

```
1 #
2 # Developed as part of YBI Foundation Internship
3 # By [Your Name]
4
5 # Step 1: Import Libraries
6 import pandas as pd
7 import numpy as np
8 import seaborn as sns
9 import matplotlib.pyplot as plt
10 from sklearn.model_selection import train_test_split
11 from sklearn.linear_model import LinearRegression
12 from sklearn.metrics import mean_squared_error, r2_score
13
14 # Step 2: Load Dataset
15 df =
16 pd.read_csv("https://raw.githubusercontent.com/YBIFoundation/Dataset/main/Fish.csv")
17 df.head()
18
19 # Step 3: Data Preprocessing
20 df.info()
21 df.describe()
22 df.isnull().sum()
23
24 # Convert 'Species' to dummy variable (optional)
25 df = pd.get_dummies(df, columns=['Species'], drop_first=True)
26
27 # Step 4: Define Features and Target
28 X = df.drop('Weight', axis=1)
29 y = df['Weight']
30
31 # Step 5: Train-Test Split
32 X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
33 random_state=42)
34
35 # Step 6: Train Model
36 model = LinearRegression()
37 model.fit(X_train, y_train)
38
39 # Step 7: Predict and Evaluate
40 y_pred = model.predict(X_test)
41
42 print("R2 Score:", r2_score(y_test, y_pred))
43 print("RMSE:", np.sqrt(mean_squared_error(y_test, y_pred)))
44
45 # Step 8: Visualization
46 plt.figure(figsize=(8,6))
47 sns.scatterplot(x=y_test, y=y_pred)
48 plt.xlabel("Actual Weight")
49 plt.ylabel("Predicted Weight")
50 plt.title("Actual vs Predicted Fish Weight")
51 plt.grid(True)
```

```
50 plt.show()
```

```
51
```