The Management University of Africa

Sponsored by the Kenya Institute of Management

UNDERGRADUATE UNIVERSITY EXAMINATIONS

SCHOOL OF MANAGEMENT AND LEADERSHIP

DEGREE OF BACHELOR OF BUSINESS INFORMATION TECHNOLOGY

BBIT 208: DATA MINING

DATE: 7TH APRIL 2026

DURATION: 2 HOURS

MAXIMUM MARKS: 70

INSTRUCTIONS:

1. This paper contains **SIX (6)** questions.
2. Question **ONE** is compulsory.
3. Answer any other **THREE** questions.
4. Question **ONE** carries **25 MARKS** and the rest carry **15 MARKS** each.
5. **Save all work in one folder named: BBIT208_APR26_AdmissionNo**
6. This is a **practical examination**. All work must be done on the computer.

7. Use **separate worksheets** for each question
8. All formulas and steps must be visible
9. Do **NOT** hard-code values

QUESTION ONE

Read the Case Study below carefully and answer the questions that follow:

You are provided with an Excel dataset containing supermarket transaction data. Management wants insights to improve sales performance and customer targeting.

Required:

A. Using Excel:

i. Identify the number of records and attributes

ii. Generate summary statistics for Quantity Sold and Unit Price

iii. Identify missing values, duplicates, and outliers **(10 Marks)**

B. Clean the dataset by:

i. Handling missing values appropriately

ii. Removing or correcting duplicate Transaction IDs

iii. Treating extreme outliers

iv. Create a new worksheet named **Clean Datac**. **(10 Marks)**

C. Create a Total Sales column AND

i. Categorize Total Sales into Low, Medium, High

ii. Briefly explain your transformation logic (in Excel) **(5 Marks)**

QUESTION TWO

A. Use PivotTables to identify:

i. Top-selling product categories **(5 Marks)**

B. Create a PivotTable showing:

i. Product Category vs Customer Type **(5 Marks)**

C. Explain two meaningful patterns discovered. **(5 Marks)**

QUESTION THREE

A. Define IF-based rules to classify customers into:

   i. High Value

   ii. Medium Value

   iii. Low Value **(5 Marks)**

B. Apply the classification rules to the dataset **(5 Marks)**

C. Evaluate classification results using counts and percentages. **(5 Marks)**

QUESTION FOUR

A. Select and justify variables used for segmentation. **(5 Marks)**

B. Use sorting, filtering, or grouping to form at least three customer clusters. **(5 Marks)**

C. Describe key characteristics of each cluster. **(5 Marks)**

QUESTION FIVE

A. Identify anomalies using:

   - Conditional formatting or statistical thresholds **(5 Marks)**

B. Create a chart highlighting anomalies. **(5 Marks)**

C. Explain possible causes of the anomalies identified**.** **(5 Marks)**

QUESTION SIX: INSIGHTS & DECISION SUPPORT

A. Summarize three key insights from the analysis. **(5 Marks)**

B. Provide data-driven recommendations to management. **(5 Marks)**

C. Using your analysis results, propose a practical action plan that management can implement over the next 3 months.

   Your plan should address:

   i. Product strategy

   ii. Customer segmentation

iii. Sales improvement initiatives **(5 Marks)**