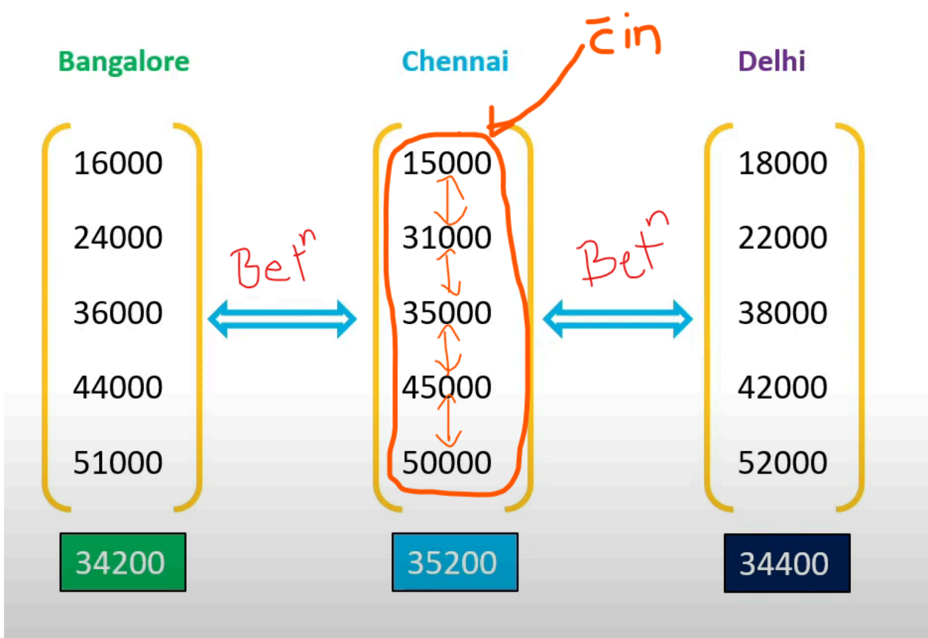
F- Test (ANOVA)

ANOVA= Analysis of Variance

- Used to compare **means** between **two or more groups** to determine if at least one differs significantly.

https://www.youtube.com/watch?v=tRqUNwEY63Y

- Used when Independent variables are categorical and dependent variables are numeric→Use ANOVA/T-test
    - 2 Categories → T-Test
    - 2+ Categories → ANOVA
- Used to **compare means** of 2 or more groups.
- Means of 2 group → Z Test, t test
- Difference between groups variables matters more than the within group variables.



$$F = \frac{\textit{Between Group Variance}}{\textit{Withing Group Variance}}$$

There is a less difference between groups

- **Less F-value= There is less diff between groups**

Samsung	Google	Sony
31000	44000	15000
35000	45000	16000
36000	50000	18000
38000	51000	22000
42000	52000	24000

ANOVA

F(2,12) = 73, P = 0

Bangalore	Chennai	Delhi
16000	15000	18000
24000	31000	22000
36000	35000	38000
44000	45000	42000
51000	50000	52000

ANOVA

F(2,12) = 0.007, P = 0.99

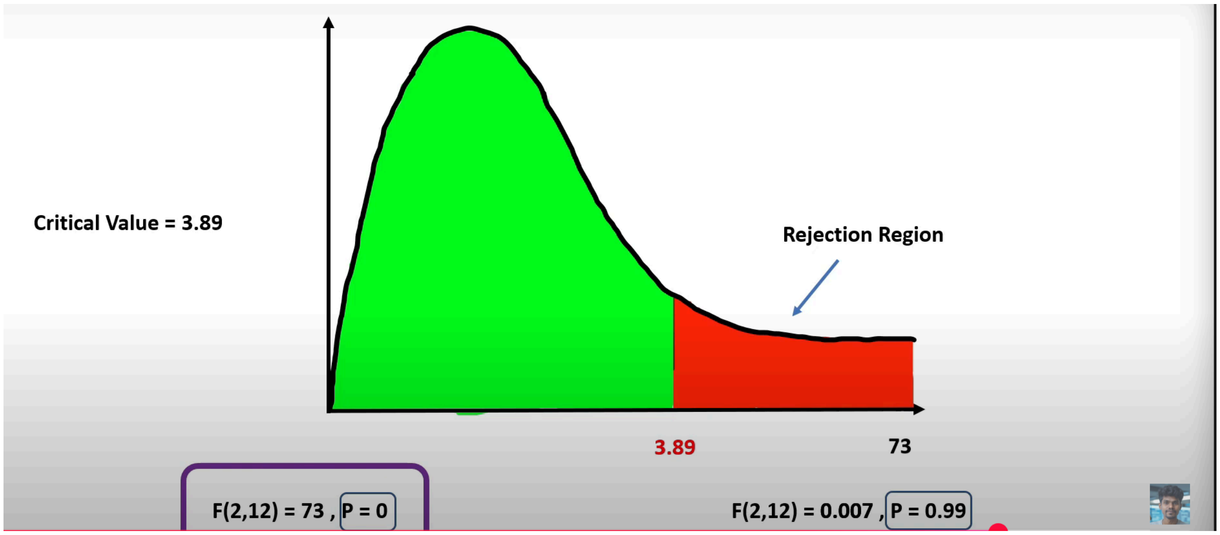
F(2,12) = 73 , P = 0

F(2,12) = 0.007 , P = 0.99

2 = No of Categories - 1

12 = No of Samples - No of Categories

Degrees of Freedom



2 Most Important parameters:

1. Factors (Variables)
2. Levels

Factors:

- If a medicine has 5 mg, 10 mg, 15 mg dosage → This is level
    - Medicine is factor
- Mode of payment → Factor
    - Gpay, Phonepe, IMPS, NEFT → Levels

Types of ANOVA

Type	Used When...	Example
One-Way ANOVA	Comparing means of one independent variable with multiple groups.	Comparing exam scores of students from three different schools.

Type	Used When...	Example
Two-Way ANOVA	Comparing means of two independent variables to see if they affect the dependent variable.	Checking if teaching method and gender affect exam scores.
Repeated Measures ANOVA	When the same individuals are measured multiple times (before and after changes).	Measuring **blood pressure** before, during, and after treatment.

Types of ANOVA

1. One way: 1 factor with 2 levels. These levels are independent
    - eg. Dr. wants to test a medication with 5mg, 10mg and 15 mg medicine
2. Repeated measure: 1 factor with at least 2 levels. Levels are **dependent**.
    - eg. A man running → Day 1, Day 2, Day 3
        - a study measuring participants' anxiety levels at three different times: before therapy, one month after therapy, and six months after therapy
3. Factorial: 2 or more factors : Each with at least 2 or more levels. Levels could be **dependent** or independent

    eg.

    1. **Study Method (Factor A)**:
        - Level 1: Visual
        - Level 2: Auditory
    2. **Time of Day (Factor B)**:
        - Level 1: Morning
        - Level 2: Evening

    ### Design

    This creates a **2 × 2 factorial design:**

    - Group 1: Visual Study Method in the Morning
    - Group 2: Visual Study Method in the Evening

- Group 3: Auditory Study Method in the Morning
- Group 4: Auditory Study Method in the Evening

ANOVA is done in **F- Distribution**

Assumptions

- **Normality:** Means are **Normally distributed**
- **No outliers**
- **Homogeneity of variance** → variance (or spread) of data is equal across different groups or samples.
    - In simpler terms, it means that the data in each group is equally spread out.
    - Same variance

$$\left[\sigma_1^2 = \sigma_2^2 = \sigma_3^2\right]$$

- Population variance in different levels of each independent variables are equal.
- Samples are **independent** and **random**.

Null Hypothesis: H0: μ1=μ2=μ3…μk

Alt Hypothesis: H1: μ1≠μ2 ≠ μ3…μk

Formula / Test Statistics (F- Test)

The ANOVA test is based on the **F-statistic**, which compares the variance between groups to the variance within groups:

$$F = \frac{\text{Variance Between Groups}}{\text{Variance Within Groups}}$$

Where:

- **Variance Between Groups**: Measures how much the group means differ from the overall mean.
- **Variance Within Groups**: Measures how much individual data points differ from their group mean.

Steps to Perform ANOVA

1. **State the Hypotheses**:
    - $H_0 : \mu_1 = \mu_2 = \mu_3$
    - H_1: At least one mean is different.
2. **Choose Significance Level (αα)**:
    - Common choice: **α=0.05**
3. **Calculate the F-Statistic**:
    - Use the formula above to compute *F*.
4. **Find the Critical Value or p-value**:
    - Use the F-distribution table or Python's `scipy.stats` to find the p-value.
        - `f_stat, p_value = stats.f_oneway(group1, group2, group3)`
5. **Make a Decision**:
    - If p<α, reject H_0
    - Otherwise, fail to reject H_0

Eg:

$\underbrace{X_1 \quad X_2 \quad X_3}$ → Variance between Sample.

Variation within Sample

X_1	X_2	X_3
1	6	5
2	7	6
4	3	3
5	2	2
3	1	4
$\sum X_1 = 15$	$\sum x_2 = 19$	$\sum x_3 = 20$
$\bar{X}_1 = 3$	$\bar{X}_2 = 19/5$	$\bar{X}_3 = 4.$

One way ANOVA Example

① Doctors want to test a new medication which reduces headache. They split the participant into 3 condition [15mg, 30mg, 45mg]. Later on the doctor ask the patient to rate the headache between [1-10]. Are there any differences between the 3 conditions using alpha = 0.05?

15mg	30mg	45mg.
9	7	4
8	6	3
7	6	2
8	7	3
8	8	4
9	7	3
8	6	2

We have to solve this with Chi Square test.

H_0: μ15=μ30=μ45

H_1: At least 1 mean is not equal

Degree of freedom

N=21 (entire sample)

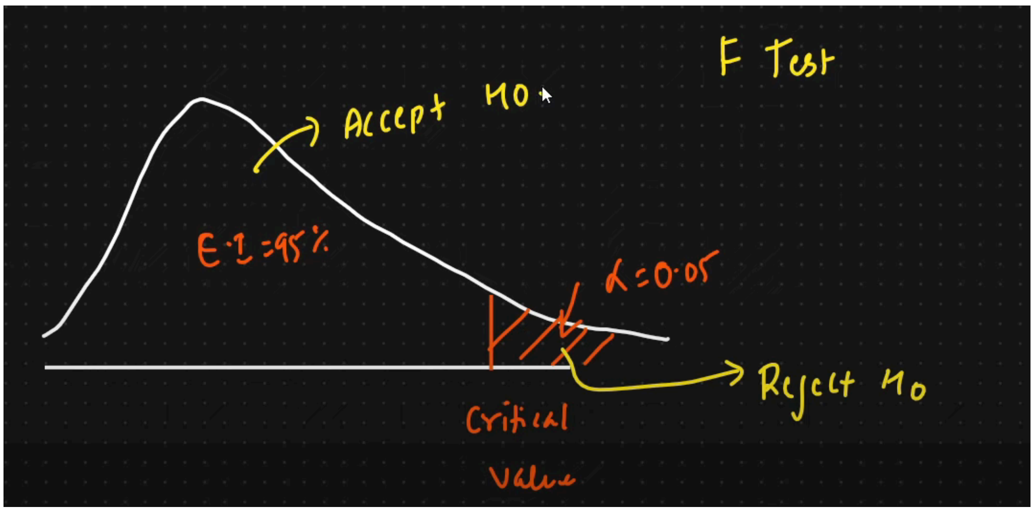
a = 3 (categories)

n=7 (sample)

df between= a-1= 3-1= **2**

df within= n-a= 21-3= **18**

df total= N-1= 21-1= 20



We see (2,18) in the f table for 0.05 significance level

Critical values of F for the 0.05 significance level:

	1	2	3	4	5	6	7	8	9	10
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.39	19.40
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.97	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.10	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.97	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.56	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	4.33	3.47	3.07	2.84	2.69	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.38	2.32	2.28

Critical value= 3.56

Decision Rule:

If F>3.56 → We reject the H0

If F≤ 3.56 → We fail to reject the H0

- **Between-group variability** (Mean Square Between, MSB):

$$MSB = \frac{SSB}{df_{\text{between}}}$$

- **Within-group variability** (Mean Square Within, MSW):

$$MSW = \frac{SSW}{df_{\text{within}}}$$

Where:

- SSB = Sum of squares between groups
- SSW = Sum of squares within groups
- df_{between} = Degrees of freedom between groups
- df_{within} = Degrees of freedom within groups

Calculate F-Test Statistics:

$$F = \frac{\text{Variance between Sample}}{\text{Variance within Sample.}}$$

$SS_{Between}$= some of squares

$$SS_{between} = \sum \frac{\left(\sum a_i\right)^2}{n} - \frac{T^2}{N}$$

$$= \frac{57^2 + 47^2 + 21^2}{7} - \frac{[57 + 47 + 21]}{21}$$

=98.67

SS_{within}=

$$SS_{within} = \sum y^2 - \frac{\sum(\sum a_1)^2}{n}$$

$$\sum y^2 = 9^2 + 8^2 + 7^2 + 8^2 + 8^2 + 9^2 + \cdots$$

$$= 853$$

SS_{within}=10.29

SS_{Total}= 98.67 + 10.29 = 108.45

	SS	df	MS	F
Between	98.67	2	49.34	
Within	10.29.	18	0.54	
Total	108.95	20.		

MS= SS/df

$$F = \frac{MS_{Between}}{MS_{Within}} = \frac{49.34}{0.54} = 86.56$$

86.56>3.56

Therefore, we reject the H0.

Hence, there is a difference and we are 95% confident about it.

Example 2:

Low Noise			Medium Noise			Loud Noise		
Student	Questions (X)		Student	Questions (X)		Student	Questions (X)	
1	10		5	8		9	4	
2	9		6	4		10	3	
3	6		7	6		11	6	
4	7		8	7		12	4	

Low Noise			Medium Noise			Loud Noise		
Student	Questions (X)	X^2	Student	Questions (X)	X^2	Student	Questions (X)	X^2
1	10	100	5	8	64	9	4	16
2	9	81	6	4	16	10	3	9
3	6	36	7	6	36	11	6	36
4	7	49	8	7	49	12	4	16
	$\sum X_1 = 32$	$\sum X_1^2 = \mathbf{266}$		$\sum X_2 = 25$	$\sum X_2^2 = \mathbf{165}$		$\sum X_3 = 17$	$\sum X_3^2 = \mathbf{77}$

$\sum X_1 = 32$	$\sum X_1^2 = \mathbf{266}$	$\sum X_2 = 25$	$\sum X_2^2 = \mathbf{165}$	$\sum X_3 = 17$	$\sum X_3^2 = \mathbf{77}$

Correction Term:

$$C_x = \frac{(\sum X)^2}{N} = \frac{(32+25+17)^2}{12} = \frac{(74)^2}{12} = \frac{5476}{12} = 456.33$$

Sum of Squares of Total:

$$SS_T = \sum X^2 - C_x = (266+165+77) - 456.33$$
$$= 508 - 456.33 = 51.67$$

Sum of Squares Among groups:

$$SS_A = \frac{(\sum X)^2}{n} - C_x = \left(\frac{32^2}{4} + \frac{25^2}{4} + \frac{17^2}{4}\right) - 456.33$$
$$= 484.5 - 456.33 = 28.17$$

Sum of Squares Within groups:

$$SS_W = SS_T - SS_A$$
$$= 51.67 - 28.17 = 23.5$$

Mean of Sum of Squares Among groups:

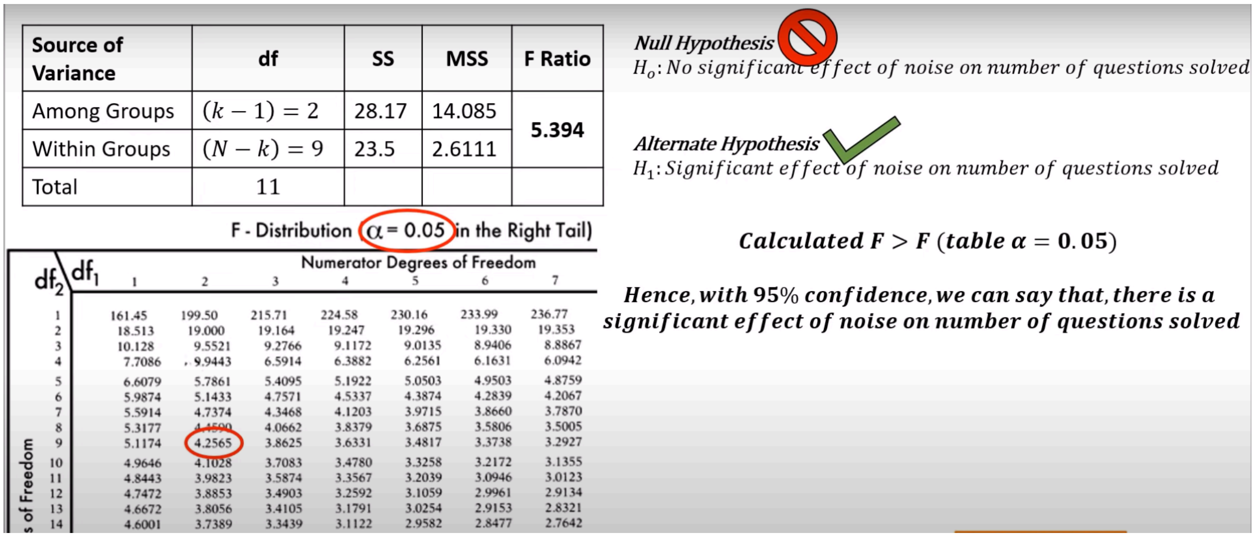
$$MSS_A = \frac{SS_A}{k-1} = \frac{28.17}{3-1} = 14.085$$

Mean of Sum of Squares Within groups:

$$MSS_W = \frac{SS_W}{N-k} = \frac{23.5}{12-3} = 2.611$$

F Ratio:

$$F\ Ratio = \frac{MSS_A}{MSS_W} = \frac{14.085}{2.611} = \mathbf{5.394}$$



Source of Variance	df	SS	MSS	F Ratio
Among Groups	$(k-1)=2$	28.17	14.085	5.394
Within Groups	$(N-k)=9$	23.5	2.6111	
Total	11			

F - Distribution ($\alpha = 0.05$ in the Right Tail)

Numerator Degrees of Freedom

$df_2 \backslash df_1$	1	2	3	4	5	6	7
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867
4	7.7086	9.9443	6.5914	6.3882	6.2561	6.1631	6.0942
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642

Null Hypothesis
$H_o: No\ significant\ effect\ of\ noise\ on\ number\ of\ questions\ solved$

Alternate Hypothesis
$H_1: Significant\ effect\ of\ noise\ on\ number\ of\ questions\ solved$

$\textbf{Calculated } F > F\ (table\ \alpha = 0.05)$

Hence, with 95% confidence, we can say that, there is a significant effect of noise on number of questions solved

Python Code

1-Way ANOVA

```
from scipy import stats

# Sample data for three groups
group1 = [25, 30, 35, 40, 45]  # Teaching Method A
group2 = [30, 35, 40, 45, 50]  # Teaching Method B
group3 = [35, 40, 45, 50, 55]  # Teaching Method C

# Perform One-Way ANOVA
f_stat, p_value = stats.f_oneway(group1, group2, group3)

print(f"F-statistic: {f_stat:.4f}")
print(f"P-value: {p_value:.4f}")

# Interpret the result
alpha = 0.05
if p_value < alpha:
    print("Reject H0: At least one group mean is significantly different.")
```

```
else:
    print("Fail to reject H₀: No significant difference between group means.")
```

Output:

F-statistic: 2.0000
P-value: 0.1780
Fail to reject H_0: No significant difference between group means.

Two-Way ANOVA Python Code

```
# Sample data
data = pd.DataFrame({
    "Score": [85, 88, 90, 86, 87, 89, 80, 82, 84, 78, 80, 81],
    "Method": ["Online"] * 6 + ["Offline"] * 6,
    "Gender": ["Male", "Male", "Male", "Female", "Female", "Female",
               "Male", "Male", "Male", "Female", "Female", "Female"]
```

	Score	Method	Gender
0	85	Online	Male
1	88	Online	Male
2	90	Online	Male
3	86	Online	Female
4	87	Online	Female
5	89	Online	Female
6	80	Offline	Male
7	82	Offline	Male
8	84	Offline	Male
9	78	Offline	Female
10	80	Offline	Female
11	81	Offline	Female

```
from statsmodels.formula.api import ols
import statsmodels.api as sm
# Two-Way ANOVA
model = ols('Score ~ C(Method) + C(Gender) + C(Method):C(Gender)', data=dat
a).fit()
```

`ols` : Stands for **Ordinary Least Squares**, which helps fit the ANOVA model.

`import statsmodels.api as sm`

This imports the **Statsmodels** library, which is used for statistical modeling, hypothesis testing, and data analysis.

- `api` is a module in Statsmodels that provides access to key functions for regression, ANOVA, and other statistical tests.

`'Score ~ C(Method) + C(Gender) + C(Method):C(Gender)'` :

`C(Method)` : Tests if **teaching method** affects scores.

`C(Gender)` : Tests if **gender** affects scores.

`C(Method):C(Gender)` : Checks **interaction effect** (e.g., Does "Online" work better for males?).

`.fit() :` Fits the ANOVA model to the data by estimating coefficients and calculating F-values & p-values.

  - Essentially "trains" the model using the given data.

```
anova_table = sm.stats.anova_lm(model, typ=2)
print(anova_table)
```

```
                     sum_sq   df          F    PR(>F)
C(Method)        133.333333  1.0  35.555556  0.000337
C(Gender)          5.333333  1.0   1.422222  0.267207
C(Method):C(Gender)  3.000000  1.0   0.800000  0.397204
Residual          30.000000  8.0        NaN       NaN
```

`typ=2` is preferred for **balanced or slightly unbalanced designs**.

Term	Sum of Squares (SS)	df	F-Value	p-value (PR(>F))
C(Method)	133.33	1	35.56	0.000337
C(Gender)	5.33	1	1.42	0.267207
C(Method):C(Gender)	3	1	0.8	0.397204
Residual (Error)	30	8	NaN	NaN

C(Method):

- **F = 35.56**, p-value = **0.000337 (< 0.05)**
- **Conclusion:** The teaching method has a **significant effect** on scores. Different teaching methods lead to different results.

C(Gender):

- **F = 1.42**, **p-value = 0.267207 (> 0.05)**
- **Conclusion:** Gender **does not** have a significant impact on scores. Male and female students perform similarly.

C(Method):C(Gender) (Interaction Effect):

- **F = 0.80**, **p-value = 0.397204 (> 0.05)**
- **Conclusion:** There is **no significant interaction** between teaching method and gender. This means the effectiveness of a teaching method is **not different** between males and females.

Post-Hoc Tests

If ANOVA rejects H_0, you can perform **post-hoc tests** to identify which specific groups differ. Common post-hoc tests include:

- **Tukey's HSD Test**: Compares all possible pairs of group means.
- **Bonferroni Correction**: Adjusts the significance level for multiple comparisons.