

1)

The correlation coefficient of 0.7 indicates a strong positive relationship between the SAT score and the college GPA. when the correlation is positive, it means if SAT scores increase, the college GPA tends to increase.

2)

$\bar{x} = 170$ cm, $\sigma = 10$ cm,
Sample size (n) = 1000,
normally distributed = True

a)

We can say, By using Empirical formula (*because the dataset is normally distributed and \bar{x} & σ is given*) the percentage of individuals in the dataset have between 160 cm and 180 cm is **68%** approx.

b)

Sample size (n) = 100
 $\bar{x} = 170$
 σ for the sample = $10 \div \sqrt{100} = 1$ cm
Z-score = $(X - \mu) / (\sigma / \sqrt{n}) = 175 - 170 / 1 = 5$
Using z-table we get $0.99 \approx 100\%$

c)

$x = 185$
Z-score = $(X - \mu) / (\sigma / \sqrt{n}) = (185 - 175) / 10 = 1.5$

d)

$1 - 0.025 = 0.975$
On checking z-table we get = -1.645
Therefore,

$$\begin{aligned} \text{Z-score} &= (X - \mu) / (\sigma / \sqrt{n}) \\ X &= z * \sigma + \mu \\ X &= -1.645 * 10 + 170 \\ X &= 153.55 \text{ cm} \end{aligned}$$

e)

$$cv = \frac{\sigma}{\mu} * 100$$

$$cv = 10/170 * 100$$

$$cv = 5.88 \%$$

f)

$$\text{Skewness} = 3 * (\text{mean} - \text{median}) / \sigma$$

Since data is normally distributed, mean = median

$$\text{Skewness} = 3 * (\text{mean} - \text{mean}) / \sigma = 0$$

3)

a)

Dispersion in "Blood Pressure Before":

Range: The range is the difference between the maximum and minimum values.

Maximum value: 148

Minimum value: 120

$$\text{Range: } 148 - 120 = 28$$

$$\text{Variance} = \sum (X_i - \bar{X})^2 / N$$

$$\text{Variance for Blood Pressure Before (mmHg)} = 43.1019$$

$$\text{Variance for Blood Pressure After (mmHg)} = 46.9704$$

$$\text{Standard deviation} = \text{variance} ^{0.5}$$

$$\text{Standard deviations for Blood Pressure Before (mmHg)} = 6.565203729$$

$$\text{Standard deviation for Blood Pressure After (mmHg)} = 6.853495459$$

Range:

The range provides a measure of how much the values spread out from the minimum to the maximum. A larger range indicates a wider dispersion of values.

Variance:

The variance measures the average squared deviation from the mean. A higher variance indicates a greater dispersion or spread of values around the mean.

Standard Deviation:

The standard deviation is the square root of the variance. It measures the dispersion or spread of values from the mean. A higher standard deviation indicates a wider spread of values.

b)

For Blood Pressure Before (mmHg)

$$\text{Mean} = (130 + 142 + 120 + \dots + 135) / 100 = 132.58$$

$$\text{Standard Deviation} = 6.565203729$$

$$\begin{aligned}\text{Standard Error} &= \text{Standard Deviation} / \text{Square root of the number of values} \\ &= 0.6598278012\end{aligned}$$

$$\begin{aligned}\text{Margin of Error} &= 1.96 * \text{Standard Error} \\ 1.96 * 0.6598278012 &= 12.93262490352 \approx 13\end{aligned}$$

To calculate the confidence interval:

$$\text{Lower Bound} = \text{Mean} - \text{Margin of Error} = 131.28$$

$$\text{Upper Bound} = \text{Mean} + \text{Margin of Error} = 133.88$$

For Blood Pressure After(mmHg)

$$\text{Mean} = (120 + 135 + 118 \dots + 130) / 100 = 128.36$$

$$\text{Standard Deviation} = 6.565203729$$

$$\begin{aligned}\text{Standard Error} &= \text{Standard Deviation} / \text{Square root of the number of values} \\ &= 6.888022103\end{aligned}$$

$$\begin{aligned}\text{Margin of Error} &= 1.96 * \text{Standard Error} \\ 1.96 * 6.888022103 &= 13.50052332188 \approx 13.5\end{aligned}$$

To calculate the confidence interval:

$$\text{Lower Bound} = \text{Mean} - \text{Margin of Error} = 114.86$$

$$\text{Upper Bound} = \text{Mean} + \text{Margin of Error} = 141.86$$

C)

$$\text{MAD} = (\sum |x_i - \bar{x}|) / n$$

Where x_i is the i th value of column

\bar{x} is mean

n is the count

For Blood Pressure Before (mmHg)

$$\text{MAD} = 3.4106051316484808e-15$$

For Blood Pressure After(mmHg)

$$\text{MAD} = 1.3642420526593923e-14$$

4)

Perfect squares = 1,4,9 &16

No of favorable outcome = 4

Total no of outcome = 20

$$\text{Probability} = \frac{\text{No of favorable outcome}}{\text{Total no of outcome}} = \frac{4}{20} = 0.2 \text{ or } 20 \%$$

5)

A → Taxi belongs to company A

B → Taxi belongs to company B

L → Taxi is Late

$$P(A/L) = ?$$

Using Baye's theorem,

$$P(A/L) = P(L/A) * P(A) / P(L)$$

$$\begin{aligned} P(L) &= P(L/A) * P(A) + P(L/B) * P(B) \\ &= 1-0.95 * 0.80 + 1-0.90 * 0.2 \\ &= 0.05 * 0.80 + 0.10 * 0.2 \\ &= 0.04 + 0.02 \\ &= 0.06 \end{aligned}$$

Therefore,

$$P(A/L) = 0.06 * 0.80 / 0.06 = 80 \%$$

Probability of Late taxi belongs to company A is 80 %

7)

$$2x + 3 - 8 = 0 \Rightarrow x = 5/2$$

$$2y + x - 5 = 0 \Rightarrow y = 5/4$$

Variance of x = 4

$$\begin{aligned} \text{Var}(y) &= \text{Var}(x) * (\text{slope of } y)^2 \\ &= 4 * (5/4)^2 \\ &= 4 * 25/16 = 25/4 \end{aligned}$$

8)

Here are the steps to perform the Wilcoxon signed-rank test:

Step 1: Calculate the absolute values of the differences and rank them from smallest to largest, preserving the sign:

Difference (absolute)	Rank
1	4
1	5
1	6
2	7
2	8
2	9
2	10
2	11
3	12
3	13

Step 3: Calculate the sum of positive ranks (W^+) and the sum of negative ranks (W^-):

$$W^+ = 4 + 5 + 6 + 7 + 8 + 9 + 10 + 11 + 12 + 13 = 85$$

$$W^- = 0 \text{ (there are no negative ranks)}$$

Step 4: Calculate the test statistic (T):

$$T = \min (W^+, W^-) = \min (85, 0) = 0$$

Step 5: Determine the critical value or p-value:

The critical value or p-value is obtained from the Wilcoxon signed-rank table based on the sample size and desired significance level (e.g., $\alpha = 0.05$). Since the sample size is 10, we look for the critical value in the row corresponding to $n = 10$.

Based on the table, if the test statistic (T) is less than or equal to the critical value, we reject the null hypothesis (H_0) and conclude that the therapy had a significant effect on anxiety levels. Otherwise, if the test statistic is greater than the critical value, we fail to reject H_0 .

Since $T = 0$ and there is no critical value less than or equal to 0, we can conclude that the therapy had a significant effect on anxiety levels.

9)

Step 1:

Set up the null and alternative hypotheses:

Null hypothesis (H_0): The mean scores of all the students are the same.

Alternative hypothesis (H_1): The mean scores of the students are not all the same.

Step 2:

Calculate the mean score for each exam:

Mean Exam 1 = $(85 + 70 + 90 + 75 + 95) / 5 = 83$

Mean Exam 2 = $(90 + 80 + 85 + 70 + 92) / 5 = 83.4$

Mean Final Exam = $(92 + 85 + 88 + 75 + 96) / 5 = 87.2$

Step 3:

Calculate the sum of squares between groups (SSG):

$$SSG = (n_1 * (\text{mean}_1 - \text{overall_mean})^2) + (n_2 * (\text{mean}_2 - \text{overall_mean})^2) + (n_3 * (\text{mean}_3 - \text{overall_mean})^2)$$

$$= (5 * (83 - 84.12)^2) + (5 * (83.4 - 84.12)^2) + (5 * (87.2 - 84.12)^2)$$

$$= 5.56 + 0.1344 + 8.7616$$

$$= 14.455$$

Step 4:

Calculate the sum of squares within groups (SSW):

$$SSW = (n_1 - 1) * \text{variance}_1 + (n_2 - 1) * \text{variance}_2 + (n_3 - 1) * \text{variance}_3$$

$$= (4 * 45.2) + (4 * 45.6) + (4 * 53.2)$$

$$= 180.8 + 182.4 + 212.8$$

$$= 575.6$$

Step 5:

Calculate the degrees of freedom between groups (dfG) and within groups (dfW):

$$\text{dfG} = \text{number of groups} - 1 = 3 - 1 = 2$$

$$\text{dfW} = \text{total number of observations} - \text{number of groups} = 15 - 3 = 12$$

Step 6:

Calculate the mean squares between groups (MSG) and within groups (MSW):

$$\text{MSG} = \text{SSG} / \text{dfG} = 14.455 / 2 = 7.2275$$

$$\text{MSW} = \text{SSW} / \text{dfW} = 575.6 / 12 = 47.9667$$

Step 7:

Calculate the F-statistic:

$$F = \text{MSG} / \text{MSW} = 7.2275 / 47.9667 = 0.1504$$

Step 8:

Determine the critical value or p-value:

To determine the critical value or p-value for the F-statistic, we need to consult an F-distribution table or use statistical software. We can compare the calculated F-value with the critical value or obtain the p-value.

If the calculated F-value is greater than the critical value or the p-value is less than the chosen significance level (e.g., $\alpha = 0.05$), we reject the null hypothesis and conclude that the mean scores of the students are not all the same. Regarding the student with the highest score, based on the given data, Jeevan has the highest scores in all three exams (Exam 1: 95, Exam 2: 92, Final Exam: 96)

10)

Using Binomial probability,

$$P(x=k) = {}^nC_k * p^k * (1-p)^{(n-k)}$$

Given,

$$n = 500$$

$$k = 20$$

$$p(x=20) = ({}^{500}C_{20}) * 0.05^{20} * (1-0.05)^{(500-20)}$$