(Knowledge for Development)

KIBABII UNIVERSITY
(KIBU)

UNIVERSITY EXAMINATIONS
2021/2022 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: 06/09/2021   **TIME: 9.00 A.M. – 11.00 A.M.**

INSTRUCTIONS TO CANDIDATES

ANSWER QUESTIONS ONE AND ANY OTHER TWO

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. Khetia group of stores has a huge sales network. To control the sales, it is divided into regions. Each region has multiple zones and each zone has different towns. Each sales person is allocated different towns. The objective is to track sales figure at different granularity levels of region and to count number of products sold.

   i.   Design a star schema by considering granularity levels for region, sales person and time.   **[6 Marks]**

   ii. Convert the star schema to snowflake schema.   **[6 Marks]**

QUESTION TWO [20 MARKS]

a. Describe any four characteristics of the computing environment needed to provide strategic information.   **[8 Marks]**

b. Explain any two reasons why operational systems are not suitable for providing strategic information.   **[4 Marks]**

c. With the aid of a diagram, discuss the typical architecture of a typical data mining system.   **[8 Marks]**

QUESTION THREE [20 MARKS]

a. Use a well labelled diagram describe the components of a data warehouse.   **[12 Marks]**

b. Discuss application and major issues in data mining.   **[6 Marks]**

c. Distinguish between slowly changing dimension and rapidly changing dimension. **[2 Marks]**

QUESTION FOUR [20 MARKS]

a. Describe any four types of activities that are part of the ETL process. **[8 Marks]**

b. Explain why is metadata vital for end-users. **[4 Marks]**

c. Discuss the differences between data mining and OLAP. **[8 Marks]**

QUESTION FIVE [20 MARKS]

a. Explain the concept of Neural Networks. **[4 Marks]**

b. Describe the functional features of a web-enabled warehouse. **[10 Marks]**

c. Describe any three different methods for information delivery. **[6 Marks]**