(Knowledge for Development)

KIBABII UNIVERSITY
(KIBU)

UNIVERSITY EXAMINATIONS
2021/2022 ACADEMIC YEAR

SPECIAL/SUPPLEMENTARY 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: 23/11/2022   TIME: 11.00 A.M. – 1.00 P.M.

INSTRUCTIONS TO CANDIDATES

ANSWER QUESTIONS ONE AND ANY OTHER TWO

QUESTION ONE [COMPULSORY] [30 MARKS]

a. Describe any three characteristics of a data warehouse. **[6 Marks]**

b. Using illustrations, explain the differences between star schema, snowflake schema and fact constellation schema. **[6 Marks]**

c. Explain three benefits of data warehouse to organizations that use them. **[6 Marks]**

d. For Khetia Supermarket Chain consider the following dimensions, Product, Store, Promotion and Time. The schema contains a central fact table named sales facts with three measures namely unit_sales, shilling_sales and shillings_cost.

   i. Design an information package for this application. **[6 Marks]**

   ii. Design a star schema for this application. **[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. **[5 Marks]**

   ii. Design the information delivery framework for this data warehouse. **[5 Marks]**

b. With the help of a sample data set, describe how the K-Means clustering algorithm works. **[6 Marks]**

QUESTION THREE [20 MARKS]

a. With the aid of a neat diagram, explain the architecture of a data warehouse. **[8 Marks]**

b. *Discuss Association Rule in Data Mining.* **[2 Marks]**

c. Explain any four types of activities that are part of ETL process and identify the activity that is time-consuming. **[10 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. "Data warehousing is the only viable means to resolve the information crisis and to provide strategic information". Justify. **[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]**