Printed Page:- 04 **Subject Code**:- **ACSDS0603**

Roll. No:

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA

(An Autonomous Institute Affiliated to AKTU, Lucknow)

B.Tech

SEM: VI - THEORY EXAMINATION (2023 - 2024)

Subject: Data Analytics

Time: 3 Hours **Max. Marks: 100**

General Instructions:

IMP: *Verify that you have received the question paper with the correct course, code, branch etc.*

1. *This Question paper comprises of* **three Sections -A, B, & C.** *It consists of Multiple Choice Questions (MCQ's) & Subjective type questions.*

2. *Maximum marks for each question are indicated on right -hand side of each question.*

3. *Illustrate your answers with neat sketches wherever necessary.*

4. *Assume suitable data if necessary.*

5. *Preferably, write the answers in sequential order.*

6. *No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked.*

SECTION A **20**

1. Attempt all parts:-

1-a. Which of the following is not a application for data science? (CO1) 1

(a) Recommendation Systems

(b) Image & Speech Recognition

(c) Online Price Comparison

(d) Privacy Checker

1-b. Identify the key data science skills among the following(CO1) 1

(a) Data Visualisation

(b) Machine learning

(c) Statistics

(d) All of the above

1-c. Which of the following statements is true about the collection of data?(CO2) 1

(a) The data that is collected from the place of origin is known as primary data.

(b) The data that is collected from the place of origin is known as secondary

data.

(c) The data that is collected from the place of origin is known as tertiary data

(d) None of the above

1-d. RDBMS follows:(CO2) 1

(a) Strucured Data

(b) Un structured Data

(c) Semi Structured Data

(d) None of the above

1-e. Examples of Nominal can be: (CO3) 1

(a) ID Numbers, eye color, zip codes

(b) Rankings, grades, height

(c) Calender dates, phone numbers

(d) length, time,counts

1-f. Why do we need feature transformation?(CO3) 1

(a) Converting non-numeric features into numeric

(b) Resizing inputs to a fixed size

(c) Both Aand B

(d) None

1-g. What is the role of exploratory graphs in data analysis? (CO4) 1

(a) They are made for formal presentations

(b) They are typically made very quickly

(c) Axes,legends, and other details are clean and exactly detailed

(d) They are used in place of formal modeling

1-h. Discriminative learning algorithm include.....(CO4) 1

(a) Continuous regression

(b) Logistic regression

(c) Linear regression

(d) None of the above

1-i. What are the file extensions in Tableau ? (CO5) 1

(a) Tableau Packaged Workbook (.twbx)

(b) Tableau Data Source(.tds)

(c) Tableau Workbook (.twb)

(d) All the above

1-j. The most popular data visualization library in python is _____(CO5) 1

(a) matinfolib

(b) matplotlib

(c) matpiplib

(d) pip

2. Attempt all parts:-

2.a. Explain the concept of Data.(CO1) 2

2.b. Describe secondary source of data.(CO2) 2

2.c. Explain the concept of data Clustering.(CO3) 2

2.d. Explain some data wrangling techniques (CO4) 2

2.e. How can we improve the performance of Tableau. justify your answer with suitable example (CO5) 2

SECTION B **30**

3. Answer any <u>five</u> of the following:-

3-a. Explain why is Big Data Important?(CO1) 6

3-b. Describe the libraries in Python used for Data Analysis and Scientific Computations.(CO1) 6

3-c. Differetiate between categorical data and numerical data. (CO2) 6

3-d. How to handle high dimensional data? (CO2) 6

3.e. Describe the various steps of data cleaning.(CO3) 6

3.f. Differentiate between feature selection and feature extraction.(CO4) 6

3.g. State some reasons for the low performance of Tableau? Explain in detail.(CO5) 6

SECTION C **50**

4. Answer any <u>one</u> of the following:-

4-a. Explain the steps use in Data Analysis. (CO1) 10

4-b. Explain how Amazon helps in transforming e-commerce with Data Science. (CO1) 10

5. Answer any <u>one</u> of the following:-

5-a. Describe dimensionality. Explain high dimensional data with examples.(CO2) 10

5-b. Why do we need data collection? Explain five common ways for the collecting data?(CO2) 10

6. Answer any <u>one</u> of the following:-

6-a.	Describe some common problems that occur during data processing? How can they be fixed? (CO3)	10
6-b.	Describe Qualitative and Quantitative attributes.(CO3)	10

7. Answer any <u>one</u> of the following:-

7-a.	Explain the Principle Component Analysis (PCA) technique. (CO4)	10
7-b.	Differentiate between mean, median and mode. How are these terms used to impute missing values in numeric variables? (CO4)	10

8. Answer any <u>one</u> of the following:-

8-a.	Why Stacked Bar chart are used in Tableau? Explain its advantages. (CO5)	10
8-b.	Explain the sorting and filtering steps in Tableau. (CO5)	10