TECHNICAL UNIVERSITY OF MOMBASA

UNIVERSITY EXAMINATION FOR:

BARCHELOR OD SCIENCE IN INFORMATION TECHNOLOGY

EIT 4415: **INTRODUCTION TO DATA WAREHOUSING**

END OF SEMESTER EXAMINATION

SERIES: 2016/2017

TIME: 2 HOURS

DATE: Pick DateSelect MonthPick Year

Instructions to Candidates
You should have the following for this examination
-Answer Booklet, examination pass and student ID
This paper consists of Choose No questions. Answer Question one(Compulsory) and any other two questions.
Do not write on the question paper.

QUESTION ONE

a) The main features of data mining operations, include predictive modeling, database segmentation, link analysis, and deviation detection. Write notes on the above data mining operations. **[Marks 12]**

b) List four main application areas of data mining **[Marks 4]**

c) State at least two mining results in each of the application areas listed in (b) above **[Marks 8]**

a) Data mining tools sometimes offer a choice of operations to implement a technique. Outline three of the criteria for selection of these tools. **[Marks 6]**

QUESTION TWO

a) Explain the different types of data ware housing techniques listed below.

      i)   Enterprise Data warehousing
      ii)   Operational Data Store
      i)   Data Mart
      ii)   Operational Data Store (ODS)   **[Marks 8]**

a) Explain the need (TWO reasons) for ETL in the data warehouse Phase
                                                 [Marks 6]
      i)   List TWO tools available for ETL   **[Marks 2]**

b) What is the difference between View and Materialized View?   **[Marks 4]**

QUESTION THREE

a) In a networked environment with series of clients and users in a database management systems, Explain Three reasons why it is necessary to implement transaction control in both database management systems and data warehouse
                                                 [Marks 6]

b) Write brief notes on the following concurrency control features
      i)   serializability.
      ii)   Deadlock and how it can be resolved.
      iii)   Timestamping
      iv)   Optimistic concurrency control.
      V)   Granularity of locking   **[Marks 14]**

QUESTION FOUR

a) As a database administrator, you are always of concern to availability, integrity and consistency of the data items in the database. Outline FIVE various types of database failures   **[Marks 5]**

b) In the event of database failures, Outline FOUR main recover techniques you will employ to ensure continuity of the business systems.   **[Marks 8]**

c) Describe what a fact table is in the context of snowflake schema and describe how it is generated   **[Marks 4]**

d) What is the difference between metadata and data dictionary?   **[Marks 3]**

QUESTION FIVE

a) You have been appointed to head the retail chain supermarket in Mombasa. The design of the applications are distributed across all the business units i.e. Point of sale, Inventory management, Staff payroll, Creditors, Debtors etc. The Business has all these data spanning over a five year period.

    i)   State how data mining would add value to this retail chain enterprise company   **[Marks 6]**

    ii)   Describe at least four application area you will use the data mining concepts

                                                             [Marks 8]

iii)   Describe the process you would undertake to identify patterns and relationships that are already intuitive   **[Marks 6]**