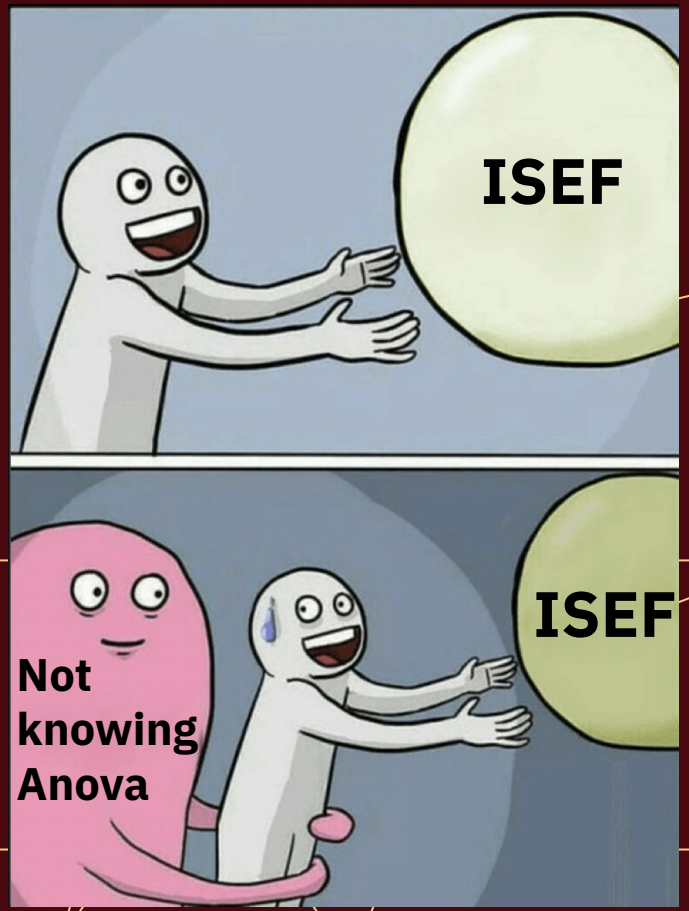
ANOVA

Analysis of Variance

Mohammad Golji ‘27
Sam Wheatley ‘27



Overview:

1. Introduction
2. What is ANOVA?
3. ANOVA Components
4. Example Test
5. ANOVA-Calculator
6. ANOVA-Excel

Introduction

What we know
... and what we don't

- **Hypothesis testing compares variables to draw a conclusion based on probability**
- **Z-tests** (prop. vocab filled before last day of term)
- **T-tests** (mean viruses on Arnav's computer)
- **Chi^2** (distr. of ST#M grades)
- These can't do multiple populations with different variables
- This is crucial for *ST*M*!

Objective

Learn what ANOVA is and be able to apply it

What is it used for

- To test if group means differ across categories
- Can be categorized into two ways:
  - One way - test differences across 1 factor
  - Two way - test difference across 2 factors + their interaction
- Example: “Tanay asks how different test fees and different prep fees affect SAT scores?”

Conditions and Assumptions

Both need:

- Independence of samples
  - Each sample is independent
- Normality of residuals
  - After computation, the leftover errors should be normally distributed
- Homogeneity of variance
  - Variance between groups should be similar

One-way ANOVA:

- Assumptions apply across groups of one factor

Two-way Anova:

- Assumptions apply across each cell

Test statistic

One way:

Purpose: to compare 3 or more group means within ONE categorical factor

Test statistic: f-ratio

Formula:

F=MS(within)/MS(between)

What it measures:

- Between group variation (how far the group means differ)
- Within-group variation (variability within groups)

How to interpret:

- Large F = the groups differ more then expected by change
- Small F = differences are likely due to randomness

Two way:

Purpose: to test two categorical factors at once

Test Statistic: (3) F-statistics:

- Main effect A
- Main effect B
- Interaction of A and B

Test formula: (use for each one)

F=MS(effect)/MS(error)

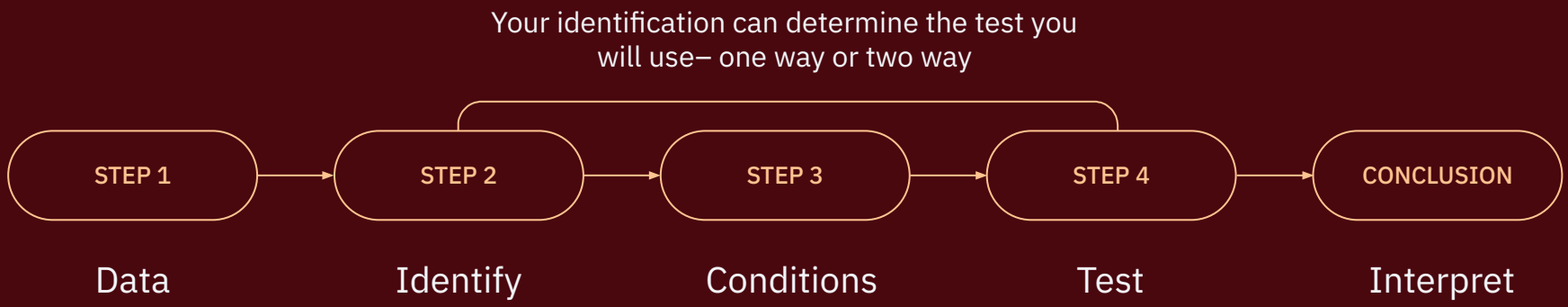
what it measures:

- Factor A effect sees if A changes the mean
- Factor B effect sees if B changes the mean
- Interaction sees if the effect of A deped on B

Interpretation:

- Large F = strong evidence the effect is real
- Small F = no meaningful effect

Process Summary

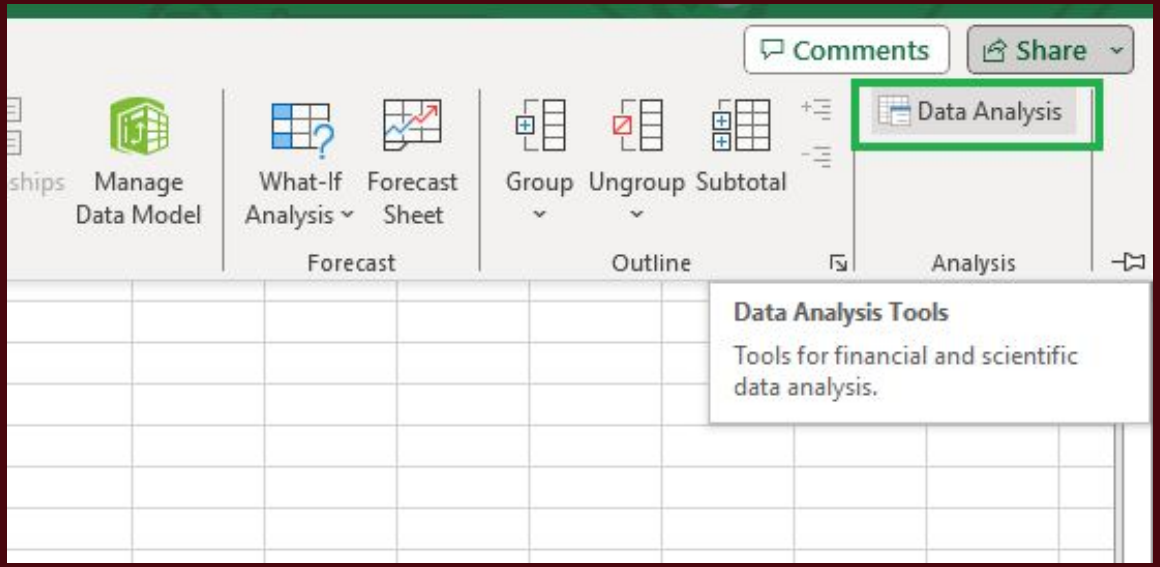


Example: Weekly Crashout Rates

	2025	2026	2027
Weekly crashout rates (crash outs/week for 10 samples students)	0.2	1.2	2.1
	1.8	1.1	0.9
	2.2	0.7	1.4
	0	1	1.2
	0	1.9	1.15
	0.4	0.6	0
	0.34	0.7	0.6
	0.7	0	0.9
	1	0.2	0.8
	0.07	0.9	3.8

Example: Excel.

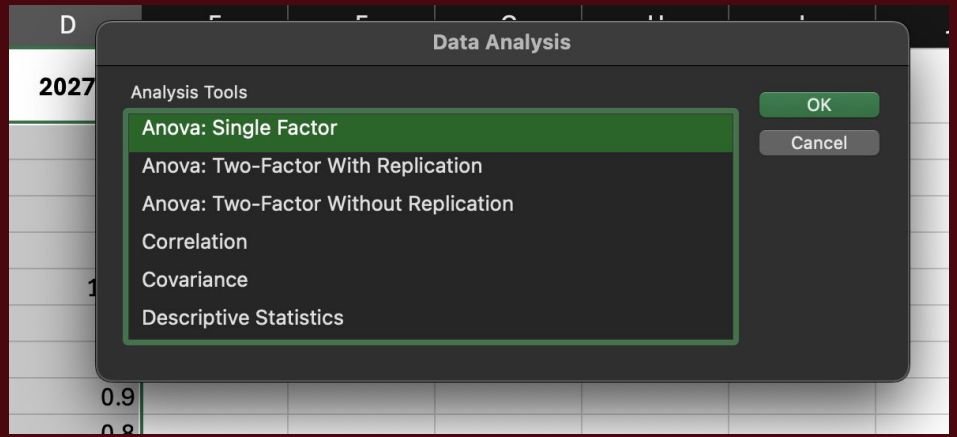
Make sure you have Data Analysis Installed:



Example: Excel.

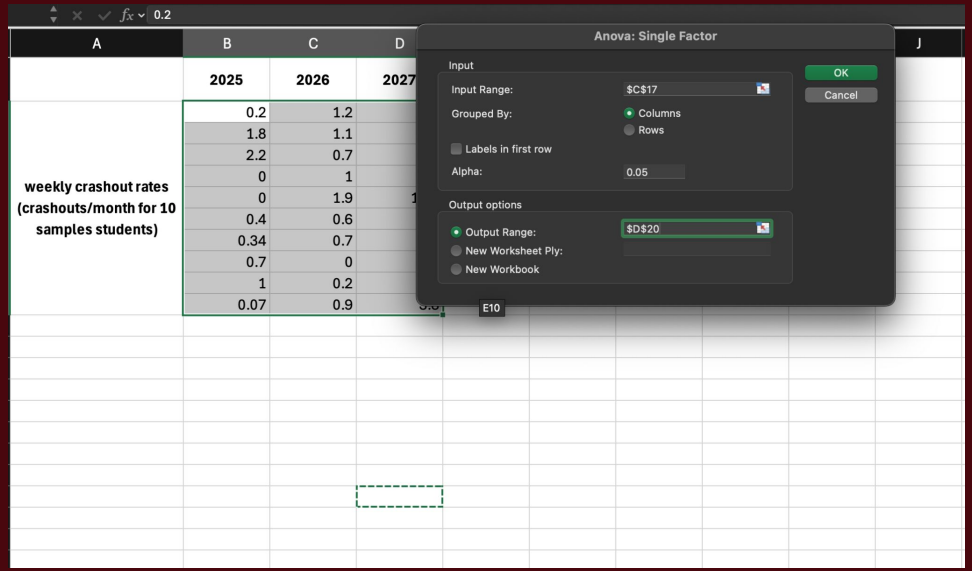
Select your data and click on Data Analysis. Select Anova: Single Factor:

A	B	C	D
	2025	2026	2027
weekly crashout rates (crashouts/month for 10 samples students)	0.2	1.2	2.1
	1.8	1.1	0.9
	2.2	0.7	1.4
	0	1	1.2
	0	1.9	1.15
	0.4	0.6	0
	0.34	0.7	0.6
	0.7	0	0.9
	1	0.2	0.8
	0.07	0.9	3.8



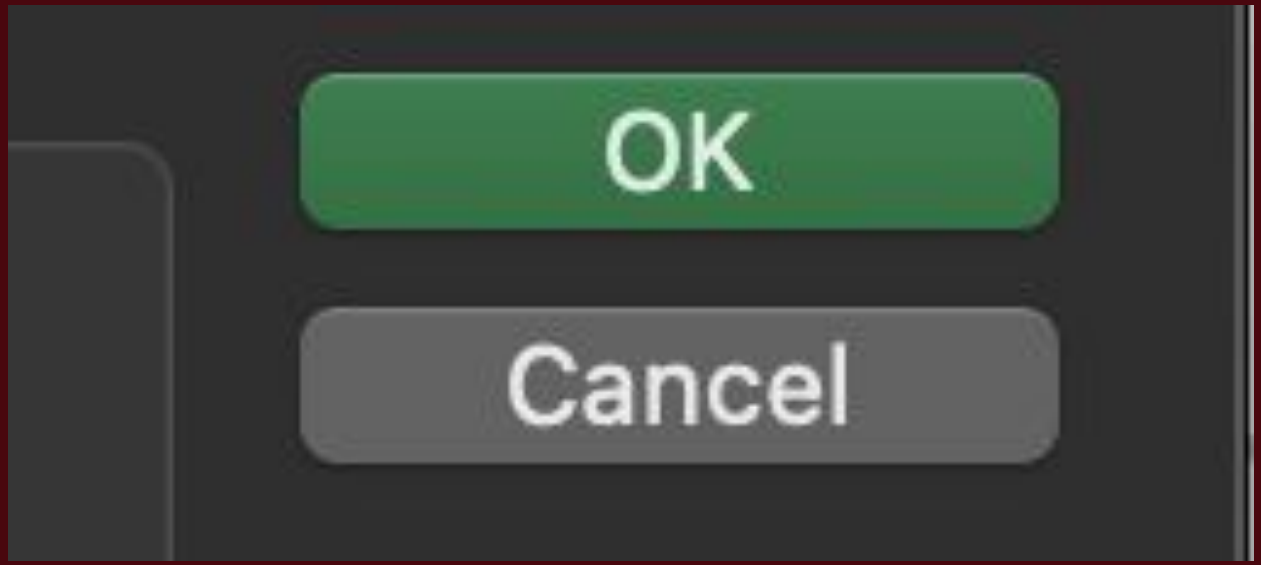
Example: Excel.

Select Columns, Output, and Data (if not already selected):



Example: Excel.

Hardest Step: Click OK



Example: Excel.

Take a look at your results and Interpret

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Column 1	10	6.71	0.671	0.59867667
Column 2	10	8.3	0.83	0.28455556
Column 3	10	12.85	1.285	1.07558333

ANOVA

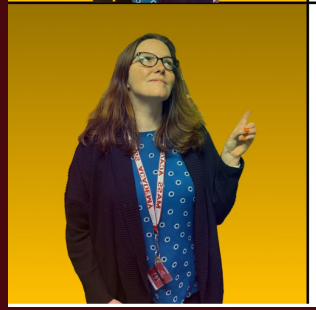
Source of Variati	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Gro	2.03100667	2	1.01550333	1.55528171	0.22945943	3.35413083
Within Group	17.62934	27	0.65293852			
Total	19.6603467	29				

Example: Calculator.

Take out your calculators!

	2025	2026	2027
Weekly crashout rates (crash outs/week for 10 samples students)	0.2	1.2	2.1
	1.8	1.1	0.9
	2.2	0.7	1.4
	0	1	1.2
	0	1.9	1.15
	0.4	0.6	0
	0.34	0.7	0.6
	0.7	0	0.9
	1	0.2	0.8
	0.07	0.9	3.8

When the example is just on the slide



When we do it in person with a real calculator