



JARAMOGI OGINGA ODINGA UNIVERSITY OF SCIENCE AND TECHNOLOGY

SCHOOL OF MATHEMATICS AND ACTUARIAL SCIENCE

UNIVERSITY EXAMINATION FOR DIPLOMA IN APPLIED STATISTICS

2ND YEAR FIRST SEMESTER 2023/2024

REGULAR (MAIN)

COURSE CODE: WAB 2215

COURSE TITLE: Electronic Data Collection And Management

EXAM VENUE: AH

**STREAM: DIPLOMA IN APPLIED
STATISTICS**

DATE: 29/4/24

EXAM SESSION: 14-16.00 PM

TIME: 3.00 HOURS

Instructions:

- 1. Answer question 1 (Compulsory) and ANY other 2 questions**
- 2. Candidates are advised not to write on the question paper.**
- 3. Candidates must hand in their answer booklets to the invigilator while in the examination room.**

QUESTION ONE (30 MARKS)

- a.) Discuss three challenges and benefits of using online and offline data capture methods in the context of a Kenyan healthcare research project. Provide practical examples to support your discussion. [6 marks]
- b.) In a survey conducted in five Kenyan counties, the number of respondents who preferred Product A were as follows: County X - 120, County Y - 90, County Z - 80, County W - 110, County V - 130. Calculate the percentage of respondents who preferred Product A in each county. [3marks]
- c.) Given a dataset with the following values representing monthly rainfall in millimeters (mm) in different Kenyan regions: [120, 80, 90, 110, 130], calculate the mean rainfall. [2 marks]
- d.) Provide three potential reasons why variables might be missing in a dataset. How can researchers address each of these reasons in the context of a Kenyan agricultural study? [6 marks]
- e.) In a survey on household income in Kenyan urban areas, the income of three households was missing. If the incomes of the remaining households were \$1,000, \$1,200, \$800, and \$1,500, calculate the mean household income. [4 marks]
- f.) A Kenyan research project generates 2 terabytes of data annually. If this data needs to be stored in a cloud platform that charges \$0.05 per gigabyte per month, calculate the estimated monthly cost of storing this data. [4 marks]
- g.) Explain the importance of template preparation in the context of manual (PAPI) data collection. Provide a specific example related to a Kenyan educational survey. [5 marks]

QUESTION TWO (20 MARKS)

- a.) Discuss the impact of missing variables on the validity and reliability of research findings, focusing on a Kenyan educational dataset. Propose strategies to handle missing variables to ensure accurate and meaningful results. [6 marks]
- b.) Imagine you are conducting a health survey in a rural Kenyan community, and the age variable for a significant portion of the respondents is missing. How would you address this missing data issue to ensure the robustness of your study? [4 marks]
- c.) Explain the concept of archiving data in the context of cloud computing. Provide a specific example of how a Kenyan research institution could benefit from cloud-based data archiving. [10 marks]

QUESTION THREE (20 MARKS)

- a.) Discuss the potential security challenges associated with storing research data in the cloud, particularly in the Kenyan context. Propose strategies to address these challenges and ensure data integrity. [10 marks]
- b.) A Kenyan NGO is considering migrating its data storage to the cloud. Currently, they have 500 gigabytes of data. If the cloud service provider offers a migration discount of 20%, calculate the total cost of migrating the NGO's data to the cloud. [10 marks]

QUESTION FOUR (20 MARKS)

- a.) In a Kenyan agricultural study, researchers have identified 10 variables to be recorded manually. If each variable has 15 possible codes, calculate the total number of unique combinations that can be generated. [6 marks]
- b.) Compare and contrast the advantages and disadvantages of manual (PAPI) and electronic (CAPI) data collection methods, highlighting their suitability in diverse Kenyan research settings. [8 marks]
- c.) A research team in Kenya is preparing a code book for a health survey. If there are 8 different codes assigned to record various medical conditions, and each code corresponds to a specific illness, calculate the percentage of codes representing non-communicable diseases. [6 marks]

QUESTION FIVE (20 MARKS)

- a.) Examine the importance of proper storage of data source files in a research project. Discuss potential risks and benefits associated with different storage methods, drawing on examples from Kenyan studies. [5 marks]
- b.) In a Kenyan education survey, researchers identified 15 variables to be cleaned. If, during the data cleaning process, they found and corrected errors in 8% of the entries, calculate the total number of entries that were cleaned. [5 marks]
- c.) Compare the advantages and disadvantages of manual (PAPI) and electronic (CAPI) data collection methods, emphasizing their applicability in different Kenyan research contexts. Consider factors such as cost, data quality, and accessibility. [5 marks]
- d.) Discuss the role of a code book in ensuring data accuracy and consistency in research. Provide examples from Kenyan research projects where a well-prepared code book has positively influenced the data collection process.

