```python
import pandas as pd
import numpy as np
import pathlib as pl
import os
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import
mean_absolute_error,mean_squared_error,explained_variance_score,r2_score

script_dir = pl.Path(__file__).parent.absolute()
os.chdir(script_dir)

df = pd.read_csv('admissions-predict.csv')

df = df.drop(['Serial No.'], axis=1)
df.columns= df.columns.str.strip().str.replace(" ","_")

df.isnull().sum()
df.head()
df.describe()
df.info()

print("Correlation Matrix", df.corr())

X = df.drop(['Chance_of_Admit'], axis=1)
Y = df['Chance_of_Admit']

X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=0)

model = LinearRegression()
model.fit(X_train, Y_train)
predictions = model.predict(X_test)

print("----------------------\nLinear Regression Model\n----------------------")
print("Coefficients:\n\t|", "\n\t|".join([f"{X.columns[i]:17}:\t{model.coef_[i]:.8f}" for
i in range(len(X.columns))]))
print("Mean Absolute Error: ", mean_absolute_error(Y_test, predictions))
print("Mean Squared Error: ", mean_squared_error(Y_test, predictions))
print("Root Mean Squared Error: ", np.sqrt(mean_squared_error(Y_test, predictions)))
print("Linear Score: ", model.score(X_test, Y_test))
print("R Score: ", r2_score(Y_test, predictions))
print("Explained Variance Score: ", explained_variance_score(Y_test, predictions))

plt.scatter(Y_test, predictions)
plt.xlabel('True Values')
plt.ylabel('Predictions')
plt.plot(Y_test, Y_test, color='red', linewidth=2)
plt.title('Linear Regression Model')
plt.savefig('a.png')
```

```
dev@fedora:~/Projects/RTMNU-SEM-6$ /bin/python "/home/dev/Projects/RTMNU-SEM-6/PS-II/Practical 05/a.py"
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 500 entries, 0 to 499
Data columns (total 8 columns):
 #   Column             Non-Null Count  Dtype
---  ------             --------------  -----
 0   GRE_Score          500 non-null    int64
 1   TOEFL_Score        500 non-null    int64
 2   University_Rating  500 non-null    int64
 3   SOP                500 non-null    float64
 4   LOR                500 non-null    float64
 5   CGPA               500 non-null    float64
 6   Research           500 non-null    int64
 7   Chance_of_Admit    500 non-null    float64
dtypes: float64(4), int64(4)
memory usage: 31.4 KB
Correlation Matrix              GRE_Score  TOEFL_Score  University_Rating       SOP       LOR      CGPA  Research  Chance_of_Admit
GRE_Score          1.000000     0.827200           0.635376  0.613498  0.524679  0.825878  0.563398         0.810351
TOEFL_Score        0.827200     1.000000           0.649799  0.644410  0.541563  0.810574  0.467012         0.792228
University_Rating  0.635376     0.649799           1.000000  0.728024  0.608651  0.705254  0.427047         0.690132
SOP                0.613498     0.644410           0.728024  1.000000  0.663707  0.712154  0.408116         0.684137
LOR                0.524679     0.541563           0.608651  0.663707  1.000000  0.637469  0.372526         0.645365
CGPA               0.825878     0.810574           0.705254  0.712154  0.637469  1.000000  0.501311         0.882413
Research           0.563398     0.467012           0.427047  0.408116  0.372526  0.501311  1.000000         0.545871
Chance_of_Admit    0.810351     0.792228           0.690132  0.684137  0.645365  0.882413  0.545871         1.000000
-----------------------
Linear Regression Model
-----------------------
Coefficients:
        | GRE_Score         :   0.00214141
        |TOEFL_Score        :   0.00277039
        |University_Rating:     0.00555693
        |SOP               :   0.00051031
        |LOR               :   0.02041338
        |CGPA              :   0.11505398
        |Research          :   0.02412567
Mean Absolute Error:  0.04825389374865476
Mean Squared Error:  0.004079680034602108
Root Mean Squared Error:  0.06387237301527247
Linear Score:  0.7664048993199383
R Score:  0.7664048993199383
Explained Variance Score:  0.7674057830249
```


Linear Regression Model