Statistics Assignment:-

* Powered by :- Raj sir (TOPS TECHNOLOGY )
* Topics :- Basic Statistics using Data Analys

Topic 1 :- Measure of the Cetral tendencies mean +median +mode +std +vaiance

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

1. Week 4: 70 units

Answer: - mean p(x):- sum weekly sales

Number weeks

(50+60+55+70)

4

235/4

P(x) =58.45

Median :- 1st) step:-set all the sales of the week sales ascending order

* 50+55+60+70

2nd ) step is the mean value which are the data set in available in data set

58.45 in (55 or 60) between the consider the this mean values

Median (n+1/2)

= 55+60/2

= 115/2

= 57.5 units

Mode :- mode is the most commonly repealtable value is the mode value

50+55+60+70

Ans. = (There are thje no repeatable value that we are sales are the no data .)

Range:-range will be defines as the highest and lowest values we known as the range .

Range = Highest – lowest

70 - 50

= 20 unit arethe range value

Standard Deviation :- Std is the stand for the find the data set will be actual (center point) will be what is the defierence in graph.

1. Mean value = 58.75
2. Diffrence between = mean value – own sales data

50-58.75 = -8.75

55-58.75 = -3.75

60-58.75= 1.25

70-58.75 = 11.25

[(58.75-50 + (60-58.75)+(55-58.75)+ (70-58.75) /4)]

[(58.75)+)(1.25)+(3.75)+(11.25) / 4]

[3456.5625 / 4]

Square root = 864.140625

29.38

Weekly 29.38 units approximaticallty decision between data .

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

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

By answering these questions using the mean+ median+ and mode+ the restaurant can gain insights into the average waiting time+ identify any common or peak waiting periods+ and make informed decisions to optimize the customer service process+ such as adjusting staffing levels+ streamlining operations+ or implementing strategies to reduce waiting times.

Answer :- mean :- 15+10+20+25+15+10+30+20+15+10+10+25020+15+10+120+13+20+30

Mean =

5+7+8+10+12+15+18+20+22+25+30+35+40+45+50

15

P(x)= 21.3333

median :- 1st step :- all the dataset the sort the ascending order curtomer service process data set

5+7+8+10+12+15+18+20+20+25+30+35+40+40+50

Then mean value is the 21.33 all the data set is the center value range is the

8+9/2

M(x) =12.5

Mode :-mode the most frequiently multi time repeatated data from the customer process service data set

20 +40 data are the most repeated waiting time

* 20+40 timesthe customer service process+ such as adjusting staffing levels+ streamlining operations+ or implementing strategies to reduce waiting times.

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?

2. Median: What is the typical or central rental duration experienced by customers?

3. Mode: Are there any recurring or most frequently occurring rental durations for customers?

By answering these questions using the mean+ median+ and mode+ the car rental company can gain insights into the average rental duration+ understand the most common rental periods+ and make informed decisions regarding pricing+ fleet size+ and availability. Additionally+ this analysis can help the company optimize resource allocation+ plan for peak demand periods+ and enhance customer satisfaction by aligning service offerings with customers' typical rental needs.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_3

Answer: - 1

mean process :-

Total rental durations customer sample

Quantity of sample

P(x) 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

50

= 3.46

1. Median :- most mean value is the value for the which range to outcome Is the dataset

3.46 is the 2 and 3 value between

M = (25th value + 26th value)

2

2+3/2

M = 2.5

Mode :- most repeatated value is the 2+3 is the mode answer

***Topic :-2 Question on Measure of The 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 .

Range = highest value - Lowest value

= 140 - 105

= 35

Mean value 120+110+130+115+125+105+135+115+125+140

= \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

10

= 1185/10

= 118.5

Variance = variance is the using for the mean centeral point value is the actually difference between mean value measure the frequence level.

Variance = Mean answer value - own dataset

(120-118.5)²+

(110 -118.5) ²+

(130-118.5) ² +

(115-118.5) ²+

(125 – 118.5) ²+

(105 – 118.5) ²+

/10 = 1269.5/10

= 126.95

Std = actual data set around the mean vaue is the actually center point is the check it the usibibtlty I use for the standard deviation function .

Standard deviation is the variance value ( square)

√126.95

= 11.27

Question :- 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?

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

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

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

$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

Range = 800- 400

= 300

Daily sales report is the 300$ inventory management.

Variance: What is the variance of the daily sales?

Variance :- 1sr step is the find the excepted value(mean)

$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

30

= 18000/30

=600$

Variance (500-600) ²+(700-600) ² ………………….+ (550 - 600) ² /30

= 10000+10000+2600

$22083.33

Standard deviation :- std is the root of the variance and is ameasure of the variance value

√ 220833

= 148

1. Range: The range of the daily sales is $400.
2. Variance: The variance of the daily sales is approximately $22083.33.
3. Standard Deviation: The standard deviation of the daily sales is approximately = $148.66

Question :- 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?

Range = maximum value – minimum value

7-1

6

2. Variance: What is the variance of the delivery times?

Mean value:- 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 / 50

=3.22

(3-3.22)²+ (5-3.22)² + …………………………………………………. (3.3.22) / 50

= 1.1564

Std values = standard deviation is the sqluare the standard deviation .

v 1.1564

question :-

1. 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?

Measure central tendency :- is the considered as the as The actual data is the center point is the exist.

Mean:- ($120+ $150+ $110+ $135+ $125+ $140+ $130+ $155+ $115+ $145+ $135+ $130)/ 12

1610/12

= 134.17

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

The range the find the maximum value – minimum values data set . we arrange the data in ascending .

110+ $115+ $120+ $125+ $130+ $130+ $135+ $135+ $140+ $145+ $150+$15

Range = Maximum - minimum

155 -110

= 45

Question :- 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:

Questions:

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

Mean = total feedback value / count of the feedback

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 / 50

= 7.54

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

((8-7.54) ² + (7-7.54) ² + …………………………………(6+7.54) ² /50

6.7048/ 50

0.1341

1. Standard deviation :- std is the square root of the variance value.

√ 0.1341

= 0.3661

Measre of the central tendency is the = 7.54

Variance value is the = 0.1341

Standard deviation = 0.3661

Question :- 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?

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

/ 100

= 16.83

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

Range value = maximum value – minimum value

27-8

=19

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

Measure of the dispersion :- 1 ) mean value:- excepted center point value

P(x) = 16.83

1. Variance = difference between mean value frequency level .

[ (10-16.83)+(15-16.83)+ (18-16.83 )………………………………(18-16.83 ) ] ² / 50

95.8031 /50

0.9580

3 ) standard deviartion :- std is the variance value root value.

√0.9580

= 0.9787

Answer:---

Range:- 19

Std:- 0.9787

Mean:- 16.03

Variance:- 0.9580

Question:- 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?

Mean :-

model-A :- 30+ 32+ 33+ 28+ 31+ 30+ 29+ 30+ 32+ 31+/10 = 30.6

model-B :- 25+ 27+ 26+ 23+ 28+ 24+ 26+ 25+ 27+ 28+/10 = 26.2

model-C :- 22+ 23+ 20+ 25+ 21+ 24+ 23+ 22+ 25+ 24+ /10 = 23.7

model-D :- 22+ 23+ 20+ 25+ 21+ 24+ 23+ 22+ 25+ 24+ /10 = 18.9

model-E :- 22+ 23+ 20+ 25+ 21+ 24+ 23+ 22+ 25+ 24+ /10 = 34.0

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

RANGE VALUE MODEL WISE = MODEL = A

MODEL = A MAXIMUM VAUE - MINIMUM VALUE = 5

MODEL = B MAXIMUM VAUE - MINIMUM VALUE = 5

MODEL = C MAXIMUM VAUE - MINIMUM VALUE = 5

MODEL = D MAXIMUM VAUE - MINIMUM VALUE = 5

MODEL = E MAXIMUM VAUE - MINIMUM VALUE = 5

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

VARIANCE = MEAN Value difference frequeny level

MODEL = A = (30-30.6)+ (32-30.6)+ (33-30.6)+ (28-30.6)+ (31-30.6)+ (30-30.6)+ (29-30.6)+ (30-30.6)+ (32-30.6)+ (31+-30.6)/10

= 1.04MPG²

MODEL = B = (25-26.2)+ (27-26.2)+ (26-26.2)+ (23-26.2)+ (28-26.2)+ (24-26.2)+ (26-26.2)+ (25-26.2)+ (27-26.2)+ (28-26.2) /10

= 2.16MP²

MODEL =C =(25-23.7)+( 27-23.7)+ (26-23.7)+ (23-23.7)+ (28-23.7)+ (24-23.7)+ (26-23.7)+ (25-23.7)+( 27-23.7)+ (28-23.7)+/10

= 2.63MP²

MODEL =D=(18-18.9)+ (17-18.9)+ (19-18.9)+( 20-18.9)+ (21-18.9)+ (18-18.9)+(19-18.9)+ (17-18.9)+ (20-18.9)+ (19-18.9)+/10

=1.21MP²

MODEL =E =(35-34.0)+ (36-34.0)+ (34-34.0)+(35-34.0)+(33-34.0)+ (34-34.0)+ (32-34.0)+ (33-34.0)+ (36-34.0)+ (34-34.0)+/10

=0.9MP²

1. STANDARD DEVIATION :- square root of the variance values

Model A = √ 1.04 = 1.01980

Model B =√2.16 = 1.46

Model C =√2.63 = 1.62

Model-D = √1.21 = 1.1

Model-E = √0.9 = 0.94

Topic no = 3 Mor e Statistics Question (covered topic is the) := [ quartiles+deciles+persantage +Frequency distribution ]

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.

Frequency table is the = all the employee data which time consider in Frequeny that is called the Frquency level

1) Descreate (continues natural number positive natural number)

2) Continues Frequeny distribution (-Range value (max- min) create range )

|  |  |
| --- | --- |
| age | Frequency(which time repeated the data ) |
| 27 | 3 |
| 28 | 5 |
| 29 | 6 |
| 30 | 6 |
| 31 | 8 |
| 32 | 4 |
| 33 | 7 |
| 34 | 4 |
| 35  36  37  38  39  40  41  42  43  44  45 | 8  4  5  6  6  6  4  3  3  4  2 |

1. Mode: What is the mode (most common age) among the employees?

Most repeated value in this Data Is the 31 age is the most repeated value frequency .\

31 = 8

Mode value is the = 31age

1. Median: What is the median age of the employees?

1 )Meadian value is find the all the ages value is the sorted the ascending order = 27+27+27+28+28+28+……………………..+4042+44

1. All the number are the even that’s why

Median = n+1 /2

50+1/2

51/2

25.5

(25and 26 value age is the meadian central point is the substituted )

25age=(31+31/2)

62/2

= 31 age value is the most center poin in this data

4. Range: what is the range of ages among the employees?

Range = Maximum value – minimum value

45-27

= 18

Frequency distribution :-

Median:- 31

Mode:- 31

Range:- 18

Problem = (9)

: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.

|  |  |
| --- | --- |
| Purchase Amount | Frequency |
| 28 | 1 |
| 35 | 1 |
| 36 | 1 |
| 38 | 1 |
| 39 | 2 |
| 40 | 3 |
| 41 | 2 |
| 42 | 2 |
| 43 | 1 |
| 44 | 1 |
| 45 | 2 |
| 47 | 3 |
| 48 | 2 |
| 49 | 1 |
| 51 | 2 |
| 52 | 3 |
| 55 | 2 |
| 56  57  58  58  59  60  61  62  63  65  68  73 | 2  1  2  2  2  1  1  2  1  3  1  1 |

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

\* Most repeated value us the 40+47+49+52+65 is the amont more than 3 times max repeated value.

1. Median: What is the median purchase amount among the customers?

Median value is the all the number are the ascending order.

=> 28+35+36+38+38+39+…………..73

Total data is the 50

Median = total data + 1 /2

(50/2)

25 or 26 value is the median value.

Median = (49+51)/2

100/2

==50

1. Interquartile Range: What is the interquartile range of the purchase amounts?

Even = number Q1 = (n+1)/2 = q1 odd number = (direct central number )

Q2 = (n+1)/2 = q3

Interquartile value is the 75% of the data all the data are the 4 part devided 25%+25%25%+25%

Inter quartile include the 75 % data is the

q3 = 75% data uppr highest value = (n+1)\* 3/4

Q2 = median center point value = (n+1)\*2/4

Q1 = 25% data lower quartilr valu = (n+1)\*1/4

Total =50 data

(first half )

(second half) q1= 38 [ 12 data point Is the ascending orderd]

[ 25 % data ] [25%data ] q3 = 58[ 37 data is the ascending ordered]

(Median value )

Inter quartile = [Q3-Q1]

= 58-39

= 19

Quartiles example :-

5+8+12+15+20+21+24

1st step – all the data is the sorted the ascending order

5+8+12+15+20+21+24

2nd step =

Q1 = (n+1)\*1/4

(7+1)\*1/4

8\*1/4

2nd value lower value = 8

Q2 = (n+1)\* 2/4

(7+1)\*2/4

8\*2/4

16/4

4th value is the medium value = 15

Q3 = (n+1)\*3/4

= (8+1)\*3/4

= 9\*3/4

27/4

6.5th value is the higher quartiles = 21

Q23 = (n+1)\*23/4

(8+23)\*23/4

178

Inter quartile = this is the one types of the range value.

Maxvalue-min value

= q3-q1

=21-8

= 13

Persantile = persantile is the all over data is the devided to the 100 the parts

Formula = p(1)= (1/100)\*(n+1)

N= total count of the data set

P()= number of the persantage

P(25)=(25/100)\*(30+1) p(50)= (50/100)\*(n+1) p(75)=(75/100)\*(n+1)

= 6.25\*31 = 0.25\*31 = 0.75\*31

= 193.75 =15.5 = 23.2

620

Finally quartile formuala = (n+1)\* (qvalue/4

Because (100/4) 25 %20 = 4 something data is th fetch the value.

Question = 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.

|  |  |
| --- | --- |
| Detect type | Frecuency |
| a | 30 |
| B | 40 |
| C | 20 |
| D | 10 |
| E | 45 |
| F | 25 |
| G | 30 |

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

Most common highest type value level frequency repeated value is the - 45

1. Histogram: Create a histogram to represent the defect frequencies. By answering these questions using a bar chart and histogram+ the manufacturing company can visually understand the g

10-19 = d

20-29 = c++f

30-39 = a+g

40-49 = b+e

distribution of defect types+ identify the most common defect+ and prioritize quality control efforts to address the prevalent issues.

Question =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.

Rating Frequency

1 0

2 10

3 30

4 40

5 20

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

Most repetitive value is the 4 number 40 time frequency .

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

Rating Frequency

1 0

2 10

3 30

4 40

5 20

Question – 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.

Sales range frequency

28-30 7

31-33 10

34-36 8

37-39 12

40-42 8

43-45 4

46-47 1

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

Measure of thr central tendency is the mean.

Mean=[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 ] / 50

= 39.96

Meadian = (n+1)/2

= (50+1)/2

= 25.5

25+26 value is the central point

31+45/2

=53.5

Mode = most repeated value is the 42

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

Sales range frequency

28-30 7

31-33 10

34-36 8

37-39 12

40-42 8

43-45 4

46-47 1

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.

Region 1 | sales(numberic value)

Reqion 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

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

Central tendency is that is the mean individual all the region wise data find the average value .

Region 1 = (45+35+40+38+42+37+39+43+44+41)/ 10

= 39.4

Region 2 = (32 +28+ 30+ 34+ 33+ 35+ 31+ 29+ 36+ 37)/10

= 32.5

Region 3 : - (40 +39 +42+ 41+ 38+ 43 +45+ 44+ 41+ 37)/10

= 41.0

What is the range of sales figures in each region?

Range = highest value – lowest value

= region 1 = 45-35 = 10

Region 2 = 37-28 = 9

Region 3 = 45-37 = 8

Topic =4 measure of the skewness and kurtosis :-

1. 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.

Skewness is the measure of the distribution negative or positive distribution (left or right side)

Skewness formula :- [n / ((n-1)(n-2))] \* Σ[((xi - mean) / standard deviation) ^ 3]

Xi= number of the each data point

Mean = average value of the data ‘

Standard deviation = standard deviation is the use for the mesure the average data is the actual center point to whch frequency far.

Mean = [-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]/50

= 0.338

Standard deviation :- = [-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] – 0.338

= 1.493

Lets calculate the skewness = [n / ((n-1)(n-2))] \* Σ[((returns - mean) / standard deviation) ^ 3]

= (50(50+1(50-2\*Σ[((returns - 0.338) /1.493 ^3))

= 0.323

1. Kurtosis: Calculate the kurtosis of the monthly returns

Kurtosis :- kurtosis is the measures the thickness of the graph tail . basicall y graph tail rate is

Three types od the kurotosis types

the 3 maximum 3 plus tail thatis for the leptokurtic .

lowest3 value tail to platykurotis and

the exjectpt graph 3 value is the l meso kurtosis

formula of the l kurotsis =[n(n+1) / (n-1)(n-2)(n-3)] \* / standard deviation) ^ 4] - [3 \* (n-1)^2 / (n-2)(n-3)]

[50(50+1)/(50-1)(50-2)(50-3))] \* Σ[((returns – 0.338)/1.493)^4]-[3\*(50-2]^2 / (50^2(50-3))]

Kurtosis = 3.288

. 3.2 Interpretation: Based on the Σ[((returns - mean) skewness and kurtosis values+ what can be said about the distribution of returns?

the skewness value of approximately -0.323+ we can say that the distribution is slightly negatively skewed+ meaning it has a longer left tail.

Question :- 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 100 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.

Skewness = (3\*(mean-median)) / standard deviation =

Find the Skeness :- is the data of the mean+median and std find is the first step

Mean = total of sum of the data / total count of the data

( 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)/ 100

= 325.9/ 100

= 3.259

Median = the middle value is the sorted in the ascending order .

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

Median = is the all the even number is the median formula is the .

Median =((n/2)+ (n/2+1))

M = (100/2 ) +(100+1/2)

(50th position value) + (51th position value is )

(50th + 51) value is the /2

(2.9 +29)/2

2.9

(50th postion value ) = 2.9

Standard Deviation is the :-

mean = total of sum of the data / total count of the data

(325.9) / 100

3.259

Step :2 Find the Numerator :- mean value ( difference value find the then Diffrence value Square)

(2.5-3.25) ^2 = 0.58281

4.8 - 3.259)^2 = 1.81081

(3.2 - 3.259)^2 = 0.00361

(2.1 - 3.259)^2 = 1.34961

( Continue as the 100 data difference value will be fin and square )

Step3 :- find the Squared differences ( Eccepted mean)

( 0.32534/100)

= 0.621001

Last Step :- take the diffrenced mean itS find the Square roots .

√(Mean of Squared Differences)

√.621001

= 0.78798

Lets Calculate the Skewness :- mean value :- 0.3259

Median value = 3.2

Std = 0.819

Skenwess FOrmulla = Skewness = (3\*(mean-median)) / standard deviation

= (3\*(3.259-3.2)) / 0.819

= 0.215

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

Lets Calculate the 4th Position of the Kurtosis = kurtosis is the uses the pickedness measures comapared.

Kurtiisis Value generated the graph if the graphs is the

( graph tail right to left = positive kurtosis)

( graph tail left to right = positive kurtosis )

:- kurtosis Formula is the Generated grapth 3 types :-

1. Lepto Kurtosis (grath tailed is the less than 3 )
2. Meso kuritsos ( normal Distriburtion is the data is >3 )
3. Platkurtosis ( graph tail flat )

Kurtosis formula :- (sum of (x- mean)^ 4 ) / (count of the point \* standard Deviation ^ 4 ) - 3

kurtosis = (sum of ( x – 3.259 ) ^ 4) / ( 100 – 0.819 ^ 4 ) -3

kurtosis (26.0235006258) / (100- 0.819) ^ 4) -3

2.907

Skewdness =

Kurtosis value is the = 2.907 is the data graph tailed is the less than 3 that is why kurtosis types is the lepto graph

Question :3 )

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 200 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

Skewness formula :- [n / ((n-1)(n-2))] \* Σ[((xi - mean) / standard deviation) ^ 3]

Xi= number of the each data point

Mean = average value of the data ‘

Standard deviation = standard deviation is the use for the mesure the average data is the actual center point to whch frequency far.

Skewness :-

1st step : mean:- mean = total of sum of the data / total count of the data

= : ( 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 )/100

= 355/100

3.55

2nd Step is the median Find

:- first then all the data in Interpreate in Ascending order .

2+ 2+ 2+ 2+ 2+ 2+ 2+ 2+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 3+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 4+ 5+ 5+ 5+ 5+ 5+ 5+ 5+ 5+ 5+ 5+ 5+ 5+ 5+ 5+ 5+ 5+ 5+ 5+ 5+ 5+ 5+ 5+ 5

All the data is the even this data middle value is the find the Median

Median =((n+1/2)

(101/2)

= 50th AND 51st value is the find this data set

3+3/2

= 3

Is that 100 and 101 value position value is the median value

100th value is the =5 .

3rd step :- is the standard Deviation

Standard Deviation is the :-

mean = total of sum of the data / total count of the data

(377) / 100

3.708666

Median = is the all the even number is the median formula is the .

Median =((n/2)+ (n/2+1))

(100/2 )

50 or 51 values

= (4+4 /2 )

= 4

Step :2 Find the Numerator :- mean value ( difference value find the then Diffrence value Square)

(5-3.7) ^2 = 1.69

(4-3.7) ^2 = 0.09

(2-3.7) ^2 = -2.89

(3-3.7) ^2 = 1.69

(4-3.7) ^2 =0.09

( Continue as the 100 data difference value will be fin and square )

Step3 :- find the Squared differences ( Eccepted mean)

( 0.3253434643/100)

=0.03400927

Last Step :- take the diffrenced mean itS find the Square roots .

√(Mean of Squared Differences)

√0.03400927

= 0.184335

Lets Find the Skewness = Skenwess FOrmulla = Skewness = (3\*(mean-median)) / standard deviation

= (3\*(3.70 -4)) / 0.184335

= -4.882415168

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

Lets Calculate the 4th Position of the Kurtosis = kurtosis is the uses the pickedness measures comapared.

Kurtiisis Value generated the graph if the graphs is the

( graph tail right to left = positive kurtosis)

( graph tail left to right = positive kurtosis )

:- kurtosis Formula is the Generated grapth 3 types :-

1. Lepto Kurtosis (grath tailed is the less than 3 )
2. Meso kuritsos ( normal Distriburtion is the data is >3 )
3. Platkurtosis ( graph tail flat )

Kurtosis formula :- (sum of (x- mean)^ 4 ) / (count of the point \* standard Deviation ^ 4 ) - 3

= ( sum of (x-3.777) ^4) / (100 \* 3.75 ^ 4) -3

= 0.00964

By answering these questions using measures of skewness and kurtosis+ the survey can assess the skewness and peakedness of the satisfaction ratings+ determine if the ratings are skewed towards positive or negative evaluations+ and understand the distribution characteristics of customer satisfaction.

Skewness= -4.882415168

Kurtosis values = 0.000964

Question = 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 150 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.

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

Skewness = (3\*(mean-median)) / standard deviation =

Find the Skeness :- is the data of the mean+median and std find is the first step

Mean = total of sum of the data / total count of the data

Mean = (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) / 100

= 28+420 / 100

= 284.20

2nd Median value Find this :- 1st all the data is the sorted Ascending order all the value set .

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

Median value = (n/2)

= 50th value and 51st value is the find this sorted List

= (290 +290) /2

= (580/2)

=290

Standard Deviation:- 1st Step is the mean value : -

Mean = total of sum of the data / total count of the data

Mean = (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) / 100

= 28+420 / 100

= 284.20

( Difffence between mean value and orginal data set . )

(280-284) ^2 = -8

(350-284) ^2 = 4356

(310-284) ^2 = 676

(270-284) ^2 = -196

Last Step :- take the diffrenced mean itS find the Square roots .

79800

Variance = Divide the sum of squared differences by the number of values (N) to get the variance:

Variance = Sum of squared differences / N

= 79800/100

√798

28.25

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Lets find the Skewness : -

Mean = 79800

Median = 290

Count of the all the data = 100

Std = 28.25

* Skewness:-

Skewness = (3\*(mean-median)) / standard deviation

= ( 3\* 79800- 290)) / 28.25

= -6.22

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

Kurtosis: Calculate the kurtosis of the satisfaction ratings.

Lets Calculate the 4th Position of the Kurtosis = kurtosis is the uses the pickedness measures comapared.

Kurtiisis Value generated the graph if the graphs is the

( graph tail right to left = positive kurtosis)

( graph tail left to right = positive kurtosis )

:- kurtosis Formula is the Generated grapth 3 types :-

1. Lepto Kurtosis (grath tailed is the less than 3 )
2. Meso kuritsos ( normal Distriburtion is the data is >3 )
3. Platkurtosis ( graph tail flat )

Mean = 79800

Median = 290

Count of the all the data = 100

Std = 28.25

Kurtosis = (Σ(xi - mean)^4 / N) / (standard deviation)^4

= 79800 - 284.20 )^4 /100) / 28.25 ) ^4

3203 / 28.25

= 113 ^4

= 4.54

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

Skewness = -6.22

Kurtosis = 4.54

Is the kurtosis values is the grater than 3 is the platy kurtosis . this graph tailed is the value grater in the right to left direction .

Question :- 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.

Skewness = (Σ(xi - mean)^3 / (N \* standard deviation)^3)

Skewness find the Three step devided the

1st mean +Standard deviation + Diffrence beween+ Variance

Mean:-

(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 )/100

= 1812 /100

= 18.12

1. DIFFRence between mean value and old current data set value .

( 12 -18.12) ^4 = -229.220928

(18 – 18.12) ^4 = -0.001728

( 15– 18.12) ^4 =-30.371328

(22– 18.12) ^4=58.41

(20 – 18.12) ^4 = 6.6446

(14 – 18.12) ^4= 69.9344

Continue all the data in mean difference value on quabe. All the diffrnece mean value ia the = 21.164

Variance = variance all the data value is the standard deviation value square root .

√Variance = √0.723944

= 0.851106

Lets caluate skewness :-

Skewness = (sum of cubbed Diffrences / (n \* Standard Deviation ) ^3

= ( -21.164 /(100 \* 0.851106 ) ^ 3 )

= 0.4117

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

Kurtosis: Calculate the kurtosis of the satisfaction ratings.

Lets Calculate the 4th Position of the Kurtosis = kurtosis is the uses the pickedness measures comapared.

Kurtiisis Value generated the graph if the graphs is the

( graph tail right to left = positive kurtosis)

( graph tail left to right = positive kurtosis )

:- kurtosis Formula is the Generated grapth 3 types :-

1. Lepto Kurtosis (grath tailed is the less than 3 )
2. Meso kuritsos ( normal Distriburtion is the data is >3 )
3. Platkurtosis ( graph tail flat )

Mean -=18.12

QUAbed Answer = 37+374.68364

N = 100

Standard Deviation :- 51.952512

Kurtosis = (Σ(xi - mean)^4 / (N \* standard deviation)^4) – 3

Kurtosis = (Sum of fourth power differences / (N \* standard deviation)^4) – 3

(37+374.68364 / (100 \* 0.51952512 ) ^4)) – 3

(37+374.68364 / 51.952512 ) -3

720.045386 -3

= 717.045386

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

Skewness answer valus is the

Kurtosis value is the 717.045386 kurtosis value is approximately 7.17.045386 . always the kurtosis value is the grater than 3 is the graph tailed is the leptokurtic .

By answering these questions using measures of skewness and kurtosis+ the company can assess the symmetry and tail behavior of the waiting time distribution+ identify any patterns or anomalies in customer waiting times+ and make improvements to streamline the service process and enhance customer satisfaction.

QUESTION := PERSANTILE AND QUEARTILES

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)+

Q1 is the all the 100% data into 25% data

Q2 is the 50 % percentage data of the all the population is this data .

1st step

Data =[ 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 ]

Meadian(q2)= average of the 100 and 101st observations

All the number all the count is the 200 and even number are the in meadian in two values

All the dat find in the 1st step is the all the data sort in asenvding order

Q2 median center point . the center of the all the population . data

Q3 is the value below whch all of the population in the 75% data.

MEDIAN = (n/2

(200/2)

100 and 101 value ia the median value .

(250+255) /2

252.5

100th value are the meadian values . 100 is the even number are the even median number formula is the . (n/2+1)

Q1 =valu ar the 25th persanrtile = (25/100)\*(n+1)

= (25/100)\*(200+1)

= (0.25)\*201

= 50.25

Q2= value is the median (q2 is the second quartile ): (q2) = (data[100]+ data[101]/2

= 200/2 = 100

100 is the even number is the median formula is the ( n/2+1)

(250+255)/2

= 252.5

Q3 =value are the 75th persanrtile = (75/100)\*(n+1)

= (75/100)\*(200+1)

= (0.75)\*201

= 50.25

Q4 =value are the 100th persanrtile = (100/100)\*(n+1)

= (1)\*(200+1)

= (1)\*201

= 201 =

Q1 = 50.25

Q2 = 252.50

Q3 = 150.75

Q4 = 201

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

Persan tiles is the all over data about of 100 % 25 +75 +90 persantage is the all over 100 % persantage data in 80 % +60% + in explain in this example .

Persantile formula = (NUMBER OF THE VALUE /TOTAL DATA PERSANTAGE \* 100 )

25TH = PERSANAGE DATA ( 25/100)

= (25/100)\*(100)

= 0.25\*100

=25.25

50TH = PERSANAGE DATA ( 50/100)

= (50/100)\*(200+1)

= 0.50\*201

=150.75

90TH = PERSANAGE DATA ( 90/100)

= (90/100)\*(200+1)

= 0.90\*201

=180.9

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

P(10)= 201 +p(25)= 50.25 +p(75)= 150.75 +p(90) = 180.9 all over ther population data into the result iiin persantage .

And quartile in 25 Q1 = 50.25 + Q2 = 252.50 +Q3 = 150.75 +Q4 = 201

Queastion – 2

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.

Quartiles: Calculate the first quartile (Q1)+

Q1 is the all the 100% data into 25% data

Q2 is the 50 % percentage data of the all the population is this data .

1st step:- 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 / 100

= 73325/100

= 733.25

Meadian = 1 st step is thefind the meadian in all the population in this sample . of the 1st step is the all the data sorting in asscending order.

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 / 100

All the population of the data is the 100 mean value.100 is the even number is the median even number formula is the

MEDIAN = (n/2

(100/2)

Is the 50 th postion and 51st position value is the median values is the consideration.

(260+265) /2

= 262.5 kg

50 and 51 value ia the median value .

100th value are the meadian values . 100 is the even number are the even median number formula is the . (n/2+1)

Quartiles is the all the population data devided to the eqyual parts .

* Q1 (25th percentile): The value that separates the lowest 25% of the data from the rest.
* Q2 (50th percentile or median): The value that separates the lowest 50% of the data from the highest 50%.
* Q3 (75th percentile): The value that separates the lowest 75% of the data from the rest.

Since you have 100 data points, the positions of the quartiles can be estimated as follows:

* Q1: 25th data point (25% of 100)
* Q2: 50th data point (50% of 100)
* Q3: 75th data point (75% of 100)

Q1 =valu ar the 25th quartiles = (25/100)\*(n+1)

= (25/100)\*(100+1)

= (0.25)\*101

= 25.25

Q2 =value ar the 50th quartiles = (50/100)\*(n+1)

= (50/100)\*(100+1)

= (0.50)\*101

= 50.5

Q1 =value are the 75th quartiles = (75/100)\*(n+1)

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

= (0.75)\*101

= 75.25

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

Persan tiles is the all over data about of 100 % 25 +75 +90 persantage is the all over 100 % persantage data in 80 % +60% + in explain in this example .

Persantile formula = (NUMBER OF THE VALUE /TOTAL DATA PERSANTAGE \* 100 )

15TH = PERSANAGE DATA ( 25/100)

= (15/100)\*(100)

= 0.15\*100

=0.`15

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

By answering these questions using quartiles and percentiles , the research study can understand the weight distribution and identify the weight ranges at different percentiles , such as underweight , normal weight , overweight , and obese categories. This information can be used for evaluating health risks , designing appropriate interventions , and providing personalized recommendations for weight management