ASIGNMENT NO: 1

Code:

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

#file_path = '/content/sample_data/Book1.csv'
file_path = '/content/sample_data/kaggle1.csv'

data = pd.read_csv(file_path)
print(data.head())
print(data.info())
print(data.describe())

if 'Height' not in data.columns or 'Weight' not in data.columns:
    raise ValueError("Dataset must contain 'Height' and 'Weight' columns.")

X = data['Height'].values.reshape(-1, 1)
y = data['Weight'].values

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

model = LinearRegression()
model.fit(X_train, y_train)
y_pred = model.predict(X_test)
mse = mean_squared_error(y_test, y_pred)
```

```
r2 = r2_score(y_test, y_pred)


print(f"Mean Squared Error: {mse:.2f}")
print(f"R-squared: {r2:.2f}")
plt.scatter(X_test, y_test, color='blue', label='Actual')
plt.plot(X_test, y_pred, color='red', label='Predicted')
plt.title('Height vs Weight Prediction')
plt.xlabel('Height')
plt.ylabel('Weight')
plt.legend()
plt.show()
sample_height = [[170]]
predicted_weight = model.predict(sample_height)
print(f"Predicted weight for height {sample_height[0][0]} cm: {predicted_weight[0]:.2f}
kg")
```

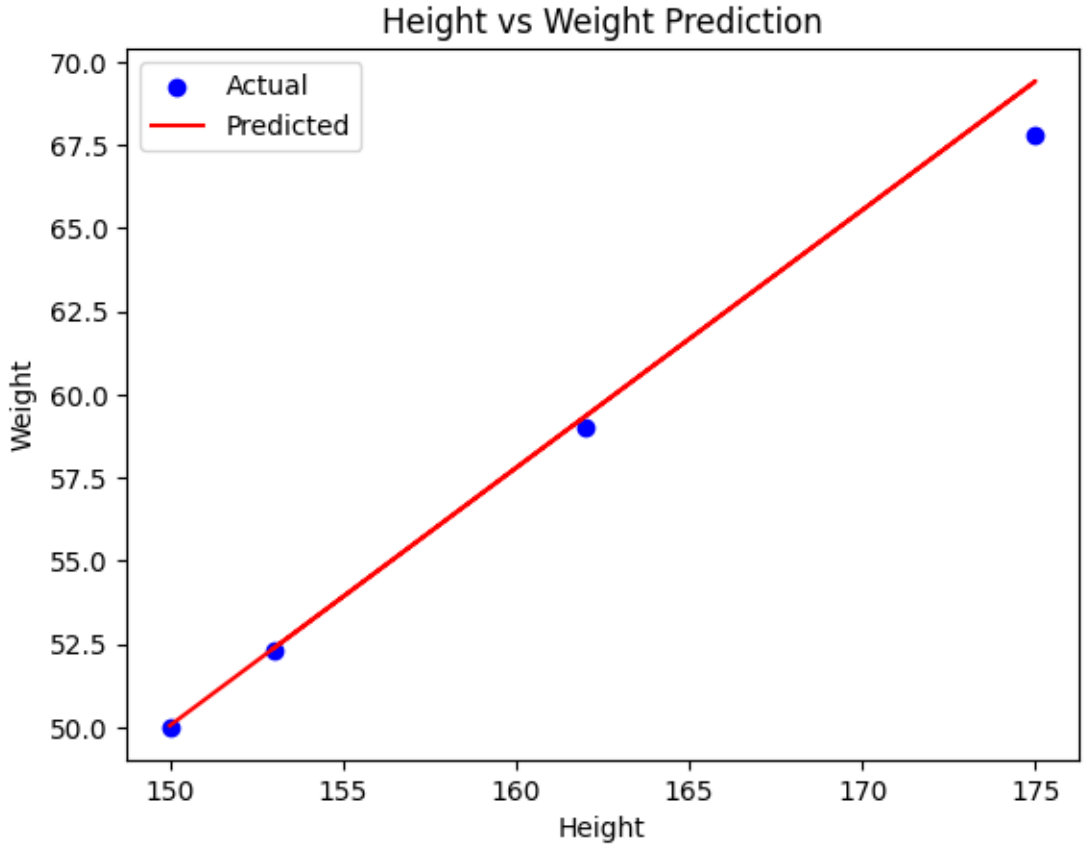
Output:

```
    Height  Weight
0       150    50.0
1       153    52.3
2       156    55.0
3       159    54.8
4       160    57.1
          Height          Weight
count   19.000000       19.000000
mean   170.263158       65.642105
std     11.807917        9.237022
min    150.000000       50.000000
25%    161.000000  58.050000
50%    170.000000  66.700000
75%    179.500000  73.550000
```

max 190.000000 78.900000

Mean Squared Error: 0.69

R-squared: 0.99



Predicted weight for height 170 cm: 65.55 kg

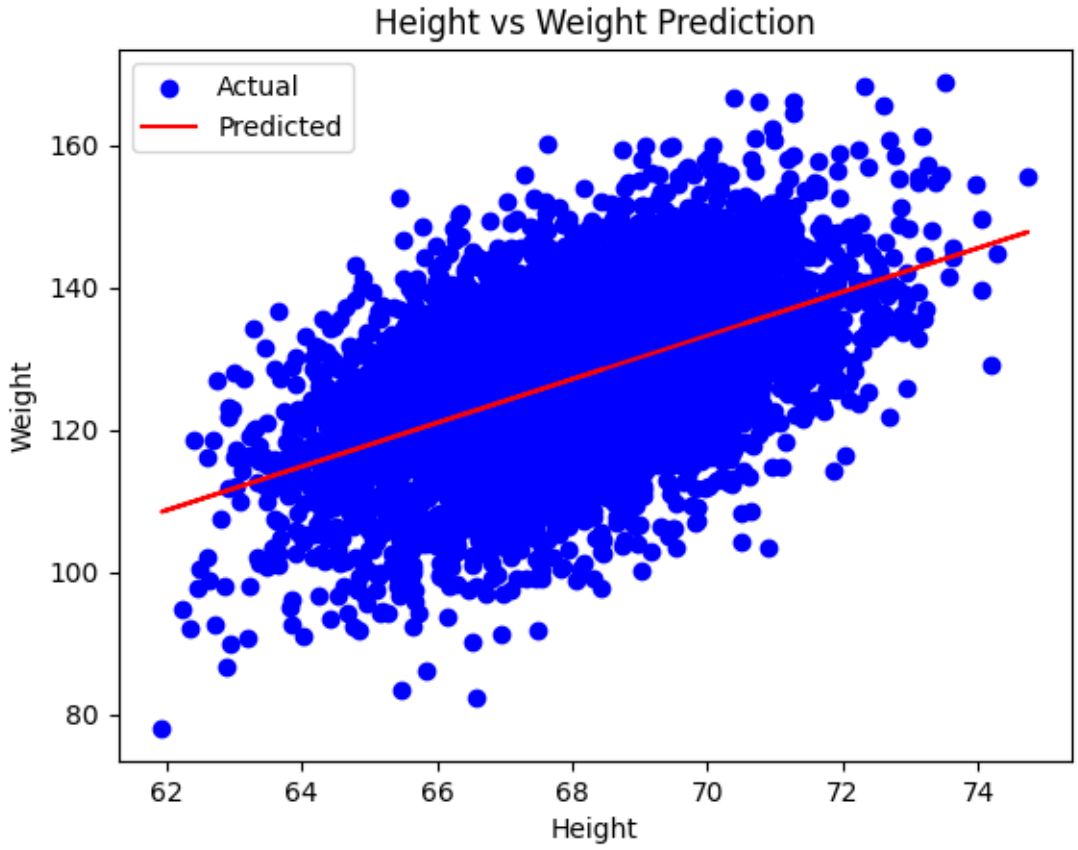
Kaggle Dataset Output:

Index	Height	Weight
0	1 65.78331	112.9925
1	2 71.51521	136.4873
2	3 69.39874	153.0269
3	4 68.21660	142.3354
4	5 67.78781	144.2971

	Index	Height	Weight
count	25000.000000	25000.000000	25000.000000
mean	12500.500000	67.993114	127.079421
std	7217.022701	1.901679	11.660898
min	1.000000	60.278360	78.014760
25%	6250.750000	66.704397	119.308675
50%	12500.500000	67.995700	127.157750
75%	18750.250000	69.272958	134.892850
max	25000.000000	75.152800	170.924000

Mean Squared Error: 102.49

R-squared: 0.26



Predicted weight for height 170 cm: 440.37 kg