False  False  False  False  False  False  False  False  False    False  False  False  False
332 False  False  False  False  False  False  False  False  False  False  False    False  False  False  False

[333 rows x 15 columns]
=======choose features for prediction=======
       RM  LSTAT     CRIM
0   6.575   4.98  0.00632
1   6.421   9.14  0.02731
2   6.998   2.94  0.03237
3   7.147   5.33  0.06905
4   6.012  12.43  0.08829
..    ...    ...      ...
328 5.569  15.10  0.17783
329 6.593   9.67  0.06263
330 6.120   9.08  0.04527
331 6.976   5.64  0.06076
332 6.030   7.88  0.04741
```

File Edit Selection View Go Run ···

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

```
330    20.6
331    23.9
332    11.9
Name: MEDV, Length: 333, dtype: float64
==========Split the data into training and testing sets==========
=======x X_train========
(266, 3)
=======Y_train========
(266,)
=======X_test========
(67, 3)
=======Y_test========
(67,)
=========Create a linear regression model=========
=========Predicted values=========
[25.94994028 22.311055   23.88319868 29.76547353 25.29344805 18.15056907
 20.58575698 24.67375192 17.15016838 27.55641588 26.44128196 23.49137878
 20.33640121 33.09518209 19.87844805 31.87604667 21.18525052 10.92973781
 24.09975381 20.85388283 35.77417739 29.8626697  22.08397568 26.30996866
 17.8597552  38.81088311  6.4206052   5.60203376 26.47038316 15.19602277
 24.29828381 21.94122353 25.9754051  12.05897902 20.43713947 13.07497207
 26.66375322  1.46075136 19.23244138 18.0402809  24.28579093 29.68792852
 29.62760924 34.71748601 36.47502647 21.02987568 12.84270019 23.95249596
 13.53595639 19.53424865 14.84890308 30.96145905 27.63769854 15.62058186
 27.44583864 -4.59860403 18.81990355 27.99778056 25.39126484 20.86869291
 20.33427369 23.81131757 20.98227479 31.74029857 22.65408244 14.96713924
 24.8508193 ]
=========Mean score=========
Mean squared error:  28.6734629608111
=========r2_score=========
R-squared:  0.681400241911525
=========Scatter plot=========
```