https://docs.google.com/document/d/1BJNnzpc5tBMBTTbPrTxIX3luLelriHijEJ6kHcGXKZY/edit?tab=t.0

To identify whether a problem involves permutations or combinations, consider the following key points:

---

1. Order Matters?

- **Permutations**: If the order of selection matters, it's a permutation problem.
  Example: Arranging 3 books on a shelf (ABC is different from CAB).
  Formula: P(n,r)=n!(n−r)!P(n, r) = \frac{n!}{(n-r)!}P(n,r)=(n−r)!n!
- **Combinations**: If the order does *not* matter, it's a combination problem.
  Example: Choosing 3 books out of 5 (ABC is the same as CAB).
  Formula: C(n,r)=n!r!(n−r)!C(n, r) = \frac{n!}{r!(n-r)!}C(n,r)=r!(n−r)!n!

---

2. Key Words

- **Permutations**: Look for terms like "arrange," "order," "sequence," or "rank."
  Example: "How many ways can you arrange 3 students in a line?"
- **Combinations**: Look for terms like "choose," "select," "pick," or "group."
  Example: "How many ways can you choose 3 team members from a group of 10?"

---

3. Real-World Scenarios

- **Permutations**:
  Situations where arrangement or order is essential.
  - Assigning roles (e.g., president, vice president, and secretary).
  - Creating passwords with distinct characters.
  - Scheduling tasks in a specific sequence.
- **Combinations**:
  Situations where only selection matters, not the order.
  - Forming teams or committees.
  - Picking lottery numbers.
  - Choosing ice cream flavors (e.g., a scoop of vanilla and chocolate is the same as chocolate and vanilla).

---

4. Example to Differentiate

- **Question 1 (Permutation)**: "In how many ways can 3 people sit in 3 chairs?"
  Order matters → Permutation → P(3,3)=6P(3, 3) = 6P(3,3)=6.

- **Question 2 (Combination)**: "How many ways can 3 people be selected from a group of 5?"
  Order doesn't matter → Combination → C(5,3)=10C(5, 3) = 10C(5,3)=10.

Qrome Yesterday at 11:23 AM
yes
tips for PnC:
- select is the keyword for Combination
- arrange is the keyword for Permutation

Permutation

Formula

$$_nP_r = \frac{n!}{(n-r)!}$$

$_nP_r$ = permutation
n = total number of objects
r = number of objects selected

Combination

Formula

$$ {}_nC_r = \frac{n!}{r!(n-r)!} $$

${}_nC_r$ = number of combinations

n = total number of objects in the set

r = number of choosing objects from the set

Geometric Mean

Formula

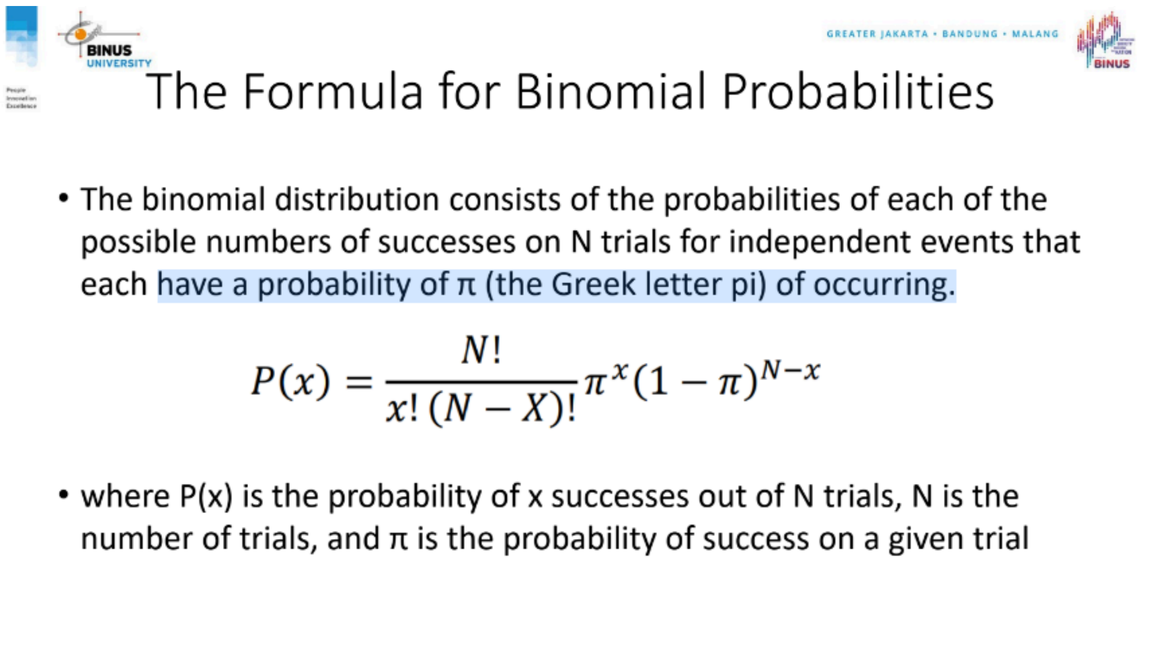
$$ \left(\prod_{i=1}^{n} x_i \right)^{\frac{1}{n}} = \sqrt[n]{x_1 x_2 \cdots x_n} $$

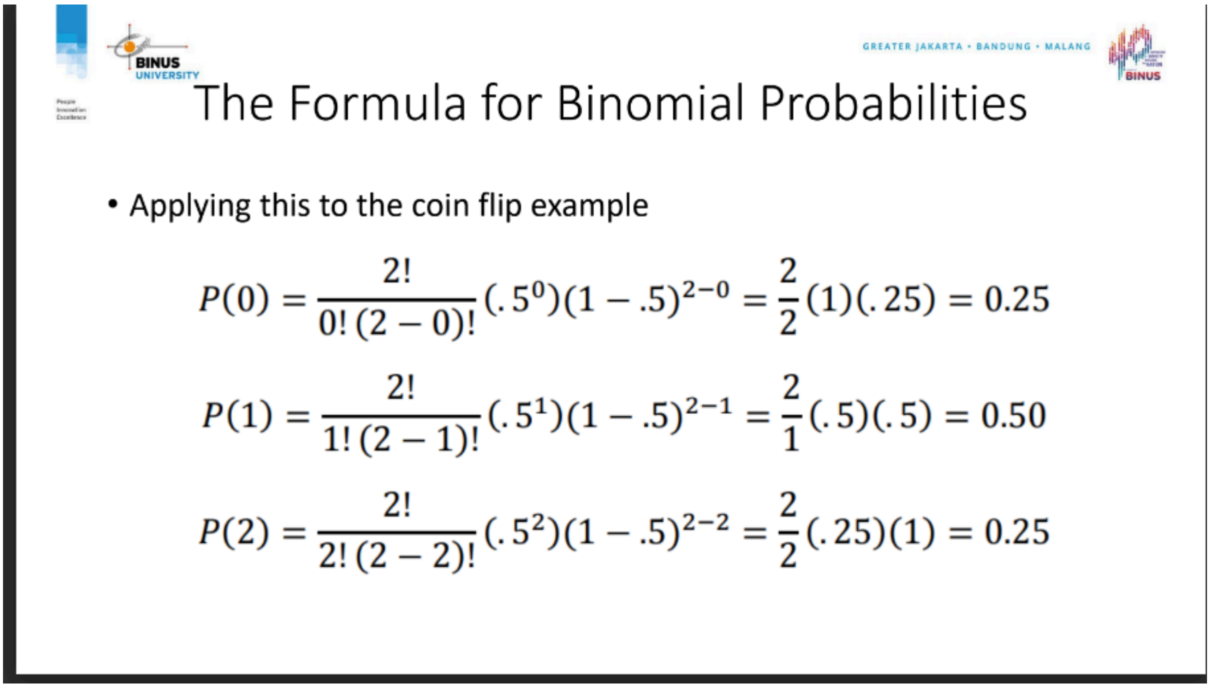
$\prod$ = geometric mean

n = number of values

x_i = values to average

Probability





This problem applies to the **binomial distribution**, which is used to model the number of successes in a fixed number of independent trials, where each trial has two possible outcomes (success or failure).

Characteristics of a Binomial Problem:

1. **Fixed Number of Trials (nnn)**:
   - In this case, there are 2 coin flips, so n=2n = 2n=2.
2. **Independent Trials**:

    - Each coin flip is independent, meaning the outcome of one flip doesn't affect the other.
3. **Two Outcomes per Trial**:
    - Each coin flip results in either heads (success) or tails (failure).
4. **Constant Probability of Success (ppp)**:
    - The probability of heads (success) is p=0.5p = 0.5p=0.5, and the probability of tails (failure) is 1−p=0.51 - p = 0.51−p=0.5.

Hypergeometric

The **hypergeometric distribution** is a specific type of probability problem where items are selected **without replacement** from a finite population, and the focus is on the probability of selecting a certain number of items with a specific characteristic.

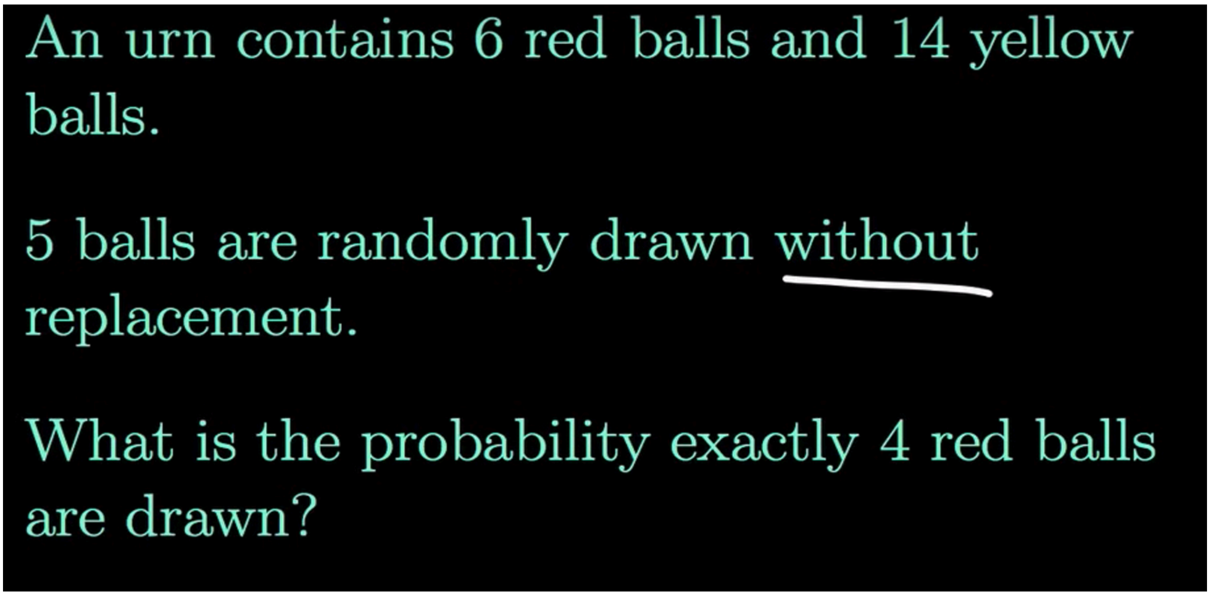
To identify if a problem involves the **hypergeometric distribution**, ask the following questions:

---

1. Key Features

- Is there a **finite population** divided into two distinct groups (e.g., "successes" and "failures")?
- Are items being drawn **without replacement** (i.e., once an item is chosen, it is not put back)?
- Is the goal to find the probability of a certain number of "successes" in the selection?

If the answer to all these questions is yes, it's a hypergeometric problem.



There are 6 red balls and 14 yellow balls

Any sample of 5 balls is equally likely, so:

$$P(\text{Exactly 4 red balls}) = \frac{\#\text{ of samples that result in 4 red balls and 1 yellow ball}}{\#\text{ of possible samples of size 5}}$$

$$= \frac{\binom{6}{4}\binom{14}{1}}{\binom{20}{5}} = \frac{15 \cdot 14}{15504} = 0.01354$$

Had the sampling been done *with* replacement, using the binomial distribution would be appropriate.

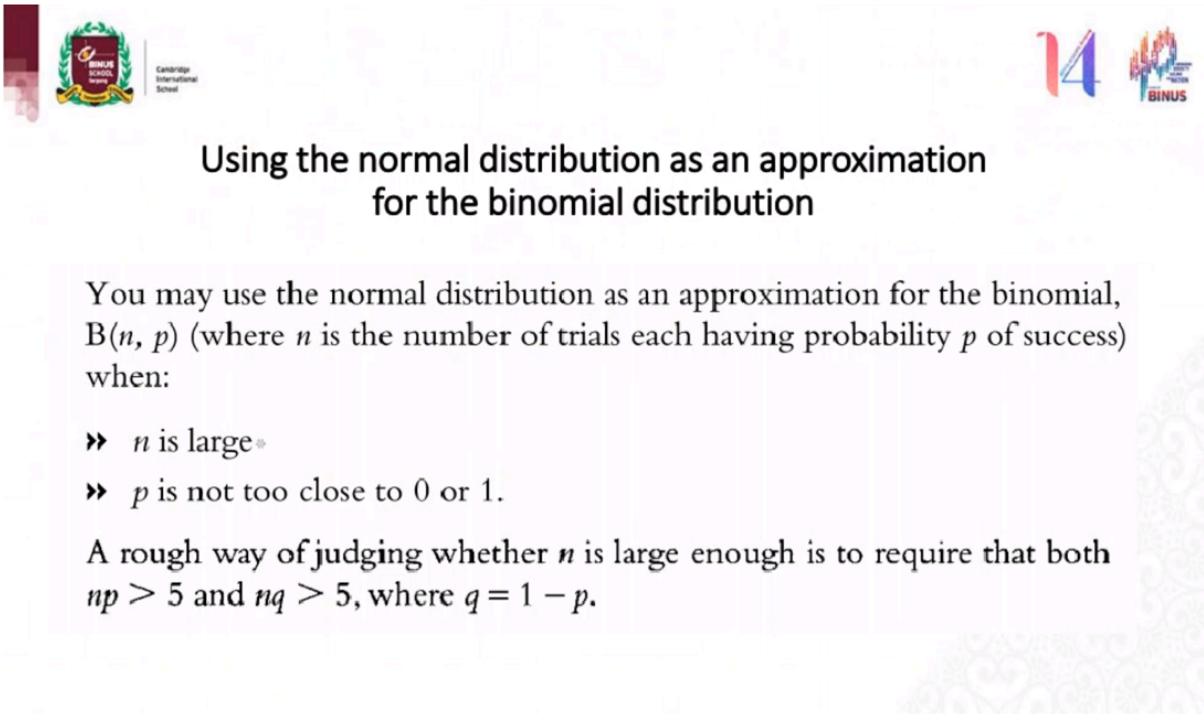
$$P(\text{4 red balls}) = \binom{5}{4}\left(\frac{6}{20}\right)^4\left(1 - \frac{6}{20}\right)^{5-4}$$

On any draw:

$$P(\text{Red}) = \frac{6}{20}$$

$$= 0.02835$$

Without replacement:
P(4 red) = 0.01354



tips for t test:

- one tailed t test -> determine if one population mean is greater than or less than another
- two tailed t test -> determine if the means are different

Pearson correlation

Hours (x)	Height (Y)
2	10
4	15
6	20
8	25
10	30

x	y	xy	x^2	y^2
2	10	20	4	100
4	15	60	16	225
6	20	120	36	400
8	25	200	64	625
10	30	300	100	900
30	100	700	220	2250

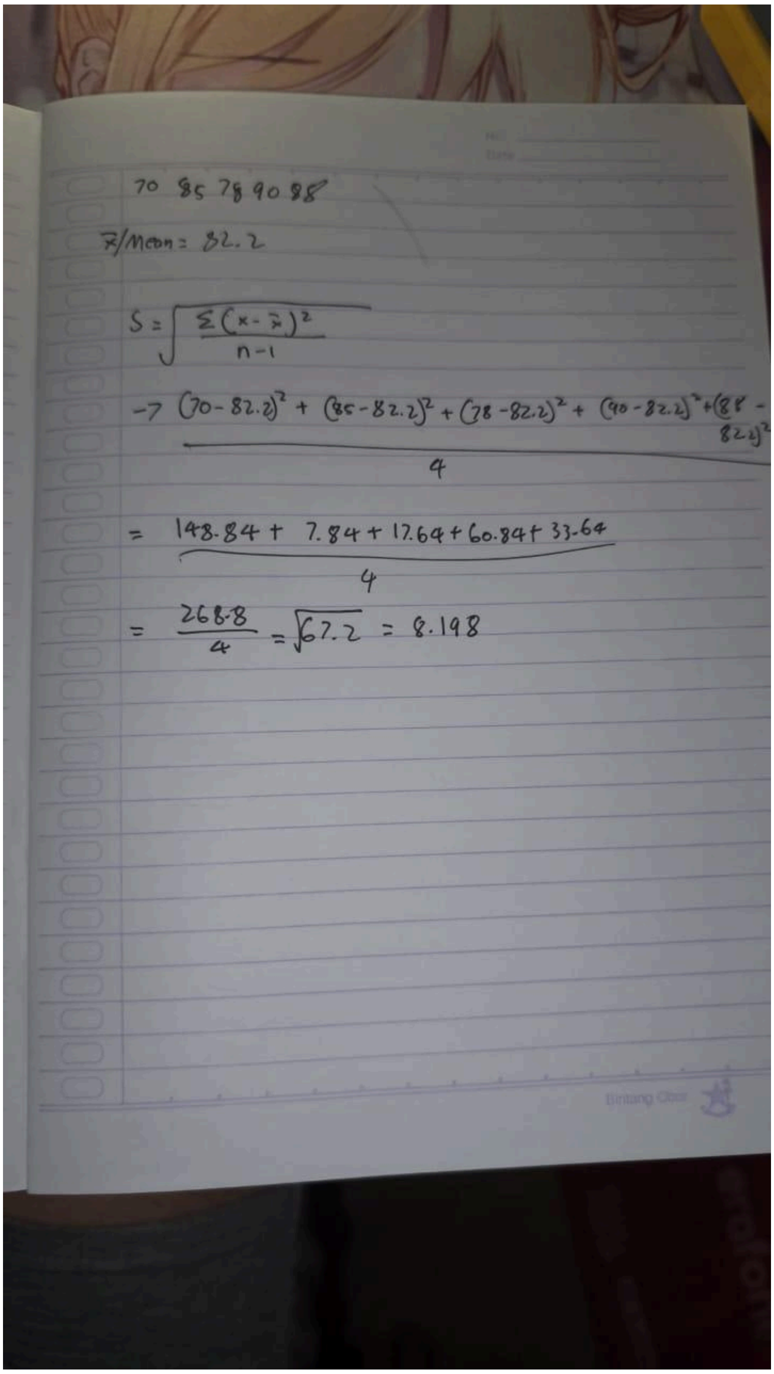
$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{\left[n\sum x^2 - (\sum x)^2\right]\left[n\sum y^2 - (\sum y)^2\right]}}$$

$$= \frac{5(700) - 30(100)}{\sqrt{\left[5(220) - (30)^2\right]\left[5(2250) - (100)^2\right]}}$$

$$= \frac{500}{\sqrt{(1100 - 900)(11250 - 10000)}} = \frac{500}{\sqrt{200(1250)}} = \frac{500}{500} = 1$$

Standard deviation

Sample method



70 85 78 90 88

$\bar{x}$/Mean = 82.2

$$S = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

$$\rightarrow \frac{(70-82.2)^2 + (85-82.2)^2 + (78-82.2)^2 + (90-82.2)^2 + (88-82.2)^2}{4}$$

$$= \frac{148.84 + 7.84 + 17.64 + 60.84 + 33.64}{4}$$

$$= \frac{268.8}{4} = \sqrt{67.2} = 8.198$$

Population method

Gatau liat bimay ex 5

Population Standard Deviation (σ\sigmaσ):

- **Formula:**
  σ=∑(xi−μ)2N\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}σ=N∑(xi−μ)2
  Where:
  - xix_ixi: Individual data points.
  - μ\muμ: Population mean.
  - NNN: Total number of data points in the population.
- **Key Characteristics:**
  - Used when you have data for the **entire population**.
  - Denominator is NNN, the total number of data points.
  - Provides the exact measure of variability within the population.

Sample Standard Deviation (sss):

- **Formula:**
  s=∑(xi−x‾)2n−1s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}s=n−1∑(xi−x‾)2
  Where:
  - xix_ixi: Individual data points.
  - x‾\bar{x}x‾: Sample mean.
  - nnn: Total number of data points in the sample.
- **Key Characteristics:**
  - Used when you have data for only a **subset (sample)** of the population.
  - Denominator is n−1n - 1n−1 (the sample size minus 1). This adjustment is called **Bessel's correction** and corrects for the bias in estimating the population variance from a sample.
  - Provides an unbiased estimate of the population standard deviation.

How to Differentiate:

1. **Look at the Dataset:**
   - If you are analyzing the **entire population**, use the **population formula**.
   - If you are analyzing a **sample** (subset of the population), use the **sample formula**.
2. **Denominator Difference:**
   - **Population formula** divides by NNN.
   - **Sample formula** divides by n−1n - 1n−1.
3. **Purpose:**
   - The **population standard deviation** measures the true variability in the population.
   - The **sample standard deviation** estimates the variability in the population based on the sample.

Example to Differentiate:

Population:

A population has 4 values: 2,4,6,82, 4, 6, 82,4,6,8.

- N=4N = 4N=4, μ=5\mu = 5μ=5.
- σ=(2−5)2+(4−5)2+(6−5)2+(8−5)24=184=2.12\sigma = \sqrt{\frac{(2-5)^2 + (4-5)^2 + (6-5)^2 + (8-5)^2}{4}} = \sqrt{\frac{18}{4}} = 2.12σ=4(2−5)2+(4−5)2+(6−5)2+(8−5)2=418=2.12.

Sample:

A sample from the population is 2,4,62, 4, 62,4,6.

- n=3n = 3n=3, x‾=4\bar{x} = 4x‾=4.
- s=(2−4)2+(4−4)2+(6−4)23−1=82=2.s = \sqrt{\frac{(2-4)^2 + (4-4)^2 + (6-4)^2}{3-1}} = \sqrt{\frac{8}{2}} = 2.s=3−1(2−4)2+(4−4)2+(6−4)2=28=2.

The **sample standard deviation** is slightly larger because of the smaller denominator (n−1n-1n−1). This accounts for the fact that we have less information about the population.

Z-table
If both n*q and n*p not >= 5, cannot use binomial

Z Distribution

$$z = \frac{x - \mu}{\sigma}$$

Binomial distribution

$n = 100$
$P = 0.3$
$q = 0.7$

Can be approximated with normal distribution if n is large & P is not too close with 0/1.

$n \cdot P \geq 5 \qquad n \cdot P = 100 \cdot 0.30 = 30$
$n \cdot q \geq 5 \qquad n \cdot q = 100 \cdot 0.7 = 70$

$\mu = n \cdot P \qquad \mu = 100 \times 0.30 = 30$
$\sigma = \sqrt{n \cdot P \cdot q} \qquad \sigma = \sqrt{100 \times 0.3 \times 0.7} = \sqrt{21} = 4.583$

Mean = 30, Sd = 4.58

Asked: $P(x < 25) \approx P(x \leq 24) \xrightarrow{\text{Continuity Correction}} P(x \leq 24.5)$

$x = 24.5$

$$z = \frac{24.5 - 30}{4.58} = \frac{-5.5}{4.58} = -1.20$$

	0
-1.2	.11507

Probability $0.11507 \times 100 = 11.507\% / 11.51\%$

$n = 100$
$p = 0.3$
$q = 0.7$

$$z = \frac{x - \mu}{\sigma}$$

$n \cdot p = 100 \times 0.3 = 30$ $\quad \mu = n \cdot p = 30$

$n \cdot q = 100 \times 0.7 = 70$ $\quad \sigma = \sqrt{n \cdot p \cdot q} = \sqrt{100 \cdot 0.3 \cdot 0.7} = \sqrt{21}$

$\cancel{\frac{100-30}{\sqrt{21}}} = \cancel{\frac{70}{\sqrt{21}}}$ Asked $\to P(X < 25) \approx P(X \leq 24)$
$= P(X \leq 24.5)$

$$\frac{24.5 - 30}{\sqrt{21}} = \frac{-5.5}{\sqrt{21}} = -1.20$$

	0
-1.2	.11507

$0.11507 \times 100 = 11.51\%$

T-test

Null hypothesis - nothing is going on

https://youtu.be/VekJxtk4BYM?si=W5dUPF345xkXNIfy

4. Look at real-world examples for context

- **One-tailed test:**
  - Testing if a new teaching method leads to higher test scores than the current method.
  - Checking if a new medication reduces blood pressure compared to a placebo.
- **Two-tailed test:**
  - Testing if there is any difference in test scores between two groups, without specifying the direction.
  - Checking if a medication has any effect on blood pressure (increase or decrease).

Summary Table

Aspect	One-Tailed Test	Two-Tailed Test
Hypothesis Direction	Specific (greater or less than)	Non-specific (different from)
Power	Higher (in one direction)	Lower (spread across both tails)
Applicability	Clear directional expectation	Uncertain or no directional expectation

---

$$t = \frac{\overset{\text{diff between mean?}}{\bar{x} - \mu}}{\underbrace{\frac{s}{\sqrt{n}}}_{\text{SD from the mean}}} \longrightarrow 1 \text{ Sample}$$

1 tail:

$$\frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$ Independent

2 tail:

→ mean value of diff

$$\frac{\bar{x}_d - 0}{\frac{s}{\sqrt{n}}}$$ → null hypo value

Dependent

2 tail

1 ~~950, 960, 970, 980~~ $a: 0.05$

Mean $= 990.5$

sd $= 25.87$

$$\frac{990.5 - 1000}{\frac{25.87}{\sqrt{10}}} = -1.16$$

units of freedom $= n-1 = 10-1 = 9$

	0.05
9	2.262

$-2.262 < -1.16 < 2.262$, fail to reject

1 tail

Diff = -25 $\quad a = 0.05$

$\mu = -25/8 \quad sd = 0.835$

$$t = \frac{d}{sd\sqrt{n}} = \frac{-25/8}{0.835/(\sqrt{8})} = \cancel{1.324}\ -10.585$$

	0.05
7	1.895

Unit of freedom = n-1 = 8-1 = 7

$-10.59 < -1.895 < 1.895$, reject

$a = 0.05$

$\bar{x}_1 = 8 \quad \sigma_1 = 2 \quad n_1 = 25$

$\bar{x}_2 = 6 \quad \sigma_2 = 2.5 \quad n_2 = 25$

Unit of freedom = 25+25-2 = 48

$$t = \frac{8-6}{\sqrt{\frac{2^2}{25} + \frac{2.5^2}{25}}} = \frac{2}{\frac{\sqrt{41}}{10}} = 3.12$$

	0.05
48	1.679

$3.12 > 1.679$, reject the null hypo

Chi square

Chi-square (χ^2)

Calculation of χ^2

Observed Values (O)	Expected Values (E)	$(O-E)$	$(O-E)^2$	$\frac{(O-E)^2}{E}$
18	11.7	6.3	39.69	3.39
36	27	9	81	3
21	25.2	-4.2	17.64	0.7
9	16.2	-7.2	51.84	3.2
6	9.9	-3.9	15.21	1.53
12	19.5	-7.5	56.25	2.88
36	45	-9	81	1.8
⋮	⋮	⋮	⋮	⋮
3	3.3	-0.3	0.09	0.02
				$\sum\frac{(O-E)^2}{E}$ $\chi^2 = 23.57$

$$\chi^2_{calculated} = 23.57$$

Table of Observed Values

Qualification / Marital Status	Middle School	High School	Bachelor's	Master's	Ph.D	Total
Never married	18	36	21	9	6	90
Married	12	36	45	36	21	150
Divorced	6	9	9	3	3	30
Widowed	3	9	9	6	3	30
Total	39	90	84	54	33	300

Table of Expected Values

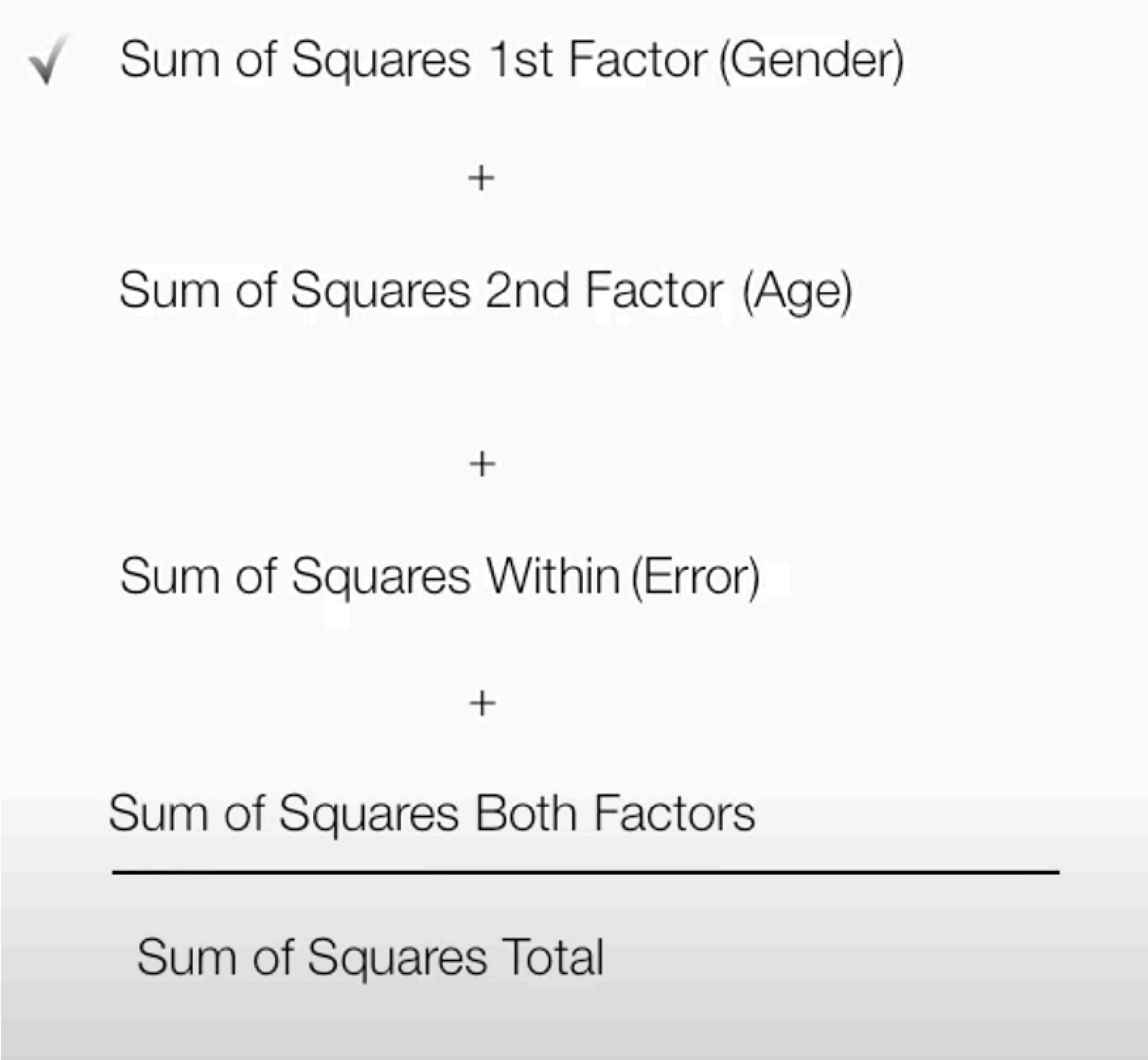
Qualification / Marital Status	Middle School	High School	Bachelor's	Master's	Ph.D
Never Married	$\frac{90 \times 39}{300} = 11.7$	$\frac{90 \times 90}{300} = 27$	25.2	16.2	9.9
Married	19.5	45	42	27	16.5
Divorced	3.9	9	8.4	5.4	3.3
Widowed	3.9	9	8.4	5.4	3.3

$df = n-1 = (5-1)(4-1) = 4 \times 3 = 12$

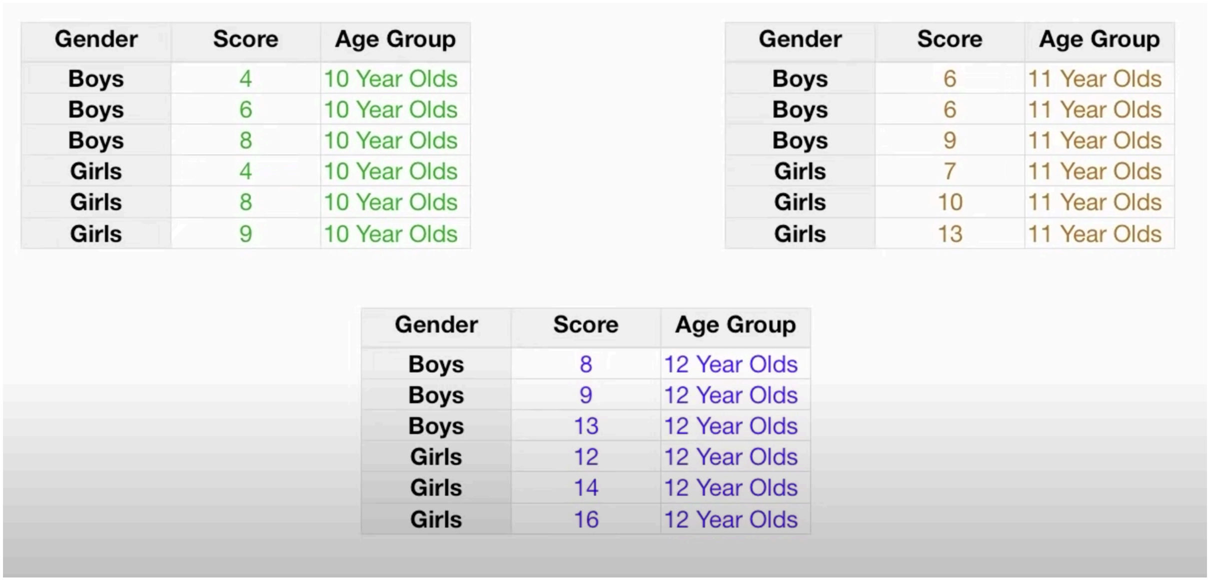
	0.05
12	21.03

23.57 > 21.03, reject the null hypo

2 way anova



Sort by age:

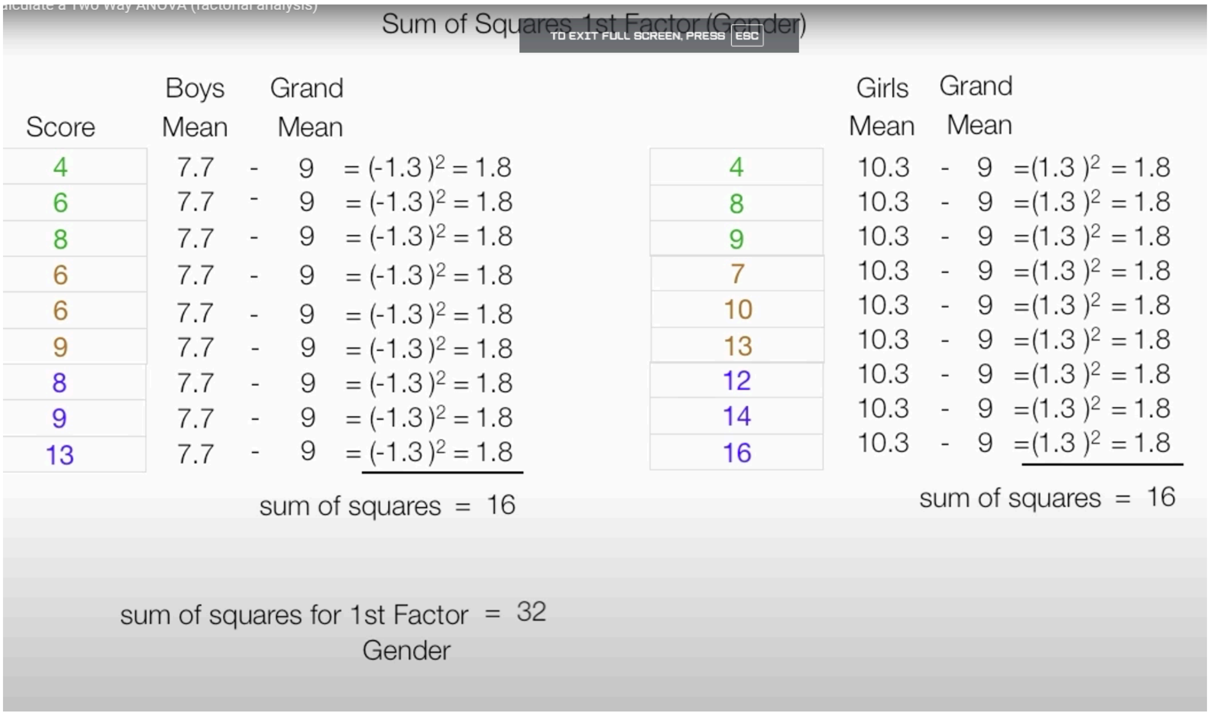


Gender	Score	Age Group
Boys	4	10 Year Olds
Boys	6	10 Year Olds
Boys	8	10 Year Olds
Girls	4	10 Year Olds
Girls	8	10 Year Olds
Girls	9	10 Year Olds

Gender	Score	Age Group
Boys	6	11 Year Olds
Boys	6	11 Year Olds
Boys	9	11 Year Olds
Girls	7	11 Year Olds
Girls	10	11 Year Olds
Girls	13	11 Year Olds

Gender	Score	Age Group
Boys	8	12 Year Olds
Boys	9	12 Year Olds
Boys	13	12 Year Olds
Girls	12	12 Year Olds
Girls	14	12 Year Olds
Girls	16	12 Year Olds

Sum of 1st



Score	Boys Mean		Grand Mean	
4	7.7	-	9	$= (-1.3)^2 = 1.8$
6	7.7	-	9	$= (-1.3)^2 = 1.8$
8	7.7	-	9	$= (-1.3)^2 = 1.8$
6	7.7	-	9	$= (-1.3)^2 = 1.8$
6	7.7	-	9	$= (-1.3)^2 = 1.8$
9	7.7	-	9	$= (-1.3)^2 = 1.8$
8	7.7	-	9	$= (-1.3)^2 = 1.8$
9	7.7	-	9	$= (-1.3)^2 = 1.8$
13	7.7	-	9	$= (-1.3)^2 = 1.8$

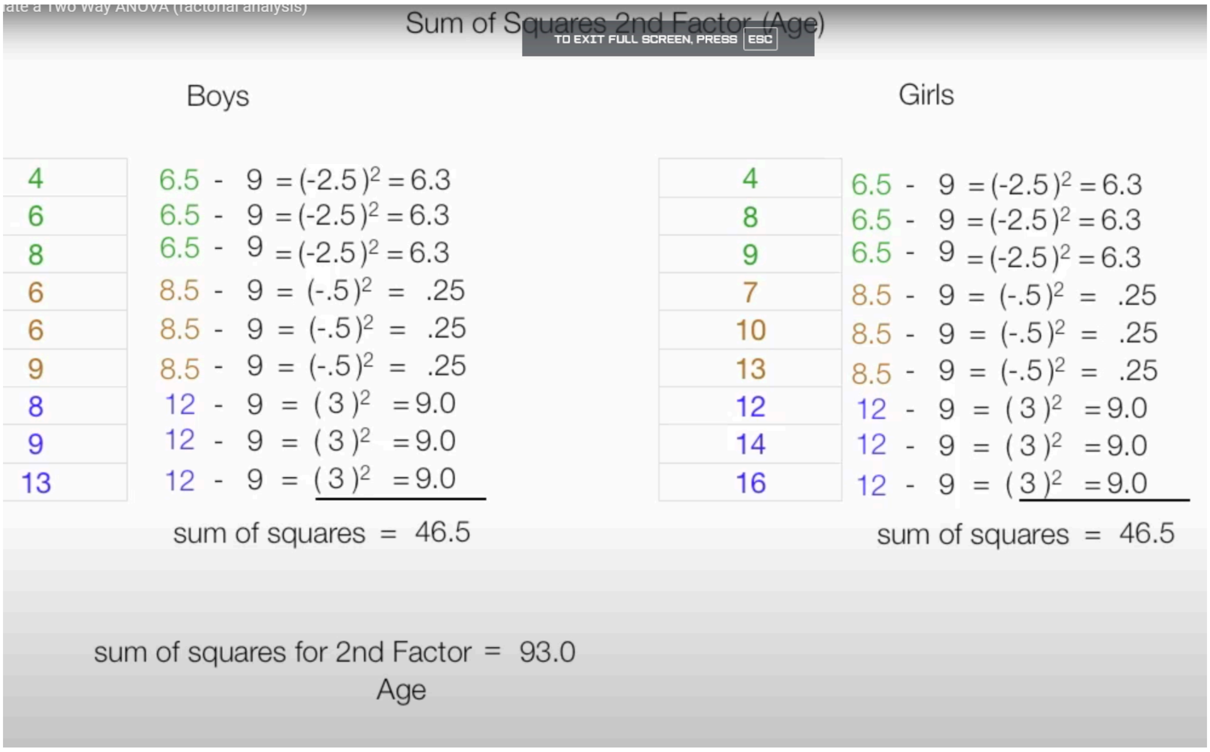
sum of squares = 16

Score	Girls Mean		Grand Mean	
4	10.3	-	9	$= (1.3)^2 = 1.8$
8	10.3	-	9	$= (1.3)^2 = 1.8$
9	10.3	-	9	$= (1.3)^2 = 1.8$
7	10.3	-	9	$= (1.3)^2 = 1.8$
10	10.3	-	9	$= (1.3)^2 = 1.8$
13	10.3	-	9	$= (1.3)^2 = 1.8$
12	10.3	-	9	$= (1.3)^2 = 1.8$
14	10.3	-	9	$= (1.3)^2 = 1.8$
16	10.3	-	9	$= (1.3)^2 = 1.8$

sum of squares = 16

sum of squares for 1st Factor = 32
Gender

Sum of second



Boys

4	6.5	-	9	$= (-2.5)^2 = 6.3$
6	6.5	-	9	$= (-2.5)^2 = 6.3$
8	6.5	-	9	$= (-2.5)^2 = 6.3$
6	8.5	-	9	$= (-.5)^2 = .25$
6	8.5	-	9	$= (-.5)^2 = .25$
9	8.5	-	9	$= (-.5)^2 = .25$
8	12	-	9	$= (3)^2 = 9.0$
9	12	-	9	$= (3)^2 = 9.0$
13	12	-	9	$= (3)^2 = 9.0$

sum of squares = 46.5

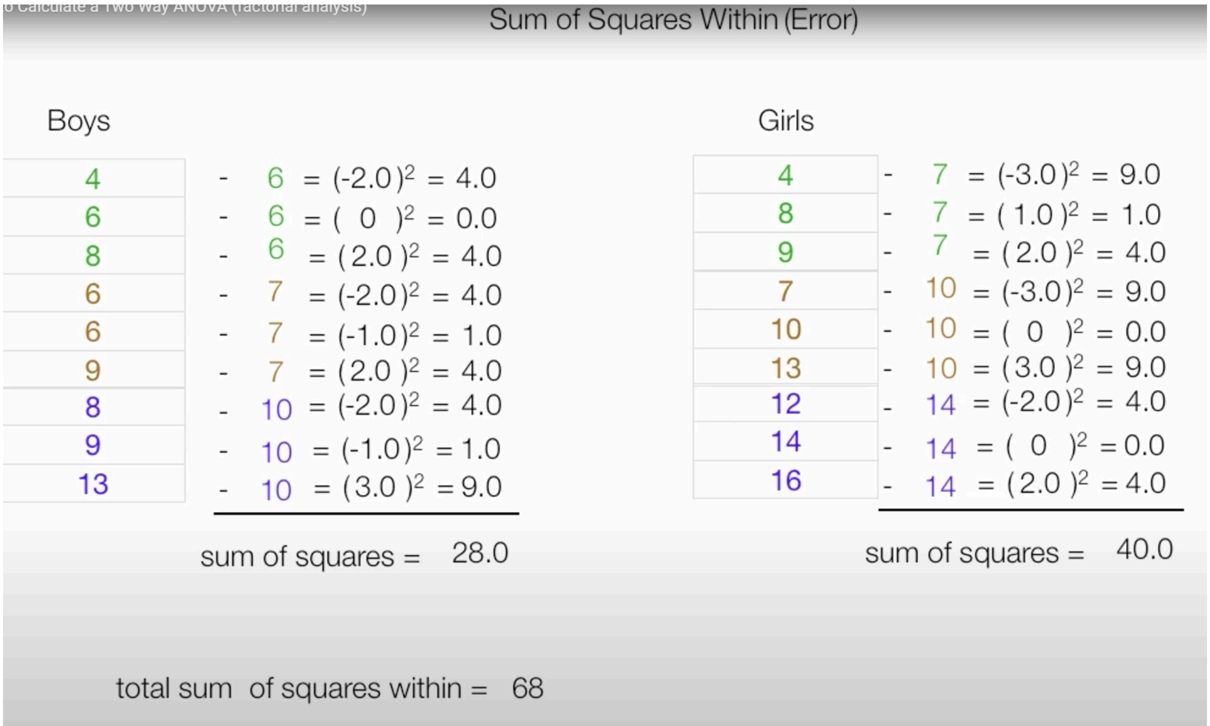
Girls

4	6.5	-	9	$= (-2.5)^2 = 6.3$
8	6.5	-	9	$= (-2.5)^2 = 6.3$
9	6.5	-	9	$= (-2.5)^2 = 6.3$
7	8.5	-	9	$= (-.5)^2 = .25$
10	8.5	-	9	$= (-.5)^2 = .25$
13	8.5	-	9	$= (-.5)^2 = .25$
12	12	-	9	$= (3)^2 = 9.0$
14	12	-	9	$= (3)^2 = 9.0$
16	12	-	9	$= (3)^2 = 9.0$

sum of squares = 46.5

sum of squares for 2nd Factor = 93.0
Age

Sum within error



o Calculate a Two Way ANOVA (factorial analysis)
Sum of Squares Within (Error)
Boys
4 - 6 = (-2.0)² = 4.0
6 - 6 = (0)² = 0.0
8 - 6 = (2.0)² = 4.0
6 - 7 = (-2.0)² = 4.0
6 - 7 = (-1.0)² = 1.0
9 - 7 = (2.0)² = 4.0
8 - 10 = (-2.0)² = 4.0
9 - 10 = (-1.0)² = 1.0
13 - 10 = (3.0)² = 9.0
sum of squares = 28.0
Girls
4 - 7 = (-3.0)² = 9.0
8 - 7 = (1.0)² = 1.0
9 - 7 = (2.0)² = 4.0
7 - 10 = (-3.0)² = 9.0
10 - 10 = (0)² = 0.0
13 - 10 = (3.0)² = 9.0
12 - 14 = (-2.0)² = 4.0
14 - 14 = (0)² = 0.0
16 - 14 = (2.0)² = 4.0
sum of squares = 40.0
total sum of squares within = 68

Total sum

Calculate a Two Way ANOVA (factorial analysis)

Score		Grand Mean	(Score - Grand Mean)^2
4	-	9	$= (-5)^2 = 25.0$
6	-	9	$= (-3)^2 = 9.0$
8	-	9	$= (-1)^2 = 1.0$
6	-	9	$= (-3)^2 = 9.0$
6	-	9	$= (-3)^2 = 9.0$
9	-	9	$= (0)^2 = 0.0$
8	-	9	$= (-1)^2 = 1.0$
9	-	9	$= (0)^2 = 0.0$
13	-	9	$= (4)^2 = 16.0$
4	-	9	$= (-5)^2 = 25.0$
8	-	9	$= (1)^2 = 1.0$
9	-	9	$= (0)^2 = 0.0$
7	-	9	$= (-2)^2 = 4.0$
10	-	9	$= (1)^2 = 1.0$
13	-	9	$= (4)^2 = 16.0$
12	-	9	$= (-3)^2 = 9.0$
14	-	9	$= (5)^2 = 25.0$
16	-	9	$= (7)^2 = 49.0$
			200

Degrees of Freedom

	Sum of Squares	d.f.	Mean Square	F Score
Sum of Squares 1st Factor (Gender)	32	1		
Sum of Squares 2nd Factor (Age)	93	2		
Sum of Squares Within (Error)	68	12		
Sum of Square Both Factors	7	2		
Sum of Squares Total	200	17		

Sum of Squares Within (Error)
degrees of freedom

Boys

10 Year Olds	11 Year Olds	12 Year Olds
4	6	8
6	6	9
8	9	13
n - 1	n - 1	n - 1
3 - 1	3 - 1	3 - 1
2	2	2

Girls

10 Year Olds	11 Year Olds	12 Year Olds
4	7	12
8	10	14
9	13	16
n - 1	n - 1	n - 1
3 - 1	3 - 1	3 - 1
2	2	2

$2 + 2 + 2 + 2 + 2 + 2 = 12$

Sum of both square factors = df1 x df2 = 1 x 2 = 2

Degrees of Freedom	Sum of Squares	d.f.	Mean Square	F Score
Sum of Squares 1st Factor (Gender)	32	1	$\frac{32}{1} = 32$	$\frac{32}{5.67} = 5.64$
Sum of Squares 2nd Factor (Age)	93	2	$\frac{93}{2} = 46.50$	$\frac{46.50}{5.67} = 8.20$
Sum of Square Both Factors	7	2	$\frac{7}{2} = 3.5$	$\frac{3.5}{5.67} = .62$
Sum of Squares Within (Error)	68	12	$\frac{68}{12} = 5.67$	
Sum of Squares Total	200	17		

$F(1,12) = 5.64$ $p < .05$ 95% confidence level

to Interpret the Results of A Two Way ANOVA (Factorial)

F Distribution $F(1,12) = 5.64$, $p<.05$

degrees of freedom numerator

degrees of freedom denominator	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	161.5	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	246.0	248.0	249.1	250.1
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25

$F(1,12)$ = 1 column, 12 row (4.75)
$F(1,12) = 5.64 > 4.75$, reject null hypothesis

For Age, $F(2,12) = 8.2$ $p<.05$
$F(2,12) = 8.2 > 3.89$, reject null hypothesis

For sum of square both factors, $F(2,12) = .62$ $p<.05$
$F(2,12) = .62 < 3.89$, accept null hypothesis

Gender	Score	Age
B	4	16
B	6	10
B	8	10
G	4	10
G	8	10
G	9	10
B	6	11
B	6	11
B	9	11
G	7	11
G	10	11
G	13	11
B	8	13
B	9	12
B	13	12
G	12	12
G	14	12
G	16	12

Factors

Gender Age

Boys

10 yo	11 yo	12 yo
4	6	8
6	6	9
8	9	13

Girls

10 yo	11 yo	12 yo
4	7	12
8	10	14
9	13	16

Mean table

	10 yo	11 yo	12 yo	Average
B	6	7	16	7.7
G	7	10	14	10.3
Average	6.5	8.5	12	9

Gender	Score	Age
B	4	16
B	6	10
B	8	10
b	6	11
b	6	11
b	9	11
b	8	12
b	9	12
b	13	12

Gender	Score	Age
G	4	
G	8	
G	9	
G	7	
G	16	
G	13	
G	12	
G	14	
G	16	

Gender

B		G
7.7		10.3
	9	

Boys Score	Mean	Grand Mean		
4	7.7	- 9	$= (-1.3)^2$	$= 1.8$
6	7.7	- 9		
8	7.7	- 9		
6	7.7	- 9		
6	7.7	- 9		
9	7.7	- 9		
8	7.7	- 9		
9	7.7	- 9		
13	7.7	- 9		
				16.2

Girls Score	Mean	Grand Mean		
4	10.3	- 9	$= (-1.3)^2$	$= 1.8$
8				
9				
7				
16				
13				
12				
14				
16				
				16.2

Sum of squares for 1st Gender = 32

Age

6.5 8.5 12

9

Boys

4	6.5	$- 9 = (-2.5)^2 = 6.3$
6	6.5	
8	6.5	
6	8.5	$= (-.5)^2 = .25$
6	8.5	
9	8.5	
8	12	$= (3)^2 = 9$
9	12	
13	12	

46.5

Girls

4	6.5	$- 9 =$
8	6.5	$-$
9	6.5	$-$
7	8.5	$-$
10	8.5	$-$
13	8.5	$-$
12	12	$-$
14	12	$-$
16	12	

46.5

Sum of Squares Of 2nd Age = 93

Sum of Squares within Error

Boys $- 6 = (-2)^2 = 4$
$4 - 6 = 0^2$ 0
$6 - 6 = 2^2$ 4
$8 - 6 = (-2)^2$ 4
$6 - 7 = (-1)^2$ 1
$6 - 7 = 2^2$ 4
$9 - 7 = (-2)^2$ 4
$8 - 10 = (-1)^2$ 1
$9 - 10 = 3^2$ 9
$13 - 10 =$

28

Girls
$4 \quad -7 = -3^2 = 9$
$8 \quad -7 \quad 1^2 \quad 1$
$9 \quad -7 \quad 2^2 \quad 4$
$7 \quad -10 \quad -3^2 \quad 9$
$10 \quad -10 \quad 0^2 \quad 0$
$13 \quad -10 \quad 3^2 \quad 9$
$12 \quad -14 \quad -2^2 \quad 4$
$14 \quad -14 \quad 0^2 \quad 0$
$16 \quad -14 \quad 2^2 \quad 4$

40

total sum = 68

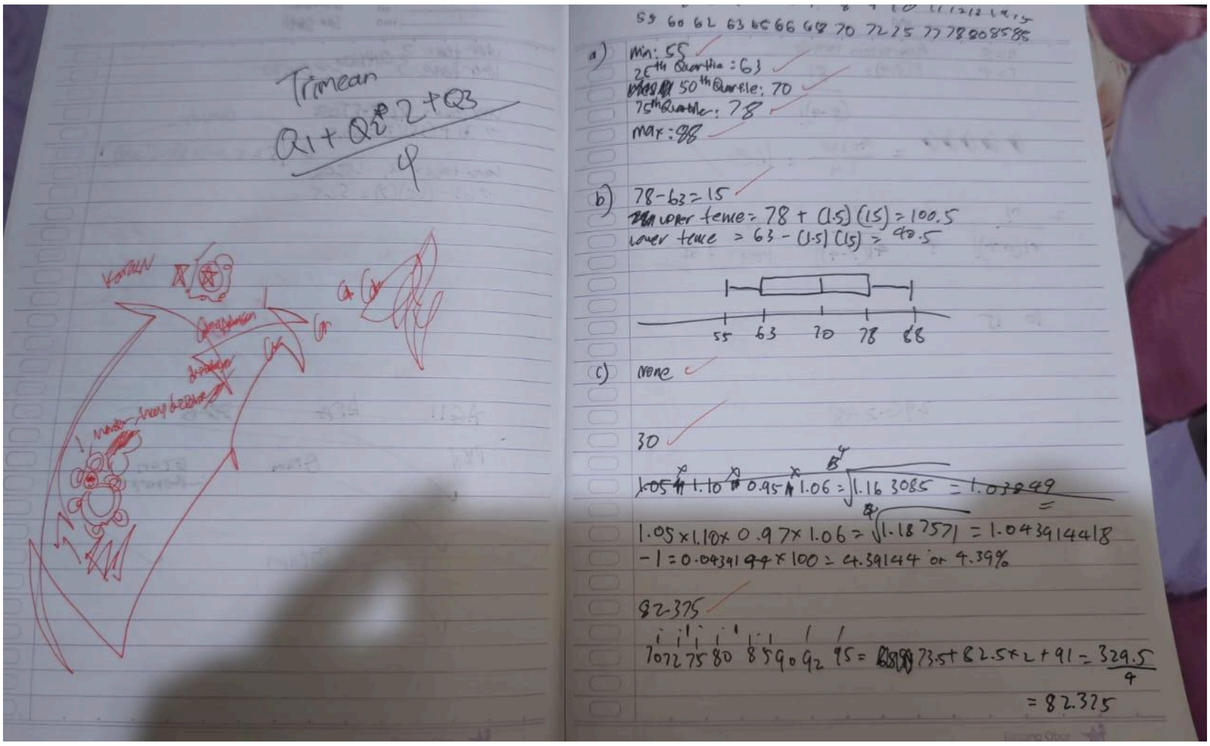
Sum of Squares total

Score		Grand Mean		(Score - GrandMean)^2	
4	-	9	=	$(-5)^2$	= 25
6			=	-3^2	9
8			=	-1^2	1
6			=	-3^2	9
6			=	-3^2	9
9			=	0^2	0
8			=	-1^2	1
9			=	0^2	0
13			=	4^2	16
4			=	-5^2	25
8			=	1^2	1
9			=	0^2	0
7			=	-2^2	4
10				1^2	1
13				4^2	16
12				-3^2	9
14				5^2	25
16				7^2	49
					200

Sum of 1st (Gender)	32
Sum of 2nd (AGG)	93
Sum of Squares within (Error)	68
Sum of Squares both factor	7
Sum of squares total	200

200 - 32 - 93 - 68 = 7

Exercise 1 & Exercise 2



Trimean

$$\frac{Q_1 + Q_2 \cdot 2 + Q_3}{4}$$

59 60 62 63 66 66 68 70 72 75 77 78 80 85 88

a) Min: 55

25th Quartile: 63

50th Quartile: 70

75th Quartile: 78

max: 88

b) 78 − 63 = 15

upper fence = 78 + (1.5)(15) = 100.5

lower fence = 63 − (1.5)(15) = 40.5

55 63 70 78 88

c) none

30

~~1.05 × 1.10 × 0.95 × 1.06 = $\sqrt[4]{1.163085}$ = 1.03849~~

1.05 × 1.10 × 0.97 × 1.06 = $\sqrt[4]{1.187571}$ = 1.043914418

−1 = 0.04391444 × 100 = 4.39144 or 4.39%

82.375

70 72 75 80 85 90 92 95 = 73.5 + 82.5 × 2 + 91 = $\frac{329.5}{4}$

= 82.325

Exercise 3

2 $n=8$ Permutation formula

$r=4$ $P(8,4)=\dfrac{8!}{(8-4)!}$

$=\dfrac{40320}{24}=1680$ ✓

2 $\dfrac{n!}{r!(n-r)!}$ $\dfrac{7}{4}$ $\dfrac{7!}{4!(7-4)!}=\dfrac{5040}{144}=35$

10 15

29<-->48

3

25

15

10

5

3

$$\frac{\binom{10}{3}\cdot\binom{15}{2}}{\binom{25}{5}} = \frac{120\cdot 105}{53130} = \frac{60}{253}\times 100 = 23.7\%$$

$$\frac{10!}{3!(7!)} = 120$$

$$\frac{15!}{2!(13!)} = 105$$

$$\frac{25!}{5!(20!)} = 53130$$

Exercise 4

1

10

15

-8

8

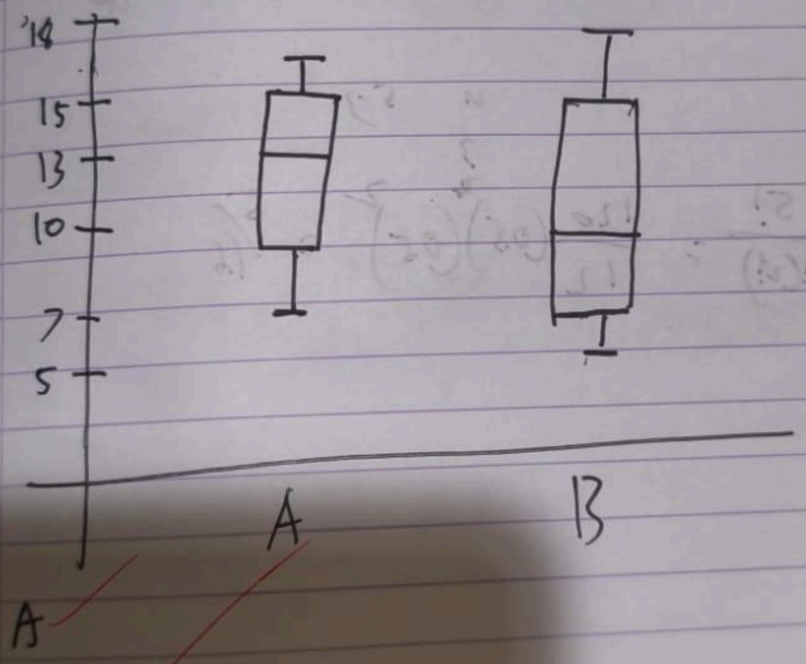
12

$1.10 \times 1.15 \times 0.95 \times 1.08 \times 1.12 = \sqrt[5]{1.4536368} = 1.077688$

$-1 = 0.077683 \times 100 = 7.767\%$

2

A	B8	6	8
min: 7	Min: 5	15+9=24	15+12=27
25: 9	25: 7	No	.No
50: 13	50: 10	9-9=0	7-12=-5
75: 15	75: 15		
Max: 16	Max: 18		



A

.NO

$4/52 = 1/13 \cdot 1/2 = 1/26$

X	Y
1
2
3

$1/2 \times 1/2 \times 1/2 = 1/16$

$\frac{5!}{3!\,2!} = \frac{120}{12} = 10\,(0.5)^3(0.5)^2 = \frac{5}{16}$

$n \quad 5-3$

3

$\frac{5!}{3!(2!)} = \frac{120}{12}(0.5)^3(0.5)^2 = 5/16$

$$\frac{15!}{12!(3!)}(0.8)^{12}(0.2)^{3} = 0.25$$

$$\frac{15!}{13!(2!)}(0.8)^{13}(0.2)^{2} = 0.23$$

0.647

$$\frac{15!}{14!(1!)}(0.8)^{14}(0.2)^{1} = 0.132$$

$$\frac{15!}{15!(0!)}(0.8)^{15}(0.2)^{0} = 0.035$$

Hours (x)	Height (Y)
2	10
4	15
6	20
8	25
10	30

x	y	xy	x^2	y^2
2	10	20	4	100
4	15	60	16	225
6	20	120	36	400
8	25	200	64	625
10	30	300	100	900
30	100	700	220	2250

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

$$= \frac{5(700) - 30(100)}{\sqrt{[5(220) - (30)^2][5(2250) - (100)^2]}}$$

$$= \frac{500}{\sqrt{(1100 - 900)(11250 - 10000)}} = \frac{500}{\sqrt{200(1250)}} = \frac{500}{500} = 1$$

Bintang Obor

Exercise 5

70 85 78 90 88

$\bar{x}$/Mean = 82.2

$$S = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$$

$$\rightarrow \frac{(70-82.2)^2 + (85-82.2)^2 + (78-82.2)^2 + (90-82.2)^2 + (88-82.2)^2}{4}$$

$$= \frac{148.84 + 7.84 + 17.64 + 60.84 + 33.64}{4}$$

$$= \frac{268.8}{4} = \sqrt{67.2} = 8.198$$

1

z Distribution

$$z = \frac{x - \mu}{\sigma}$$

Binomial distribution

$n = 100$
$p = 0.3$
$q = 0.7$

Can be approximated with normal distribution if n is large & p is not too close with 0/1.

$n \cdot p \geq 5$ $\quad n \cdot p = 100 \cdot 0.30 = 30$
$n \cdot q \geq 5$ $\quad n \cdot q = 100 \cdot 0.7 = 70$

$\mu = n \cdot p$ $\quad \mu = 100 \times 0.30 = 30$
$\sigma = \sqrt{n \cdot p \cdot q}$ $\quad \sigma = \sqrt{100 \times 0.3 \times 0.7} = \sqrt{21} = 4.583$

Mean = 30, Sd = 4.58

Asked: $P(x < 25) \approx P(x \leq 24) \xrightarrow{\text{continuity correction}} P(x \leq 24.5)$

$x = 24.5$

$$z = \frac{24.5 - 30}{4.58} = \frac{-5.5}{4.58} = -1.20$$

	0
-1.2	.11507

Probability $0.11507 \times 100 =$ 11.507% / 11.51%

$n = 100$

$p = 0.4$

$q = 0.6$

$$z = \frac{x - \mu}{\sigma}$$

$\mu = n \cdot p = 100 \times 0.4 = 40$

$\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{100 \cdot 0.4 \cdot 0.6} = \sqrt{24} = 2\sqrt{6}$

$P(x > 44) \approx P(x \geq 45) = P(x \geq 44.5)$

$$z = \frac{44.5 - 40}{2\sqrt{6}} = \frac{4.5}{2\sqrt{6}} = 0.9148$$

	0.02
0.9	.82121

$0.82121 \times 100 = 82.12\%$

$1 - 0.82121 = 0.17879 \times 100 = 17.88\%$

at least 45 success → > not <

~~0.82121×100~~

Exercise 6

1 ~~950, 960, 970, 980~~ a: 0.05

mean = 990.5

sd = 25.87

$$\frac{990.5 - 1000}{\frac{25.87}{\sqrt{10}}} = -1.16$$

units of freedom = n-1 = 10-1 = 9

	0.05
9	2.262

-2.262 < -1.16 < 2.262 , fail to reject

Date

Diff = -25 $\quad a = 0.05$

$\mu = -25/8 \quad s_d = 0.835$

$$t = \frac{d}{s_d\sqrt{n}} = \frac{-25/8}{0.835/(\sqrt{8})} = -10.585$$

Unit of freedom = n-1 = 8-1 = 7

	0.05
7	1.895

$-10.59 < -1.895 < 1.895$, reject

$a = 0.05$

$\bar{x}_1 = 8 \quad \sigma_1 = 2 \quad n_1 = 25$

$\bar{x}_2 = 6 \quad \sigma_2 = 2.5 \quad n_2 = 25$

Unit of freedom = 25+25-2 = 48

$$t = \frac{8-6}{\sqrt{\frac{2^2}{25} + \frac{2.5^2}{25}}} = \frac{2}{\frac{\sqrt{41}}{10}} = 3.12$$

	0.05
48	1.679

3.12 > 1.679, reject the null hypo

Exercise 7

Answer

①	②	③
1	2	2
2	4	3
5	2	4

① $H_0 : M_1 = M_2 = M_3$

$H_a \; M_1 \neq M_2 \neq M_3$

$a = 0.05$

② Degrees of freedom between
$= k\,(\text{groups}) - 1 = 3 - 1 = 2$

Degrees of freedom within
$= N - k = 9 - 3 = 6$

$df_{total} = 6 + 2 = 8$

$F_{crit} = 5.14$

	2
6	5.14
	10.92

③ $\bar{x}_1 = 8/3$ or 2.67

$\bar{x}_2 = 2.67$

$\bar{x}_3 = 3$

$G = \frac{25}{9} = 2.78$

$SS_{Total} = \sum (x - G)^2$

$= (1-2.78)^2 + (2-2.78)^2 + (5-2.78)^2 + (2-2.78)^2 + (4-2.78)^2$

$+ (2-2.78)^2 + (2-2.78)^2 + (3-2.78)^2 + (4-2.78)^2 = 13.6$

$SS_{within} = \sum (x_1 - \bar{x}_1)^2 + (x_2 - \bar{x}_2)^2 + (x_3 - \bar{x}_3)^2$

$= (1-2.67)^2 + (2-2.67)^2 + (5-2.67)^2 + (2-2.67)^2 + (4-2.67)^2$
$+ (2-2.67)^2 + (2-3)^2 + (3-3)^2 + (4-3)^2 = 13.3$

$SS_{between} = 13.6 - 13.3 = 0.23$

A	B	C
15	20	25
16	22	27
14	19	26
15	21	28
17	20	29

$a = 0.05$

$df_{between} = 3-1 = 2$

$df_{within} = 15-3 = 12$

$df_{total} = 12+2 = 14$

$F_{crit} = 3.89$

	2
12	3.89

$\bar{x}_1 = 15.4$

$\bar{x}_2 = 20.4$ $\quad G = 20.6$

$\bar{x}_3 = 26$

$SS_{total} = \sum(x-G)^2 = (15-20.6)^2 + (16-20.6)^2 + (14-20.6)^2 + (15-20.6)^2 + (17-20.6)^2 + (20-20.6)^2 + (22-20.6)^2 + (19-20.6)^2 + (21-20.6)^2 + (20-20.6)^2 + (25-20.6)^2 + (27-20.6)^2 + (26-20.6)^2 + (28-20.6)^2 + (29-20.6)^2 = 301.6$

$SS_{within} = (15-15.4)^2 + (16-15.4)^2 + (14-15.4)^2 + (15-15.4)^2 + (17-15.4)^2 + (20-20.4)^2 + (22-20.4)^2 + (19-20.4)^2 + (21-20.4)^2 + (20-20.4)^2 + (25-26)^2 + (27-26)^2 + (26-26)^2 + (28-26)^2 + (29-26)^2$
$= 20.4$

$SS_{between} = 301.6 - 20.4 = 281.2$

$MS_{between} = \frac{281.2}{2} = 140.6$

$MS_{within} = \frac{20.4}{12} = 1.7$

$F = \frac{140.6}{1.7} = 82.7$

OBServed

Fertilizer	Plant A	Plant B	Plant C	Total	a=0.05
x	10	20	10	40	
Y	15	10	5	30	
Z	5	5	10	20	
Total	30	35	25	90	

~~df = n-1 = 5-1 = 4~~ df = n-1 = 2×2 = 4

~~0.05 4 9.49 → crit~~

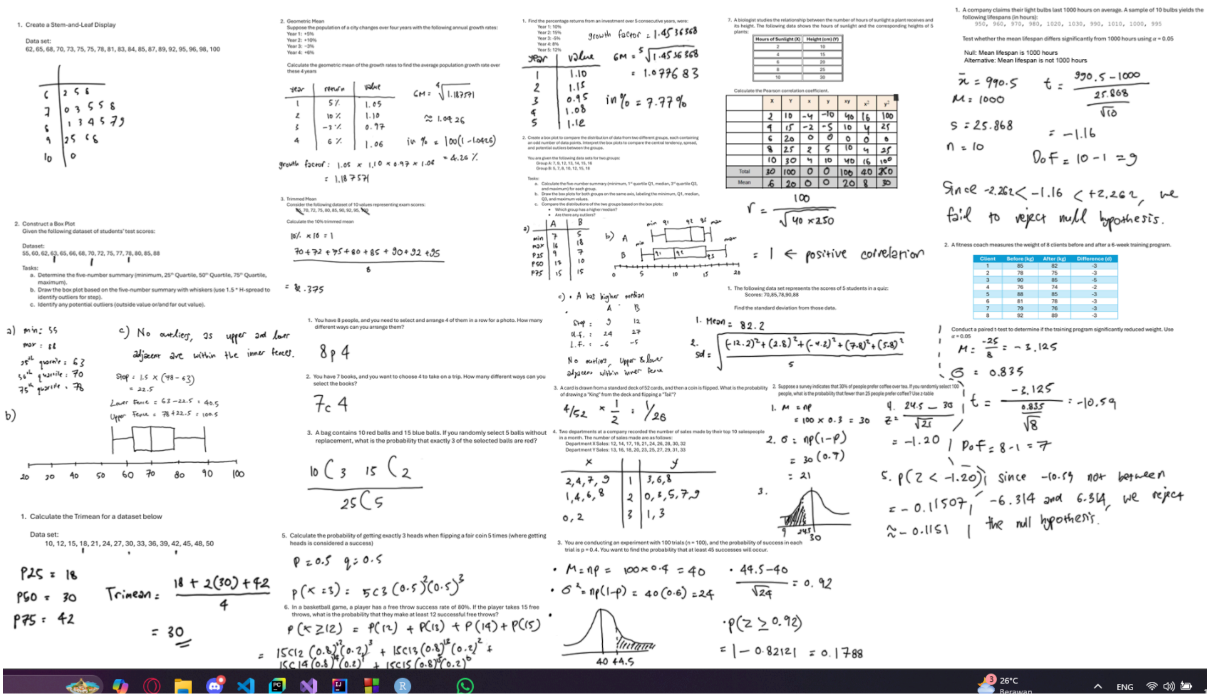
	.05
4	9.49 → crit

Expected

Fertilizer	Plant A	Plant B	Plant C
x	13.3	15.6	11.1
Y	10	11.7	8.3
Z	6.67	7.8	5.6

O	E	O-E	$(O-E)^2$	$\frac{(O-E)^2}{E}$
10	13.3	-3.3	10.89	0.82
15	10	5	25	2.5
5	6.7	-1.7	2.89	0.43
20	15.6	4.4	19.36	1.24
10	11.7	-1.7	2.89	0.25
5	7.8	-2.8	7.84	1.01
10	11.1	-1.1	1.21	0.11
5	8.3	-3.3	10.89	1.31
10	5.6	4.4	19.36	3.46
				11.13

11.13 > 9.44, reject



To perform a Two-Way ANOVA, we'll follow these steps:

Step 1: Define the Null Hypotheses

1. **H$_{01}$**: There is no significant effect of the programming language on test scores.
2. **H$_{02}$**: There is no significant effect of the study method on test scores.
3. **H$_{03}$**: There is no significant interaction effect between the programming language and the study method on test scores.

Step 2: Data Entry

We have the following data from the table:

- **Python (Self-Study)**: 78, 82, 85
- **Python (Instructor-Led)**: 90, 88, 92
- **Java (Self-Study)**: 72, 75, 74
- **Java (Instructor-Led)**: 85, 80, 84
- **C++ (Self-Study)**: 65, 68, 70
- **C++ (Instructor-Led)**: 78, 75, 80

Step 3: Perform the Two-Way ANOVA

Let's compute the Two-Way ANOVA using this data.

Step 4: Interpret the Results

From the ANOVA table:

- **Effect of Language**: F(2,12)=40.97F(2, 12) = 40.97F(2,12)=40.97, p=0.000004p = 0.000004p=0.000004
    - The p-value is much less than 0.05, so we reject the null hypothesis H01H_{01}H01. This indicates a significant effect of programming language on test scores.
- **Effect of Study Method**: F(1,12)=59.90F(1, 12) = 59.90F(1,12)=59.90, p=0.000005p = 0.000005p=0.000005
    - The p-value is much less than 0.05, so we reject the null hypothesis H02H_{02}H02. This indicates a significant effect of the study method on test scores.
- **Interaction Effect (Language * Study Method)**: F(2,12)=0.17F(2, 12) = 0.17F(2,12)=0.17, p=0.849605p = 0.849605p=0.849605
    - The p-value is greater than 0.05, so we fail to reject the null hypothesis H03H_{03}H03. This suggests there is no significant interaction effect between the programming language and the study method on test scores.

Conclusion:

Both the programming language and the study method have significant main effects on test scores. However, there is no significant interaction effect between these two factors.