ANOVA

Analysis of Variance

MUKESH KUMAR

AGENDA

- What is ANOVA
- Types of ANOVA
- One-Way ANOVA
- Two-Way ANOVA

What is ANOVA

- **Definition:** ANOVA (Analysis of Variance) is a statistical method used to compare means among three or more groups.

- **Purpose:** To determine if at least one group mean is significantly different from others.

Types of ANOVA

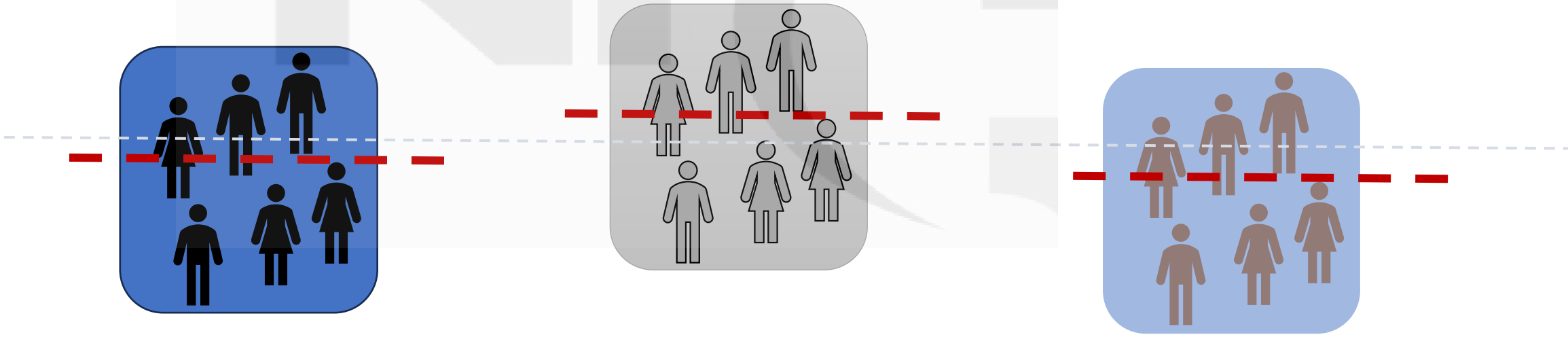
- One-Way ANOVA:
  - Compares means across one independent variable with multiple levels.
  - Example: Comparing test scores across different teaching methods.
- Two-Way ANOVA:
  - Compares means across two independent variables.
  - Example: Examining the effect of teaching method and study time on test scores.

One Way ANOVA:

Assumptions:

- The samples are independent.
- The data is normally distributed.
- The variances are equal across groups (homogeneity of variance).

- An ANOVA tests whether there are statistically significant differences between three or major groups
- More precisely , it tests whether there is a significant difference between the mean values of groups.

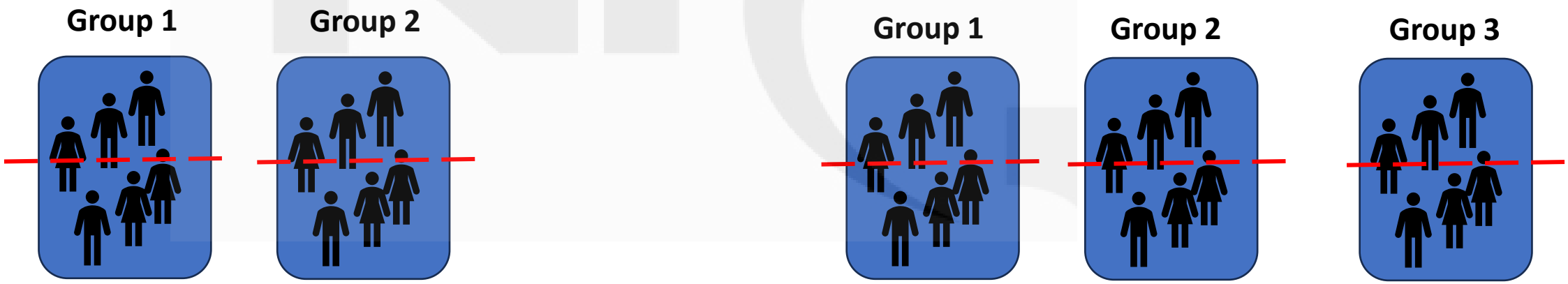


Extension of t-Test

- t-Test

- ANOVA

- ANOVA is and extension of t-Test



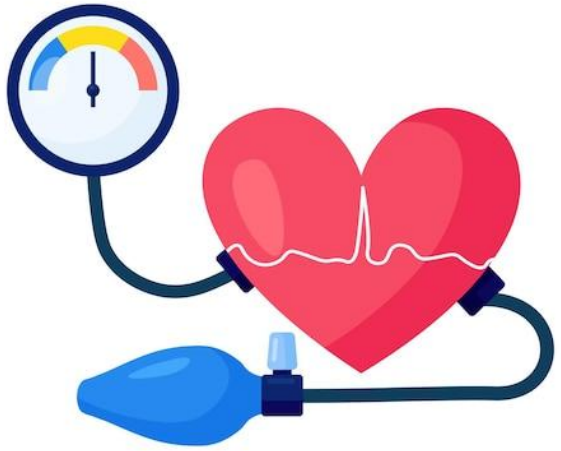
What are the hypothesis?

- Null Hypothesis : The means of the groups are equal

- Alternate Hypothesis: Mean values of groups are not equal

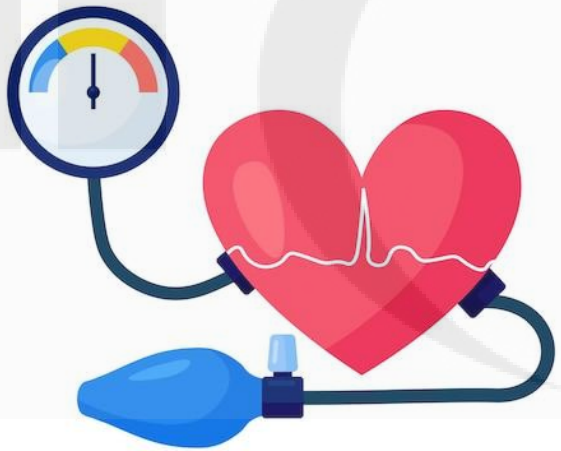
Example

- Three different drugs for lowering BP



- **Research Question:**

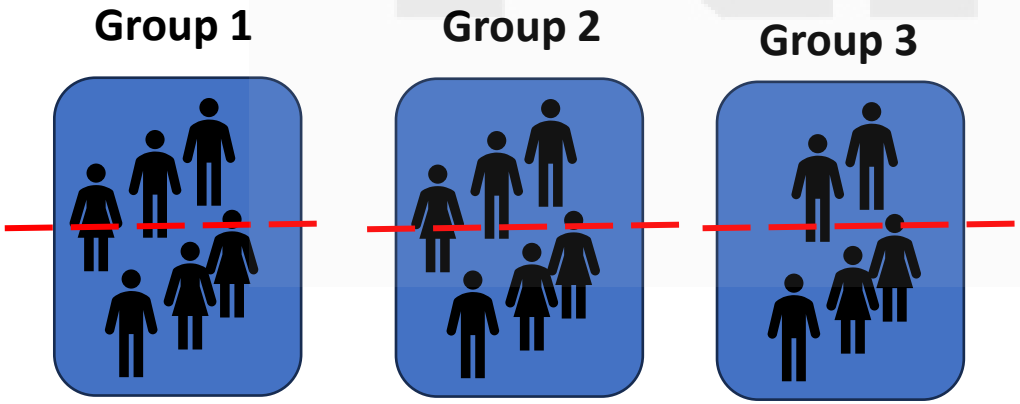
Do the three drugs have different effects on BP?



Null and Alternate Hypothesis

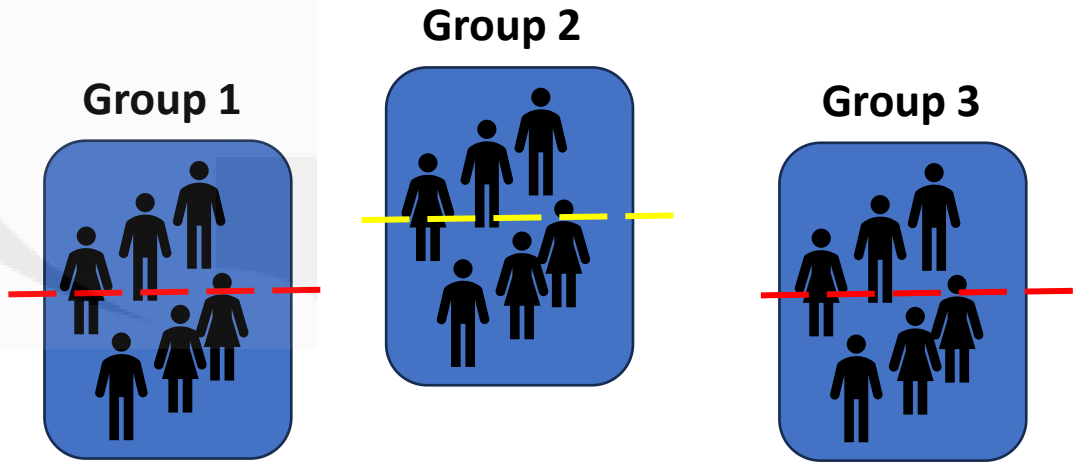
- **Null Hypothesis:**

There is no difference btn three drugs in terms of BP



- **Alternate Hypothesis:**

There is a difference btn three drugs in terms of BP

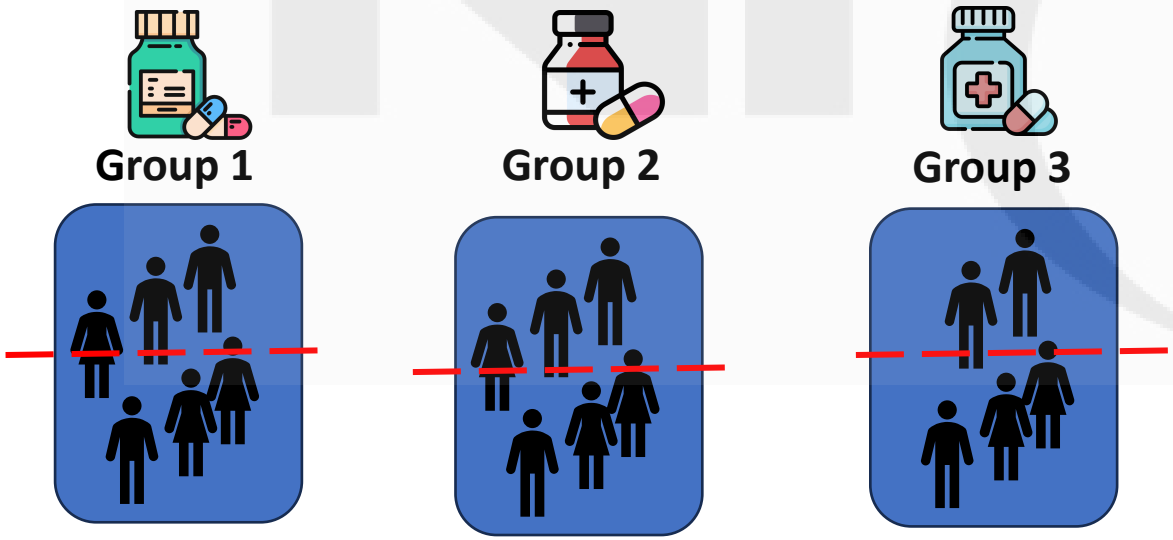


What next?

- We get the data now

Population and Sample

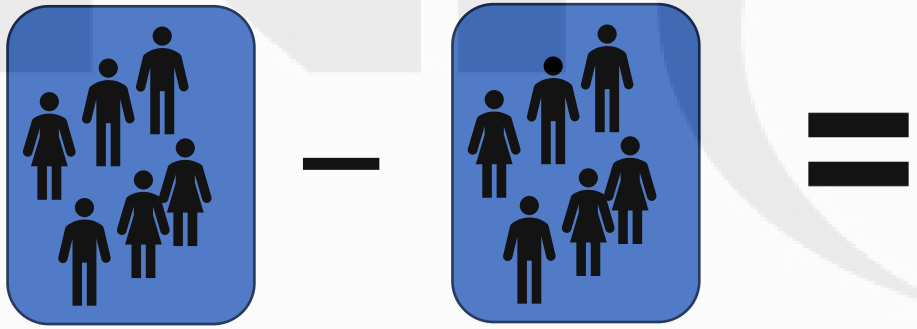
- Who is our population?
- All the people with high BP say across India



- sample
- So, we draw a sample of say 1500 people and divide them in 3 groups
- Each group is then given a different medicine

Measure the BP for each group

- We measure the BP for each group before and after medication
- We find the difference in BP for each group



5
7
8
34
3
2
6
9

Assumptions

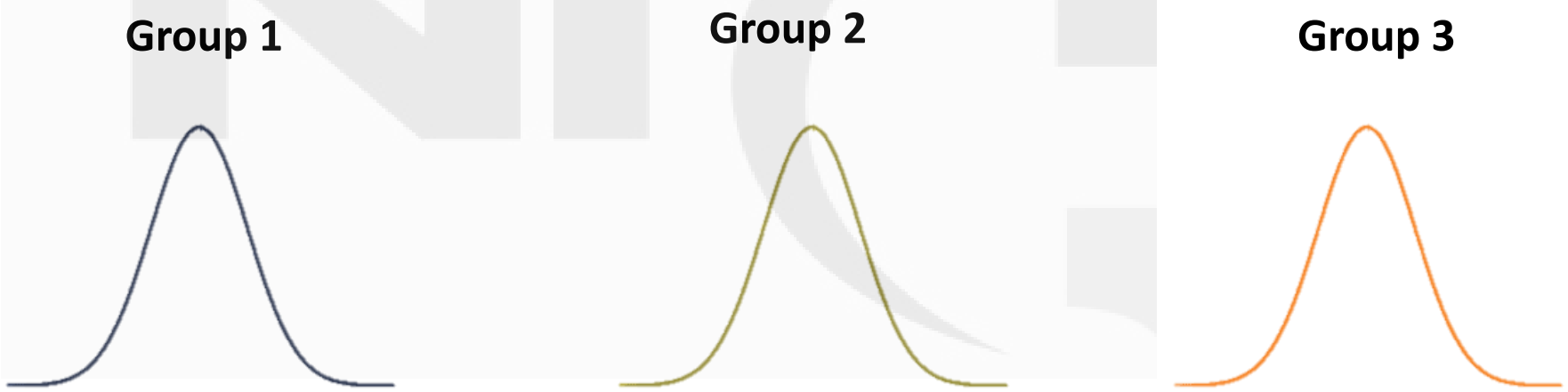
- The samples are independent.
- The data is normally distributed.
- The variances are equal across groups (homogeneity of variance).

Samples are independent.

- Measure value of one group should not be influenced by the measured value of another group

Normality

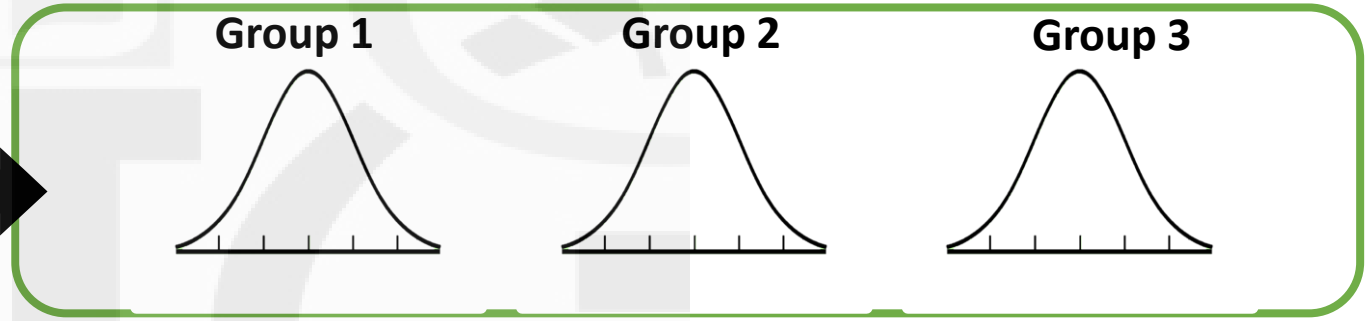
- Data across the three groups should be normally distributed
- This assumption is less important when the sample size is big. For smaller



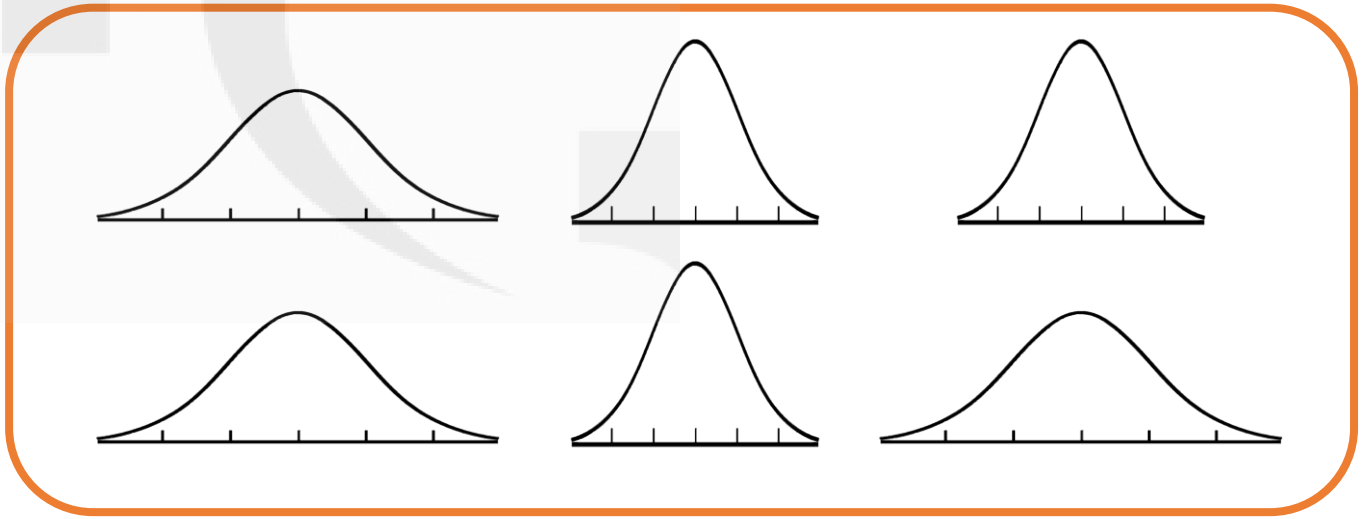
Homogeneity

- Variances in each group should be roughly the same (Levene Test)

Similar VARIANCES across groups



Different variances across groups



How ANOVA works?

ANOVA uses a ratio called **F-Statistics**:

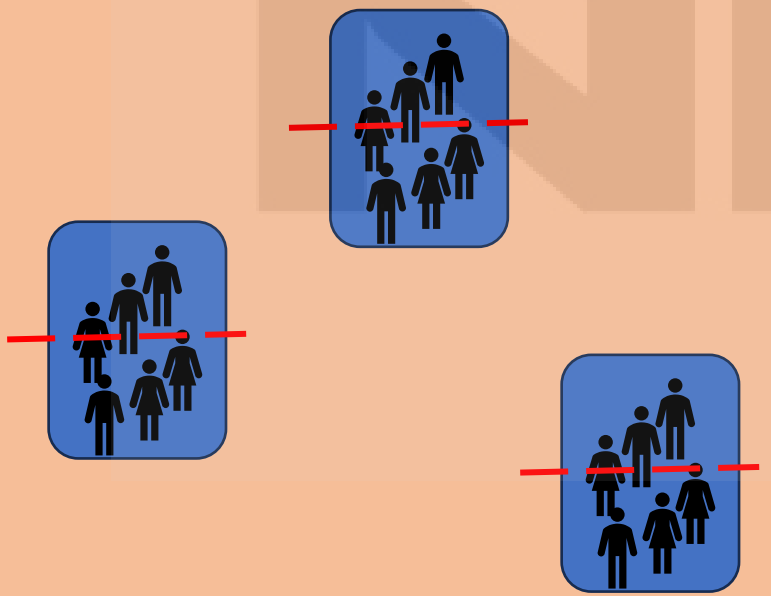
$$F - \text{Statistics} = \frac{\text{Variance Between Groups}}{\text{Variance within Groups}}$$

If F is significantly **large, we reject H_0** and conclude that at least one group is different

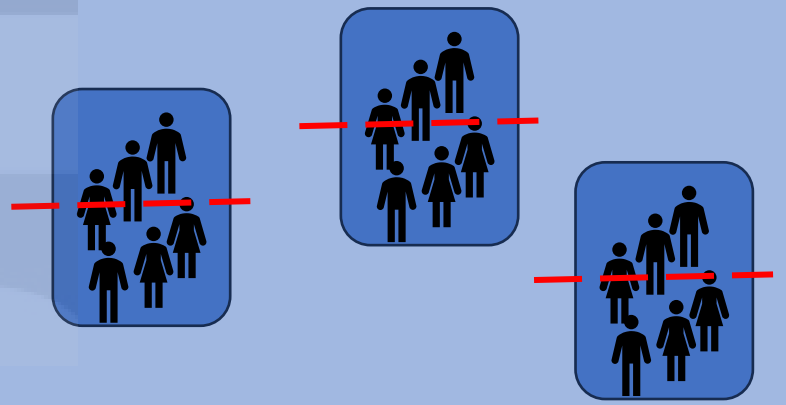
Variance Between Groups

- It measures how much the mean values of the group differ from each other

1 . High Variance between groups



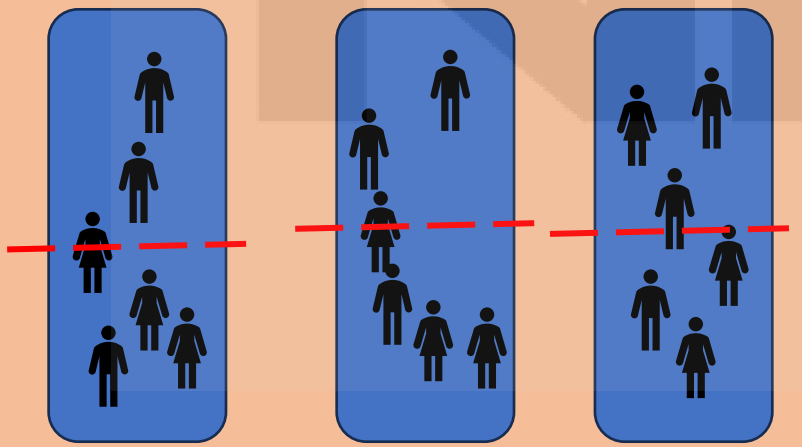
2. Low Variance between the Groups



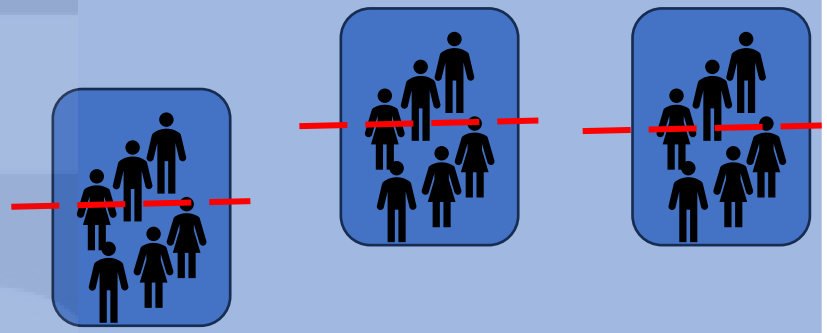
Variance within Groups

- It measures how much each individual data point in each group fluctuate

1 . High Variance **within** groups



2. Low Variance within the Groups



Interpreting One-Way ANOVA Results

- **F-ratio:** The larger the ratio, the more likely the differences are statistically significant.
- **p-value:** If $p < 0.05$, we reject H_0.
- **Post hoc tests:** If ANOVA is significant, use Tukey's HSD or Bonferroni correction to identify which specific groups differ

Summary

Steps in Conducting ANOVA

1. Determine the appropriate test and collect data.
2. Establish the level of significance (α).
3. Calculate the F-statistic and degrees of freedom.
4. Calculate p-Value

Interpreting Results

- **P-value Analysis**: If $p<\alpha$, reject H_0; conclude that significant differences exist among group means.

- **Post-Hoc Tests**: If significant, use tests like Tukey's HSD to identify which groups differ.

Practical Applications of ANOVA

- Used in various fields such as:
  - Medicine (comparing treatment effects)
  - Business (evaluating customer preferences)
  - Agriculture (assessing crop yields under different conditions)

TWO WAY ANOVA

- Will be covered after LR model as it depends on other topics