****

**NATIONAL OPEN UNIVERSITY OF NIGERIA**

**University Village, 91 Cadastral Zone, Nnamdi Azikwe Expressway, Jabi, Abuja**

**FACULTY OF SCIENCE**

**APRIL, 2019 EXAMINATIONS**

**COURSE CODE: DAM301**

**COURSE TITLE: Data Mining and Data Warehousing**

**CREDIT: 3 Units**

**TIME ALLOWED: 2½ Hours**

**INSTRUCTION: Answer Question 1 and any other FOUR (4) Questions**

**QUESTIONS**

1a) State three (3) main reasons for data pre-processing ***(3 marks)***

b) Outline the roles of data mining in Science and Engineering ***(2 marks)***

c) List and explain the common types of data preparation method ***(1½ marks)***

d) **Enumerate the stages of Data warehousing? *(2½ Marks)***

e) **In a tabular form, compare OLTP and OLAP? *(4 Marks)***

**f**) Mention any three (3) data mining techniques. ***(3 Marks)***

g) State four ‘(4) fields in which clustering techniques are used. ***(4 Marks)***

h) List any four requirements of cluster analysis? ***(2 Marks)***

2a) Write short notes on data mining ***(4 marks)***

b) Briefly explain the scope of data mining under the following headings and give at least two illustrative examples of each:

i) Automated prediction of trends and behaviours (***3 marks)***

 ii) Automated discovery of previously unknown patterns. (***2 marks)***

**c**). Within the context of data mining, state three ‘(3) reasons to justify the statement “Real world

 data are dirty”.  ***(3 marks)***

 3) Describe the following concepts of data mining functionalities:

1. Supervised classification ***(3 marks)***
2. Characterization ***(3 marks)***
3. Clustering ***(3 marks)***
4. Association analysis ***(3 marks)***

 4a) Write short notes on OLAP ***(5 marks)***

b) List four (4) statistical techniques ***(2 marks)***

c) State six (6) advantages and four (4) disadvantages of Data Warehousing ***(5 marks)***

5a) Briefly describe the following with respect to the Component of the Data Warehouse

 Architecture:

 i) Data Warehouse Database ***(4 marks)***

 ii) Data Marts ***(4 marks)***

b) Describe the challenges to data mining regarding data mining methodology and user interaction issues. ***(2 Marks)***

c) What is a Decision tree? ***(2 Marks)***

6a) Constraint-Based Data Mining is a method of implementing some forms of human involvement into data mining. List and briefly explain any four of the various types of constraints that are used. ***(6 marks)***

b)Summarised Data is classified into two. State and describe them. ***(6 marks)***