



(University of Choice)

**MASINDE MULIRO UNIVERSITY OF
SCIENCE AND TECHNOLOGY
(MMUST)**

MAIN CAMPUS

**UNIVERSITY EXAMINATIONS
2021/2022 ACADEMIC YEAR**

SECOND YEAR SECOND TRIMESTER EXAMINATIONS

**FOR THE DEGREE
OF
BACHELOR OF SCIENCE IN NURSING (UPGRADING)**

COURSE CODE: NCN 218

COURSE TITLE: BASIC STATISTICS (MAIN)

DATE: 22ND APRIL 2022

TIME: 11.30AM - 02:30PM

INSTRUCTIONS TO CANDIDATES

- (i) WRITE YOUR UNIVERSITY REGISTRATION NUMBER ON EVERY PIECE OF PAPER YOU USE.
- (ii) DO NOT WRITE YOUR NAME ON ANY PIECE OF PAPER YOU USE.
- (iii) THIS MCQ, SAQs AND LAQs
- (iv) ANSWER **ALL THE QUESTIONS** .
- (v) READ CAREFULLY THE ADDITIONAL INSTRUCTIONS PRECEDING EACH SECTION.

TIME: 3 Hours

MMUST observes ZERO tolerance to examination cheating

SECTION A: SHORT ANSWER QUESTIONS (20 POINTS)

1. Which scale is the simplest form of measurement?
 - A. Nominal
 - B. Ordinal
 - C. Interval
 - D. Ratio

2. What is the median of the following set of numbers: 7, 6, 10, 9, 5?
 - A. 7.4
 - B. 7
 - C. 10
 - D. 6

3. Which of the following sets of four numbers has the largest standard deviation?
 - A. 100, 101, 102, 103
 - B. 7, 8, 9, 10
 - C. 5, 5, 5, 5
 - D. 0, 0, 10, 10

4. For this question and the next (Q5), what graphical display would be most appropriate for the variable described? The number of hours per week students study during a semester?
 - A. Pie chart
 - B. Scatterplot
 - C. Bar chart
 - D. Histogram

5. State the reason of your choice _____

6. Which radio stations are the students' favorites?
 - A. Stem plot
 - B. Bar chart
 - C. Boxplot

D. Scatterplot

7. A data set about some students is shown below. Use the data set for this question and the following question.

Student Class Registration Number	Gender	Specialization	GPA
1	Female	Psychology	3.8
2	Male	Biology	3.2
3	Female	Chemistry	2.9
4	Female	Sociology	3.6
5	Male	Biology	2.7

GPA is which of the following?

- A. a categorical variable
- B. a quantitative variable
- C. an individual
- D. Discrete variable

8. In the data set of Question 7, Student registration number is which of the following?

- A. qualitative variable
- B. a categorical variable
- C. a quantitative variable
- D. Ratio variable

9. The people on the ship Titanic when it sank were passengers, either in first class, second class or third class, or crew members. Records were kept so that we know how many people were in each group. Use this information for this question and the one following:

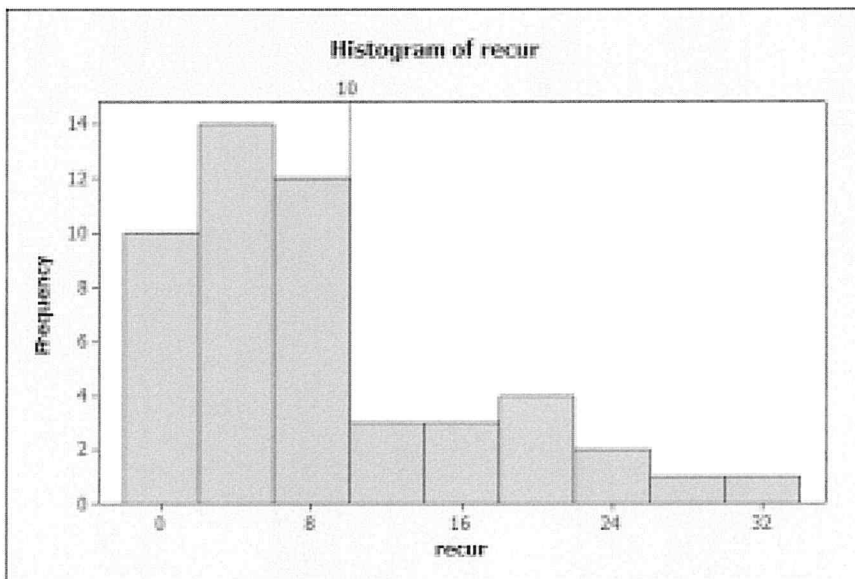
Suppose we wanted to know whether the third class passengers made up more or less than a quarter of all the people on the ship. Which graph would be most appropriate for showing this?

- A. a histogram
- B. a pie chart
- C. a stem plot
- D. a bar chart

10. Using the information in Question 9 above, suppose that we wanted to know which group (first class passengers, second class passengers, third class passengers, crew) had fewest people in it. Which graph would be most appropriate for showing this?

- A. a bar chart
- B. a histogram
- C. a stemplot
- D. a pie chart

11. The length of time (in months) was recorded for 50 patients between the onset of a particular illness and its recurrence. A histogram of the results is shown below. Use this histogram to answer this question and the two questions following.



How would you describe the shape of the distribution of recurrence times?

- A. Skewed to the right
- B. Bell-shaped
- C. The distribution has outliers
- D. Symmetric

12. Look again at Question 11. The value 10 months is marked (at the top of the plot). What percentage of patients had a recurrence time less than (or equal to) 10 months?
- A. 36
 - B. 72
 - C. 48
 - D. 24
13. If population A has a larger standard deviation than population B, which of the following is **NOT** true?
- A. Population B has a smaller variance than population A.
 - B. The mean of a sample of 20 from population A has a larger standard deviation than the mean of a sample of 20 from population B.
 - C. A typical observation from population A will be farther from the mean of population A than a typical observation from B will be from the mean of population B.
 - D. The mean of a sample from population A will on average be larger than the mean of a sample from population B.
14. A researcher is studying students in college in California. She takes a sample of 400 students from 10 colleges. The average age of all college students in California is
- A. a statistic.
 - B. a parameter
 - C. the median.
 - D. a population
15. Half of the observations in a data set are greater than the
- A. mean.
 - B. median.
 - C. mode.
 - D. standard deviation.

15. Since the population size is always larger than the sample size, then the sample statistic
- can never be larger than the population parameter
 - can never be equal to the population parameter
 - can never be zero
 - can never be smaller than the population parameter
16. The mean of a sample is
- always equal to the mean of the population
 - always smaller than the mean of the population
 - computed by summing the data values and dividing the sum by $(n - 1)$
 - computed by summing all the data values and dividing the sum by the number of items
17. The difference between the largest and the smallest data values is the
- variance
 - interquartile range
 - range
 - coefficient of variation
18. If a data set has an even number of observations, the median
- cannot be determined
 - is the average value of the two middle items
 - must be equal to the mean
 - is the average value of the two middle items when all items are arranged in ascending order
19. The value that has half of the observations above it and half the observations below it is called the
- range
 - median
 - mean
 - mode
20. In a sample of 800 students in a university, 160, or 20%, are Business majors. Based on the above information, the school's paper reported that "20% of all the students at the university are Business majors." This report is an example of
- a sample
 - a population
 - statistical inference
 - descriptive statistics

SECTION B: SHORT ANSWER QUESTIONS (40 POINTS)

1. i) Paul has data, for each of the 47 counties, on the percentage of the population which smokes, and the lung cancer rate. She wants to display the connection between the two variables. What sort of graph should she use? **(3 pts)**
 _____ histogram _____ line graph. State the reason for your choice _____
- ii) Jane has data on the number of students enrolled in 10 largest major universities in Kenya. What sort of graph should he use to display the data? **(3 pts)**
 _____ bar chart. State the reason for your choice _____
- iii) Charles has data on the mean December temperature in Kakamega County, for the past fifty years. He wants to draw a graph that will display the trend in the data. What sort of graph should he use? **(3 pts)**
 _____ line graph. State the reason for your choice _____
2. You are tabulating data that classifies a sample of 100 incidents of domestic violence according to the Western Province in which each incident occurs. You number the districts from west to east with Kakamega being number 1 and Mumias being number 5. In your tabulation you write down next to each incident the assigned number of the district in which it occurred. Is the resulting column of district numbers a quantitative or qualitative variable? Why? **(4 points)**
3. A manufacturer of car batteries claims that the average length of life for its grade A battery is 60 months. But the guarantee on this brand is for just 36 months. Suppose that the frequency distribution of the life-length data is unimodal and symmetrical and that the standard deviation is known to be 10 months. Suppose further that your battery lasts 37 months. What could you infer, if anything, about the manufacturer's claim? **(10 points)**
4. Explain how the relationship between the mean and the median provides information about the symmetry or skewness of the data's distribution. **(5 pts)**
5. Calculate the variance and standard deviation for samples where **(6 pts)**
 - a) $n = 10$, $\Sigma X^2 = 84$, and $\Sigma X = 20$.
 - b) $n = 40$, $\Sigma X^2 = 380$, and $\Sigma X = 100$
 - c) $n = 20$, $\Sigma X^2 = 18$, and $\Sigma X = 17$.

6. Illustrate by drawing negatively and positively skewed frequency polygon and estimate the location of the mode, median and mean on the curve **(6 points)**

SECTION C: LONG ANSWER QUESTIONS (40 POINTS)

1. Medical scientists at a pharmaceutical company are working hard to develop a cure for influenza. Preliminary research indicates that the disease may be linked to levels of the polycyclic hydrocarbon dimethylbenzanthracene (DMBA) in a person's bloodstream. Given below are DMBA levels (in parts per million, or ppm) found in blood samples of thirty healthy volunteers. **(20 pts)**

21 24 25 27 30 22 24 25 27 32 23 24 26 28 34
23 24 26 28 36 23 25 26 29 37 23 25 27 30 40

- i) Sketch a histogram of these data. (Make sure you label the axes appropriately.) - **6 pts**
- ii) State in a sentence or two what your graph tells you about the distribution of DMBA in healthy individuals. **2 pts**
- iii) Based upon these data, what are "typical" and "abnormal" levels of DMBA? **2 pts**
- iv) Re-organize your data into class intervals with a class width of 5 then calculate the following (grouped data): **10 pts**
- a. Mode
 - b. Median
 - c. Mean
2. A HPE student randomly selects a sample of five nursing students. He obtains data on how often each has fallen asleep in class, so far this trimester. The data are below. **(20 pts)**

3 9 1 4 8

[Remember to show your work]

- a) Give the sample mean for the data. **5 pts**
- b) Find the sample median for the data. **5 pts**
- c) Find the sample variance for the data. **5 pts**
- d) Find the sample standard deviation for the data. **5 pts**