ANOVA [Analysis Of Variance]

ANOVA (Analysis of Variance) is a statistical test used to compare the means of **two** or **more** groups. It determines if there are statistically significant differences between the group means

Assumptions of ANOVA
- **Independence of observations**
- **Normality** of the residuals - can test using Shapiro-Wilk Test
- **Homogeneity of variances** (equal variance across groups) – can test using **Levene's Test**

H_0 (Null Hypothesis): All groups have equal variances.

H_1 (Alternate Hypothesis): At least one group has a different variance.

Types	One Way ANOVA	Two Way ANOVA
Definition	a statistical method used to Compares the means of two or more groups/columns based on a single independent variable/column	a statistical method used to analyze the effects of two or more independent variables/columns on a dependent variable/column
Example	Comparing the effectiveness of three different types of fertilizers on plant growth, where the independent variable is the type of fertilizer	Comparing worker productivity based on two independent variables, such as salary and skill set.
If assumptions are violated	Use **Kruskal-Wallis Test** instead of One-Way ANOVA	Use **Friedman Test** instead of Two-Way ANOVA
which group significantly differ ?	Tukey's HSD Test	Tukey's HSD Test

Use when you have **repeated measures or related groups** (e.g., same subjects under different treatments)

If dataset have independent samples Friedman Test is not appropriate Use a **Generalized Linear Model (GLM)** as it can handle Non-

A test to determine which group means are significantly different from each other

🧠 What Is GLM?

GLM = Generalized Linear Model

It's a flexible version of linear regression that can model:

Type	Model Family	Example Use Case
Continuous data	Gaussian	Predicting sales, conversion rates
Count data	Poisson	Number of visits, clicks
Binary outcomes	Binomial	Yes/No decisions (e.g., converted or not)