SUBJECT CODE: 15CS/AE/DA45

B.Sc. /B.Com./B.B.A. DEGREE EXAMINATION APRIL 2018
FOURTH SEMESTER

COURSE   :   **ALLIED ELECTIVE**
PAPER   :   **DATA ANALYSIS**
TIME   :   **3 HOURS**   **MAX. MARKS: 100**

SECTION A

Answer all the questions: **20 X 1=20 MARKS**
Choose the best answer:

1. The information/data collected by a person directly is known as _____.
    a) primary   b) secondary   c) modified   d) research
2. _____ is a categorical measurement expressed not in terms of numbers.
    a) primary   b) secondary   c) qualitative   d) quantitative
3. _____ data can state the difference or 'interval' between any two data
    values for a particular variable.
    a) primary   b) quantitative   c) interval   d) cross
4. Selecting the option Treat Consecutive Delimiters which ensures that consecutive spaces
    will result in only a _____break.
    a) single column   b) single row   c) double column   d) double row
5. If the correlation coefficient is zero, then the variables are _____
    a) perfectly independent   b) perfectly dependent
    c) partially dependent   d) partially independent
6. _____ is a commonly used tool to summarize data.
    a) histogram   b) paper   c) excel   d) word
7. The correlation between any two variables is always between _____ and _____.
    a) +2 -2   b) -1 -1   c) -1 +1   d)+1 +1
8. _____ indicates that there is a statistically significant difference between the
    groups' means.
    a) alternative hypothesis b) null hypothesis   c) anova   d) all the above
9. SPSS is a _____.
    a) statistical software   b) software   c) program   d) hardware
10. The other name for categorization is _____.
    a) grouping   b) condensation   c) hierarchy   d) node

Fill in the Blanks:

11. _____ diagram is also known as divided circle.
12. _____ is middle post positional average.
13. _____refer to data whose values cannot be measured numerically but can be either
    classified into sets (categories) according to the characteristics.
14. To display message when an invalid data is entered _____tab is used in data
    validation dialog box.
15. A _____ enables you to quantify the strength of the linear relationship between
    two ranked or numerical variables.
16. _____in Pivot table use the sum of other items in the same field.
17. Column in a database table is called _____.

18. Each row of the database table is called _____.
19. _____ is number of words.
20. _____is defined broadly as an account of an experience that is told in a sequenced way.

SECTION B

ANSWER All QUESTIONS: **5 X 2=10 MARKS**

21. Define primary data.
22. Define discrete data.
23. What is meant by positive correlation?
24. What is meant by weak linear relationship?
25. What is data cleaning?

SECTION C

ANSWER ANY EIGHT QUESTIONS: **8 X 5=40 MARKS**

26. Differentiate between qualitative and quantitative data with example.
27. What are the important points to remember for interpreting the information?
28. Explain briefly how to check for the errors in the data.
29. What are the four main reasons for missing data?
30. Differentiate between type I and type II errors.
31. Explain neatly with an example how to check the association of two variables.
32. Describe a) correl function usage b) relationship between correlation and R^2 .
33. Explain briefly on randomized blocks and two-way anova with an example.
34. What are the necessary checks to be performed before data is ready in the electronic format?
35. Explain briefly on categorizing the data.

SECTION D

ANSWER ANY THREE QUESTIONS: **3 X 10=30 MARKS**

36. Explain briefly about processing of data.
37. Write short notes on importing data into Microsoft excel.
38. Explain with an example the different shapes of histogram.
39. Write short notes on filtering data in excel with example.
40. Explain about analytical aids with example.

****