**1. Scenario: A company wants to analyze the sales performance of its products in different regions. They have collected the following data:**

**Region A: [10, 15, 12, 8, 14]**

**Region B: [18, 20, 16, 22, 25]**

**Calculate the mean sales for each region.**

Ans-To calculate the mean sales for each region, we sum up all the sales values for each region and divide by the number of values.

For Region A: mean = (10 + 15 + 12 + 8 + 14) / 5 = 59 / 5 = 11.8

For Region B: mean = (18 + 20 + 16 + 22 + 25) / 5 = 101 / 5 = 20.2

So the mean sales for Region A is 11.8 and for Region B is 20.2.

**2. Scenario: A survey is conducted to measure customer satisfaction on a scale of 1 to 5. The data collected is as follows:**

**[4, 5, 2, 3, 5, 4, 3, 2, 4, 5]**

**Calculate the mode of the survey responses.**

Ans-The mode is the value that appears most frequently in a data set. To calculate the mode of the survey responses, we can count the frequency of each value in the data set.

The frequency of each value is as follows: 2: 2, 3: 2, 4: 3, 5: 3

Since the values 4 and 5 both appear 3 times in the data set, they are both modes of the survey responses. So the mode of the survey responses is 4 and 5.

**3. Scenario: A company wants to compare the salaries of two departments. The salary data for Department A and Department B are as follows:**

**Department A: [5000, 6000, 5500, 7000]**

**Department B: [4500, 5500, 5800, 6000, 5200]**

**Calculate the median salary for each department.**

Ans- For Department A: First we need to order the data from least to greatest: [5000, 5500, 6000, 7000] Since there are an even number of observations (4), we take the mean of the two middle values (5500 and 6000).

So, median = (5500 + 6000) / 2 = 5750

For Department B: First we need to order the data from least to greatest: [4500, 5200, 5500, 5800, 6000] Since there are an odd number of observations (5), the median is the middle value (5500). median = 5500

So, the median salary for Department A is 5750 and for Department B is 5500.

**4.** **Scenario: A data analyst wants to determine the variability in the daily stock prices of a company. The data collected is as follows:**

**[25.5, 24.8, 26.1, 25.3, 24.9]**

**Calculate the range of the stock prices.**

Ans- The range is a measure of variability that is calculated by subtracting the smallest value in a data set from the largest value. For this data set: The smallest value is 24.8 and the largest value is 26.1.

Range = largest value - smallest value range = 26.1 - 24.8 range = 1.3

So, the range of the stock prices is 1.3.

**5. Scenario: A study is conducted to compare the performance of two different teaching methods. The test scores of the students in each group are as follows:**

**Group A: [85, 90, 92, 88, 91]**

**Group B: [82, 88, 90, 86, 87]**

**Perform a t-test to determine if there is a significant difference in the mean scores between the two groups**.

Ans- To perform a t-test and determine if there is a significant difference in the mean scores between the two groups, we can use the independent two-sample t-test. This test compares the means of two independent groups to determine if there is a statistically significant difference between them.

The null hypothesis (H0) for the t-test is that there is no significant difference in the means of the two groups. The alternative hypothesis (H1) is that there is a significant difference between the means of the two groups.

Let's perform the t-test for the given data:

Group A: [85, 90, 92, 88, 91]

Group B: [82, 88, 90, 86, 87]

Calculate the means of each group.

Mean of Group A = (85 + 90 + 92 + 88 + 91) / 5 = 88.4

Mean of Group B = (82 + 88 + 90 + 86 + 87) / 5 = 86.6

Then, Calculate the standard deviations of each group.

Standard deviation of Group A = sqrt(((85-88.4)^2 + (90-88.4)^2 + (92-88.4)^2 + (88-88.4)^2 + (91-88.4)^2) / (5-1)) = 2.302

Standard deviation of Group B = sqrt(((82-86.6)^2 + (88-86.6)^2 + (90-86.6)^2 + (86-86.6)^2 + (87-86.6)^2) / (5-1)) = 2.607

Then, Calculate the t-value.

t = (mean of Group A - mean of Group B) / sqrt((standard deviation of Group A^2 / nA) + (standard deviation of Group B^2 / nB))

t = (88.4 - 86.6) / sqrt((2.302^2 / 5) + (2.607^2 / 5))

t = 1.8 / sqrt(0.946 + 1.076)

t ≈ 1.8 / sqrt(2.022)

t ≈ 1.8 / 1.422

t ≈ 1.264

Also, Determine the degrees of freedom.

Degrees of freedom = nA + nB - 2

Degrees of freedom = 5 + 5 - 2

Degrees of freedom = 8

Determine the critical t-value.

The critical t-value depends on the desired significance level and degrees of freedom. Let's assume a significance level of 0.05 (5%) for a two-tailed test. Using a t-table or a t-distribution calculator, the critical t-value for 8 degrees of freedom at a 0.05 significance level is approximately 2.306.

Compare the calculated t-value with the critical t-value.

Since the calculated t-value (1.264) is smaller than the critical t-value (2.306), we fail to reject the null hypothesis.

Therefore, based on the t-test, there is no significant difference in the mean scores between Group A and Group B at a 5% significance level.

**6. Scenario: A company wants to analyze the relationship between advertising expenditure and sales. The data collected is as follows:**

**Advertising Expenditure (in thousands): [10, 15, 12, 8, 14]**

**Sales (in thousands): [25, 30, 28, 20, 26]**

**Calculate the correlation coefficient between advertising expenditure and sales.**

Ans- The correlation coefficient is a measure of the strength and direction of the linear relationship between two variables. It can be calculated using the formula:

r = 1/n-1 \* Σ[(xi - x̄) (yi - ȳ) / (sx \* sy)], where n is the number of observations, xi and yi are the individual x and y values, x̄ and ȳ are the means of x and y, respectively, and sx and sy are the standard deviations of x and y, respectively.

Given the data you provided, we can calculate the correlation coefficient between advertising expenditure and sales as follows:

First, we need to calculate the means of advertising expenditure and sales:

x̄ = (10 + 15 + 12 + 8 + 14) / 5 = 11.8

ȳ = (25 + 30 + 28 + 20 + 26) / 5 = 25.8

Next, we need to calculate the standard deviations of advertising expenditure and sales:

sx = √[(10-11.8)² + (15-11.8)² + (12-11.8)² + (8-11.8)² + (14-11.8)²] / √(5-1) = √(14) / √4 = √3.5 = 1.87

sy = √[(25-25.8)² + (30-25.8)² + (28-25.8)² + (20-25.8)² + (26-25.8)²] / √(5-1) = √(50) / √4 = √12.5 = 3.54

Now we can calculate the correlation coefficient using the formula above:

r = 1/4 \* [(10-11.8)(25-25.8)+(15-11.8)(30-25.8)+(12-11.8)(28-25.8)+(8-11.8)(20-25.8)+(14-11.8)(26-25.8)] / (1.87\*3.54)

r = 0.94

The correlation coefficient between advertising expenditure and sales is approximately 0.94.

**7. Scenario: A survey is conducted to measure the heights of a group of people. The data collected is as follows:**

**[160, 170, 165, 155, 175, 180, 170]**

**Calculate the standard deviation of the heights.**

Ans-The standard deviation is a measure of the spread of the data. It can be calculated using the formula: s = √[Σ(xi-x̄)²/(n-1)], where n is the number of observations, xi are the individual x values, and x̄ is the mean of x.

Given the data you provided, we can calculate the standard deviation of the heights as follows:

First, we need to calculate the mean of the heights:

x̄ = (160 + 170 + 165 + 155 + 175 + 180 + 170) / 7 ≈ 167.86

Next, we can calculate the standard deviation using the formula above:

s = √[(160-167.86)² + (170-167.86)² + (165-167.86)² + (155-167.86)² + (175-167.86)² + (180-167.86)² + (170-167.86)²] / (7-1)

s ≈ 8.55

The standard deviation of the heights is approximately 8.55.

**8. Scenario: A company wants to analyze the relationship between employee tenure and job satisfaction. The data collected is as follows:**

**Employee Tenure (in years): [2, 3, 5, 4, 6, 2, 4]**

**Job Satisfaction (on a scale of 1 to 10): [7, 8, 6, 9, 5, 7, 6]**

**Perform a linear regression analysis to predict job satisfaction based on employee tenure.**

Ans- Linear regression is a method used to model the relationship between a dependent variable (in this case, job satisfaction) and one or more independent variables (in this case, employee tenure). The goal of linear regression is to find the line of best fit that can accurately predict the value of the dependent variable based on the values of the independent variables.

Given the data you provided, we can perform a linear regression analysis to predict job satisfaction based on employee tenure as follows:

First, we need to calculate the means of employee tenure and job satisfaction:

x̄ = (2 + 3 + 5 + 4 + 6 + 2 + 4) / 7 ≈ 3.71

ȳ = (7 + 8 + 6 + 9 + 5 + 7 + 6) / 7 ≈ 6.86

Next, we need to calculate the slope of the line of best fit using the formula: b = Σ[(xi-x̄)(yi-ȳ)] / Σ[(xi-x̄)²], where xi and yi are the individual x and y values, and x̄ and ȳ are the means of x and y, respectively.

b = [(2-3.71)(7-6.86) + (3-3.71)(8-6.86) + (5-3.71)(6-6.86) + (4-3.71)(9-6.86) + (6-3.71)(5-6.86) + (2-3.71)(7-6.86) + (4-3.71)(6-6.86)] / [(2-3.71)² + (3-3.71)² + (5-3.71)² + (4-3.71)² + (6-3.71)² + (2-3.71)² + (4-3.71)²]

b ≈ -0.26

Now that we have the slope of the line of best fit, we can calculate the y-intercept using the formula: a = ȳ - bx̄, where x̄ and ȳ are the means of x and y, respectively, and b is the slope of the line of best fit.

a = 6.86 - (-0.26)(3.71)

a ≈ 7.83

The equation of the line of best fit is therefore: ŷ = a + bx, where ŷ is the predicted value of y (job satisfaction), a is the y-intercept, b is the slope of the line of best fit, and x is the value of the independent variable (employee tenure).

ŷ = 7.83 - 0.26x

This equation can be used to predict job satisfaction based on employee tenure.

**9. Scenario: A study is conducted to compare the effectiveness of two different medications. The recovery times of the patients in each group are as follows:**

**Medication A: [10, 12, 14, 11, 13]**

**Medication B: [15, 17, 16, 14, 18]**

**Perform an analysis of variance (ANOVA) to determine if there is a significant difference in the mean recovery times between the two medications.**

Ans- To perform an analysis of variance (ANOVA) and determine if there is a significant difference in the mean recovery times between Medication A and Medication B, we can use a one-way ANOVA test. This test compares the means of three or more groups to determine if there is a statistically significant difference between them.

Let's perform the ANOVA for the given data:

Medication A: [10, 12, 14, 11, 13]

Medication B: [15, 17, 16, 14, 18]

1. Calculate the means of each group.

Mean of Medication A = (10 + 12 + 14 + 11 + 13) / 5 = 12

Mean of Medication B = (15 + 17 + 16 + 14 + 18) / 5 = 16

2. Calculate the sum of squares between groups (SSB).

SSB = (nA \* (meanA - meanTotal)^2) + (nB \* (meanB - meanTotal)^2)

= (5 \* (12 - 14)^2) + (5 \* (16 - 14)^2)

= 40 + 40

= 80

3. Calculate the sum of squares within groups (SSW).

SSW = Σ(xi - meanA)^2 + Σ(xi - meanB)^2

= (10 - 12)^2 + (12 - 12)^2 + (14 - 12)^2 + (11 - 12)^2 + (13 - 12)^2

+ (15 - 16)^2 + (17 - 16)^2 + (16 - 16)^2 + (14 - 16)^2 + (18 - 16)^2

= 2^2 + 0^2 + 2^2 + 1^2 + 1^2 + 1^2 + 1^2 + 0^2 + 2^2 + 2^2

= 4 + 0 + 4 + 1 + 1 + 1 + 1 + 0 + 4 + 4

= 20

4. Calculate the degrees of freedom.

Degrees of freedom between groups (dfB) = Number of groups - 1 = 2 - 1 = 1

Degrees of freedom within groups (dfW) = Total number of data points - Number of groups = 10 - 2 = 8

5. Calculate the mean squares.

Mean square between groups (MSB) = SSB / dfB = 80 / 1 = 80

Mean square within groups (MSW) = SSW / dfW = 20 / 8 = 2.5

6. Calculate the F-value.

F = MSB / MSW = 80 / 2.5 = 32

7. Determine the critical F-value.

The critical F-value depends on the desired significance level and the degrees of freedom. Assuming a significance level of 0.05 (5%) and dfB = 1, dfW = 8, the critical F-value is approximately 5.32.

8. Compare the calculated F-value with the critical F-value.

Since the calculated F-value (32) is greater than the critical F-value (5.32), we reject the null hypothesis.

Therefore, based on the ANOVA, there is a significant difference in the mean recovery times between Medication A and Medication B at a 5% significance level.

**10. Scenario: A company wants to analyze customer feedback ratings on a scale of 1 to 10. The data collected is**

**as follows:**

**[8, 9, 7, 6, 8, 10, 9, 8, 7, 8]**

**Calculate the 75th percentile of the feedback ratings.**

Ans- To calculate the 75th percentile of the feedback ratings, we need to arrange the data in ascending order and find the value that corresponds to the position denoted by the 75th percentile.

The data collected is as follows:

[8, 9, 7, 6, 8, 10, 9, 8, 7, 8]

Arrange the data in ascending order:

6, 7, 7, 8, 8, 8, 8, 9, 9, 10

Calculate the position of the 75th percentile:

Position = (75 / 100) \* (n + 1)

= (75 / 100) \* (10 + 1)

= (75 / 100) \* 11

= 8.25

Since the position is not a whole number, we take the ceiling value to find the nearest whole number greater than or equal to 8.25, which is 9.

The 75th percentile corresponds to the value in the 9th position after arranging the data in ascending order.

So, the 75th percentile of the feedback ratings is 9

**11. Scenario: A quality control department wants to test the weight consistency of a product. The weights of a sample of products are as follows:**

**[10.2, 9.8, 10.0, 10.5, 10.3, 10.1]**

**Perform a hypothesis test to determine if the mean weight differs significantly from 10 grams.**

Ans- To perform a hypothesis test and determine if the mean weight of the product differs significantly from 10 grams, we can use a one-sample t-test. This test compares the sample mean to a hypothesized population mean to assess if there is a statistically significant difference.

Let's perform the hypothesis test for the given data:

Weights of the sample of products: [10.2, 9.8, 10.0, 10.5, 10.3, 10.1]

Hypothesized population mean: 10 grams

1. Calculate the sample mean.

Sample Mean = (10.2 + 9.8 + 10.0 + 10.5 + 10.3 + 10.1) / 6 = 10.167 (rounded to three decimal places)

2. Calculate the sample standard deviation.

Sample Standard Deviation = sqrt(((10.2-10.167)^2 + (9.8-10.167)^2 + (10.0-10.167)^2 + (10.5-10.167)^2 + (10.3-10.167)^2 + (10.1-10.167)^2) / (6-1))

= sqrt((0.002689 + 0.002689 + 0.000224 + 0.004641 + 0.000224 + 0.000002) / 5)

= sqrt(0.010469 / 5)

= sqrt(0.0020938)

= 0.0458 (rounded to four decimal places)

3. Calculate the t-value.

t-value = (Sample Mean - Hypothesized Mean) / (Sample Standard Deviation / sqrt(n))

= (10.167 - 10) / (0.0458 / sqrt(6))

= 0.167 / (0.0458 / 2.449)

= 0.167 / 0.0187

= 8.942 (rounded to three decimal places)

4. Determine the degrees of freedom.

Degrees of Freedom = Sample Size - 1 = 6 - 1 = 5

5. Determine the critical t-value.

The critical t-value depends on the desired significance level and degrees of freedom. Assuming a significance level of 0.05 (5%) and 5 degrees of freedom, the critical t-value is approximately 2.571.

6. Compare the calculated t-value with the critical t-value.

Since the calculated t-value (8.942) is greater than the critical t-value (2.571), we reject the null hypothesis.

So, based on the t-test, there is a statistically significant difference between the mean weight of the product and the hypothesized mean of 10 grams

**12. Scenario: A company wants to analyze the click-through rates of two different website designs. The number of clicks for each design is as follows:**

**Design A: [100, 120, 110, 90, 95]**

**Design B: [80, 85, 90, 95, 100]**

**Perform a chi-square test to determine if there is a significant difference in the click-through rates between the two designs.**

**13. Scenario: A survey is conducted to measure customer satisfaction with a product on a scale of 1 to 10. The data collected is as follows:**

**[7, 9, 6, 8, 10, 7, 8, 9, 7, 8]**

**Calculate the 95% confidence interval for the population mean satisfaction score.**

Ans-To calculate the 95% confidence interval for the population mean satisfaction score based on the given data, we can use the following formula:

1. Calculate the sample mean:

Sample mean = (7 + 9 + 6 + 8 + 10 + 7 + 8 + 9 + 7 + 8) / 10 = 7.9

2. Calculate the sample standard deviation:

Sample standard deviation = √((7 - 7.9)² + (9 - 7.9)² + ... + (8 - 7.9)²) / (10 -1)) = √1.211111111 = 1.101

3. Determine the sample size:

Sample size = 10

4. Determine the z-value for a 95% confidence level:

z-value = 1.96

5. Calculate the margin of error using the formula:

margin of error = z-value \* (sample standard deviation / √sample size):

Margin of error = 1.96 \* (1.101 / √10) = 0.684

6. Calculate the lower and upper bounds of the confidence interval using the formula:

confidence interval = sample mean +/- margin of error:

Lower bound = 7.9 - 0.684 = 7.216

Upper bound = 7.9 + 0.684 = 8.584

Therefore, the 95% confidence interval for the population mean satisfaction score is [7.216, 8.584].

**14. Scenario: A company wants to analyze the effect of temperature on product performance. The data collected is as follows:**

**Temperature (in degrees Celsius): [20, 22, 23, 19, 21]**

**Performance (on a scale of 1 to 10): [8, 7, 9, 6, 8]**

**Perform a simple linear regression to predict performance based on temperature.**

Ans- To perform a simple linear regression and predict performance based on temperature, we can use the least squares method to find the best-fit line that represents the linear relationship between the two variables.

The simple linear regression analysis for the given data:

Temperature (in degrees Celsius): [20, 22, 23, 19, 21]

Performance (on a scale of 1 to 10): [8, 7, 9, 6, 8]

1.Calculate the means of each variable.

Mean of Temperature = (20 + 22 + 23 + 19 + 21) / 5 = 21

Mean of Performance = (8 + 7 + 9 + 6 + 8) / 5 = 7.6

2. Calculate the deviations from the means for each variable.

Deviation of Temperature = [20-21, 22-21, 23-21, 19-21, 21-21] = [-1, 1, 2, -2, 0]

Deviation of Performance = [8-7.6, 7-7.6, 9-7.6, 6-7.6, 8-7.6] = [0.4, -0.6, 1.4, -1.6, 0.4]

3. Calculate the product of the deviations.

Product of Deviations = [-1 \* 0.4, 1 \* -0.6, 2 \* 1.4, -2 \* -1.6, 0 \* 0.4] = [-0.4, -0.6, 2.8, 3.2, 0]

4.Calculate the squared deviations of Temperature.

Squared Deviations of Temperature = [(-1)^2, 1^2, 2^2, (-2)^2, 0^2] = [1, 1, 4, 4, 0]

5. Calculate the sum of the product of the deviations.

Sum of Product of Deviations = -0.4 + -0.6 + 2.8 + 3.2 + 0 = 5

6. Calculate the sum of the squared deviations of Temperature.

Sum of Squared Deviations of Temperature = 1 + 1 + 4 + 4 + 0 = 10

7. Calculate the slope (b) of the regression line.

b = Sum of Product of Deviations / Sum of Squared Deviations = 5 / 10 = 0.5

8. Calculate the intercept (a) of the regression line.

a = Mean of Performance - (b \* Mean of Temperature) = 7.6 - (0.5 \* 21) = -1

9. The equation of the regression line is:

Performance = a + b \* Temperature

Performance = -1 + 0.5 \* Temperature

Therefore, the simple linear regression analysis predicts performance based on temperature using the equation:

Performance = -1 + 0.5 \* Temperature

**15. Scenario: A study is conducted to compare the preferences of two groups of participants. The preferences are measured on a Likert scale from 1 to 5. The data collected is as follows:**

**Group A: [4, 3, 5, 2, 4]**

**Group B: [3, 2, 4, 3, 3]**

**Perform a Mann-Whitney U test to determine if there is a significant difference in the median preferences between the two groups.**

**16. Scenario: A company wants to analyze the distribution of customer ages. The data collected is as follows:**

**[25, 30, 35, 40, 45, 50, 55, 60, 65, 70]**

**Calculate the interquartile range (IQR) of the ages.**

Ans- To Calculate the interquartile range (IQR) of the ages, We can use following formula

1. Order the data from smallest to largest:

[25, 30, 35, 40, 45, 50, 55, 60, 65, 70]

2. Find the median (Q2), which is the middle value of the dataset. Since there are an even number of observations in this case, the median is the average of the two middle values:

Q2 = (40 + 45) / 2 = 42.5

3. Find the first quartile (Q1), which is the median of the lower half of the dataset (not including the median if the number of observations is odd). Since there are an even number of observations in this case, Q1 is the average of the two middle values of the lower half:

Q1 = (30 + 35) / 2 = 32.5

4. Find the third quartile (Q3), which is the median of the upper half of the dataset (not including the median if the number of observations is odd). Since there are an even number of observations in this case, Q3 is the average of the two middle values of the upper half:

Q3 = (55 + 60) / 2 = 57.5

5. Calculate the interquartile range (IQR) using the formula:

IQR = Q3 - Q1

IQR = 57.5 - 32.5 = 25

Therefore, the interquartile range (IQR) of customer ages is 25.

**17. Scenario: A study is conducted to compare the performance of three different machine learning algorithms. The accuracy scores for each algorithm are as follows:**

**Algorithm A: [0.85, 0.80, 0.82, 0.87, 0.83]**

**Algorithm B: [0.78, 0.82, 0.84, 0.80, 0.79]**

**Algorithm C: [0.90, 0.88, 0.89, 0.86, 0.87]**

**Perform a Kruskal-Wallis test to determine if there is a significant difference in the median accuracy scores between the algorithms.**

**18. Scenario: A company wants to analyze the effect of price on sales. The data collected is as follows:**

**Price (in dollars): [10, 15, 12, 8, 14]**

**Sales: [100, 80, 90, 110, 95]**

**Perform a simple linear regression to predict**

**sales based on price.**

Ans- To perform a simple linear regression and predict sales based on price, we can use the least squares method to find the best-fit line that represents the linear relationship between the two variables.

Let's perform the simple linear regression analysis for the given data:

Price (in dollars): [10, 15, 12, 8, 14]

Sales: [100, 80, 90, 110, 95]

Step 1: Calculate the means of each variable.

Mean of Price = (10 + 15 + 12 + 8 + 14) / 5 = 11.8

Mean of Sales = (100 + 80 + 90 + 110 + 95) / 5 = 95

Step 2: Calculate the deviations from the means for each variable.

Deviation of Price = [10-11.8, 15-11.8, 12-11.8, 8-11.8, 14-11.8] = [-1.8, 3.2, 0.2, -3.8, 2.2]

Deviation of Sales = [100-95, 80-95, 90-95, 110-95, 95-95] = [5, -15, -5, 15, 0]

Step 3: Calculate the product of the deviations.

Product of Deviations = [-1.8 \* 5, 3.2 \* -15, 0.2 \* -5, -3.8 \* 15, 2.2 \* 0] = [-9, -48, -1, -57, 0]

Step 4: Calculate the squared deviations of Price.

Squared Deviations of Price = [(-1.8)^2, 3.2^2, 0.2^2, (-3.8)^2, 2.2^2] = [3.24, 10.24, 0.04, 14.44, 4.84]

Step 5: Calculate the sum of the product of the deviations.

Sum of Product of Deviations = -9 + -48 + -1 + -57 + 0 = -115

Step 6: Calculate the sum of the squared deviations of Price.

Sum of Squared Deviations of Price = 3.24 + 10.24 + 0.04 + 14.44 + 4.84 = 32.8

Step 7: Calculate the slope (b) of the regression line.

b = Sum of Product of Deviations / Sum of Squared Deviations = -115 / 32.8 = -3.506

Step 8: Calculate the intercept (a) of the regression line.

a = Mean of Sales - (b \* Mean of Price) = 95 - (-3.506 \* 11.8) ≈ 134.942

Step 9: The equation of the regression line is:

Sales = a + b \* Price

Sales = 134.942 - 3.506 \* Price

Therefore, the simple linear regression analysis predicts sales based on price using the equation:

Sales ≈ 134.942 - 3.506 \* Price

**19. Scenario: A survey is conducted to measure the satisfaction levels of customers with a new product. The data collected is as follows:**

**[7, 8, 9, 6, 8, 7, 9, 7, 8, 7]**

**Calculate the standard error of the mean satisfaction score.**

Ans-To calculate the standard error of the mean satisfaction score, you can use the formula:

Standard Error = Standard Deviation / sqrt(n)

Given the data collected: [7, 8, 9, 6, 8, 7, 9, 7, 8, 7]

1. Calculate the mean of the satisfaction scores.

Mean = (7 + 8 + 9 + 6 + 8 + 7 + 9 + 7 + 8 + 7) / 10 = 7.7

2. Calculate the deviations from the mean for each satisfaction score.

Deviations = [(7 - 7.7), (8 - 7.7), (9 - 7.7), (6 - 7.7), (8 - 7.7), (7 - 7.7), (9 - 7.7), (7 - 7.7), (8 - 7.7), (7 - 7.7)]

= [-0.7, 0.3, 1.3, -1.7, 0.3, -0.7, 1.3, -0.7, 0.3, -0.7]

3. Calculate the sum of the squared deviations.

Sum of Squared Deviations = (-0.7)^2 + 0.3^2 + 1.3^2 + (-1.7)^2 + 0.3^2 + (-0.7)^2 + 1.3^2 + (-0.7)^2 + 0.3^2 + (-0.7)^2

= 0.49 + 0.09 + 1.69 + 2.89 + 0.09 + 0.49 + 1.69 + 0.49 + 0.09 + 0.49

= 8.01

4.Calculate the sample variance.

Variance = Sum of Squared Deviations / (n - 1) = 8.01 / (10 - 1) = 0.890

5. Calculate the standard deviation.

Standard Deviation = sqrt(Variance) = sqrt(0.890) ≈ 0.943

6. Calculate the standard error of the mean.

Standard Error = Standard Deviation / sqrt(n) = 0.943 / sqrt(10) ≈ 0.298

Therefore, the standard error of the mean satisfaction score is approximately 0.298

**20. Scenario: A company wants to analyze the relationship between advertising expenditure and sales. The data collected is as follows:**

**Advertising Expenditure (in thousands): [10, 15, 12, 8, 14]**

**Sales (in thousands): [25, 30, 28, 20, 26]**

**Perform a multiple regression analysis to predict sales based on advertising expenditure.**