**HYPOTHESIS TESTING**

Keerthika Kanagaraj

College of Professional Studies, Northeastern University

November 16, 2024

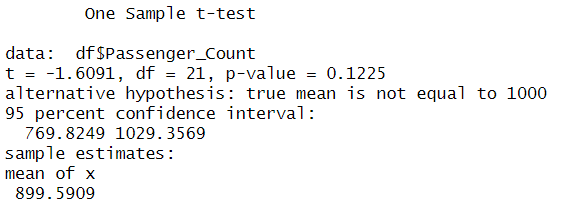
# INTRODUCTION

The dataset used for this practice is the Passenger count of LA International Airport. This dataset contains variables like Data extract date, Report period, Arrival\_Departure, Domestic\_International, Flight\_Type, and Passenger\_Count. For this practice, I have selected Passenger\_Count for hypothesis testing with data filtered for Arrival\_Departure = ‘Departure,’ Domestic\_International = ‘Domestic,’ Flight\_Type = ‘Charter’, and year = 2022 and 2023.

The dataset is sourced from <https://catalog.data.gov/dataset/los-angeles-international-airport-passenger-count-by-carrier-type>

# HYPOTHESIS TESTS:

### Analysis of One-Sample t-Test Results:

****

**Hypotheses:**

* **Null Hypothesis (H0):** The mean number of passengers is 1000.
* **Alternative Hypothesis (H1):** The mean number of passengers is not 1000.

**Test Statistic: t = -1.6091**

* + When the value is negative, the mean of the sample is lower than the predicted mean.

**Degrees of Freedom: df = 21**

* + This is the count of independent values that can vary in the dataset, calculated as the sample size minus one.

**p-Value: (α) = 0.1225**

* + If the null hypothesis is true, the p-value represents the probability of obtaining a test statistic as extreme as, or more extreme than, the one observed. The frequently accepted level of significance of 0.05 is less than the p-value of 0.1225.

**95% Confidence Interval: [769.8249, 1029.3569]**

* + This range means we are 95% confident that the actual population mean lies within the interval. Since 1000 is within the range, it further supports the null hypothesis.

**Sample Mean:**

* + The mean of sample passenger count is 899.5909.

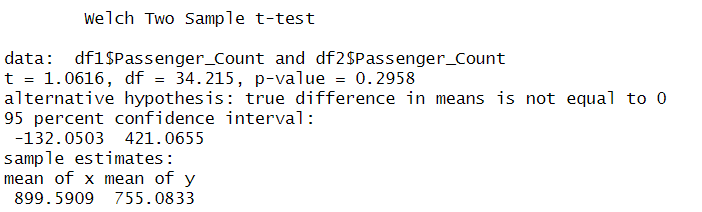
**Interpretation:**

Since the p-value (0.1225) exceeds the level of significance (α = 0.05), we fail to reject the null hypothesis. This means there is not sufficient evidence to conclude that the mean number of passengers significantly differs from 1000. Additionally, the confidence interval contains 1000, indicating that the true mean could reasonably be 1000.

We do not have enough evidence, based on the data and the one-sample t-test, to conclude that the average number of passengers differs much from 1000.

### Two Sample t-test (or) Hypothesis test for P-value:

Two different groups of passenger datasets were considered for carrying out a two-sample t-test. Group 1 contains the passenger count of LA International Airport for 2022 and 2023, while Group 2 contains the passenger count of LA International Airport for 2021 and 2020.



**Hypotheses:**

* **Null Hypothesis (H0):** Mean number of passengers in both groups is equal (μ1=μ2).
* **Alternative Hypothesis (H1):** Mean number of passengers in both groups is unequal (μ1≠μ2).

**Test Statistic: t = 1.0616**

* + This figure shows how the sample means vary regarding the samples' variability. A positive number indicates that the first group's mean exceeds the second group's.

**Degrees of Freedom: df = 34.215**

* + The degrees of freedom for the Two Sample t-test, which accounts for the sizes of the sample and variances of the two groups.

**p-Value: p = 0.2958**

* + The p-value indicates the probability of obtaining a test statistic at least as extreme as the one observed, assuming the null hypothesis is true. The conventional significance level of 0.05 is surpassed with a p-value of 0.2958.

**95% Confidence Interval: [-132.0503, 421.0655]**

* + This range means we are 95% confident that the actual difference in means lies between -132.0503 and 421.0655. Since this interval includes zero, the true difference could be zero, indicating no significant difference between the means of the group.

**Sample Means:**

* **mean of x = 899.5909**
  + The sample mean for the first group is 899.5909 passengers.
* **mean of y = 755.0833**
  + The sample mean for the second group is 755.0833 passengers.

**Interpretation:**

Since the p-value (0.2958) exceeds the level of significance (α = 0.05), we fail to reject the null hypothesis. This means there is not enough evidence to conclude that there is a significant difference between the mean passenger counts of the groups. Additionally, the 95% confidence interval includes zero, further supporting the possibility of no true difference in the means.

# In conclusion, we do not have enough evidence to conclude that the average number of passengers between the two groups is significantly different based on the data and the results of the Two-Sample t-test.

# REFERENCES:

* Bluman, A. G. (2018). *Elementary statistics: A step by step approach* (10th ed.). McGraw-Hill Education.
* Moore, D. S., Notz, W. I., & Flinger, M. A. (2015). *The basic practice of statistics* (7th ed.). W. H. Freeman and Company.
* R Core Team. (n.d.). *One-sample t-test with distributions3*. The Comprehensive R Archive Network (CRAN). Retrieved November 16, 2024, from <https://cran.r-project.org/web/packages/distributions3/vignettes/one-sample-t-test.html>
* Kassambara, A. (n.d.). *How to do two-sample t-test in R*. Datanovia. Retrieved November 16, 2024, from <https://www.datanovia.com/en/lessons/how-to-do-a-t-test-in-r-calculation-and-reporting/how-to-do-two-sample-t-test-in-r/>