1. Define Null Hypothesis (H0) and Alternative Hypothesis (H1)

2. Choose Significance Level (α)

3. Select Appropriate Test Based on Data Type and Assumptions:

a. Parametric Tests:

- Data follows a normal distribution

- Known population parameters or large sample size

i. Z-Test:

- Test for population mean when population variance is known

ii. T-Test:

- Independent T-Test:

\* Test for difference in means between two independent groups

\* Equal variance assumption (use Levene's Test for variance equality)

- Paired T-Test:

\* Test for difference in means within the same group (dependent samples)

b. Non-Parametric Tests:

- Data does not follow a normal distribution or sample size is small

i. Chi-Square Test:

- Test for independence or goodness of fit for categorical variables

ii. Kruskal-Wallis (KW) Test:

- Non-parametric alternative to one-way ANOVA for comparing more than two independent groups

iii. Kolmogorov-Smirnov (KS) Test:

- Test for normality of data distribution

4. Perform Test and Obtain Test Statistic and P-Value

5. Compare P-Value with Significance Level:

- If P-Value < α, reject Null Hypothesis (H0)

- If P-Value ≥ α, fail to reject Null Hypothesis (H0)

6. Interpret Results

**Python code for mean and median**

# Import necessary libraries

import numpy as np

import scipy.stats as stats

**# Define your data**

data\_group1 = np.array([...])

data\_group2 = np.array([...])

data\_paired1 = np.array([...])

data\_paired2 = np.array([...])

data\_categorical = np.array([...])

data\_distribution = np.array([...])

**# Z-Test for population mean (mean\_known)**

z\_stat, p\_value = stats.ztest(data\_group1, value=mu\_population, sigma=sigma\_population)

**If you want to compare the mean of a sample to a known population mean**

# Import necessary library

import scipy.stats as stats

# Define your data

sample\_data = [...] # Your sample data

population\_mean = ... # The known population mean you want to compare against

# Perform one-sample t-test

t\_stat, p\_value = stats.ttest\_1samp(sample\_data, population\_mean)

**# Independent T-Test**

t\_stat, p\_value = stats.ttest\_ind(data\_group1, data\_group2, equal\_var=stats.levene(data\_group1, data\_group2)[1]<0.05)

**# Paired T-Test**

t\_stat, p\_value = stats.ttest\_rel(data\_paired1, data\_paired2)

**# Chi-Square Test**

chi2\_stat, p\_value = stats.chisquare(data\_categorical)

**# Kruskal-Wallis Test**

kw\_stat, p\_value = stats.kruskal(data\_group1, data\_group2, ...)

**# Kolmogorov-Smirnov Test**

ks\_stat, p\_value = stats.kstest(data\_distribution, 'norm')

**Python test for proportion**

# Import necessary library

import scipy.stats as stats

# Define your data

successes1 = [...] # Number of successes in Group 1

total1 = [...] # Total number of observations in Group 1

successes2 = [...] # Number of successes in Group 2

total2 = [...] # Total number of observations in Group 2

**# Perform Z-test for population proportion**

z\_stat, p\_value = stats.proportions\_ztest([successes1, successes2], [total1, total2])