(Knowledge for Development)

KIBABII UNIVERSITY

UNIVERSITY EXAMINATIONS
2020/2021 ACADEMIC YEAR

END OF SEMESTER EXAMINATIONS
YEAR THREE SEMESTER TWO EXAMINATIONS

FOR THE DEGREE OF
(COMPUTER SCIENCE)

COURSE CODE : CSC 367E

COURSE TITLE : DATA WAREHOUSING & MINING

DATE: 13/10/2021 **TIME: 02.00 P.M – 04.00 P.M**

INSTRUCTIONS TO CANDIDATES

ANSWER QUESTIONS ONE AND ANY OTHER TWO.

QUESTION ONE (COMPULSORY) [30 MARKS]

a) Why would Kibabii University Datawarehouse consultant prefer Bottom-Up approach in the implementation of a data warehouse. **[2 marks]**

b) All the past attempts by Information Technology to provide strategic information failed. Explain why. **[3 marks]**

c) Suppose you are an IT Specialist in your company, discuss five characteristics of the computing environment needed to provide strategic information for your esteemed organization. **[5 marks]**

d) Analyze the differences between operational systems and informational systems computing environments. **[5 marks]**

e) Discuss the types of processing that take place in a data warehouse **[4 marks]**

f) The authors in the field of data warehousing and mining have a proposition that a data warehouse is considered an environment, and not a product. Justify this statement. **[5 marks]**

g) In the design and implementation of a data warehouse, it is believed that every data structure in the data warehouse contains the time element. Explain why? **[3 marks]**

h) Explain data granularity and how it is applicable to the data warehouse. **[3 marks]**

QUESTION TWO [20 MARKS]

a) Define data mining? **[1 mark]**

b) Differentiate between the following terms as used in data warehousing and mining.

   i. Descriptive and predictive mining tasks **[2 marks]**

   ii. Database vs. data warehouse **[2 marks]**

c) Explain any four data mining techniques. **[8 marks]**

d) Explain the steps involved in knowledge discovery process **[7 marks]**

QUESTION THREE [20 MARKS]

a) What is meant by data fusion? Where does it fit in data warehousing? **[2 marks]**

b) Explain what is meant by agent technology? How can this technology be used in a data warehouse? **[5 marks]**

c) Describe any one of the options available to integrate ERP with data warehousing. What is CRM? How can you make your data warehouse CRM-ready? **[8 marks]**

d) What do we mean by a Web-enabled data warehouse? Describe three of its functional features **[5 marks]**

QUESTION FOUR [20 MARKS]

a) What is a data cube? **[1 mark]**

b) Differentiate between the following terms as used in data warehousing.

i. Facts and dimensions **[2 marks]**

ii. Base cuboid and apex cuboid **[2 marks]**

c) Just as relational query languages like SQL can be used to specify relational queries, a data mining query language can be used to specify data mining tasks. Define a multidimensional schema for these data using SQL-based data mining query language (DMQL) using two language primitives, one for cube definition and one for dimension definition syntax for each of the following:

Location = "Chicago" "Vancouver"					**Location** = "New York"				**Location** = "Toronto"				**Location** =			
Time (quarter)	**Item**				**Item**				**Item**				**Item**			
	Home Ent.	Comp.	Phone	Security	Home ent.	Comp.	Phone	Security	Home ent.	Comp.	Phone	Sec	Home ent.	Comp.	Phone	Sec
Q1	854	882	89	623	1087	968	38	872	818	746	43	591	605	825	14	400
Q2	943	890	64	698	1130	1024	41	925	894	769	52	682	680	952	31	512
Q3	1032	924	59	789	1034	1048	45	1002	940	795	58	728	812	1023	30	501
Q4	1129	992	63	870	1142	1091	54	984	978	864	59	784	952	1038	38	580

i) Star schema definition **[5 marks]**

ii) Snowflake schema definition. **[5 marks]**

iii) Fact constellation schema definition **[5 marks]**

QUESTION FIVE [20 Marks]

a) Define Cluster Analysis **[2 marks]**

b) Identify the typical requirements of clustering in data mining **[5 marks]**

c) Discus the following clustering methods

i. Partitioning Methods **[3 marks]**

ii. Hierarchical Methods **[3 marks]**

iii. Density-Based Methods **[3 marks]**

d) Outline the types of outlier detection in mining **[4 marks]**