Questions on measure of central tendency

1) Business Problem: A retail store wants to analyse the sales data of a particular product category to understand the typical sales performance and make strategic decisions.

Data:

Let's consider the weekly sales data (in units) for the past month for a specific product category:

Week 1: 50 units
Week 2: 60 units
Week 3: 55 units
Week 4: 70 units

Question:

1. Mean: What is the average weekly sale of the product category?
   Answer: 58.75

2. Median: What is the typical or central sales value for the product category?
   Answer: 57.5

3. Mode: Are there any recurring or most frequently occurring sales figures for the product category?
   Answer: No recurring sales figure in given data.

2) Business Problem: A restaurant wants to analyze the waiting times of its customers to understand the typical waiting experience and improve service efficiency.

Data: Let's consider the waiting times (in minutes) for the past 20 customers:

15, 10, 20, 25, 15, 10, 30, 20, 15, 10,
10, 25, 15, 20, 20, 15, 10, 10, 20, 25

Question:

1. Mean: What is the average waiting time for customers at the restaurant?
   Answer: 17

2. Median: What is the typical or central waiting time experienced by customers?
   Answer:15

3. Mode: Are there any recurring or most frequently occurring waiting times for customers?
   Answer:10

3) Business Problem: A car rental company wants to analyze the rental durations of its customers to understand the typical rental period and optimize its pricing and fleet management strategies.

Data: Let's consider the rental durations (in days) for a sample of 50 customers:

3, 2, 5, 4, 7, 2, 3, 3, 1, 6,
4, 2, 3, 5, 2, 4, 2, 1, 3, 5,
6, 3, 2, 1, 4, 2, 4, 5, 3, 2,
7, 2, 3, 4, 5, 1, 6, 2, 4, 3,
5, 3, 2, 4, 2, 6, 3, 2, 4, 5

Question:

1. Mean: What is the average rental duration for customers at the car rental company?
   Answer: 3.44
2. Median: What is the typical or central rental duration experienced by customers?
   Answer: 3
3. Mode: Are there any recurring or most frequently occurring rental durations for customers?
   Answer: 2

Questions on measure of dispersion

1) Problem: A manufacturing company wants to analyze the production output of a specific machine to understand the variability or spread in its performance.

Data: Let's consider the number of units produced per hour by the machine for a sample of 10 working days:

Day 1: 120 units
Day 2: 110 units
Day 3: 130 units
Day 4: 115 units
Day 5: 125 units
Day 6: 105 units
Day 7: 135 units
Day 8: 115 units
Day 9: 125 units
Day 10: 140 units

Question:

1. Range: What is the range of the production output for the machine?
   Answer: 35

2. Variance: What is the variance of the production output for the machine?
   Answer: 111

3. Standard Deviation: What is the standard deviation of the production output for the machine?
   Answer: 10.53

2) Problem: A retail store wants to analyze the sales of a specific product to understand the variability in daily sales and assess its inventory management.

Data: Let's consider the daily sales (in dollars) for the past 30 days:

$500, $700, $400, $600, $550, $750, $650, $500, $600, $550,
$800, $450, $700, $550, $600, $400, $650, $500, $750, $550,
$700, $600, $500, $800, $550, $650, $400, $600, $750, $550

Questions:

1. Range: What is the range of the daily sales?
   Answer: 350

2. Variance: What is the variance of the daily sales?
   Answer: 12725

3. Standard Deviation: What is the standard deviation of the daily sales?
   Answer: 112.80

3) Problem: An e-commerce platform wants to analyze the delivery times of its shipments to understand the variability in order fulfillment and optimize its logistics operations.

Data: Let's consider the delivery times (in days) for a sample of 50 shipments:
3, 5, 2, 4, 6, 2, 3, 4, 2, 5,
7, 2, 3, 4, 2, 4, 2, 3, 5, 6,
3, 2, 1, 4, 2, 4, 5, 3, 2, 7,
2, 3, 4, 5, 1, 6, 2, 4, 3, 5,
3, 2, 4, 2, 6, 3, 2, 4, 5, 3

Questions:

1. Range: What is the range of the delivery times?
   Answer: 6
2. Variance: What is the variance of the delivery times?
   Answer: 2.289

3. Standard Deviation: What is the standard deviation of the delivery times?

Answer: 1.513

4) Problem : A company wants to analyze the monthly revenue generated by one of its products to understand its performance and variability.

Data: Let's consider the monthly revenue (in thousands of dollars) for the past 12 months:
$120, $150, $110, $135, $125, $140, $130, $155, $115, $145, $135, $130

Questions:

1. Measure of Central Tendency: What is the average monthly revenue for the product?

Answer: 132.5

2. Measure of Dispersion: What is the range of monthly revenue for the product?

Answer: 45

5) Problem : A survey was conducted to gather feedback from customers regarding their satisfaction with a particular service on a scale of 1 to 10.

Data: Let's consider the satisfaction ratings from 50 customers:
8, 7, 9, 6, 7, 8, 9, 8, 7, 6,
8, 9, 7, 8, 7, 6, 8, 9, 6, 7,
8, 9, 7, 6, 7, 8, 9, 8, 7, 6,
9, 8, 7, 6, 8, 9, 7, 8, 7, 6,
9, 8, 7, 6, 7, 8, 9, 8, 7, 6

Questions:

1. Measure of Central Tendency: What is the average satisfaction rating?

Answer: 7.5

2. Measure of Dispersion: What is the standard deviation of the satisfaction ratings?

Answer: 1.024

6) Problem :A company wants to analyze the customer wait times at its call center to assess the efficiency of its customer service operations.

Data:
Let's consider the wait times (in minutes) for a sample of 100 randomly selected customer calls:
10, 15, 12, 18, 20, 25, 8, 14, 16, 22,
9, 17, 11, 13, 19, 23, 21, 16, 24, 27,
13, 10, 18, 16, 12, 14, 19, 21, 11, 17,
15, 20, 26, 13, 12, 14, 22, 19, 16, 11,
25, 18, 16, 13, 21, 20, 15, 12, 19, 17,
14, 16, 23, 18, 15, 11, 19, 22, 17, 12,
16, 14, 18, 20, 25, 13, 11, 22, 19, 17,
15, 16, 13, 14, 18, 20, 19, 21, 17, 12,
15, 13, 16, 14, 22, 21, 19, 18, 16, 11,
17, 14, 12, 20, 23, 19, 15, 16, 13, 18

Questions:

1. Measure of Central Tendency: What is the average wait time for customers at the call center?

Answer: 16.74

2. Measure of Dispersion: What is the range of wait times for customers at the call center?

Answer: 19

3. Measure of Dispersion: What is the standard deviation of the wait times for customers at the call center?

Answer: 4.122

7) Problem : A transportation company wants to analyze the fuel efficiency of its vehicle fleet to identify any variations across different vehicle models.

Data: Let's consider the fuel efficiency (in miles per gallon, mpg) for a sample of 50 vehicles:

Model A: 30, 32, 33, 28, 31, 30, 29, 30, 32, 31,
Model B: 25, 27, 26, 23, 28, 24, 26, 25, 27, 28,
Model C: 22, 23, 20, 25, 21, 24, 23, 22, 25, 24,
Model D: 18, 17, 19, 20, 21, 18, 19, 17, 20, 19,
Model E: 35, 36, 34, 35, 33, 34, 32, 33, 36, 34

Questions:

1. Measure of Central Tendency: What is the average fuel efficiency for each vehicle model?

Answer:

MODEL	7.1
Model A	30.6
Model B	25.9
Model C	22.9
Model D	18.8
Model E	34.2

2. Measure of Dispersion: What is the range of fuel efficiency for each vehicle model?

Answer:

MODEL	7.2
Model A	5
Model B	5
Model C	5
Model D	4
Model E	4

3. Measure of Dispersion: What is the variance of the fuel efficiency for each vehicle model?

Answer:

MODEL	7.3
Model A	2.04
Model B	2.49
Model C	2.49
Model D	1.56
Model E	1.56

More Statistics Questions

8) Problem : A company wants to analyze the ages of its employees to understand the age distribution and demographics within the organization.

Data:

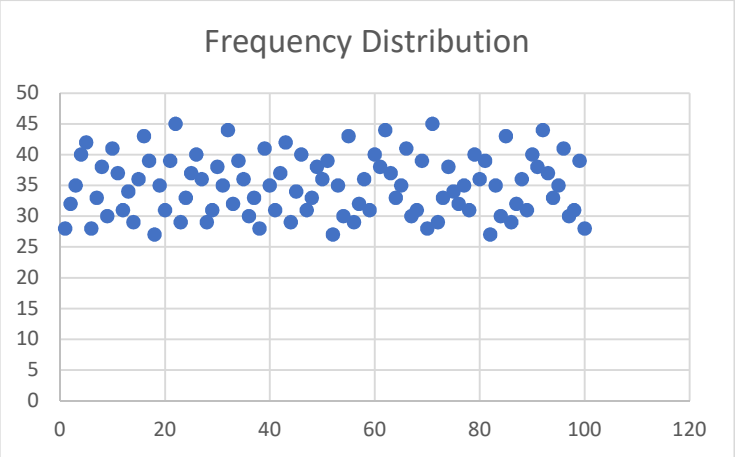
Let's consider the ages of 100 employees:

28, 32, 35, 40, 42, 28, 33, 38, 30, 41,
37, 31, 34, 29, 36, 43, 39, 27, 35, 31,
39, 45, 29, 33, 37, 40, 36, 29, 31, 38,
35, 44, 32, 39, 36, 30, 33, 28, 41, 35,
31, 37, 42, 29, 34, 40, 31, 33, 38, 36,
39, 27, 35, 30, 43, 29, 32, 36, 31, 40,
38, 44, 37, 33, 35, 41, 30, 31, 39, 28,
45, 29, 33, 38, 34, 32, 35, 31, 40, 36,
39, 27, 35, 30, 43, 29, 32, 36, 31, 40,
38, 44, 37, 33, 35, 41, 30, 31, 39, 28

Questions:

1. Frequency Distribution: Create a frequency distribution table for the ages of the employees.

Answer:



2. Mode: What is the mode (most common age) among the employees?
Answer: 31

3. Median: What is the median age of the employees?
Answer: 35

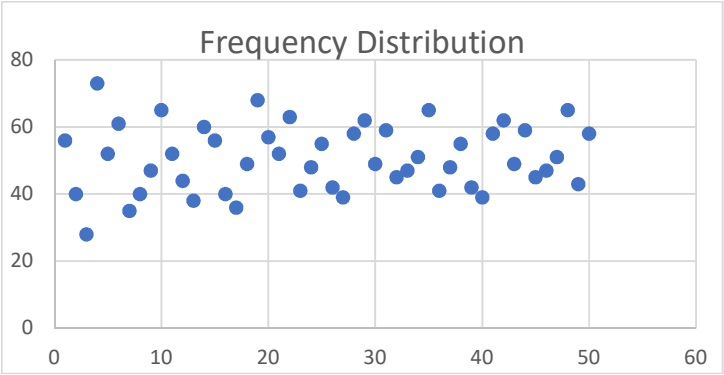
4. Range: What is the range of ages among the employees?
Answer: 18

9) Problem :A retail store wants to analyze the purchase amounts made by customers to understand their spending habits.

Data: Let's consider the purchase amounts (in dollars) for a sample of 50 customers:
56, 40, 28, 73, 52, 61, 35, 40, 47, 65,
52, 44, 38, 60, 56, 40, 36, 49, 68, 57,
52, 63, 41, 48, 55, 42, 39, 58, 62, 49,
59, 45, 47, 51, 65, 41, 48, 55, 42, 39,
58, 62, 49, 59, 45, 47, 51, 65, 43, 58

Questions:

1. Frequency Distribution: Create a frequency distribution table for the purchase amounts.
Answer:



2. Mode: What is the mode (most common purchase amount) among the customers?
Answer: 40

3. Median: What is the median purchase amount among the customers?
Answer: 50

4. Interquartile Range: What is the interquartile range of the purchase amounts?
Answer: Range = 15.75 [Q3:58 – Q1:42.25]

10) Problem : A manufacturing company wants to analyze the defect rates of its production line to identify the frequency of different types of defects.

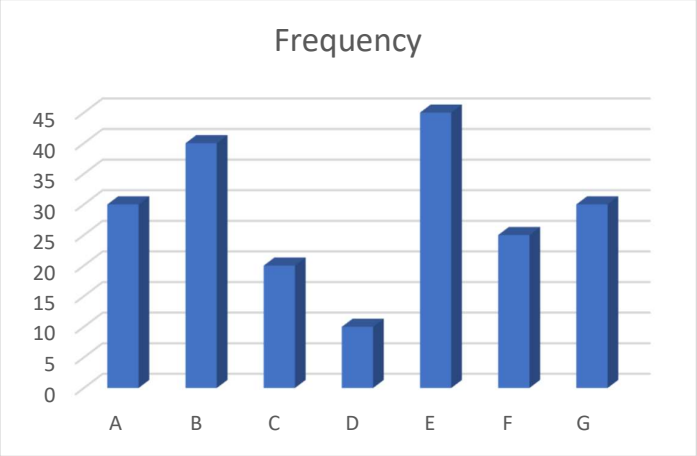
Data: Let's consider the types of defects and their corresponding frequencies observed in a sample of 200 products:

Defect Type: A, B, C, D, E, F, G
Frequency: 30, 40, 20, 10, 45, 25, 30

Questions:

1. Bar Chart: Create a bar chart to visualize the frequency of different defect types.

Answer :

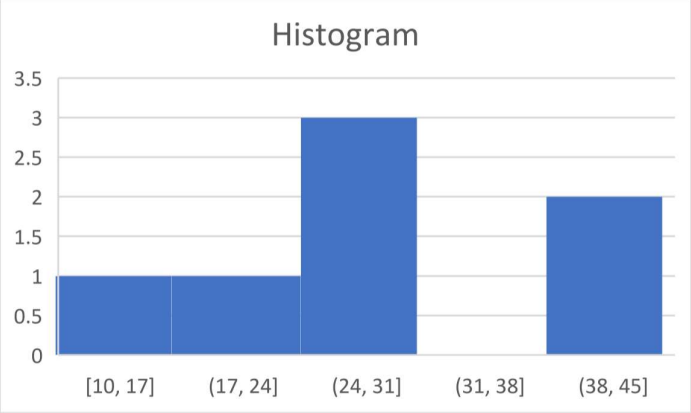


2. Most Common Defect: Which defect type has the highest frequency?

Answer : Type E = 45

3. Histogram: Create a histogram to represent the defect frequencies.

Answer :



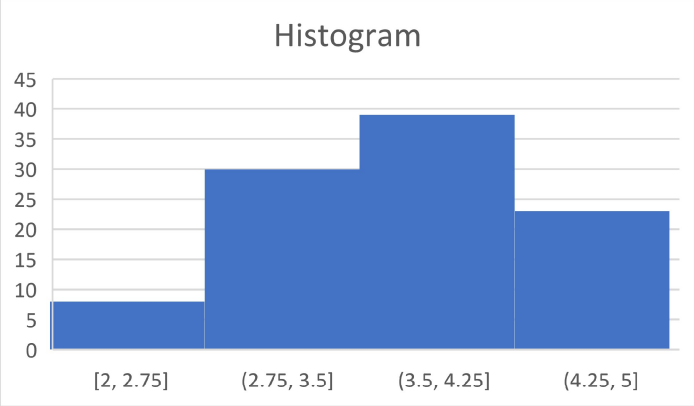
11) Problem : A survey was conducted to gather feedback from customers about their satisfaction levels with a specific service on a scale of 1 to 5.

Data: Let's consider the satisfaction ratings from 100 customers: Ratings:
4, 5, 3, 4, 4, 3, 2, 5, 4, 3,
5, 4, 2, 3, 4, 5, 3, 4, 5, 3,
4, 3, 2, 4, 5, 3, 4, 5, 4, 3,
3, 4, 5, 2, 3, 4, 4, 3, 5, 4,
3, 4, 5, 4, 2, 3, 4, 5, 3, 4,
5, 4, 3, 4, 5, 3, 4, 5, 4, 3,
3, 4, 5, 2, 3, 4, 4, 3, 5, 4,
3, 4, 5, 4, 2, 3, 4, 5, 3, 4,
5, 4, 3, 4, 5, 3, 4, 5, 4, 3,
3, 4, 5, 2, 3, 4, 4, 3, 5, 4

Questions:

1. Histogram: Create a histogram to visualize the distribution of satisfaction ratings.

Answer :

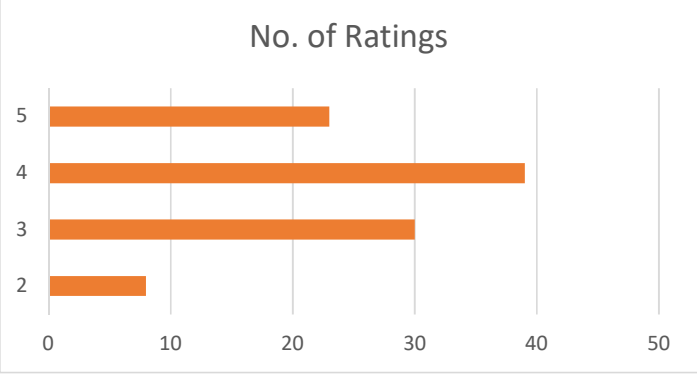


2. Mode: Which satisfaction rating has the highest frequency?

Answer: 4

3. Bar Chart: Create a bar chart to display the frequency of each satisfaction rating.

Answer :



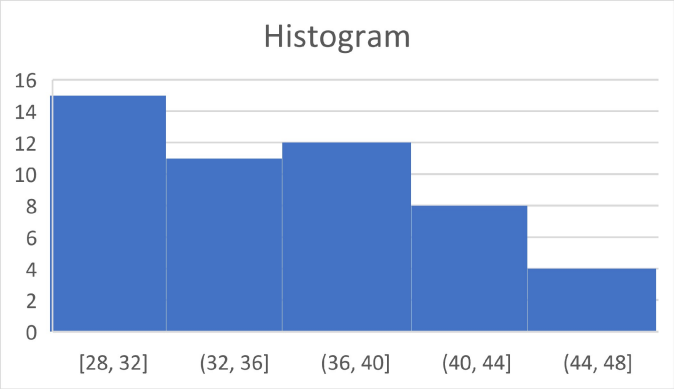
12) Problem : A company wants to analyze the monthly sales figures of its products to understand the sales distribution across different price ranges.

Data: Let's consider the monthly sales figures (in thousands of dollars) for a sample of 50 products:
Sales:
35, 28, 32, 45, 38, 29, 42, 30, 36, 41,
47, 31, 39, 43, 37, 30, 34, 39, 28, 33,
36, 40, 42, 29, 31, 45, 38, 33, 41, 35,
37, 34, 46, 30, 39, 43, 28, 32, 36, 29,
31, 37, 40, 42, 33, 39, 28, 35, 38, 43

Questions:

1. Histogram: Create a histogram to visualize the sales distribution across different price ranges.

Answer :

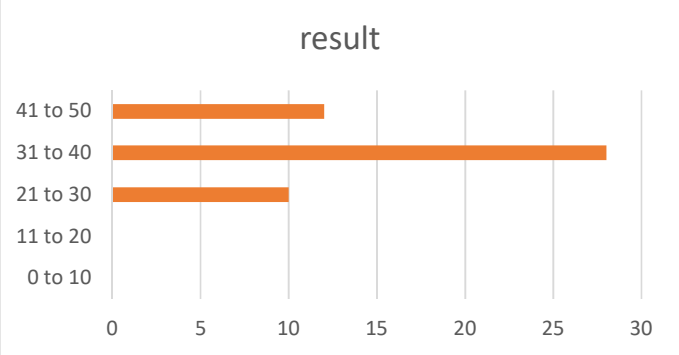


2. Measure of Central Tendency: What is the average monthly sales figure?

Answer : 36.14

3. Bar Chart: Create a bar chart to display the frequency of sales in different price ranges.

Answer :

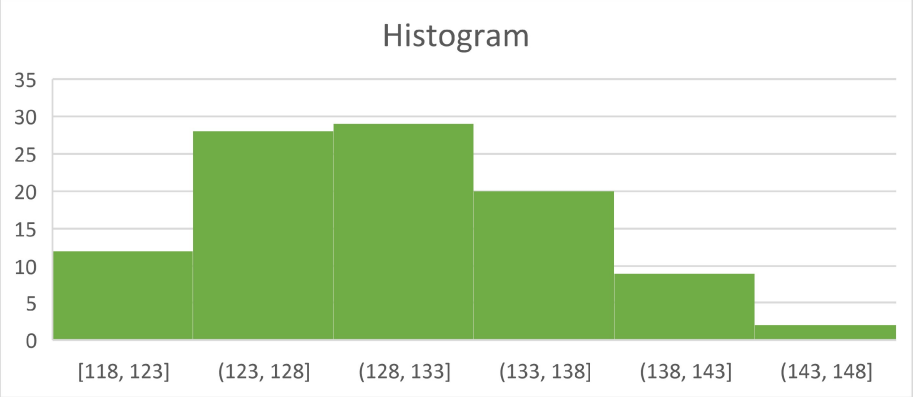


13) Problem : A study was conducted to analyze the response times of a website for different user locations.

Data: Let's consider the response times (in milliseconds) for a sample of 200 user requests:
Response Times:
125, 148, 137, 120, 135, 132, 145, 122, 130, 141,
118, 125, 132, 136, 128, 123, 132, 138, 126, 129,
136, 127, 130, 122, 125, 133, 140, 126, 133, 135,
130, 134, 141, 119, 125, 131, 136, 128, 124, 132,
136, 127, 130, 122, 125, 133, 140, 126, 133, 135,
130, 134, 141, 119, 125, 131, 136, 128, 124, 132,
136, 127, 130, 122, 125, 133, 140, 126, 133, 135,
130, 134, 141, 119, 125, 131, 136, 128, 124, 132,
136, 127, 130, 122, 125, 133, 140, 126, 133, 135,
130, 134, 141, 119, 125, 131, 136, 128, 124, 132
Questions:

1. Histogram: Create a histogram to visualize the distribution of response times.

Answer:

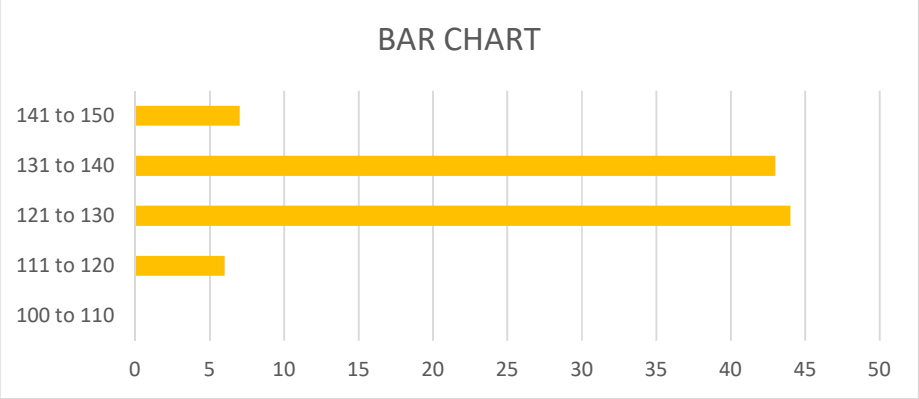


2. Measure of Central Tendency: What is the median response time?

Answer: 130.5

3. Bar Chart: Create a bar chart to display the frequency of response times within different ranges.

Answer:



14) Problem : A company wants to analyze the sales performance of its products across different regions.

Data: Let's consider the sales figures (in thousands of dollars) for a sample of 50 products in three regions:

Region 1: 45, 35, 40, 38, 42, 37, 39, 43, 44, 41,

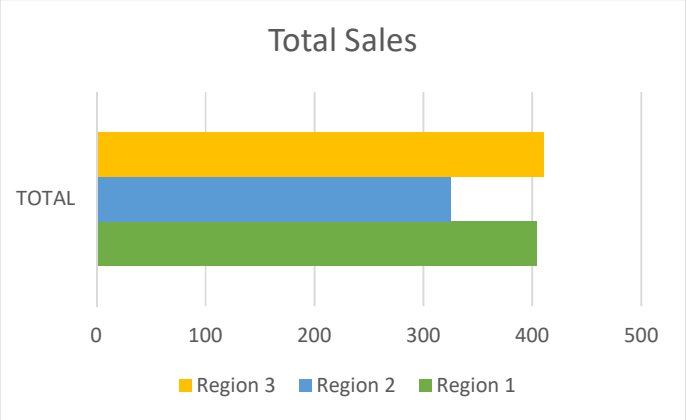
Region 2: 32, 28, 30, 34, 33, 35, 31, 29, 36, 37,

Region 3: 40, 39, 42, 41, 38, 43, 45, 44, 41, 37

Questions:

1. Bar Chart: Create a bar chart to compare the sales figures across the three regions.

Answer:



2. Measure of Central Tendency: What is the average sales figure for each region?

Answer:

R	Average
Region 1	40.4
Region 2	32.5
Region 3	41

3. Measure of Dispersion : What is the range of sales figures in each region?

Answer:

R	RANGE
Region 1	10
Region 2	9
Region 3	8

END OF TOPIC

Questions on Percentile and Quartiles

1) Question: A company wants to analyze the salary distribution of its employees to determine the income levels at different percentiles.

Data: Let's consider the monthly salaries (in thousands of dollars) of a sample of 200 employees:
Salaries:
40, 45, 50, 55, 60, 62, 65, 68, 70, 72,
75, 78, 80, 82, 85, 88, 90, 92, 95, 100,
105, 110, 115, 120, 125, 130, 135, 140, 145, 150,
155, 160, 165, 170, 175, 180, 185, 190, 195, 200,
205, 210, 215, 220, 225, 230, 235, 240, 245, 250,
255, 260, 265, 270, 275, 280, 285, 290, 295, 300,
305, 310, 315, 320, 325, 330, 335, 340, 345, 350,
355, 360, 365, 370, 375, 380, 385, 390, 395, 400,
405, 410, 415, 420, 425, 430, 435, 440, 445, 450,
455, 460, 465, 470, 475, 480, 485, 490, 495, 500

Questions:

1. Quartiles: Calculate the first quartile (Q1), median (Q2), and third quartile (Q3) of the salary distribution.

Answer:

Q1	**128.75**
Q2	**252.50**
Q3	**376.25**

2. Percentiles: Calculate the 10th percentile, 25th percentile, 75th percentile, and 90th percentile of the salary distribution.

Answer:

10th	**74.7**
25th	**128.75**
75th	**376.25**
90th	**450.5**

3. Interpretation: Based on the quartiles and percentiles, what can be inferred about the income distribution of the employees?

Answer: By looking at the Quartiles and percentiles we can say the income is almost equally distributed in the given range.

2) Question : A research study wants to analyze the weight distribution of a sample of individuals to assess their health and body composition.

Data: Let's consider the weights (in kilograms) of a sample of 100 individuals:

Weights:
55, 60, 62, 65, 68, 70, 72, 75, 78, 80,
82, 85, 88, 90, 92, 95, 100, 105, 110, 115,
120, 125, 130, 135, 140, 145, 150, 155, 160, 165,
170, 175, 180, 185, 190, 195, 200, 205, 210, 215,
220, 225, 230, 235, 240, 245, 250, 255, 260, 265,
270, 275, 280, 285, 290, 295, 300, 305, 310, 315,
320, 325, 330, 335, 340, 345, 350, 355, 360, 365,
370, 375, 380, 385, 390, 395, 400, 405, 410, 415,
420, 425, 430, 435, 440, 445, 450, 455, 460, 465,
470, 475, 480, 485, 490, 495, 500, 505, 510, 515

Questions:

1. Quartiles: Calculate the first quartile (Q1), median (Q2), and third quartile (Q3) of the weight distribution.

Answer:

Q1	**143.75**
Q2	**267.5**
Q3	**391.25**

2. Percentiles: Calculate the 15th percentile, 50th percentile, and 85th percentile of the weight distribution.

Answer:

15th	**94.55**
50th	**252.5**
85th	**425.75**

3. Interpretation: Based on the quartiles and percentiles, what can be inferred about the weight distribution of the individuals?

Answer: No Major conclusion, but we can see slight difference between Q2 and 50th percentile.

Name – Kaushik Khodiya
Last Updated on 13 Nov 2023

3) Question : A retail store wants to analyze the distribution of customer purchase amounts to identify their spending patterns.

Data: Let's consider the purchase amounts (in dollars) of a sample of 150 customers:
Purchase Amounts:

20, 25, 30, 35, 40, 45, 50, 55, 60, 65,
70, 75, 80, 85, 90, 95, 100, 105, 110, 115,
120, 125, 130, 135, 140, 145, 150, 155, 160, 165,
170, 175, 180, 185, 190, 195, 200, 205, 210, 215,
220, 225, 230, 235, 240, 245, 250, 255, 260, 265,
270, 275, 280, 285, 290, 295, 300, 305, 310, 315,
320, 325, 330, 335, 340, 345, 350, 355, 360, 365,
370, 375, 380, 385, 390, 395, 400, 405, 410, 415,
420, 425, 430, 435, 440, 445, 450, 455, 460, 465,
470, 475, 480, 485, 490, 495, 500, 505, 510, 515,
520, 525, 530, 535, 540, 545, 550, 555, 560, 565

Questions:

1. Quartiles: Calculate the first quartile (Q1), median (Q2), and third quartile (Q3) of the purchase amount distribution.

Answer:

Q1	**156.25**
Q2	**292.5**
Q3	**428.75**

2. Percentiles: Calculate the 20th percentile, 40th percentile, and 80th percentile of the purchase amount distribution.

Answer:

20th	**129**
40th	**238**
80th	**456**

3. Interpretation: Based on the quartiles and percentiles, what can be inferred about the spending patterns of the customers?

Answer: By looking at the results we can say, the upper 20% customers spends approx. 4x times compared to lower 20% customers.

4) Question : A study wants to analyze the distribution of commute times of employees to determine the average time spent traveling to work.

Data:
Let's consider the commute times (in minutes) of a sample of 250 employees:
Commute Times:

15, 20, 25, 30, 35, 40, 45, 50, 55, 60,
65, 70, 75, 80, 85, 90, 95, 100, 105, 110,
115, 120, 125, 130, 135, 140, 145, 150, 155, 160,
165, 170, 175, 180, 185, 190, 195, 200, 205, 210,
215, 220, 225, 230, 235, 240, 245, 250, 255, 260,
265, 270, 275, 280, 285, 290, 295, 300, 305, 310,
315, 320, 325, 330, 335, 340, 345, 350, 355, 360,
365, 370, 375, 380, 385, 390, 395, 400, 405, 410,
415, 420, 425, 430, 435, 440, 445, 450, 455, 460,
465, 470, 475, 480, 485, 490, 495, 500, 505, 510,
515, 520, 525, 530, 535, 540, 545, 550, 555, 560,
565, 570, 575, 580, 585, 590, 595, 600, 605, 610

Questions:

1. Quartiles: Calculate the first quartile (Q1), median (Q2), and third quartile (Q3) of the commute time distribution.

Answer:

Q1	**163.75**
Q2	**312.5**
Q3	**461.25**

2. Percentiles: Calculate the 30th percentile, 50th percentile, and 70th percentile of the commute time distribution.

Answer:

30th	**193.5**
50th	**312.5**
70th	**431.5**

3. Interpretation: Based on the quartiles and percentiles, what can be inferred about the average commute time of the employees?

Answer: Given Data is equally distributed.

5) Question : A manufacturing company wants to analyze the defect rates in its production process to evaluate product quality.

Data:
Let's consider the defect rates (in percentage) for a sample of 120 products:

Defect Rates:
0.5, 1.0, 0.2, 0.7, 0.3, 0.9, 1.2, 0.6, 0.4, 1.1,
0.8, 0.5, 0.3, 0.6, 1.0, 0.4, 0.5, 0.7, 0.9, 1.3,
0.8, 0.6, 0.4, 0.7, 0.9, 0.5, 0.2, 1.0, 0.8, 0.3,
0.6, 0.4, 0.7, 0.9, 1.2, 0.8, 0.3, 0.6, 0.5, 0.4,
0.7, 0.9, 1.1, 0.3, 1.4, 0.9, 0.6, 0.2, 1.5, 1.0
0.6, 0.4, 0.7, 1.0, 0.8, 0.3, 0.5, 0.8, 0.6, 0.3,
0.4, 0.7, 0.9, 1.0, 0.8, 0.3, 0.5, 0.6, 0.4, 0.7,
0.9, 1.1, 0.8, 0.3, 0.5, 0.6, 0.4, 0.7, 0.9, 1.0,
0.8, 0.3, 0.5, 0.6, 0.4, 0.7, 0.9, 1.1, 0.8, 0.3,
0.5, 0.6, 0.4, 0.7, 0.9, 1.0, 0.8, 0.3, 0.5, 0.6,
0.4, 0.7, 0.9, 1.1, 0.8, 0.3, 0.5, 0.6, 0.4, 0.7,
0.9, 1.0, 0.8, 0.3, 0.5, 0.6, 0.4, 0.7, 0.9, 1.1

Questions:

1. Quartiles: Calculate the first quartile (Q1), median (Q2), and third quartile (Q3) of the defect rate distribution.

Answer:

Q1	**0.4**
Q2	**0.7**
Q3	**0.9**

2. Percentiles: Calculate the 25th percentile, 50th percentile, and 75th percentile of the defect rate distribution.

Answer:

25th	**0.4**
50th	**0.7**
75th	**0.9**

3. Interpretation: Based on the quartiles and percentiles, what can be inferred about the quality of the products?

Answer: By the results, we can say most of the time defect rate of the product remains below 1.

Questions on Correlation and Covariance

1) Question : A marketing department wants to understand the relationship between advertising expenditure and sales revenue to assess the effectiveness of their advertising campaigns.

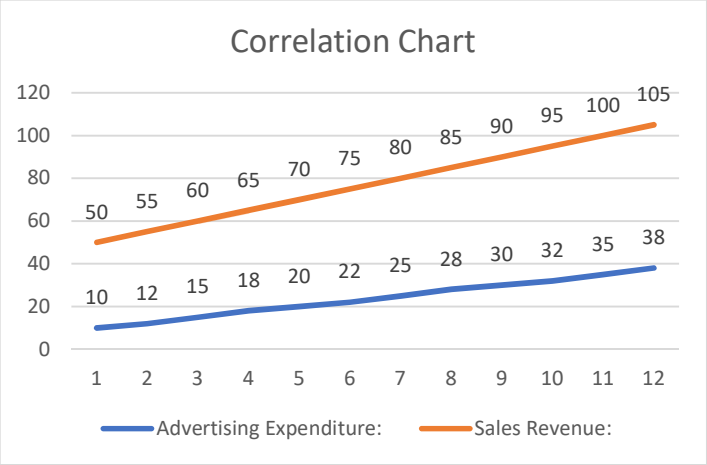
Data:
Let's consider the monthly advertising expenditure (in thousands of dollars) and corresponding sales revenue (in thousands of dollars) for a sample of 12 months:

Advertising Expenditure: 10, 12, 15, 18, 20, 22, 25, 28, 30, 32, 35, 38
Sales Revenue: 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105

Question:

Calculate the correlation coefficient between advertising expenditure and sales revenue. Interpret the value of the correlation coefficient and explain the nature of the relationship between advertising expenditure and sales revenue.

Answer: Correlation Coefficient: 0.999



Conclusion: Increase in advertising expenditure will lead to increase in sales revenue.

By analyzing the correlation coefficient, the marketing department can determine the strength and direction of the relationship between advertising expenditure and sales revenue. This information can help them make informed decisions about allocating their advertising budget and optimizing their marketing strategies.

2) Question : An investment analyst wants to assess the relationship between the stock prices of two companies to identify potential investment opportunities.

Data:

Let's consider the daily closing prices (in dollars) of Company A and Company B for a sample of 20 trading days:

Company A: 45, 47, 48, 50, 52, 53, 55, 56, 58, 60, 62, 64, 65, 67, 69, 70, 72, 74, 76, 77

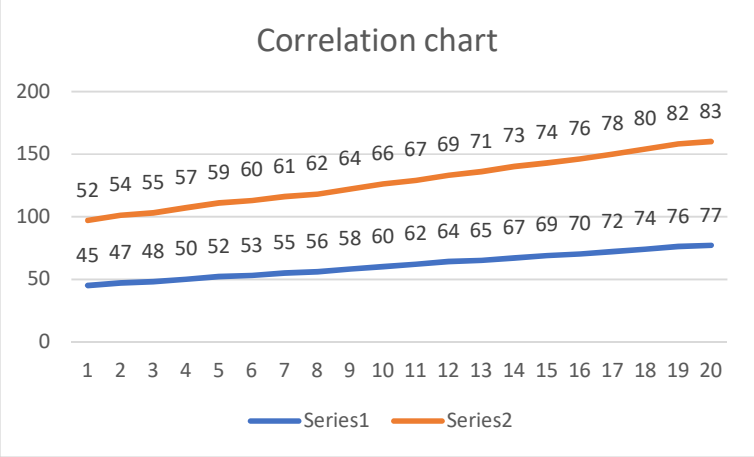
Company B: 52, 54, 55, 57, 59, 60, 61, 62, 64, 66, 67, 69, 71, 73, 74, 76, 78, 80, 82, 83

Question:

Calculate the covariance between the stock prices of Company A and Company B. Interpret the value of the covariance and explain the nature of the relationship between the two stocks.

Answer :

Covariance = 92.65



Conclusion: Positive Covariance indicates that the stock prices of company A and B will move up or down at the same time

By analyzing the covariance, the investment analyst can determine whether the stock prices of Company A and Company B move together (positive covariance) or in opposite directions (negative covariance). This information can assist in identifying potential investment opportunities and understanding the diversification benefits of combining these stocks in a portfolio.

3) Question : A researcher wants to examine the relationship between the hours spent studying and the exam scores of a group of students.

Data:
Let's consider the number of hours spent studying and the corresponding exam scores for a sample of 30 students:

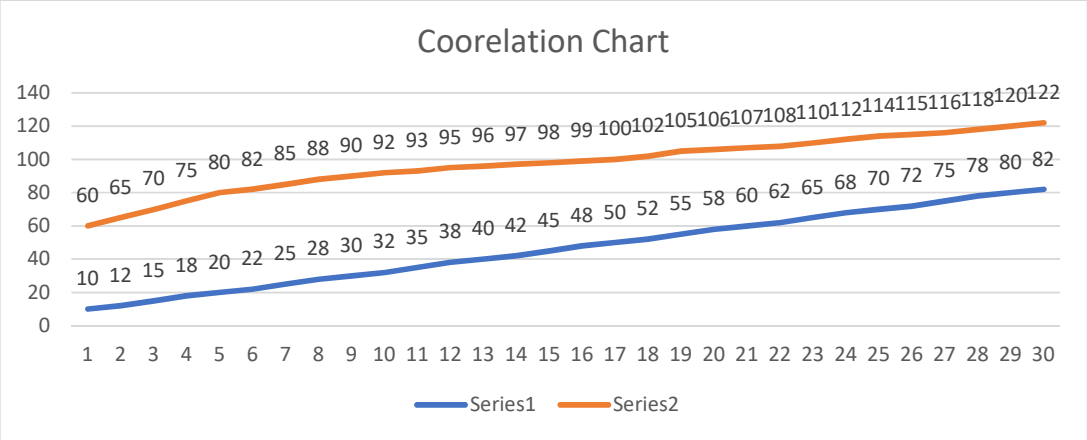
Hours Spent Studying:
10, 12, 15, 18, 20, 22, 25, 28, 30, 32, 35, 38, 40, 42, 45, 48, 50, 52, 55, 58, 60, 62, 65, 68, 70, 72, 75, 78, 80, 82

Exam Scores:
60, 65, 70, 75, 80, 82, 85, 88, 90, 92, 93, 95, 96, 97, 98, 99, 100, 102, 105, 106, 107, 108, 110, 112, 114, 115, 116, 118, 120, 122

Question:

Calculate the correlation coefficient between the hours spent studying and the exam scores. Interpret the value of the correlation coefficient and explain the nature of the relationship between studying hours and exam scores.

Answer:
Correlation coefficient : 0.997



Conclusion: Correlation Coefficient of 0.997 indicates nearby perfect linear relationship between two datasets. Hence, more hours spend studying will lead to more marks in exam.

By analyzing the correlation coefficient, the researcher can determine the strength and direction of the relationship between studying hours and exam scores. This information can provide insights into the effectiveness of studying and help students and educators make informed decisions about study habits and academic performance.

END OF TOPIC

Questions on Measure of Skewness and Kurtosis

1) Question : A company wants to analyze the monthly returns of its investment portfolio to understand the distribution and risk associated with the returns.

Data:

Let's consider the monthly returns (%) for the portfolio over a one-year period:

Returns:

-2.5, 1.3, -0.8, -1.9, 2.1, 0.5, -1.2, 1.8, -0.5, 2.3,
-0.7, 1.2, -1.5, -0.3, 2.6, 1.1, -1.7, 0.9, -1.4, 0.3,
1.9, -1.1, -0.4, 2.2, -0.9, 1.6, -0.6, -1.3, 2.4, 0.7,
-1.8, 1.5, -0.2, -2.1, 2.8, 0.8, -1.6, 1.4, -0.1, 2.5,
-1.0, 1.7, -0.9, -2.0, 2.7, 0.6, -1.4, 1.1, -0.3, 2.0

Questions:

1. Skewness: Calculate the skewness of the monthly returns.

   Answer: 0.054546

2. Kurtosis: Calculate the kurtosis of the monthly returns.

   Answer: -1.30425

3. Interpretation: Based on the skewness and kurtosis values, what can be said about the distribution of returns?

   Answer: 0.05 Skewness indicates equal distribution of data and -1.30425 Kurtosis indicates no outliers in data.

2) Question : A research study wants to analyze the income distribution of a population to understand the level of income inequality.

Data:

Let's consider the monthly incomes (in thousands of dollars) of a sample of 96 individuals:

Incomes:

2.5, 4.8, 3.2, 2.1, 4.5, 2.9, 2.3, 3.1, 4.2, 3.9,
2.8, 4.1, 2.6, 2.4, 4.7, 3.3, 2.7, 3.0, 4.3, 3.7,
2.2, 3.6, 4.0, 2.7, 3.8, 3.5, 3.2, 4.4, 2.0, 3.4,
3.1, 2.9, 4.6, 3.3, 2.5, 4.9, 2.8, 3.0, 4.2, 3.9,
2.8, 4.1, 2.6, 2.4, 4.7, 3.3, 2.7, 3.0, 4.3, 3.7,
2.2, 3.6, 4.0, 2.7, 3.8, 3.5, 3.2, 4.4, 2.0, 3.4,
3.1, 2.9, 4.6, 3.3, 2.5, 4.9, 2.8, 3.0, 4.2, 3.9,
2.8, 4.1, 2.6, 2.4, 4.7, 3.3, 2.7, 3.0, 4.3, 3.7,
2.2, 3.6, 4.0, 2.7, 3.8, 3.5, 3.2, 4.4, 2.0, 3.4,
3.1, 2.9, 4.6, 3.3, 2.5, 4.9

Questions:

1. Skewness: Calculate the skewness of the income distribution.

   Answer: 0.224025

2. Kurtosis: Calculate the kurtosis of the income distribution.

   Answer: -0.93121

3. Interpretation: Based on the skewness and kurtosis values, what can be inferred about the income inequality?

   Answer: 0.224 Skewness indicates equal distribution of data and -0.93121 Kurtosis indicates no outliers in data.

3) Question : A survey was conducted to analyze the satisfaction ratings of customers on a scale of 1 to 5 for a specific product.

Data:

Let's consider the satisfaction ratings from 100 customers:

Ratings:

4, 5, 3, 4, 4, 3, 2, 5, 4, 3,
5, 4, 2, 3, 4, 5, 3, 4, 5, 3,
4, 3, 2, 4, 5, 3, 4, 5, 4, 3,
3, 4, 5, 2, 3, 4, 4, 3, 5, 4,
3, 4, 5, 4, 2, 3, 4, 5, 3, 4,
5, 4, 3, 4, 5, 3, 4, 5, 4, 3,
3, 4, 5, 2, 3, 4, 4, 3, 5, 4,
3, 4, 5, 4, 2, 3, 4, 5, 3, 4,
5, 4, 3, 4, 5, 3, 4, 5, 4, 3,
3, 4, 5, 2, 3, 4, 4, 3, 5, 4

Questions:

1. Skewness: Calculate the skewness of the satisfaction ratings.

   Answer: -0.21091

2. Kurtosis: Calculate the kurtosis of the satisfaction ratings.

   Answer: -0.74526

3. Interpretation: Based on the skewness and kurtosis values, what can be inferred about the satisfaction ratings distribution?

   Answer: -0.210 Skewness indicates equal distribution of data and -0.745 Kurtosis indicates no outliers in data.

4) Question : A study wants to analyze the distribution of house prices in a specific city to understand the market trends.

Data:

Let's consider the house prices (in thousands of dollars) for a sample of 100 houses:

House Prices:

280, 350, 310, 270, 390, 320, 290, 340, 310, 380,
270, 350, 300, 330, 370, 310, 280, 320, 350, 290,
270, 350, 300, 330, 370, 310, 280, 320, 350, 290,
270, 350, 300, 330, 370, 310, 280, 320, 350, 290,
270, 350, 300, 330, 370, 310, 280, 320, 350, 290,
270, 350, 300, 330, 370, 310, 280, 320, 350, 290,
270, 350, 300, 330, 370, 310, 280, 320, 350, 290,
270, 350, 300, 330, 370, 310, 280, 320, 350, 290,
270, 350, 300, 330, 370, 310, 280, 320, 350, 290,
270, 350, 300, 330, 370, 310, 280, 320, 350, 290

Questions:

1. Skewness: Calculate the skewness of the house price distribution.

   Answer: 0.209219

2. Kurtosis: Calculate the kurtosis of the house price distribution.

   Answer: -1.03742

3. Interpretation: Based on the skewness and kurtosis values, what can be inferred about the distribution of house prices?

   Answer: 0.209 Skewness indicates equal distribution of data and -1.037 Kurtosis indicates no outliers in data.

5) Question : A company wants to analyze the waiting times of customers at a service center to improve operational efficiency.

Data:

Let's consider the waiting times (in minutes) for a sample of 100 customers:

Waiting Times:

12, 18, 15, 22, 20, 14, 16, 21, 19, 17,
22, 19, 13, 16, 21, 22, 17, 19, 22, 18,
14, 20, 19, 17, 22, 18, 15, 21, 20, 16,
12, 18, 15, 22, 20, 14, 16, 21, 19, 17,
22, 19, 13, 16, 21, 22, 17, 19, 22, 18,
14, 20, 19, 17, 22, 18, 15, 21, 20, 16,
12, 18, 15, 22, 20, 14, 16, 21, 19, 17,
22, 19, 13, 16, 21, 22, 17, 19, 22, 18,
14, 20, 19, 17, 22, 18, 15, 21, 20, 16,
12, 18, 15, 22, 20, 14, 16, 21, 19, 17

Questions:

1. Skewness: Calculate the skewness of the waiting time distribution.

   Answer: -0.33501

2. Kurtosis : Calculate the kurtosis of the waiting time distribution.

   Answer: -0.88101

3. Interpretation: Based on the skewness and kurtosis values, what can be inferred about the waiting time distribution?

   Answer: -0.335 Skewness indicates equal distribution of data and -0.881 Kurtosis indicates no outliers in data.

END OF TOPIC

Questions on Confidence Interval and Hypothesis Testing

Confidence Interval Problems:

1. Problem: A study is conducted to estimate the mean height of a population. A random sample of 100 individuals is selected, and their heights are measured. Calculate a 95% confidence interval for the population mean height, given that the sample mean height is 170 cm and the sample standard deviation is 8 cm.

Data: Sample size (n) = 100, Sample mean ($\bar{x}$) = 170 cm, Sample standard deviation (s) = 8 cm, Confidence level = 95%

Answer :

Confidence Value : 1.567

Confidence interval : 168.432 to 171.432

Explanation: In this problem, we use a sample to estimate the population mean height. By calculating a confidence interval, we provide a range of plausible values for the population mean. The 95% confidence level indicates that we are 95% confident that the true population mean height falls within the calculated interval.

2. Problem: A survey is conducted to estimate the proportion of people in a city who support a particular policy. A random sample of 500 individuals is surveyed, and 320 of them express support for the policy. Calculate a 90% confidence interval for the population proportion, given the sample proportion.

Data: Sample size (n) = 500, Number of successes (x) = 320, Confidence level = 90%

Answer:

Population Proportion (p) = x / n = 320 / 500 = 0.64

Confidence Value = confidence level * p * n = 0.10 * 0.64 * 320 = 20.48

Confidence Interval : 300 to 340

Explanation: In this problem, we aim to estimate the population proportion based on the sample proportion. By constructing a confidence interval, we provide a range of plausible values for the population proportion. The 90% confidence level indicates that we are 90% confident that the true population proportion falls within the calculated interval.

Hypothesis Testing Problems:

3. Problem: A researcher wants to test whether a new teaching method improves student performance. A random sample of 50 students is divided into two groups: one group taught using the new method and the other using the traditional method. The average test scores of the two groups are compared. State the null and alternative hypotheses for this study.

Data: Sample size (n) = 50, Test scores of the two groups

Explanation: In this problem, we are interested in comparing the means of two groups (new method vs. traditional method). The null hypothesis (H0) states that there is no significant difference between the means, while the alternative hypothesis (Ha) suggests that there is a significant difference.

4. Problem: A manufacturing company claims that the average weight of its product is 500 grams. To test this claim, a random sample of 25 products is selected, and their weights are measured. The sample mean weight is found to be 510 grams with a sample standard deviation of 20 grams. Perform a hypothesis test to determine if there is evidence to support the company's claim.

Data: Sample size (n) = 25, Sample mean ($\bar{x}$) = 510 grams, Sample standard

deviation (s) = 20 grams, Population mean (μ) = 500 grams

Explanation: In this problem, we are conducting a hypothesis test to assess whether the sample mean weight provides evidence to support the company's claim about the population mean weight. The null hypothesis (H0) assumes that the population mean weight is equal to the claimed value, while the alternative hypothesis (Ha) suggests otherwise.