QUESTION BANK FOR DESCRIPTIVE STATISTICS

Classification, Tabulation, Frequency distribution and Graphical representation

Fill in the Blanks

1. Classification can be done according to_____________.
2. Year wise recording of data of food production will be called _________________ classification.
3. The difference between the upper and lower limit of a class is called_____________.
4. The average of upper and lower limits of a class is known as _____________
5. Formula for determining the number of classes was given by _____________
6. Sturges formula for determining the number of classes is _____________
7. An arrangement of data in rows and columns is known as _____________
8. The graphs of less than and more than ogive intersect at _____________
9. The distribution of frequencies according to individual variate values is called _____________distribution.

1. Frequency distributions are often constructed with the help of_____. Statistics is the study of _____ and their interpretation.
2. The _____ of statistics includes collecting, organizing, analy zing, and interpreting data.
3. A _____ consists of all the items or individuals under consideration in a statistical study.
4. A _____ is a subset of the population selected for analysis.
5. Data can be categorized as _____ or _____.
6. _____ data are those that can be measured and expressed numerically.
7. _____ data are descriptive and cannot be measured numerically.
8. Attributes are the characteristics or _____ of the items being studied.
9. Variables are properties that can change or _____ among the items being studied.
10. The _____ scale of measurement involves categories without any inherent order.
11. An _____ scale of measurement has categories with a meaningful order but no fixed interval.
12. The _____ scale of measurement has meaningful order and consistent intervals but no true zero point.
13. The _____ scale of measurement has meaningful order, consistent intervals, and a true zero point.
14. _____ presentation of data involves presenting it in tables.
15. _____ presentation of data involves representing it visually through charts or graphs.
16. A _____ is a graphical representation of frequency distribution using bars.
17. An _____ is a graphical representation of cumulative frequency distribution.
18. Consistency of data refers to the degree of _____ among measurements.
19. Independence of data means that one observation does not _____ another in the dataset.
20. When dealing with attributes, _____ refers to the situation when data items belong to one and only one category.
21. The primary purpose of statistics is to _____ data to gain insights and make informed decisions.
22. A _____ is a characteristic of interest that can vary among the subjects in a population.
23. In a _____ scale, data can be ranked, but the differences between values are not meaningful.
24. The _____ scale of measurement has all the properties of the other scales and a true zero point.
25. A _____ displays data using a series of connected data points.
26. The _____ of data refers to the spread or variability of the data values.
27. An example of a _____ variable is the color of a car.
28. _____ data consist of categories that have a natural order but no consistent difference between them.
29. A _____ is a graphical representation of data that uses lines to connect data points.
30. The _____ of data measures the tendency of data values to cluster around a central point.

31. A _____ is a table that shows the frequency of each category in a dataset.
32. When two variables are _____, the presence or value of one does not affect the other.
33. A _____ variable is one that can take any value within a certain range.
34. The _____ of data refers to the middle value in a set of ordered data.
35. An example of a _____ variable is the number of people in a household.

MULTIPLE CHOICES

1. Numerical data presented in descriptive form is called
   a. Classified presentation
   b. Tabular presentation
   c. Graphical presentation
   d. Textual presentation
2. For the mid value given 25, 34, 43, 53, 61,70. The first class of the distribution isa.
   24.5-34.5
   b. 25-35
   c. 20-30
   d. 20.5-29.5
3. In an exclusive type distribution, the limits excluded are
   a. Lower limit
   b. Upper limit
   c. Either of lower or upper limit
   d. Both a and b
4. If the lower limit and upper limit of a class are 10 and 40 respectively, the midpointof the class is:
   a. 25
   b. 12.5
   c. 15
   d. 30
5. A frequency distribution can be :
   a. Discrete
   b. Continuous
   c. Both a and b
   d. None of these
6. Classification is applicable in case of :
   a. Quantitative classification
   b. Qualitative classification
   c. Both a and b
   d. None of these
7. Graphs and charts facilitate:
   a. Comparison of values
   b. To know the trend
   c. To know the relationship
   d. All the above

8. Ogive curve occur for
   a. More than type distribution

b. Less than type distribution
c. Both a and b
d. None of these

9. Mode can be calculated from
   a. Ogive
   b. Histogram
   c. Bar diagram
   d. Pie-chart
10. In bar diagram, bars are:
   a. Horizontal
   b. Vertical
   c. Slanting
   d. None of above
11. In a component bar diagram the length of the bar
   a. Will be same for all
   b. will not be same
   c. Depends on the total
   d. none of these
12. The length of the bar will be same for all categories in
   a. Multiple bar diagram
   b. component bar diagram
   c. Percentage bar diagram
   d. none of these

Short Questions (each carry 2 Marks)

1. Define statistics.
2. Define primary and secondary data
3. Give the advantages of tabulation
4. Write a detail note on the types of classification
5. What are the essential characteristics of a good table?
6. What is the purpose of data presentation?
7. Explain the concept of a sample.
8. Write the limitations of Statistics.
9. Difference between qualitative and quantitative data.
10. Differentiate between qualitative and quantitative data.
11. What does the term "scale of measurement" refer to?
12. Define consistency of data.
13. What is the central tendency of data?
14. Explain the difference between mean and median.
15. Define the term "attribute" in statistics.
16. Why is independence of data important?
17. What does statistics aim to study?
18. Give an example of qualitative data.
19. Define the term "variable."
20. Differentiate between nominal and ordinal scales.
21. What is the purpose of graphical data representation?
22. Define the concept of central tendency.

23. Explain the difference between consistency and independence of data.
24. What does a histogram display?
25. State the significance of a sample in statistics.
26. Define the term "population" in statistical terms.
27. Explain the difference between a census and a sample survey.
28. Describe the concept of a discrete variable.
29. How does a pie chart represent data?
30. Define the term "skewness" in statistics.
31. Differentiate between a frequency polygon and a histogram.
32. What is the purpose of using a line graph?
33. Explain the concept of "range" in data analysis.
34. Define the coefficient of variation.
35. Describe the key features of a bar chart.
36. What is a quartile in statistics?
37. Differentiate between a stem-and-leaf plot and a box-and-whisker plot.
38. Explain the concept of positive correlation.
39. Define the term "margin of error" in relation to surveys.
40. Describe the concept of "bivariate data."
41. What is the significance of the interquartile range?
42. Differentiate between a line graph and a scatter plot.
43. Explain the concept of a representative sample.
44. Define "cumulative frequency" in statistics.
45. Describe the purpose of a Pareto chart.

Long Questions (each carry 6 Marks)

1. Construct a frequency distribution table for the following data
2. 32,45,8, 24, 42, 22, 12, 9, 15, 26, 35, 23, 41, 47, 18, 44, 37, 27, 46, 38, 24,
3. 43,10,21,36, 45, 22, 18.

4. How diagrams are useful in representing statistical data?

5. How to draw a pie chart?

6. Explain how to draw simple and multiple bar diagrams.

7. Explain how to draw Component and percentage bar diagrams.

8. Draw frequency curve for the following data
9. Explain the concept of a frequency distribution.
10. What is the difference between a bar chart and a histogram?
11. Describe the purpose of an Ogive graph in statistics.
12. Differentiate between continuous and discrete data.
13. What is the key distinction between interval and ratio scales of measurement?
14. Explain the concept of a stratified sample.
15. Give an example of data that could be measured on a ratio scale.
16. Define the term "mode" in statistics.
17. Why is the geometric mean used instead of the arithmetic mean in certain situations?
18. What is a percentile rank and how is it calculated?
19. Describe the purpose of a scatter plot.

20. Define the term "outlier" in statistical analysis.
21. Explain the concept of data normality and its importance in statistical analysis.
22. Differentiate between cross-sectional and longitudinal data.
23. What is a frequency polygon, and how is it constructed?
24. Describe the concept of correlation in statistics.
25. Define the term "confidence interval" and its significance.
26. Explain the difference between a bar chart and a Pareto chart.
27. Describe the purpose of a box-and-whisker plot.
28. Differentiate between systematic and random sampling methods.
29. How diagrams are useful in representing statistical data?

30. How to draw a pie chart?

31. Explain how to draw simple and multiple bar diagrams.

32. Explain how to draw Component and percentage bar diagrams.

33. Draw frequency curve for the following data

Seed Yield (gms)	**No. of Plants**
2.5-3.5	4
3.5-4.5	6
4.5-5.5	10
5.5-6.5	26
6.5-7.5	24
7.5-8.5	15
8.5-9.5	10
9.5-10.5	5

MEASURES OF CENTRAL TENDENCY AND DISPERISION

Fill in the blanks

1. Arithmetic mean is very much affected by_____________.
2. In open end distribution______________cannot be determined.
3. The mean of the values 11,12,13,14, and 15 is_____________.
4. The sum of deviations from mean is _____________
5. Sum of square of the deviation from mean is always _____________
6. Geometric mean cannot be calculated if any value of the set is _____________
7. If the AM of certain value is 9 and GM is 6, the HM is_____________.
8. Relation between mean median and mode _____________
9. 2^{nd} quartile is same as _____________
10. For measuring average shoe size in a market_____________is used.
11. Range is a__________ measures of dispersion.
12. Inter-quartile range is equal to __________
13. Standard deviation is the___________of variance.
14. Best measures of dispersion is __________
15. __________deviation can be obtained in case of open end intervalas.
16. __________moment represents variances.
17. The moments about mean are called___________moments.
18. Moments about origin are called__________moments.
19. For symmetric distribution coefficient of skewness is __________
20. Formula for coefficient of variation is __________
21. For a mesokurtic curve, β_2 value is ________
22. Unitless measures of dispersion is __________
23. Peakness of a frequency curve is measured by___________curve.

Multiple choice Questions

1. mean is a measure of

   a. central values
   b. dispersion
   c. correlation
   d. none of above
2. which of the following represents median?
   a. First quartile
   b. Fiftieth percentiles
   c. Sixth decile
   d. None of above
3. if each observation of a set is divided by 2, then the mean of new values:

a. is two times the original mean.
b. Is decreased by 2
c. Is half of the original mean
d. Remain the same

4. Harmonic mean is better than other means if the data are for:
   a. Speed or rate
   b. Heights or lengths
   c. Binary values like 0 and 1
   d. None of above
5. Extreme value have no effect on:
   a. Average
   b. Median
   c. Geometric mean
   d. Harmonic mean
6. Correct relationship between A.M. , G.M., and H.M. is :
   a. A.M.=G.M.=H.M.
   b. G.M $\geq$ A.M $\geq$ H.M.
   c. A.M $\geq$ G.M. $\geq$ H.M.
   d. None of above
7. What percentage of values is greater than 3^{rd} quartile?
   a. 75percent
   b. 50percent
   c. 25percent
   d. 0percent
8. A frequency distribution having two modes is said to be:
   a. Unimodal
   b. Bimodal
   c. Trimodal
   d. Without mode
9. The average of n natural numbers isa.
   n(n+1)/2
   b. n+1/2
   c. $n^2(n+1)/2$
   d. none of above
10. for deciles, the total number of pertition values are
    a. 5
    b. 8
    c. 9
    d. 10
11. Which of the following is not a measure of dispersion?
    a. Mean deviation
    b. Quartile deviation
    c. Standard deviation
    d. average deviation from mean

12. Which of the following is a unitless measure of dispersion?
    a. Standard deviation
    b. Mean deviation

c. Coefficient of variance
d. Range

13. Correct formula for mean deviation from a constant A of a Series in which the variatevalues $x_1, x_2, x_3 \ldots\ldots x_k$ have frequencies $f_1, f_2, \ldots\ldots f_k$ respectively is:
   a. $1/N \sum(f_i x_i - A)$
   b. $1/N \sum(f_i(x_i - A))$
   c. $1/N \sum_i |f_i(x_i - A)|$
   d. None of aboveAns.

14. Quartile deviation is given by the formula:a.
   Q.D.=Q_3+Q_1/2
   b. Q.D.= Q_3 - Q_1
   c. Q.D.= (Q_3-Q_1)/2
   d. None of above

15. Sum of squares of the deviations is minimum when deviations are taken from
   a. Mean
   b. Median
   c. Mode
   d. Zero

16. Which measures of dispersion ensures highest degree of reliability?
   a. Standard deviation
   b. Mean deviation
   c. Coefficient of variance
   d. Range

17. Average wages of workers of a factory are Rs. 550 per month and SD of wages is110.The coefficient of variation is:
   a. 30 percent
   b. 15percent
   c. 500 percent
   d. 20 percent

18. Quartile deviation is equal to :
a. Interquartile range
b. Double the Interquartile range
c. Half of Interquartile range
d. None of above

19. For a positive skewed distribution, which of the following inequalities holds?
a. Median> mode
b. Mode> mean
c. Mean>median
d. Mean>mode

20. For a symmetrical distribution, the coefficient of skewness:
   a. $B_1 = 1$
   b. $B_2 = 3$
   c. $B_3 = 0$
   d. $B_4 = -1$

Ans. c

21. For a leptokurtic frequency curve, the measure of kurtosis.
    a. $\beta_1 = 0$
    b. $\beta_2 = -3$
    c. $\beta_3 < 1$
    d. $\beta_2 > 3$

22. in case of positive skewed distribution, the relation between mean, median and modethat holds is :
    a. median>mean>mode
    b. mean>median>mode
    c. mean= median= mode
    d. none of the above

23. in case of positive skewed distribution, the extreme values lie in the
    a. left tail
    b. right tail
    c. middle
    d. anywhere

24. the standard deviation of a set of values will be:
    a. positive when the values are positive
    b. positive when the values are negative
    c. always positive
    d. all the above

25. All values in a sample are same. Then their variance is:
    a. Zero
    b. One
    c. Not calculable
    d. All the above

Short Questions (each carry 2 Marks)

1. What is the main purpose of using measures of central tendency in statistics?
2. Define the median of a dataset and explain when it is preferable over the mean.
3. Calculate the mean of the following dataset: 12, 15, 18, 20, 22.
4. Find the median of the dataset: 25, 30, 35, 40, 45, 50, 55.
5. Given the dataset: 5, 7, 8, 9, 10, find the mode.
6. Calculate the weighted mean of the following dataset with respective weights: 10 (weight 2), 15 (weight 3), 20 (weight 4).
7. Calculate the range of the dataset: 25, 30, 35, 40, 45.
8. Find the interquartile range (IQR) for the dataset: 14, 17, 18, 20, 21, 22, 25, 29.
9. Compute the standard deviation for the dataset: 8, 10, 12, 14, 16.
10. Calculate the mean deviation for the dataset: 5, 8, 10, 12, 15.
11. If a dataset has an odd number of values, how do you determine the median?
12. How does an outlier affect the mean and the median differently?
13. What is the mode of a dataset and when might a dataset have multiple modes?
14. Explain why the mean of a dataset can be influenced by extreme values.
15. Define the first moment of a distribution and explain its significance.
16. How do higher moments provide more detailed information about a distribution?
17. Calculate the third central moment (skewness) for a dataset using the formula.

18. Define kurtosis and explain how it describes the shape of a distribution.
19. What does positive kurtosis indicate about the tails of a distribution?
20. Briefly describe Sheppard's corrections and their role in improving kurtosis estimation
21. Define the range of a dataset and explain its limitations as a measure of dispersion.
22. Calculate the interquartile range (IQR) given the values of the first and third quartiles.
23. How is the standard deviation different from the mean deviation?
24. Calculate the mean deviation for the dataset: 10, 12, 15, 18, 20.
25. What does the coefficient of variation (CV) indicate about a dataset?
26. Explain the concept of skewness and how it relates to the distribution of data.
27. Calculate the third central moment (skewness) for the dataset: 18, 20, 22, 23, 24, 25, 26, 28.
28. Determine the kurtosis for the dataset: 10, 10, 15, 15, 20, 20, 25, 25, 30, 30.
29. Given a distribution with a mean of 50 and a third central moment of -120, interpret the skewness.
30. Calculate the fourth central moment for the dataset: 2, 4, 6, 8, 10, using the formula for the raw moments.
31. What do you mean by measures of dispersion? Briefly explain different measures of dispersion
32. Write down the characteristics of good measures of dispersion.
33. The weights of 5 ear-heads of sorghum are 100, 102,118,124,126 gms. Find thestandard deviation
34. Compute the coefficient of variation (CV) for the dataset: 12, 14, 16, 18, 20, given that the mean is 16.

Long Questions (each carry 6 Marks)

1. Given the following frequency distribution, calculate the arithmetic mean

Marks	64	63	62	61	60	59
Number of students:	08	18	12	09	07	06

2. The following data pertaining to the number of insects per plant. Find median number of insects per plant.

Number of insects per plant (x)	1	2	3	4	5	6	7	8	9	10	11	12
No. of plants(f)	2	3	5	6	10	13	9	5	3	2	2	1

3. Find the mode for the following

Weight of sorghum ingms (x)	50	65	75	80	95	100
No. of earhead(f)	06	08	16	08	12	04

4. The yield and plant height of a paddy variety are given, he mean and standarddeviation for yield are 50 kg and 10 kg respectively. The mean and standard deviation for plant height are 55 am and 5 cm respectively. Find the variability among yield and plant height.
5. Define C.V. What are its uses?

6. What are the differences between absolute measure and relative measure of dispersion?

7. Find Variance of the marks from the following table.

Length of pods	0-10	10-20	20-30	30-40	40-50	50-60
No of pods	12	18	27	20	17	6

8. Find the variance.

Seed yield in gms (x)	2.5-35	3.5-4.5	4.5-5.5	5.5-6.5	6.5-7.5
No. of plants (f)	4	6	15	15	10

9. If the weights of sorghum ear heads are 45, 60, 48,100, 65 gms. Find theGeometric mean.
10. Compute quartiles for the data given below (grains/panicles) 25, 18, 30, 8, 15, 5,20, 40, 45
11. For which type of data mode can be calculated.

12. Write down the merit and demerit of Arithmetic mean.

13. Write the characteristics of good measures of central tendency.

14. Explain the difference between the mean and the median. Provide examples to illustrate situations where each measure is more appropriate in representing central tendency. Highlight the advantages and limitations of using mean and median in data analysis.

15. Compare the range and the standard deviation as measures of dispersion. Describe how each measure quantifies data spread. Discuss situations where one measure might be preferred over the other and explain why. Provide a dataset example to demonstrate the calculation and interpretation of both range and standard deviation.

16. Discuss the concept of skewness and kurtosis in probability distributions. Explain how skewness indicates the asymmetry of a distribution, while kurtosis measures its tail behavior. Illustrate the differences between positively and negatively skewed distributions, as well as leptokurtic and platykurtic distributions, using graphical representations and practical scenarios.

17. Explain the significance of quartiles and percentiles in statistics. Define what quartiles and percentiles represent and how they divide a dataset. Describe how they are used to identify data points at specific positions within the dataset. Provide an example to illustrate the calculation and interpretation of quartiles and percentiles.

18. Elaborate on the process of calculating the coefficient of variation (CV). Define the coefficient of variation and explain how it relates to the standard deviation and mean. Discuss the advantages of using CV to compare variability between datasets with different units of measurement. Provide a numerical example to demonstrate the calculation and interpretation of CV.

19. Compare and contrast the absolute moments and factorial moments. Explain the concept of moments in statistics and discuss how absolute moments and factorial moments are calculated. Highlight the differences in their formulas and applications. Provide practical examples to illustrate the calculation and interpretation of both types of moments.

20. Explain the difference between a scatter plot and a line graph. Detail the purposes and characteristics of each type of graph. Provide examples of scenarios where one would be more suitable than the other and explain the insights gained from each.

21. Compare the concepts of variance and standard deviation. Define both terms and describe their roles in measuring data variability. Discuss how they are related mathematically and explain when it's advantageous to use one over the other in different contexts.

22. Discuss the characteristics and applications of a histogram and a frequency polygon. Explain how each type of graphical representation displays frequency distribution patterns. Provide examples of datasets that would be effectively represented by each type and explain the benefits of using each in data visualization.

23. Elaborate on the differences between skewness and kurtosis in probability distributions. Define both skewness and kurtosis and explain their significance in describing distribution shapes. Provide graphical representations of distributions with different skewness and kurtosis values to demonstrate how these measures impact data visualization and analysis.

24. Compare the concepts of absolute moments and factorial moments in statistical analysis. Define both terms and explain their significance in characterizing data distributions. Discuss how these moments differ in terms of calculation methods and the insights they provide about the data.

25. Discuss the characteristics and applications of a cumulative frequency graph and an ogive graph. Explain how each type of graph depicts cumulative frequency distributions. Provide real-world examples to illustrate the use of these graphs and explain how they aid in understanding data trends

Correlation and Regression

Fill in the blanks

1. The correlation between two variables can be shown graphically by a ____________
2. The spearman rank order correlation is used when the variables to becorrelated are measured on______________________________scale.
3. When increase in one variable is associated with decrease in other variable, the correlation between these variables is ________________________________
4. Bivariate data involves the analysis of relationships between ______________ variables.
5. A scatter diagram visually represents the relationship between two ______________ variables.
6. Simple correlation measures the ______________ of the linear relationship between two variables.
7. Partial correlation measures the relationship between two variables while controlling for the effect of ______________ variable(s).
8. Rank correlation, also known as ______________ correlation, assesses the relationship between variables using their ranks.
9. Simple linear regression aims to find the best-fitting ______________ line for a set of data points.
10. The principle of ______________ involves minimizing the sum of squared differences between observed and predicted values.
11. Polynomial regression involves fitting a curve of ______________ degree to the data points.
12. Exponential regression models data that exhibits ______________ growth or decay.
13. The process of finding a line or curve that best fits data points is often referred to as ______________.
14. Bivariate data involves the simultaneous analysis of two ______________ variables.
15. A scatter diagram consists of points that represent paired ______________ observations.
16. The coefficient of correlation ranges between -1 and ______________.
17. In simple correlation, the Pearson correlation coefficient measures ______________ relationships.
18. Rank correlation is used when the data isn't suitable for ______________ correlation.
19. In simple linear regression, the slope of the line represents the ______________ between variables.
20. The principle of least squares aims to minimize the ______________ of the residuals.
21. Polynomial regression uses ______________ to fit curves to data.
22. Exponential regression is used for data with an ______________ growth or decay pattern.
23. Fitting a curve to data involves finding the best ______________ representation.
24. Bivariate data analysis deals with the relationship between ______________ variables.
25. A scatter diagram displays the ______________ between two variables.
26. In simple correlation, the value of the Pearson correlation coefficient lies between ______________ and ______________.
27. If the correlation coefficient is close to -1, it indicates a ______________ relationship.
28. Rank correlation, such as the Spearman rank correlation coefficient, is used when data is ______________ or ______________.
29. Simple linear regression aims to find the best-fitting ______________ line.
30. The principle of least squares minimizes the sum of squared ______________.
31. Polynomial regression fits a curve to data using ______________ functions.
32. Exponential regression is suitable for data that follows an ______________ growth or decay pattern.
33. Fitting a curve to data involves finding the optimal ______________ to represent the relationship.
34. The range of correlation coefficient is ______________________
35. Correlation coefficient is independent of ______________________
36. Correlation can be calculated when the variables have______________________unit.
37. If correlation coefficient value is +1 then it indicates ______________________
38. The regression line is also called a ____________

39. The slope of the regression line is represented by ____________
40. In regression, the independent variable is also called ____________________
41. The geometric mean of two regression coefficient is ____________________
42. If one regression coefficient is more than unity then other is ____________
43. Regression coefficient is____________________on change of origin and ____________on change of scale.
44. If r=0 , then the angle between two regression line is ____________________
45. ________gave the term Regression.

Multiple choices

1. The term regression was introduced by:
   a. R.A. Fisher
   b. Sir Francis Galton
   c. Karl Pearson
   d. None of above

2. In a regression line Y on X , the variable X is known as:
   a. Independent variable
   b. Regressor
   c. Explanatory variable
   d. All the above

3. The formula for the estimate of β in the regression equation $Y = \alpha + \beta X + \epsilon$ is:
   a. Cov(X,Y)/V(X)
   b. r$\frac{\sigma x}{\sigma y}$
   c. $\sum(X_i - \bar{X})(Y_i - \bar{Y})/(X_i - \bar{X})^2$
   d. $all\ the\ above$

4. If β_{xy} and β_{yx} are two regression coefficient, they have:
   a. Same sign
   b. Opposite sign
   c. Either same or opposite
   d. Nothing can be said

5. If $\beta_{yx} > 1$, then β_{xy} is:
   a. Less than 1
   b. Greater than 1
   c. Equal to 1
   d. Equal to 0

   Ans. a

6. Scatter diagram of the variate values (X,Y) gives the idea about:
   a. Functional relationship
   b. Regression model
   c. Distribution of errors
   d. None of the above

7. In the regression line Y= α + β X, β is called the:

a. slope of the line
b. intercept of the line
c. neither (a) nor (b)
d. both (a) and (b)

8. The idea of product moment correlation was given by:

a. R.A. Fisher
b. Sir Francis Galton
c. Karl Pearson
d. Spearman

9. The formula for simple correlation co-efficient between the variable X and Ywith usual notations is:

a. Cov(X,Y)/ $\sqrt{V(X)V(Y)}$
b. $\mu_{XY} / \sqrt{\mu_{xx}\, \mu_{yy}}$
c. $\sigma xy/\sigma x\ \sigma y$
d. all the aboveAns.

10. the unit of correlation coefficients is:

a. Kg/cc
b. Per cent
c. Non-existing
d. None of the above

11. The range of simple correlation coefficient is:

a. 0 to ∞
b. -∞ to ∞
c. 0 to 1
d. -1 to 1

12. If ρ=1, the relation between the two variables X and Y is:

a. Y is proportional to X
b. Y is inversely proportional to X
c. Y is equal to X
d. None of the above

13. The geometric mean of the two regression coefficient b_{yx} is equal to:

a. Correlation co-efficient
b. Co-efficient determination
c. Regression co-efficient
d. None of the above

14. Homogencity of three or more population correlation coefficients can betested by:

a. t- test
b. Z- test
c. Chi-square test
d. F- test

15. Regression coefficient is independent of the change of:

a. Scale
b. Origin
c. Both origin and scale
d. Neither origin nor scale

16. A positive significant correlation between the number of shoes produced andthe steel produced per year is:

a. A nonsense correlation
b. A spurious correlation
c. A meaningless correlation
d. All the above

17. If the correlation between the two variables X and Y is negative, theregression coefficient of Y on X is:

a. Positive
b. Negative
c. Not certain
d. None of the above

Short Questions (each carry 2 Marks)

1. What is a scatter diagram? Mention its uses
1. Define correlation. Write its type. Define bivariate data and provide an example of a real-world scenario where bivariate data analysis is applicable.
2. Explain the purpose of a scatter diagram in bivariate data analysis.
3. Differentiate between simple, partial, and multiple correlation. Give an example scenario for each.
4. State the principle of least squares and explain how it's used in simple linear regression.
5. Given a set of data points, calculate the slope and intercept of the regression line using the least squares method.
6. Describe the concept of rank correlation and its advantages over Pearson's correlation coefficient.
7. Explain what bivariate data is and provide an example of a real-world situation where it can be observed.
8. Describe the purpose and visual representation of a scatter diagram in bivariate data analysis.
9. Differentiate between simple, partial, and multiple correlations. Provide an example scenario for each type.
10. How does rank correlation differ from Pearson's correlation coefficient? When might rank correlation be more appropriate?
11. Two variables X and Y have a correlation coefficient of -0.8. Interpret this value in terms of the relationship between X and Y.
12. Define simple linear regression and explain how it's used to model the relationship between two variables.
13. Define regression
14. Explain the principle of least squares and its significance in the context of linear regression.
15. When might polynomial regression be preferred over simple linear regression? Provide an example.
16. Describe exponential regression and provide a real-world scenario where it could be applied.
17. Given a dataset, how can you determine whether linear or exponential regression is more appropriate for modeling the relationship between two variables?
18. Calculate the correlation coefficient for the following dataset: X (5, 10, 15, 20) and Y (8, 12, 18, 22).
19. Given the dataset: X (4, 5, 6, 7) and Y (15, 18, 20, 23), calculate the rank correlation coefficient using Spearman's formula.
20. Compute the coefficient of determination (R-squared) for a simple linear regression with an R-value of 0.75.
21. If the coefficient of correlation between variables A and B is -0.6, what percentage of the variability in B can be explained by A?

22. For a dataset X (3, 6, 9, 12) and Y (20, 40, 60, 80), calculate the predicted value of Y using simple linear regression when X is 8.
23. 6. Given a set of data points X (2, 4, 6, 8) and Y (5, 12, 15, 18), find the equation of the linear regression line using the least squares method.
24. Calculate the quadratic polynomial that best fits the dataset X (1, 2, 3, 4) and Y (6, 8, 10, 14) using the principle of least squares.
25. Fit an exponential curve to the data points X (1, 2, 3, 4) and Y (10, 20, 40, 80), and provide the equation of the curve.
26. If a linear regression model yields the equation Y = 2X + 5, what is the predicted value of Y when X is 7?
27. Determine the coefficient of determination (R-squared) for a quadratic regression model if the total sum of squares is 300 and the explained sum of squares is 220.

Long Questions (each carry 6 Marks)

1. Given two variables X and Y: X (10, 15, 20) and Y (25, 30, 35), calculate the Pearson correlation coefficient.
2. Calculate the rank correlation coefficient (Spearman's rho) for the dataset: X (15, 20, 25) and Y (30, 40, 50).
3. For a simple linear regression, if the slope is 0.5 and the intercept is 10, predict the value of Y when X is 8.
4. Fit a linear regression line to the data points: X (2, 4, 6, 8) and Y (12, 18, 22, 28), and find the equation of the line.
5. Using the principle of least squares, find the equation of the quadratic polynomial that best fits the data: X (1, 2, 3) and Y (3, 7, 12).
6. Fit an exponential curve to the data: X (1, 2, 3, 4) and Y (5, 10, 20, 40), and determine the equation of the curve.
7. Explain the concept of bivariate data with an example from real life.
8. How does a scatter diagram help in understanding the relationship between two variables? Provide a hypothetical scenario.
9. Describe the key differences between simple, partial, and multiple correlations with examples.
10. In a rank correlation analysis, what does a positive correlation coefficient indicate? Provide a brief interpretation.
11. Compare and contrast Pearson correlation coefficient and Spearman's rank correlation coefficient.
12. A dataset has three variables: X, Y, and Z. Explain how you would calculate the partial correlation coefficient between X and Y, controlling for Z.
13. Define simple linear regression. How is the regression line determined using the least squares method?
14. What is the principle of least squares in regression analysis? How does it help in finding the best-fitting line?
15. Compare polynomial regression with linear regression. When might one be more suitable than the other?
16. Explain the exponential regression model. Give an example of a situation where exponential regression is applicable.
17. A dataset has X (1, 2, 3) and Y (5, 8, 12). Calculate the slope and intercept of the linear regression line using the least squares method.
18. Given a dataset X (2, 4, 6) and Y (16, 32, 54), fit an appropriate curve and explain your choice..
19. Mention the properties of the correlation coefficient?
20. Find correlation coefficient between plant height and number of pods.

X	15	20	17	22	25	29	12
Y	18	17	21	23	20	19	22

23. Properties of regression coefficient.

24. Find Correlation coefficient.

Plant Height (in cm)	45	49	42	51	62	53	58
No. Of branches	30	25	20	32	38	30	29

25. Briefly write the significance of correlation coefficient.

26. From a paddy field, 36 plants were selected at random. The length of panicles(x) and the number of grains per panicle (y) of the selected plants were recorded. The results are given below. Fit a regression liney on x.

S.No.	Y	X
1	95	22.4
2	109	23.3
3	133	24.1
4	132	24.3
5	136	23.5
6	116	22.3
7	126	23.9
8	124	24.0
9	137	23.9
10	90	20.0

From the following data, find the regression equation $\sum X = 21$, $\sum Y = 20$, $\sum X2 = 91$, $\sum XY = 74$, $n = 10$.

INDEX NUMBER

1. An index number is a statistical measure that quantifies changes in a _______________ over time.
2. Weighted index numbers assign different _______________ to different items based on their importance.
3. Laspeyre's index is a weighted index using _______________ prices as weights.
4. Paasche's index is a weighted index using _______________ prices as weights.
5. Chain index numbers are used to account for changes in _______________ over time.
6. Index numbers quantify the _______________ of changes in a variable over time.
7. Weighted index numbers assign _______________ to different items based on their importance.
8. The Laspeyres index uses _______________ prices as weights.
9. The Paasche index uses _______________ prices as weights.
10. Chain index numbers adjust for changes in a _______________ set of items.
11. Conversion of fixed base to chain index numbers involves _______________ the index values.
12. Consumer price index numbers measure changes in the cost of _______________ goods and services.
13. Index numbers can help in comparing _______________ of different periods.
14. The Fisher's Ideal Index seeks to eliminate the _______________ bias in index numbers.
15. One limitation of index numbers is the _______________ choice of base year.
16. Index numbers provide a way to measure _______________ in a set of related values.
17. Weighted index numbers assign different _______________ to different items based on importance.
18. Laspeyres' index uses _______________ period quantities as weights.
19. Paasche's index uses _______________ period quantities as weights.
20. Errors in index numbers arise due to _______________ or _______________ biases.
21. Conversion of fixed base to chain index involves updating the index using _______________ data.
22. Consumer price index numbers measure changes in the cost of _______________ over time.
23. Index numbers are useful for _______________ trends and making comparisons.
24. Fisher's Ideal Index aims to eliminate _______________ bias in index numbers.
25. A limitation of index numbers is that they might not account for _______________ changes.

Multiple choice Questions

1. Which type of average is best to construct an index number?
   a) Arithmetic mean
   b) Harmonic mean
   c) Geometric mean
   d) All of the above

2. Index number for the base year is always _________?
   a) 100
   b) 101
   c) 198
   d) 1000

3. Qon means ________ for the given year?
   a) price
   b) quantity in current
   c) commodity
   d) quantity index

4. Index number plays an important role in the field of _________?
   a) Economic and science
   b) Economic and education
   c) Economic and business
   d) Economic and intelligence

5. Index number used to measure the change in the number of goods is known as __________?

a) Quantity index number
b) Simple 1. No
c) Aggregate 1. No
d) Appropriate

6. The period for which the index number is determined is known as ______?
a) Current period
b) Base period
c) Normal period
d) None of the above

7. Widely used weighted index is known as ________?
a) Fisher's ideal index
b) Paasche's index
c) Laspeyres index
d) Marshall- Edgeworth index

8. Weighted index number can be classified into how many categories _____?
a) Two
b) Three
c) Four
d) Five

9. An index number is called a simple index when it is computed from _____?
a) Single variable
b) Bi-variable
c) Multivariable
d) None of the above

10. An index number that can serve many purposes is called ______?
a) General-purpose index
b) Cost of living index
c) Specific purpose index
d) None of the above

11. Factor reversal test is satisfied by which index ______?
a) Simple GM of price relatives
b) Fisher's ideal index
c) Marshall- Edgeworth index
d) All of the above

12. Index number was first ever constructed in which year _____?
a) 1750
b) 1760
c) 1770
d) 1764

13. Circular test an extension of ______?
a) Unit test
b) Circular test
c) Time reversal test
d) Factor reversal test

14. The cost of living (CLI) is always the ________ index
a) Price index
b) Quantity index
c) Weighted index
d) Value index

15. Fisher's index number is based on _____ of Laspeyres index and Paasche's index?
a) Arithmetic mean
b) Geometric mean
c) Harmonic mean
d) None of the above

16. _____________ is the other name of the consumer’s price index?
   a) Wholesale price index number
   b) Cost of living index
   c) Sensitive index
   d) Composite index
17. To calculate dearness allowance, which of the following is necessarily required?
   a) Correlation
   b) Regression
   c) Index number
   d) None of the above
18. To construct an index number, which one can be the problem from the options below?
   a) Big formula
   b) Tedious job
   c) Purpose of index number
   d) None of the above
19. What does this statement mean “price index of the base year concerning 125”?
   a) 25% of the price has increased in the current year in comparison to the base year
   b) 125% of the price has increased in the current year in comparison to the base year
   c) 100% of the price has increased in the current year in comparison to the base year
   d) None of the above
20. In which term can the index number be categorized?
   a) Variable
   b) Constant
   c) Number
   d) None of the above
21. Purchasing power of money is stated as ______ price index?
   a) Equal to
   b) Unequal to
   c) Reciprocal of
   d) None
22. Index numbers are known as a specific type of ____?
   a) Average
   b) Correlation
   c) Dispersion
   d) None
23. Index number is also known as economic ____________?
   a) Parameter
   b) Barometer
   c) Constant
   d) None
24. Which of the following is known as the ideal index number?
   a) Fisher’s
   b) Paasches
   c) Laspeyres
   d) None
25. Weight in Laspeyres price index number is known as _____?
   a) Quantity during the current year
   b) Quantity in the base year
   c) Price during the current year
   d) Price in the base year
26. Weight in Paasche’s price index number is known as ____?
   a) Quantity during the current year
   b) Quantity in the base year
   c) Price during the current year
   d) Price in the base year

27. In case the values are of equal importance, then the index number is known as _______?

a) Weighted
b) Composite
c) Unweighted
d) Value index

28. Calculate fisher's ideal index, given Laspeyres index = 110, Paasche's index = 108.

a) 100
b) 108
c) 109
d) 110

29. In the consumer price index, the household budget method is also known as _____?

a) Simple average of relatives
b) Average of weights
c) The weighted average of relatives
d) All of the above

30. Which one of the following is the use of an index number?

a) To measure changes in quantity
b) To measure changes in demand
c) To measure changes in price
d) To measure changes in variables over a period of time.

31. What is the use of a simple aggregate quantity index?

a) Measure the changes in the quantity of product
b) Measure the changes in the quantity of a range of product
c) Measure the overall changes in the price of a range of a product
d) Measure the changes in the price of a product

32. The weighted system changes with _______?

a) Quarter
b) Year
c) Time
d) None

33. Price index numbers measure the changes in _________?

a) Single variable
b) Relative changes in prices of a commodity between two periods
c) The physical quantity of the produced goods
d) None of the above

34. Which of the following index measures the change from month to month in the cost of a representative "basket" of goods and services of the type which are brought by a typical household?

a) Retails price index
b) Laspeyres price index
c) Financial time index
d) Paasche's price index

35. Which of the below is another object to determine the index number?

a) Index number
b) Reliable
c) Scope
d) Special purpose

36. When the ratio of a sum of prices in the current period to the sum of prices in the base period is expressed in the form of a percentage, it is known as :

a) Simple price index number
b) Simple aggregate price index number
c) Quantity index number
d) The weighted aggregative price index

37. To measure the relative change in purchasing a specific basket of goods and services between two periods of a certain locality with a group of people with fixed incomes, which of the below can be used :

a) Consumer price index
b) Pasche's price index
c) Cost of living index
d) Both (A) and (C)

38. Commodities that show a considerable price fluctuation can be measured by __________ index?
    a) Value
    b) Price
    c) Quantity
    d) None
39. Which of the following are the limitations of using an index number?
    a) It is only useful for short term comparison
    b) It ignores the quantity of the commodity
    c) The use of each of the indexes is restricted for a specific purpose
40. All of the above Index number for base year is always considered as------
    a. 100
    b. 101
    c. 201
    d. 1000
41. Index number is a special type of ----------
    a. Average
    b. dispersion
    c. correlation
    d. None of the above
42. Index number is always expressed in ----------
    a. Percentage
    b. ratio
    c. proportion
    d. None of the above
43. Index number is also called as-----------
    a. Economic barometer
    b. Parameter
    c. Constant
    d. None of the above
44. Which index number is called as ideal index numbr.
    a. Lasperys
    b. Paasches
    c. Fisher
    d. None of the above

Short Questions (each carry 2 Marks)

1. Define index number.
2. Why index numbers are called as economic barometer ?
3. What is a simple index number?
4. Define weighted index number.
5. Explain price relative.
6. Define consumer price index number.
7. What is the wholesale price index?
8. State the two types of price index numbers.
9. What should be the base year like? What is the purpose of an index number?
10. How is a simple index number calculated?
11. Explain the concept of a weighted index number.

12. What does "price relative" refer to in the context of index numbers?
13. Define the consumer price index number.
14. How is the wholesale price index different from the consumer price index?
15. State the two main types of index numbers used in measuring price changes.
16. What role does the base year play in constructing index numbers?
17. What is the key idea behind the Laspeyres' index?
18. Describe the concept of the chain index method.
19. How does the Paasche's index differ from the Laspeyres' index?
20. What does the Edgeworth-Marshall index attempt to address?
21. Explain the concept of substitution bias in index numbers.
22. What is the Fisher's Ideal Index and why is it used?
23. Describe the concept of time reversal in index number errors.
24. How does the consumer price index measure inflation?
25. Why are chain index numbers used in place of fixed base index numbers?
26. What are some limitations of using index numbers?
27. How are index numbers used in adjusting wages for inflation?
28. Give an example of a situation where index numbers are applied outside of economics.

Long Questions (each carry 6 Marks)

1. Calculate cost of living index number using Family Budget method from the followingdata.

Items	Weight	Price in Base year	Price in Current year
Food	10	150	225
House Rent	5	50	150
Clothing	2	30	60
Fuel	3	30	75
Others	5	50	75

2. Explain briefly the steps in the construction of consumer price index number.

3. Compute Laspeyre's , Paasche's,Marshall- Edgeworth, Dorbish – Bowley,and Fisher'sIndex numbers for 2000 from the following data.

Items	1995		2000	
	Price	Quantity	Price	Quantity
A	6	50	10	56
B	2	100	2	120
C	4	60	6	60
D	10	30	12	24
E	8	40	12	36

4. Compute Laspeyre's , Paasche's,Marshall- Edgeworth, Dorbish – Bowley,and Fisher's Indexnumbers for 2000 from the following data.

Items	Base Year		Current Year	
	Price	Expenditure	Price	Expenditure
A	50	100	60	180
B	40	120	40	200
C	100	100	120	12
D	20	80	25	100

5. Compute Laspeyre's , Paasche's,Marshall- Edgeworth, Dorbish – Bowley,and Fisher'sQuantity Index numbers for 2000 from the following data.

Items	Price		Quantity	
	Base year	Current year	Base year	Current year
A	400	85	100	120
B	320	690	20	60
C	720	1600	10	10
D	720	2100	10	20

6. Determine the price index number from the following data using weighted arithmetic mean of price relatives

Commodity	Unit	Weight	Price per unit	
			Base Year	Current Year
A	Quintal	14	90	120
B	Kilogram	20	10	17
C	Dozen	35	40	60
D	Litre	15	50	95

7. Provide examples of situations where index numbers are used beyond economic contexts. Determine the price index number from the following data using weighted arithmetic mean of price relatives.

Commodity	Unit	1979		1984	
		Price(In Rs)	Quantity	Price(In Rs)	Quantity
Rice	Kilogram	8	4	120	10
Ghee	Kilogram	25	2	17	29.50
Egg	Dozen	5	5	60	6.50
Milk	Litre	2	3	95	4

8. With the help of the following data calculate price index number using Fisher9s

Commodity	1979		1984	
	Price(In Rs)	Quantity	Price(In Rs)	Quantity
Rice	32	50	32	50
Barley	30	35	25	40
Maize	16	55	18	50

9. Define index numbers and explain their significance in economics and statistics.
10. Differentiate between weighted and unweighted index numbers, providing examples of situations where each is applicable.
11. Describe Laspeyres' index and Paasche's index, highlighting their formulas and when they are used.
12. Explain the concept of the "base year" in index numbers and its role in constructing index values.

13.Compare and contrast the Laspeyres' and Paasche's indices in terms of their biases and applications. What are the potential sources of errors in index numbers? How do these errors affect the accuracy of the index?

14.Describe the concept of "time reversal" and explain how it can lead to a reversal of sign in index number errors.

15. Explain the purpose of using chain index numbers in comparison to fixed base index numbers.
16. How does the chain index approach help in addressing the issue of changing consumption patterns over time?
17. Define the Consumer Price Index (CPI) and explain its role in measuring inflation.
18. What is the basket of goods and services in the context of the CPI? How is it determined?
19. Discuss some potential limitations of using the CPI as a measure of inflation.
20. Explain how index numbers are utilized in adjusting wages, pensions, and social benefits for inflation.
    a.