



(Knowledge for Development)

**KIBABII UNIVERSITY
(KIBU)**

**UNIVERSITY EXAMINATIONS
2022/2023 ACADEMIC YEAR**

**END OF SEMESTER EXAMINATIONS
YEAR THREE SEMESTER TWO EXAMINATIONS
FOR THE DEGREE OF BACHELORS OF SCIENCE
(INFORMATION TECHNOLOGY)**

COURSE CODE : BIT 326

COURSE TITLE : DATA WAREHOUSING AND MINING

DATE: 25/04/2023 TIME: 09.00 A.M.-11.00 A.M. 2HRS

INSTRUCTIONS TO CANDIDATES

ANSWER QUESTION ONE AND ANY OTHER TWO (2) QUESTIONS

QUESTION ONE [COMPULSORY] [30 MARKS]

- a. Suppose a data warehouse for Kibabii University consists of the following four dimensions: student, course, semester, and lecturer, and two measures count and average_grade. When at the lowest conceptual level (e.g., for a given student, course, semester, and lecturer combination), the average_grade measure stores the actual course grade of a student. At higher conceptual levels, average_grade stores the average grade for the given combination.
- i. Design an information package for the data warehouse. [6 Marks]
 - ii. Draw a star schema diagram for the data warehouse. [6 Marks]
 - iii. Starting with the base cuboid [student, course, semester, lecturer], what specific OLAP operations (e.g., roll-up from semester to year) should one perform in order to list the average grade of courses for each Kibabii University Student. [6 Marks]
- b. Describe any three characteristics of a data warehouse. [6 Marks]
- c. Using illustrations, explain the differences between star schema, snowflake schema and fact constellation schema. [6 Marks]

QUESTION TWO [20 MARKS]

- a. Consider an online travel agency that helps customers to plan and schedule their holidays. The agency maintains all past history in a data warehouse.
- i. Describe the different classes of users who could access this data warehouse. [6 Marks]
 - ii. Design the information delivery framework for this data warehouse. [4 Marks]
- b. With the help of a sample data set, describe how the K-Means clustering algorithm works. [6 Marks]
- c. Identify the advantages and disadvantages of the top-down approach of data warehouse architecture. [4 Marks]

QUESTION THREE [20 MARKS]

- a. "Data warehousing is the only viable means to resolve the information crisis and to provide strategic information". Justify. [2 Marks]
- b. What is data purging in the context of data warehousing? [2 Marks]
- c. With the aid of a neat diagram, explain the architecture of a data warehouse. [8 Marks]

- d. Explain any four types of activities that are part of ETL process and identify the activity that is time-consuming. **[8 Marks]**

QUESTION FOUR [20 MARKS]

- a. Describe a factless fact table and identify a situation where a factless fact table may be used. **[4 Marks]**
- b. Explain why is data mining being used in today's businesses. **[8 Marks]**
- c. Using examples to support your answers, discuss the types of metadata stored in a data warehouse. **[8 Marks]**

QUESTION FIVE [20 MARKS]

- a. Discuss Association Rule in Data Mining. **[2 Marks]**
- b. Consider a data warehouse for Kibabii University health centre where three dimensions; Doctor, Patient, Time and two measures; Count and Charge, where charge is the fee, a patient is charged by the doctor per visit.
- i. Design a star schema for the data warehouse. **[8 Marks]**
- ii. Create a cube and illustrate the following OLAP operations; Rollup, Drill down, Slice, Dice and Pivot. **[10 Marks]**