DEDAN KIMATHI UNIVERSITY OF TECHNOLOGY

University Examinations 2015/2016

FOURTH YEAR **SEMESTER** EXAMINATION FOR THE DEGREE OF **CRIMINOLOGY AND SECURITY MANAGEMENT**

BSM 2419: BUSINESS DATA MINING AND WAREHOUSING

DATE: 17TH DECEMBER 2015   **TIME: 11.00 AM – 1.00 PM**

INSTRUCTIONS: Attempt QUESTION ONE (Compulsory) and any other two questions.

QUESTION ONE (30 MARKS)
- a) Describe the term Data Mining. (2 marks)
- b) Describe a Data warehouse. (4 marks)
- c) Describe different types of data warehouse. (6 marks)
- d) In the context of data mining define the term business intelligent. (4 marks)
- e) List and explain any six challenges of data mining. (6 marks)
- f) Given the following set of records produce dependency rules which will predict occurrence of an item based on occurrences of other items. (4 marks)

TID	Items
1	Bread, Coke, Milk
2	Beer, Bread
3	Beer, Coke, Diaper, Milk
4	Beer, Bread, Diaper, Milk
5	Coke, Diaper, Milk

- g) State the reasons why data mining is important. (4 marks)

QUESTION TWO (20 MARKS)
- a) Describe the term OLAP (On-Line Analytical Processing). (2 marks)
- b) Describe the following OLAP server architecture. (6 marks)
- c) Describe the following data mining problems and in each case state the area of its application. (12 marks)
    - i. Association rules
    - ii. Sequential. Patterns
    - iii. Classification/Regression
    - iv. Clustering

QUESTION THREE (20 MARKS)

a) What is Business Intelligence?   (2 marks)

a) Given a set of sequences, find the complete set of frequent subsequences. (4 marks)
Sequence database

SID	Sequence
10	<(a)(abc)(ac)(d)(cf)>
20	<(ad)(c)(bc)(ae)>
30	<(ef)(ab)(df)(c)(b)>
40	<(e)(g)(af)(c)(b)(c)>

b) Differentiate between OLTP and OLAP.   (8 marks)

c) Define the following terms as applied in data warehouse.   (6 marks)
   i.   Cube
   ii. Slice
   iii. Dice

QUESTION FOUR (20 MARKS)

a) Compare the differences between fact and dimension tables.   (8 marks)

b) Describe the different approaches to handling missing values.   (6 marks)

c) With the aid of examples describe the following terms.   (6 marks)
   i.   Closedness
   ii. Maximality

QUESTION FIVE (20 MARKS)

a) Describe a data warehouse.   (4 marks)

b) Describe any three types of data warehouse.   (6 marks)

a) Data are organized around one or more fact tables. Each Fact Table collects a set of homogeneous events (facts) characterized by dimensions and dependent attributes. With the aid of the data in the table below design a conceptual star schema. (10 marks)

Product	Supplier	Store	Period	Units	Sales
P1	S1	St1	1qtr	30	1500
P2	S1	St3	2qtr	100	9000