Experiment 1: Devise a program to import, load, and view a dataset.

RA2311003010359

M.SUNAYANA

```
import pandas as pd import matplotlib.pyplot as plt
import numpy as np from sklearn.linear_model
import LinearRegression

# Read uploaded file data
= []

with open("dataset1.txt", "r") as f:
    for line in f:
        parts = line.strip().split()
if len(parts) >= 25:
            try:
                weight = float(parts[22])  # 23rd column
height = float(parts[23])  # 24th column
gender = int(parts[24])    # 25th column
data.append((height, weight, gender))            except
ValueError:
                continue

df = pd.DataFrame(data, columns=["Height", "Weight", "Gender"])
```

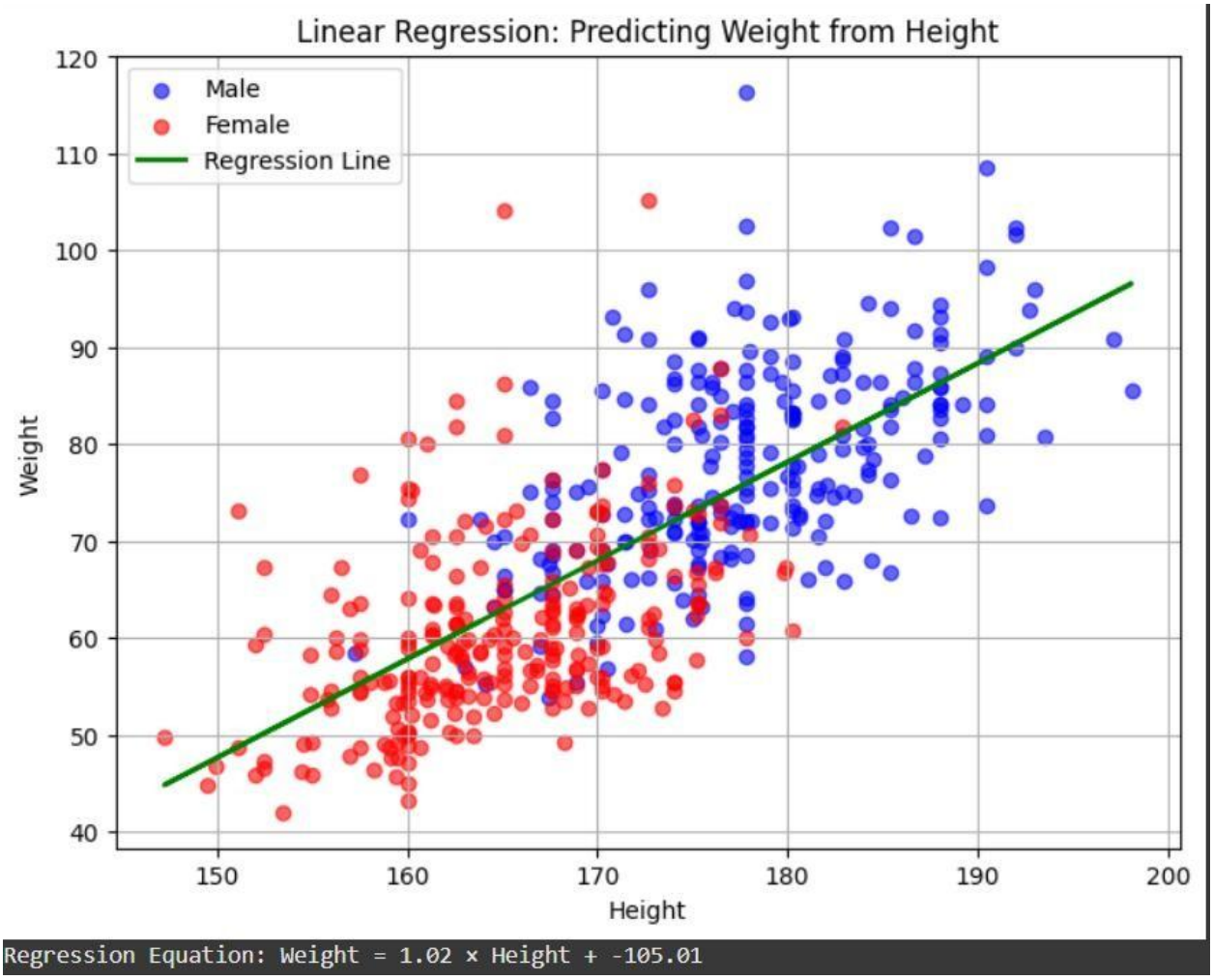
```
# Scatter plot by gender
plt.figure(figsize=(8, 6))
plt.scatter(df[df.Gender == 1].Height,
df[df.Gender == 1].Weight, color='blue',
label='Male', alpha=0.6)
plt.scatter(df[df.Gender == 0].Height, df[df.Gender == 0].Weight, color='red', label='Female',
alpha=0.6)

# Fit least-squares regression line X
= df["Height"].values.reshape(-1, 1)
y = df["Weight"].values

model =
LinearRegression()
model.fit(X, y) y_pred =
model.predict(X)

plt.plot(X, y_pred, color='green', label='Regression Line',
linewidth=2) plt.xlabel("Height") plt.ylabel("Weight")
plt.title("Linear Regression: Predicting Weight from Height")
plt.legend() plt.grid(True) plt.show()

# Print regression equation slope = model.coef_[0] intercept =
model.intercept_ print(f"Regression Equation: Weight = {slope:.2f} ×
Height + {intercept:.2f}")
```




```
Regression Equation: Weight = 1.02 × Height + -105.01
```