significant effect on student performance.

ANOVA

* In simple terms ANOVA is a statistical method that helps data scientists compare the means of several groups to find out if there are significant differences among them. Its like a tool to see if different groups are truly different or if their differences could be due to random chance.

Scenario - Imagine you are a data scientist working for a school, and your goal is to figure out whether three different teaching methods (A, B, and C) have a significant impact on students tests scores

Method A - Traditional lectures
Method B - Online Tutorials
Method C - Small Group Discussions

Considering n=15.

A	B	C
78	70	65
85	72	58
88	68	63
92	75	70
80	74	72
78	70	68
84	73	74
90	78	62
85	69	70
88	71	69
75	75	63
82	72	68
85	70	70
80	76	67
85	72	75

Step 1. Hypotheses

H_0 :- No significant difference in test scores b/w 3 methods.

$(\mu A = \mu B = \mu C)$

H_1 :- Significant diff. in test scores b/w 3 methods.

$(\mu A \neq \mu B \neq \mu C)$

Step 2. Table

	A	B	C
1.	78	70	65
2.	85	72	58
3.	88	68	63
4	92	75	70
5.	80	74	72
6	78	70	68
7.	84	73	74
8.	90	78	62
9.	85	68	70
10.	88	71	69
11.	75	75	63
12.	82	72	68
13.	85	70	70
14	80	76	67
15.	85	72	75
Step 3. Mean	83.66	72.26	67.6

Step 4. Grand Mean → 74.506 → μG

Step 5. Calculate Sum of Squares b/w Groups (SSB)

$$SSB = (\text{No. of obs}) \left[(\mu G - \mu A)^2 + (\mu G - \mu B)^2 + (\mu G - \mu C)^2 \right]$$

$$= 15 (83.79 + 5.044 + 47.69)$$

$$= 2047.86$$

Step 6- Sum of Squares within Groups (SSW)

$SSW = \sum_{i=1}^{15} (x_i - \bar{x})^2$

$= 32.03 + 1.79 + 18.83 + 69.55 + 13.39 + 32.03 + 0.1156 + 40.19 + 1.79 + 18.83 + 74.99 + 2.7556 + 1.79 + 13.39 + 1.79$

$= 323.2612$

$SSW_B = 5.1076 + 0.0676 + 18.14 + 7.5076 + 3.0276 + 5.1076 + 0.5476 + 32.94 + 18.14 + 1.5876 + 7.5076 + 0.0676 + 5.1076 + 13.98 + 0.0676 = 118.9036$

$SSW_C = 6.76 + 92.16 + 21.16 + 5.76 + 19.36 + 0.16 + 40.96 + 31.36 + 5.76 + 1.96 + 21.16 + 0.16 + 5.76 + 0.36 + 54.76$

$= 307.6$

$SSW = 749.76$

Step 7- Degrees of freedom = No. of groups - 1 = 2 → df_B (degree of freedom Between)

Degree of freedom with, $df_W = 45 - 3 = 42$

Step 8- Mean Squares.

MSB (Mean Square Between) = $SSB / df_B = \frac{2047.86}{2} = 1023.93$

MSW (Mean Square Within) = $SSW / df_W = 17.5514$

Step 9- Calculate F-statistic

F-statistic = $\frac{MSB}{MSW} = 57.3585$

Step 10- Determine Critical F-value

$df_B \Rightarrow 2 \rightarrow$ row

$df_W = 42 \rightarrow$ column

F-statistic table = 3.2199

Critical F-value = 3.2199

Step 11: Conclusion

- F-statistic > Critical F-value

We reject null hypothesis.