

10



(Knowledge for Development)

KIBABII UNIVERSITY

(KIBU)

**UNIVERSITY EXAMINATIONS
2019/2020 ACADEMIC YEAR**

**SUPPLEMENTARY/SPECIAL EXAMINATIONS
YEAR TWO SEMESTER TWO EXAMINATIONS**

**FOR THE DEGREE OF
BACHELOR OF SCIENCE
(INFORMATION TECHNOLOGY)**

COURSE CODE : BIT 225

COURSE TITLE : DATA ANALYSIS

DATE: 08/02/2021

TIME: 2.00 P.M. – 4.00P.M.

INSTRUCTIONS TO CANDIDATES

ANSWER QUESTIONS ONE AND ANY OTHER TWO.

QUESTION ONE (COMPULSARY) [30 MARKS]

1. Which of the following provides a measure of central location for the data?
a. standard deviation b. mean c. variance d. range
2. Since the population size is always larger than the sample size, then the sample statistic
a. can never be larger than the population parameter
b. can never be equal to the population parameter
c. can be smaller, larger, or equal to the population parameter
d. can never be smaller than the population parameter
3. The variance of a sample of 169 observations equals 576. The standard deviation of the sample equals
a. 13 b. 24 c. 576 d. 28,461
4. Which of the following is a measure of dispersion?
a. percentiles b. quartiles c. interquartile range d. all of the above are measures of dispersion
5. When data are positively skewed, the mean will usually be
a. greater than the median b. smaller than the median c. equal to the median d. positive
6. The measure of dispersion that is influenced most by extreme values is
a. the variance b. the standard deviation c. the range d. the interquartile range
7. The numerical value of the standard deviation can never be
a. larger than the variance b. zero c. negative d. smaller than the variance
8. A numerical measure of linear association between two variables is the
a. variance b. covariance c. standard deviation d. coefficient of variation
9. The coefficient of correlation ranges between
a. 0 and 1 b. -1 and +1 c. minus infinity and plus infinity d. 1 and 100
10. During a cold winter, the temperature stayed below zero for ten days (ranging from -20 to -5). The variance of the temperatures of the ten day period
a. is negative since all the numbers are negative b. must be at least zero
c. cannot be computed since all the numbers are negative d. can be either negative or positive
11. Which of the following is not a measure of dispersion?
a. mode b. standard deviation c. range d. interquartile range
12. The denominator (bottom) of the z-score formula is
a. The standard deviation b. The difference between a score and the mean
c. The range d. The mean
13. The goal of _____ is to focus on summarizing and explaining a specific set of data.
a. Inferential statistics b. Descriptive statistics c. None of the above d. All of the above
14. Which measure of central tendency takes into account the magnitude of scores?
a. Mean b. Median c. Mode d. Range
15. If a distribution is skewed to the left, then it is _____.
a. Negatively skewed b. Positively skewed c. Symmetrically skewed d. Symmetrical
16. _____ is the set of procedures used to explain or predict the values of a dependent variable based on the values of one or more independent variables.
a. Regression analysis b. Regression coefficient c. Regression equation d. Regression line
17. _____ are used when you want to visually examine the relationship between two quantitative variables.
a. Bar graphs b. Pie graphs c. Line graphs d. Scatterplots
18. The _____ is often the preferred measure of central tendency if the data are severely skewed.
a. Mean b. Median c. Mode d. Range
19. Confidence intervals around a mean value give you a range within which the _____ mean is likely to fall.
a. Population b. sample c. interval d. point estimate
20. When interpreting your results you should consider:
a. effect sizes b. significance levels c. confidence levels d. all of the above.

21. What is the advantage of using SPSS over calculating statistics by hand?
 - a. It equips you with a useful transferable skill.
 - b. It reduces the chance of making errors in your calculations.
 - c. Many researchers use SPSS as it is a recognised software package.
 - d. All of the above
22. In SPSS, what is the "Data Viewer"?
 - a. A table summarizing the frequencies of data for one variable.
 - b. A spreadsheet into which data can be entered.
 - c. A dialog box that allows you to choose a statistical test.
 - d. A screen in which variables can be defined and labeled.
23. How is a variable name different from a variable label?
 - a. It is shorter and less detailed
 - b. It is longer and more detailed.
 - c. It is abstract and unspecific
 - d. It refers to codes rather than variables.
24. What does the operation "Recode Into Different Variables" do to the data?
 - a. Replaces missing data with some random scores.
 - b. Reverses the position of the independent and dependent variable on a graph.
 - c. Redistributes a range of values into a new set of categories and creates a new variable.
 - d. Represents the data in the form of a pie chart.
25. How would you use the drop-down menus in SPSS to generate a frequency table?
 - a. Open the Output Viewer and click: Save As; Pie Chart
 - b. Click on: Analyze; Descriptive Statistics; Frequencies
 - c. Click on: Graphs; Frequencies; Pearson
 - d. Open the Variable Viewer and recode the value labels
26. Why might you tell SPSS to represent the "slices" of a pie chart in different patterns?
 - a. Because the program tends to crash if you ask it to use colour.
 - b. Because the patterns form symbolic visual images of different social groups.
 - c. In order to make full use of the facilities that SPSS can offer.
 - d. If you do not have a colour printer, it makes the differences between slices clearer.
27. When cross-tabulating two variables, it is conventional to:
 - a. represent the independent variable in rows and the dependent variable in columns.
 - b. assign both the dependent and independent variables to columns.
 - c. represent the dependent variable in rows and the independent variable in columns.
 - d. assign both the dependent and independent variables to rows.
28. In which sub-dialog box can the Chi Square test be found?
 - a. Frequencies: Percentages
 - b. Crosstabs: Statistics
 - c. Bivariate: Pearson
 - d. Gender: Female
29. To generate a Spearman's rho test, which set of instructions should you give SPSS?
 - a. Analyze; Crosstabs; Descriptive Statistics; Spearman; OK
 - b. Graphs; Frequencies; [select variables]; Spearman; OK
 - c. Analyze; Compare Means; Anova table; First layer; Spearman; OK
 - d. Analyze; Correlate; Bivariate; [select variables]; Spearman; OK
30. How would you print a bar chart that you have just produced in SPSS?
 - a. In Output Viewer, click File, Print, select the bar chart and click OK
 - b. In Variable Viewer, open bar chart, click File, Print, OK
 - c. In Chart Editor, click Descriptive Statistics, Print, OK
 - d. In Data Editor, open Graphs dialog box, click Save, OK

QUESTION TWO [20 MARKS]

- a. Under what conditions might a **median** be a better measure of the centre of your data set than the **mean**? [4 marks]
- b. Use the following data set to answer the next three questions

Value	1	2	3	4	5	6	7	8
Frequency	2	0	0	3	4	10	12	9

- i. Calculate the mean [3 marks]
- ii. Calculate media [1 mark]
- iii. Calculate the interquartile range [2 marks]
- iv. Describe the distribution of the data [2 marks]
- c. Compare and contrast IBM SPSS and Microsoft Excel as tools for research data analysis. [8 marks]

QUESTION THREE [20 MARKS]

- a. Describe how to Insert a variable in SPSS file. [4 marks]
- b. Explain the process of Selecting cases in SPSS [4 marks]
- c. Briefly explain the importance of recoding feature in SPSS [4 marks]
- d. Describe the steps you will take supposing you want to create a new variable "FinalGrade" which takes "A" if the average mark ≥ 90 , "B" if the average mark ≥ 85 and < 90 , "C" for the other values of the average mark. [8 marks]

QUESTION FOUR [20 MARKS]

- a. Differentiate between the data editor window and the output viewer window [4 marks]
- b. Describe the following view as used in SPSS [3 marks]
- i. Data view [5 marks]
- ii. Variable view [8 marks]
- c. Discuss Four steps you will perform in analysing data using SPSS.

QUESTION FIVE [20 MARKS]

- a. What Is Data Analysis and Why Is It Important? [3 marks]
- b. Distinguish between quantitative and qualitative data [2 marks]
- c. What is the importance of descriptive statistics in data analysis [3 marks]
- d. In your own opinion, what is the importance of inferential statistics in data analysis [4 marks]
- e. Imagine that a tutor asks 100 students to complete a Maths test. The tutor wants to know why some students perform better than others. Whilst the tutor does not know the answer to this, she thinks that it might be because of two reasons: (1) some students spend more time revising for their test; and (2) some students are naturally more intelligent than others. As such, the tutor decides to investigate the effect of revision time and intelligence on the test performance of the 100 students.
- i. With reasons identify the dependent and independent variables for the study. [3 marks]
- ii. Give reasons for each of your answers in (a) above. [3 marks]
- iii. Suggest two possible confounding variables in the above case. [2 marks]