# Hypothesis Testing

**Instructions:**

Please share your answers filled in-line in the word document. Submit code separately wherever applicable.

Please ensure you update all the details:

**Name: \_\_\_\_\_\_\_\_\_\_\_\_\_ Batch ID:** \_\_\_\_\_\_\_\_\_\_\_

**Topic: Hypothesis Testing**

**Guidelines:**

**1. An assignment submission is considered complete only when the correct and executable code(s) and documentation explaining the method and results are submitted. Failing to submit either of those will be considered an invalid submission and not a correct submission.**

**2. Ensure that you submit your assignments correctly and in full. Resubmission is not allowed.**

**3. Post the submission you can evaluate your work by referring to the keys provided. (will be available only post the submission).**

**Hints:**

1. Hypothesis Testing Assignments, explanation of the solutions along with Business Objectives & Business Constraints should be documented in black and white along with the codes.
2. All the codes (executable programs) should run without errors
3. Python code, Documentation must be submitted
4. All the tests should be explained in detail (ex: Normality test, Variance test, etc.)

**Problem Statement:**

1. A F&B manager wants to determine whether there is any significant difference in the diameter of the cutlet between two units. A randomly selected sample of cutlets was collected from both units and measured. Analyze the data and draw inferences at a 5% significance level. Please state the assumptions and tests that you carried out to check the validity of the assumptions.

Data File: **Cutlets.csv**

1. A hospital wants to determine whether there is any difference in the average Turn Around Time (TAT) of reports of the laboratories on their preferred list. They collected a random sample and recorded TAT for reports from 4 laboratories. TAT is defined as a sample collected to report dispatch. Analyze the data and determine whether there is any difference in average TAT among the different laboratories at a 5% significance level.

Data File: **LabTAT.csv**

1. Sales of products in four different regions are tabulated for males and females. Find if male-female buyer rations are similar across regions.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **East** | **West** | **North** | **South** |
| Males | 50 | 142 | 131 | 70 |
| Females | 550 | 351 | 480 | 350 |

* Ho 🡪 All proportions are equal
* Ha 🡪 Not all Proportions are equal

Hint:

Check p-value

If p-Value < alpha, we reject Null Hypothesis

Data file: Buyer Ratio.csv

1. Telecall uses 4 centers around the globe to process customer order forms. They audit a certain % of the customer order forms. Any error in the order form renders it defective and must be reworked before processing. The manager wants to check whether the defective % varies by center. Please analyze the data at a 5% significance level and help the manager draw appropriate inferences.

File: **Customer OrderForm.csv**

1. Fantaloons Sales managers commented that % of males versus females walking into the store differs based on the day of the week. Analyze the data and determine whether there is evidence at a 5 % significance level to support this hypothesis.

File: **Fantaloons.csv**