

Quiz (Data Pre-processing)

Enroll No.

Q1. Which of the following are “Measures of Central Tendency”?

- | | |
|-----------------------|--|
| A. Mean, Range, Mode | B. Mean, Standard Deviation, Range |
| C. Mode, Mean, Median | D. Range, Standard Deviation, Variance |

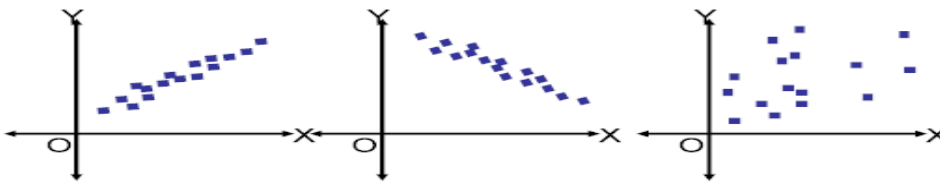
Q2. Which of the following of a random variable is a measure of spread?

- | | |
|-------------------|-------------------------|
| A. Variance | B. Standard deviation |
| C. Empirical mean | D. All of the mentioned |

Q3. Which of the following diagram is used to view correlation?

- | | | | |
|-----------------|------------|--------------|--------------|
| A. Scatter plot | B. Boxplot | C. Rectangle | D. Histogram |
|-----------------|------------|--------------|--------------|

Q4. Given below are three scatter plots for two features (Image 1, 2 & 3 from left to right).



- | |
|--|
| A. 1 shows +ve correlation, 2 shows +ve correlation, 3 shows -ve correlation |
| B. 1 shows -ve correlation, 2 shows +ve correlation, 3 shows no correlation |
| C. 1 shows +ve correlation, 2 shows -ve correlation, 3 shows no correlation |
| D. 1 shows -ve correlation, 2 shows -ve correlation, 3 shows +ve correlation |

Q5. Suppose you are given 7 Scatter plots 1-7 (left to right) and you want to compare Pearson correlation coefficients between variables of each scatterplot.

Which of the following is in the right order?



1. $1 < 2 < 3 < 4$
2. $1 > 2 > 3 > 4$
3. $7 < 6 < 5 < 4$
4. $7 > 6 > 5 > 4$

- | | | | |
|------------|------------|------------|------------|
| A) 1 and 3 | B) 2 and 3 | C) 1 and 4 | D) 2 and 4 |
|------------|------------|------------|------------|

Q6. The statistical technique that can show whether and how strongly pairs of variables are related

- | | | | |
|---------------|----------------|-------------|---------|
| A. Covariance | B. Correlation | C. Variance | D. Mean |
|---------------|----------------|-------------|---------|

Q7. The measure of how changes in one variable are associated with changes in a second variable.

- | | | | |
|---------------|----------------|-------------|---------|
| A. Covariance | B. Correlation | C. Variance | D. Mean |
|---------------|----------------|-------------|---------|

Q8. Numerosity reduction can be performed by:

- | | | |
|--------|----------------------|----------------------|
| A. PCA | B. Linear Regression | C. DWT D. Chi-Square |
|--------|----------------------|----------------------|

Q9. Dimensionality reduction can be performed by:

- | | | |
|--------|----------------------|--------------------------------------|
| A. PCA | B. Linear Regression | C. Multiple Regression D. Chi-Square |
|--------|----------------------|--------------------------------------|

Q10. Correlation also determines causality.

- | | |
|---------|----------|
| A. True | B. False |
|---------|----------|

Q11. Find the median of the set of numbers: 1,2,3,4,5,6,7,8,9 and 10.

- | | | | |
|-------|-------|------|--------|
| a. 55 | b. 10 | c. 1 | d. 5.5 |
|-------|-------|------|--------|

Q12. The following represents age distribution of students in an elementary class. Find the mode of the values: 7, 9, 10, 13, 11, 7, 9, 19, 12, 11, 9, 7, 9, 10, 11.

- | | | | |
|------|------|-------|-------|
| a. 7 | b. 9 | c. 10 | d. 11 |
|------|------|-------|-------|

Q13. “Students’ scores on a biology test” is an example of which scale of measurement?

- | | | | |
|----------|-------------|------------|---------|
| A. Ratio | B. Interval | C. Nominal | D. None |
|----------|-------------|------------|---------|

Q 14. Indicate which level of measurement is being used in the given scenario.

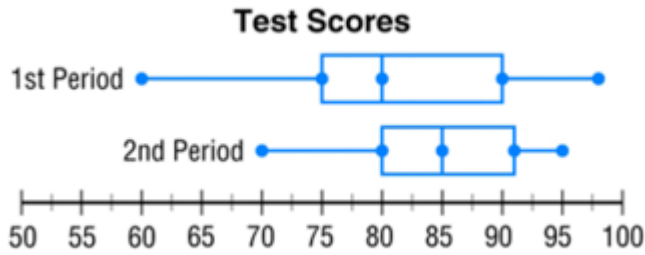
The teacher of a class of third graders records the height of each student.

- | | | | |
|----------|-------------|------------|----------|
| A. Ratio | B. Interval | C. Nominal | D. Order |
|----------|-------------|------------|----------|

Q 15. 33, 25, 42, 25, 31, 37, 46, 29, 38. What is the interquartile range of the data?

- | | | | |
|------|------|-------|-------|
| A. 8 | B. 9 | C. 13 | D. 21 |
|------|------|-------|-------|

Q16. Which class did better on the test?



- A. 1st period B. 2nd period C. They both performed the same

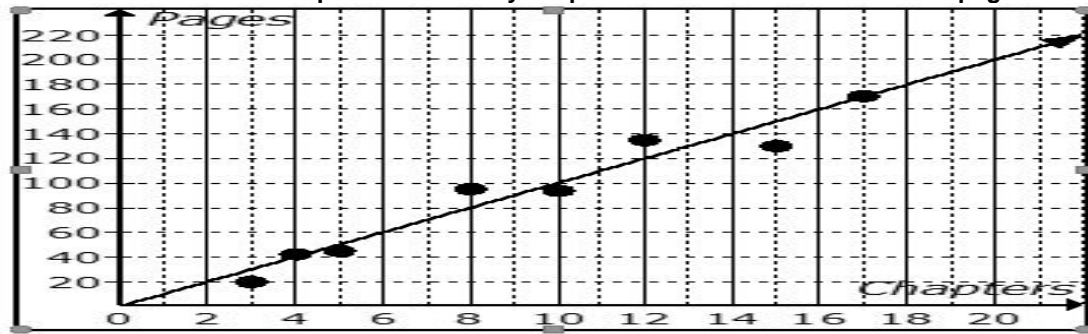
Q17. What percent of data lies within 1 standard deviation from the mean?

- A. 95% B. 99% C. 68% D. 100%

Q18. Which of the following statements about variables is correctly paired?

- A. A qualitative variable describes or categorizes an element of population, while a quantitative variable quantifies an element of a population
- B. A quantitative variable describes or categorizes an element of population, while a qualitative variable quantifies an element of a population
- C. A qualitative variable describes an element of the sample population, while a quantitative variable quantifies an element of total population
- D. A quantitative variable describes an element of the sample population, while a qualitative variable quantifies an element of total population

Q 19. The scatter plot shows the relationship between the number of chapters and the total number of pages for several books. Use the trend line to predict how many chapters would be in a book with 180 pages.



- A. 12 chapters B. 15 chapters C. 18 chapters D. 21 chapters

Q 20. Rich is a member of a gym. He pays a monthly fee plus a per-visit fee. The equation below represents the monthly amount Rich pays for his membership to the gym per month for x visits.

$$y = 10 + 3x$$

What does the y-intercept of the graph of this equation represent?

- A. The monthly fee
B. The number of times Rich goes to the gym
C. How much Rich pays per visit
D. The total cost at the end of the month

Q21. Which example shows CAUSATION? (Which one CAUSED the other to happen?)

- A. High social media usage and reduced grades. C. Car ran out of gas and being stranded on the road.
B. Recess time and number of friends. D. A person's age and what kind of cereal they eat.

Q22. How do you find the residual Error?

- A. predicted value - actual value

Q23. What are the expected counts of a female who likes Pepsi?

| | | Softdrink Choice | | |
|--------|---------------|------------------|--------------|-------|
| | | <i>Coke</i> | <i>Pepsi</i> | Total |
| Gender | <i>Male</i> | 19 | 6 | 25 |
| | <i>Female</i> | 10 | 15 | 25 |
| | Total | 29 | 21 | 50 |

- A. 10.5 B. 11 C. 14.5 D. 6.3