namibia university
OF SCIENCE AND TECHNOLOGY

FACULTY OF COMPUTING AND INFORMATICS

DEPARTMENT OF COMPUTER SCIENCE

QUALIFICATION: BACHELOR OF COMPUTER SCIENCE	
QUALIFICATION CODE: 07BACS	LEVEL: 7
COURSE: DATA AND WEB MINING	COURSE CODE: DWM710S
DATE: JUNE 2022	SESSION: 1
DURATION: 3 HOURS	MARKS: 100

FIRST OPPORTUNITY EXAMINATION QUESTION PAPER	
EXAMINER:	MR GEREON KOCH KAPUIRE
MODERATOR:	MRS JOY ANURIOHA

THIS QUESTION PAPER CONSISTS OF 4 PAGES
(Excluding this front page)

INSTRUCTIONS

1. Answer all questions.
2. When writing take the following into account: The style should inform than impress, it should be formal, in third person, paragraphs set out according to ideas or issues and the paragraphs flow in a logical order. The information provided should be brief and accurate.
3. Please, ensure that your writing is legible, neat and presentable.

Question 1   [4 Marks]

During a workshop, a Senior Data Analyst says, *"Every enterprise benefits from collecting and analysing its data"*. Would you agree with that or not and why? Explain the why using a practical example.

Question 2   [7 Marks]

What are the essential steps in the process of knowledge discovery in various areas where data mining is drawn?

Question 3   [8 Marks]

There are several data mining functionalities. Explain the difference between data characterization and data discrimination. Give a practical example for each.

Question 4   [4 Marks]

Explain the term classification and how we would predict a class label of objects for which class labels are unknown.

Question 5   [6 Marks]

Are all patterns of interest to potential users? What makes a pattern interesting?

Question 6   [4 Marks]

By illustration with a practical example, please explain Outlier analysis and Cluster analysis.

Question 7   [4 Marks]

Explain the difference between Binary Attributes and Nominal Attributes. Give an example for each.

Question 8   [2 Marks]

As a Junior Data Miner for a Marketing Company, you are tasked to visualize data. You have decided to use a Quantile-Quantile plot. How would you explain your decision as to why you would use that?

Question 9   [4 Marks]

Describe various methods for Noisy Data.

Question 10   [6 Marks]

Select a report or visualization that provides some analyses of social web data and explain the steps to accomplish the activity. First, have a look at the 5 great examples of data visualization to get inspired.

Several reports or visualization analyses that you could use are: "Where in the World Are the Hottest Social Networking Countries?", "Visualizing Sochi2014", "Twitter Topography", "Obama vs. McCain in New Media use", "Analyzing Wikipedia Authors", etc.

Question 11   [6 Marks]

What are the various strategies used for data transformation?

Question 12   [2 Marks]

By illustration with a practical example, please explain the concept of Market basket analysis.

Question 13   [8 Marks]

What are the major challenges of mining a huge amount of data (such as billions of tuples) in comparison with mining a small amount of data (such as a few hundred tuple data set)?

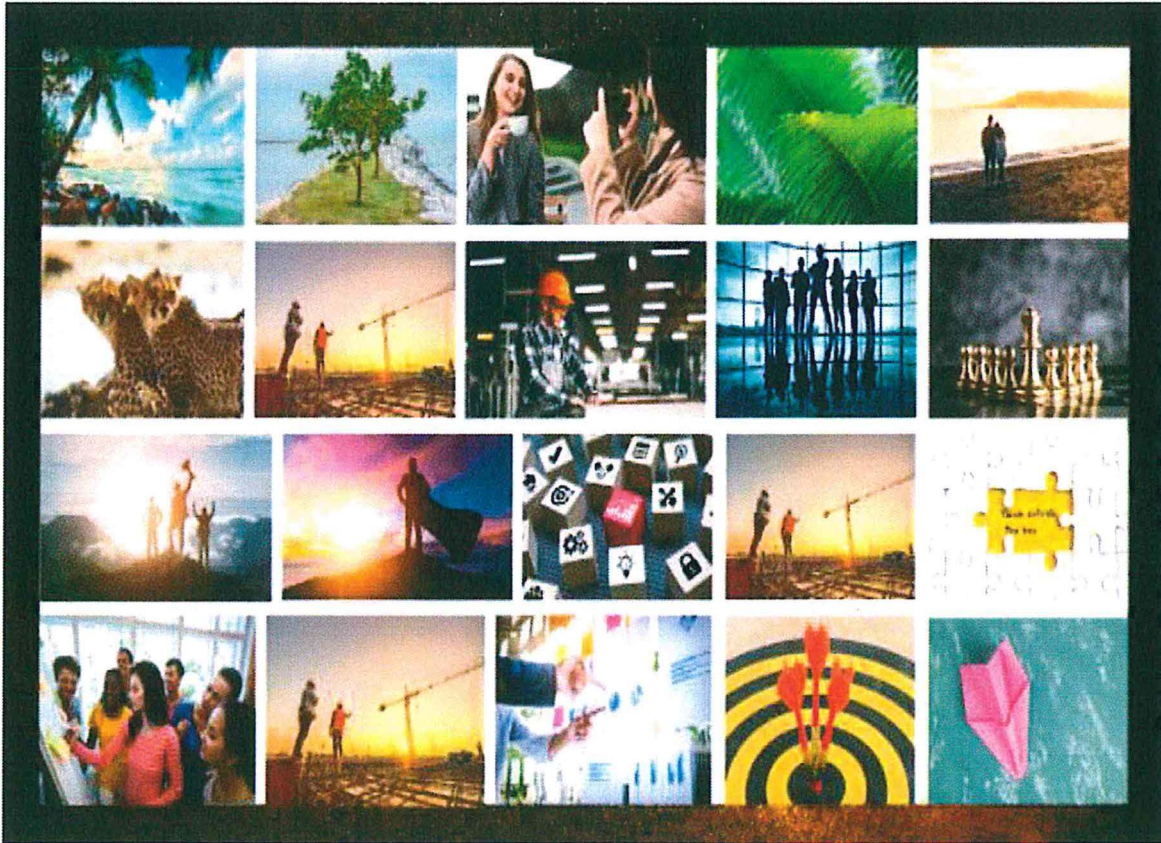Question 14   [2 Marks]

What are the pros and cons of the top-down and bottom-up approaches to data warehouse development?

Question 15   [5 Mark]
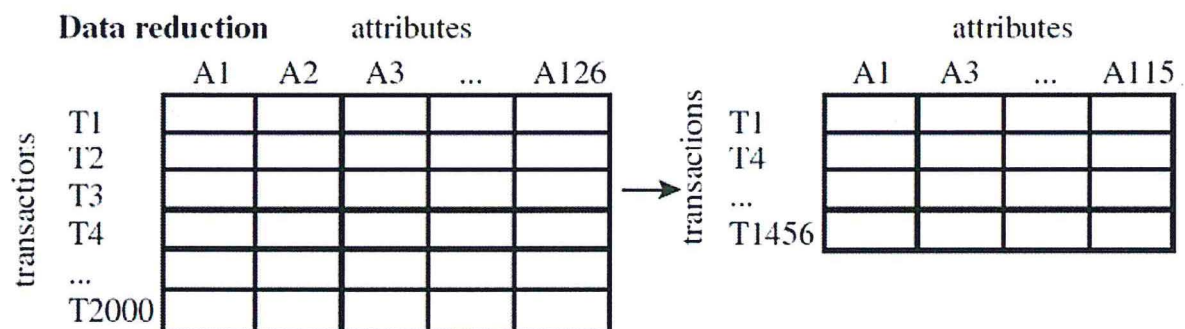
A large amount of additional meaningless data is called noise. By looking at the diagram below, can you please illustrate the concept of noise? Name any three. In addition, can you please provide a practical example demonstrating noise?

Question 16 [4 Marks]

You are preparing for a Data Mining exam at the University of Science and Technology. While you are preparing, you come across the below diagram. Please explain how you would answer what the diagram illustrates during the exam.

Question 17 [12 Marks]

During a press conference, a junior Data Miner asks you to explain to him the life cycle of the data mining project. You are employed as a senior Data Miner at the Ministry of Education. Can you please discuss the life cycle of data mining projects?

Question 18   [12 Marks]

You are employed at Tech Mining Consultants as a Mining Engineer. Part of your duties is to teach Data Mining Interns data mining concepts. During a meeting, one intern wants you to list the most important data mining techniques. What are the different techniques used for data mining? Please list and explain them.

<<<<<<End of Exam >>>>>