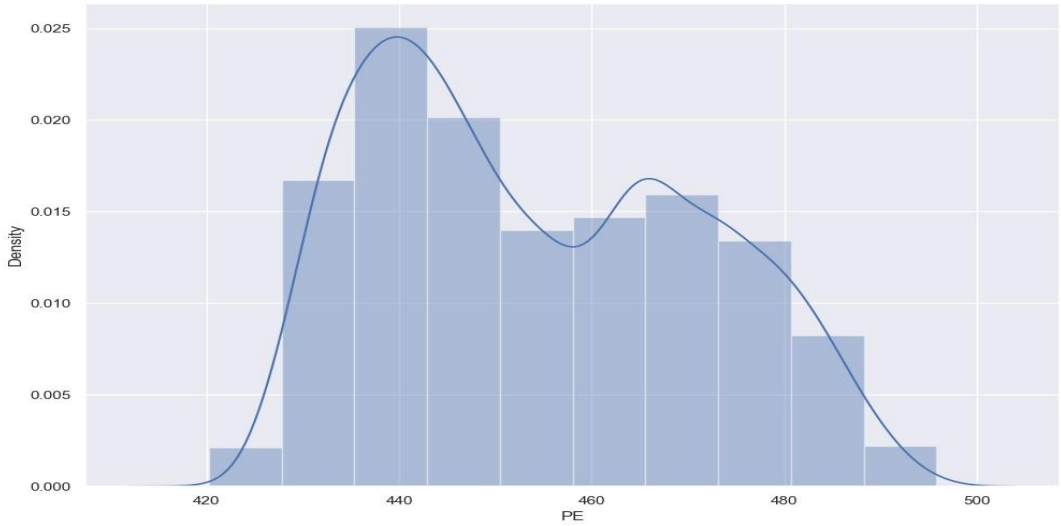
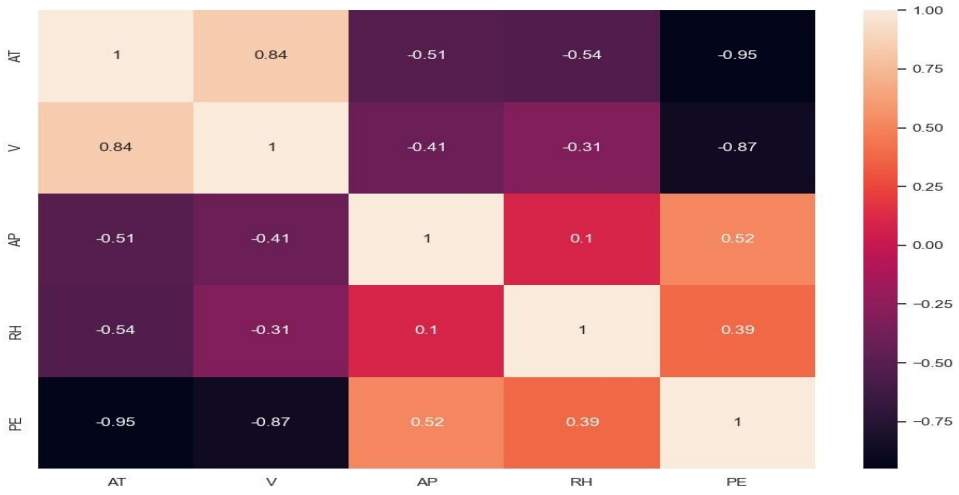
Linear Regression using EXEL DATA i.e., DATAML file

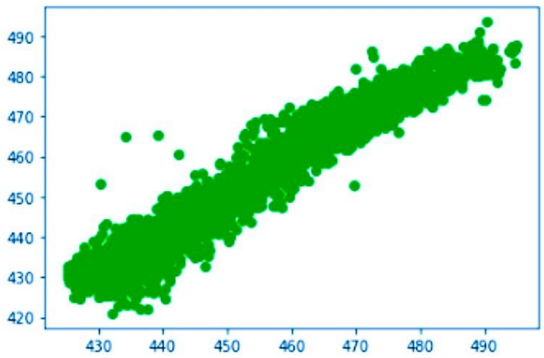
CODE:

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
from sklearn.datasets import fetch_openml
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error
from sklearn.metrics import r2_score
df = pd.read_excel('C:/Users/Zafar Ali/Desktop/Downloads/DATAML.xlsx')
df.head()
df.isnull().sum()
sns.set(rc={'figure.figsize':(11.7,8.27)})
sns.distplot(df['PE'],bins =10)
plt.show()
correlation_matrix = df.corr().round(2)
sns.heatmap(data=correlation_matrix, annot=True)
x = df.drop(['PE'],axis =1).values
y = df['PE'].values
print(x)
print(y)
x_train, x_test, y_train, y_test=train_test_split(x,
y,test_size=0.3,random_state=0)
lin_model = LinearRegression()
lin_model.fit(x_train, y_train)
y_pred = lin_model.predict(x_test)
print(y_pred)
y_train_predict = lin_model.predict(x_train)
testpred = lin_model.predict(x_test)
rmse= (np.sqrt(mean_squared_error(y_train, y_train_predict)))
r2 =r2_score(y_train,y_train_predict)
print("the model perf")
print("<<<<<<<<<<<_____________>>>>>>>>>>")
print('RMSE is {}'.format(rmse))
print('RMSE score is {}'.format(r2))
y_test_predict = lin_model.predict(x_test)
rmse= (mean_squared_error(y_test, y_test_predict))
r2 =r2_score(y_test,y_test_predict)
print("the model perf")
print("<<<<<<<<<<<_____________>>>>>>>>>>")
print('RMSE is {}'.format(rmse))
print('RMSE score is {}'.format(r2))
plt.scatter(y_test,y_test_predict,color='g')
plt.show()
```

OUTPUT:


0.025
0.020
0.015
0.010
0.005
0.000
Density
420
440
460
480
500
PE


AT
V
AP
RH
PE
1
0.84
-0.51
-0.54
-0.95
0.84
1
-0.41
-0.31
-0.87
-0.51
-0.41
1
0.1
0.52
-0.54
-0.31
0.1
1
0.39
-0.95
-0.87
0.52
0.39
1
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75


490
480
470
460
450
440
430
420
430
440
450
460
470
480
490

```
[[  14.96   41.76 1024.07   73.17]
 [  25.18   62.96 1020.04   59.08]
 [   5.11   39.4  1012.16   92.14]
 ...
 [  31.32   74.33 1012.92   36.48]
 [  24.48   69.45 1013.86   62.39]
 [  21.6    62.52 1017.23   67.87]]
[463.26 444.37 488.56 ... 429.57 435.74 453.28]
[431.40245096 458.61474119 462.81967423 ... 432.47380825 436.16417243
 439.00714594]
the model perf
<<<<<<<<<<_____________>>>>>>>>>>
RMSE is 4.576195321087113
RMSE score is 0.9279441090358653
the model perf
<<<<<<<<<<_____________>>>>>>>>>>
RMSE is 20.36838300288591
RMSE score is 0.9304112159477683
```