

# ¿QUE ES ANOVA?

Aansh, Anh, Charlotte, Noah

# What is ANOVA

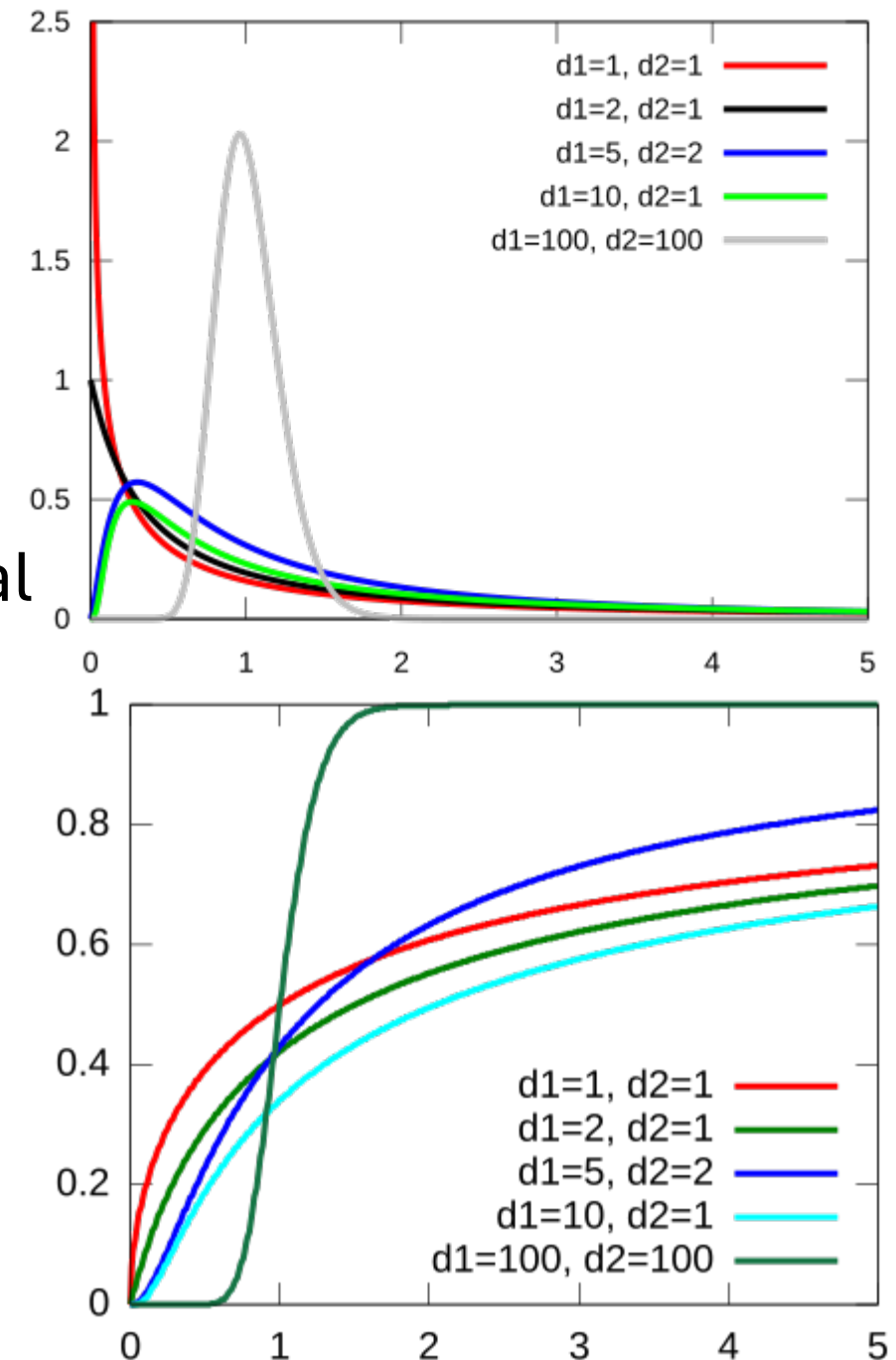
- Anova (Analysis of variance) is a method to compare 3 or more means
- Independent variable is categorical while dependent is continuous
- One way Anova: Used for when only one categorical factor affects the variable of interest. Ex: Effect of type of water on plant growth.
- Two way Anova: When there are 2 independent variables.
- Ex: Saltwater vs freshwater and to be watered in the morning or afternoon on plant growth

# What does it calculate

- Variability between groups / degrees of freedom between groups = MSB
- Variability within groups / degrees of freedom within groups = MSW
- F ratio =  $MSB/MSW$
- When you do an ANOVA test a F statistic is calculated for each hypothesis.

# ANOVA Hypothesis

- The population must be close to normal
  - Samples must be independent
  - Population variances must be roughly equal
  - Groups must have equal sample sizes
- 
- $H_0$ : all group means are equal
  - $H_a$ : at least one group means differs
  - P-value found in f-distribution
  - top probability, bottom cumulative



The image shows three damselfly nymphs, which are aquatic larvae, perched on their mud nests. The nymphs are green with large, prominent eyes and long antennae. The mud nests are brown and have a textured, layered appearance. The background is a soft, out-of-focus green, suggesting a natural aquatic environment. The text "Differences between 1 and 2 way" is overlaid in the center in a white, sans-serif font, with a white horizontal line underneath it.

# Differences between 1 and 2 way

---

# One way ANOVA Excel

Go to Data, and then click on data analysis

Select "ANOVA: Single Factor"

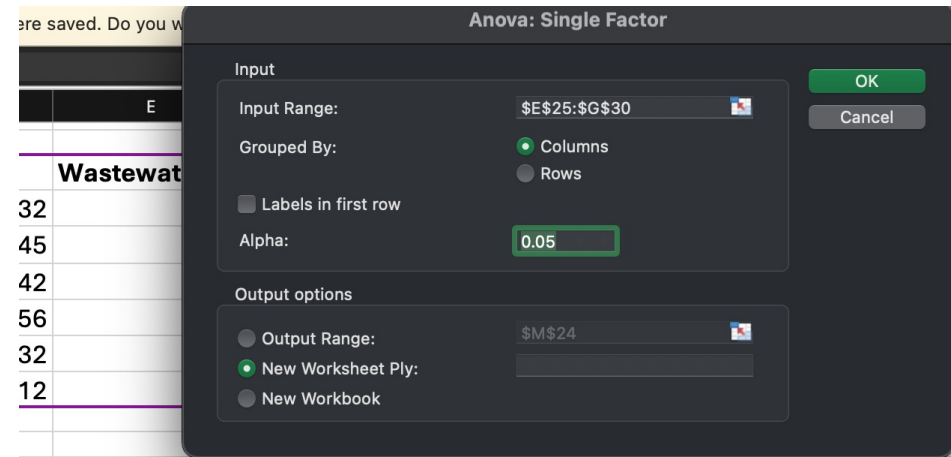
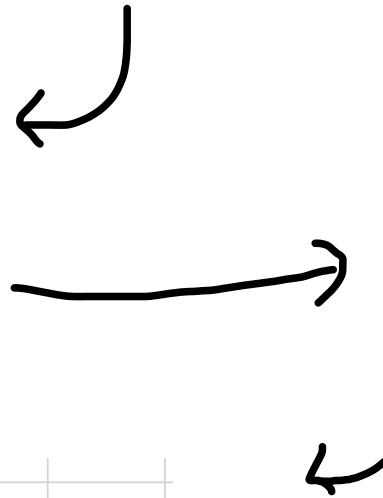
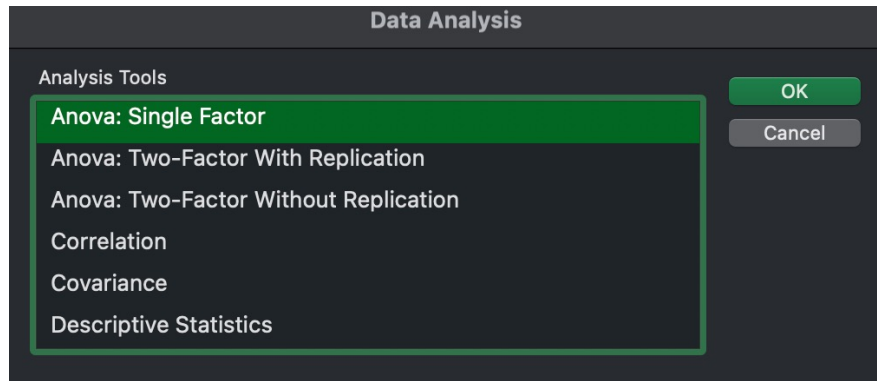
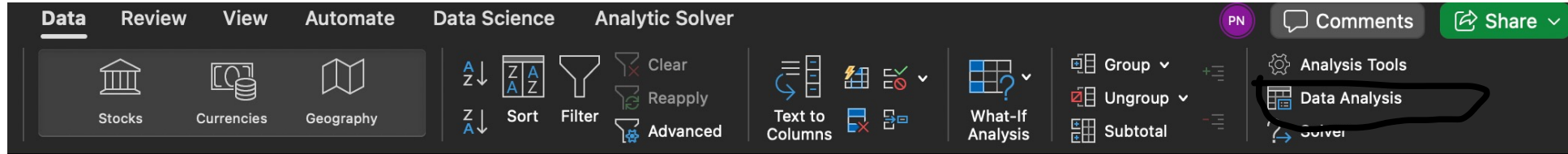
Type the number of rows

Highlight Desired Data (Do not highlight the categories of the data)  
to fill the value for the input range

Press ok and the result will be in a new sheet

The p value depends on if your data is the column or the row

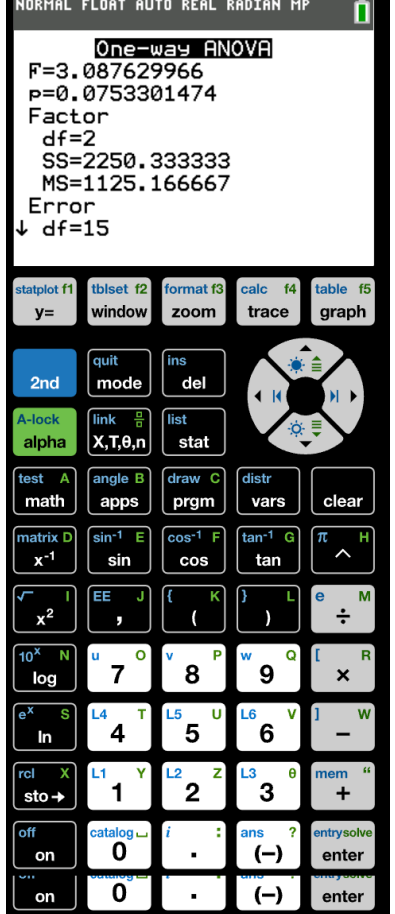
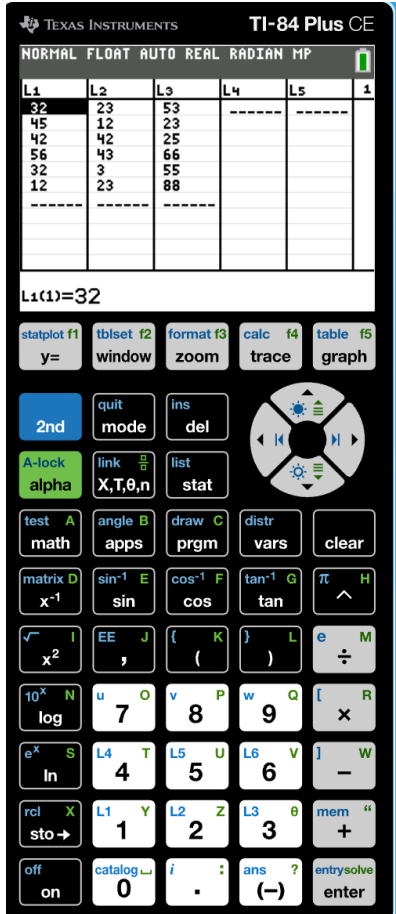
# One way ANOVA Excel



Anova: Single Factor						
SUMMARY						
Groups	Count	Sum	Average	Variance		
Column 1	6	219	36.5	224.7		
Column 2	6	146	24.33333333	254.266667		
Column 3	6	310	51.66666667	614.266667		
ANOVA						
Source of Variati	SS	df	MS	F	P-value	F crit
Between Gro	2250.333333	2	1125.16667	3.08762997	0.07533015	3.68232034
Within Group	5466.16667	15	364.411111			
Total	7716.5	17				

	Salt water	Wastewater	Freshwater
	32	23	53
	45	12	23
	42	42	25
	56	43	66
	32	3	55
	12	23	88

# Calculator Commands for 1 way ANOVA





# Excel commands for one way

Effects of types of water (Salt, fresh, wastewater) and the states (Oregon, Wyoming) in which the plant is grown on plant growth

*Fabricated Data:*

Plant growth (km)

	<b>Salt water</b>	<b>Wastewater</b>	<b>Freshwater</b>
<b>Oregon</b>	32	23	53
	45	12	23
	42	42	25
<b>Wyoming</b>	56	43	66
	32	3	55
	12	23	88

# Excel Commands for 2 way ANOVA

Highlight Desired Data

Go to Data, and then click on data analysis

Select "ANOVA: Two-Factor With Replication"

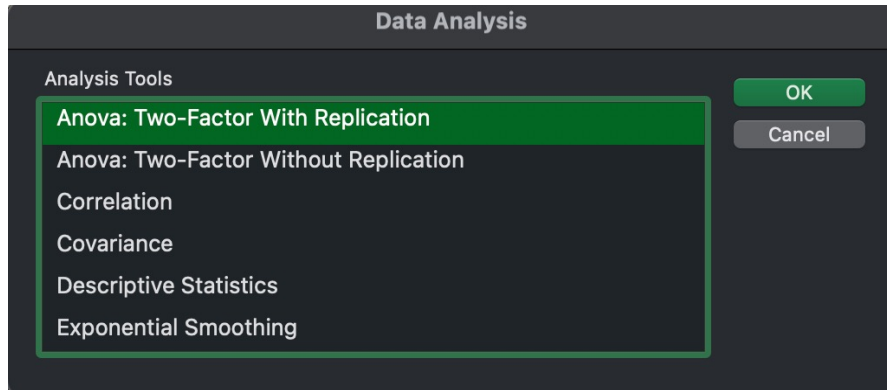
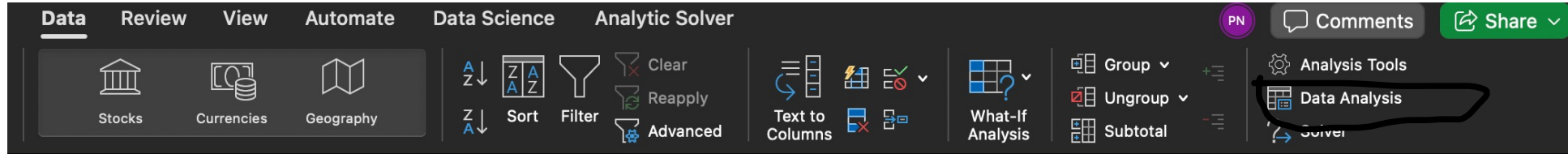
Highlight Desired Data (Highlight the categories of the data) to fill the value for the input range

Type the number of rows for the category

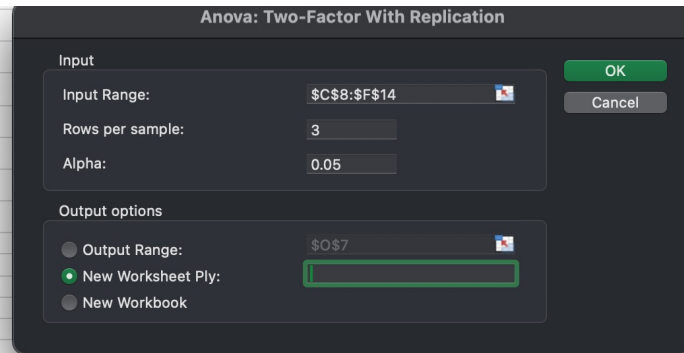
Press ok and the result will be in a new sheet

The p value depends on if your data is the column or the row

# Two way ANOVA Excel



	Salt water	Wastewater	Freshwater
<b>Oregon</b>	32	23	53
	45	12	23
	42	42	25
<b>Wyoming</b>	56	43	66
	32	3	55
	12	23	88



Anova: Two-Factor With Replication				
SUMMARY	Salt water	Wastewater	Freshwater	Total
<i>Oregon</i>				
Count	3	3	3	9
Sum	119	77	101	297
Average	39.6666667	25.6666667	33.6666667	33
Variance	46.3333333	230.333333	281.333333	176.5
<i>Wyoming</i>				
Count	3	3	3	9
Sum	100	69	209	378
Average	33.3333333	23	69.6666667	42
Variance	485.333333	400	282.333333	742.5
<i>Total</i>				
Count	6	6	6	
Sum	219	146	310	
Average	36.5	24.3333333	51.6666667	
Variance	224.7	254.266667	614.266667	

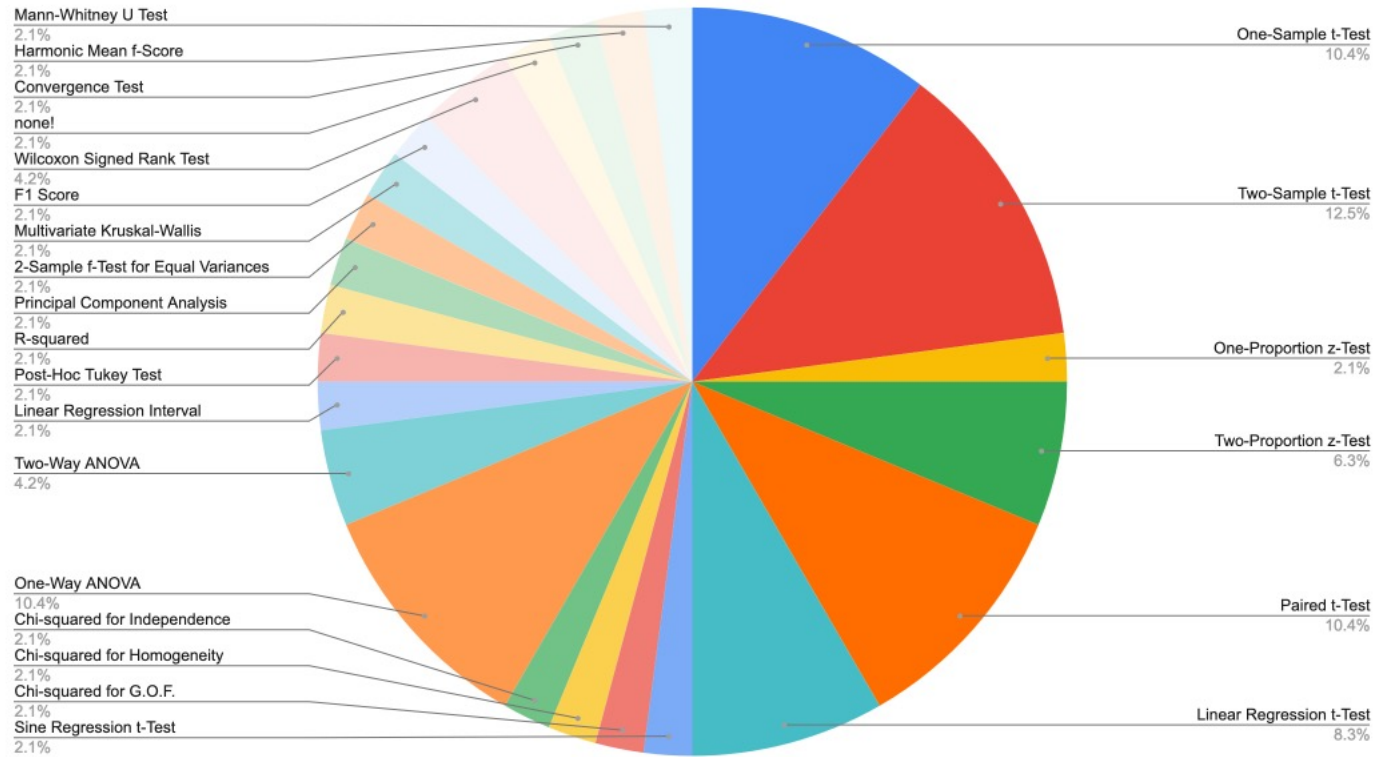
ANOVA						
Source of Variati	SS	df	MS	F	P-value	F crit
Sample	364.5	1	364.5	1.2673363	0.28227849	4.74722535
Columns	2250.33333	2	1125.16667	3.91211126	0.04919381	3.88529383
Interaction	1650.33333	2	825.16667	2.86903612	0.09586255	3.88529383
Within	3451.33333	12	287.611111			
<b>Total</b>	<b>7716.5</b>	<b>17</b>				

# P-value interpretation

- Sample: Your rows, so the categorical on the side of the matrix independent from each other or not.
- Columns: Your columns, so the categorical on the top or bottom of the matrix independent from each other or not.
- Interaction : If your columns and rows are independent from each other or not

# Tests Used Last Year

Types of Statistical Testing Used by MAMS Seniors During Their IRPs



# Conclusion

Reject the null hypothesis if  $p \text{ value} < \alpha \text{ level}$

- Accept the alternative hypotheses that is at least one of the mean is different due to significant statistical evidence.

Fail to reject the null hypothesis if  $p \text{ value} > \alpha \text{ level}$

- Do not have convincing statistical evidence for the alternative.
- Do tests after to compare categories to find which one is significantly different

# Reference

<https://www.investopedia.com/terms/a/anova.asp>

<https://education.ti.com/en/customer-support/knowledge-base/ti-83-84-plus-family/product-usage/34611>

<https://www.statology.org/two-way-anova-excel/>

<https://en.wikipedia.org/wiki/F-distribution> (the goat)