Question23:

23.Scenario:

You are a researcher working in a medical lab, investigating the effectiveness of a new treatment

for a specific disease. You have collected data from a clinical trial with two groups: a control group

receiving a placebo, and a treatment group receiving the new drug.Your goal is to analyze the data

using hypothesis testing and calculate the p-value to determine if the new treatment has a

statistically significant effect compared to the placebo. You will use the matplotlib library to

visualize the data and the p-value.

Answer:

import pandas as pd

import scipy.stats as stats

import matplotlib.pyplot as plt

import seaborn as sns

# Load the dataset

df = pd.read\_csv(r"C:\Users\jampa\Downloads\clinical\_trial\_data.csv")

# Separate data into control and treatment groups

control = df[df['group'] == 'control']['recovery\_time']

treatment = df[df['group'] == 'treatment']['recovery\_time']

# Perform an independent t-test

t\_stat, p\_value = stats.ttest\_ind(control, treatment)

# Print hypothesis testing results

print("T-statistic:", t\_stat)

print("P-value:", p\_value)

# Interpret the result

alpha = 0.05

if p\_value < alpha:

print("Result: Statistically significant difference (reject H0)")

else:

print("Result: Not statistically significant (fail to reject H0)")

# Visualization (Future-proofed for Seaborn v0.14.0+)

plt.figure(figsize=(10, 6))

sns.boxplot(x='group', y='recovery\_time', hue='group', data=df, palette='Set2', legend=False)

plt.title('Recovery Time by Group')

plt.xlabel('Group')

plt.ylabel('Recovery Time (days)')

plt.text(0.5, max(df['recovery\_time']) - 1, f'P-value: {p\_value:.4f}',

ha='center', fontsize=12, color='red')

plt.tight\_layout()

plt.show()

Output:



