.

Printed Page:-3 **Subject Code:- CCSE0204**

Roll. No:

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA
(An Autonomous Institute Affiliated to AKTU, Lucknow)
B.Tech
SEM: II - THEORY EXAMINATION (2025 - 2026)
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:-

(1.1) Recognize an example of data from the following. (CO1, K1) 1

(a) Final report
(b) Student marks
(c) Business strategy
(d) Decision making

(1.2) State when data becomes meaningful. (CO1, K1) 1

(a) Storage
(b) Collection
(c) Processing and organization
(d) Deletion

(1.3) Identify an important characteristic of CSV files. (CO2, K1) 1

(a) They are complex in structure
(b) They are binary files
(c) They are lightweight and simple
(d) They are encrypted

(1.4) Select the file extension used for Microsoft Excel files. (CO2, K1) 1

(a) .doc
(b) .xlsx
(c) .txt
(d) .json

(1.5) Recall the meaning of mode. (CO3, K1) 1

(a) Average
(b) Middle value
(c) Most frequent value
(d) Lowest value

(1.6) Recognize how data is arranged to calculate median. (CO3, K1) 1

(a) Random order
(b) Ascending or descending order
(c) Reverse order
(d) None

(1.7) Identify where normalization is mainly used. (CO4, K1) 1

(a) Data collection
(b) Data preprocessing
(c) Data storage
(d) Data deletion

(1.8) Select the technique that rescales data between 0 and 1. (CO4, K1) 1

(a) Standardization
(b) Min-Max normalization
(c) Log transformation
(d) PCA

(1.9) Recognize the purpose of PROSITE. (CO5, K1) 1

(a) Protein patterns
(b) DNA storage
(c) RNA storage
(d) Image data

(1.10) State the use of KEGG database. (CO5, K1) 1

(a) Text storage
(b) Images
(c) Pathway analysis
(d) Audio

2. Attempt all parts:-

(2.1) Describe data analytics in brief. (CO1, K2) 2

(2.2) Describe CSV format. (CO2, K2) 2

(2.3) Describe median. (CO3, K2) 2

(2.4) Describe standardization. (CO4, K2) 2

(2.5) Describe the type of data stored in biological databases. (CO5, K2) 2

SECTION-B 30

Answer any one of the following:-

(3.1) Discuss qualitative data with its types. (CO1, K2) 6

(3.2) Outline structured data and its advantages. (CO1, K2) 6

Answer any one of the following:-

(3.3)	Outline Excel data format. (CO2, K2)	6
(3.4)	Describe the data cleaning process. (CO2, K2)	6

Answer any one of the following:-

(3.5)	Outline median with steps. (CO3,K3)	6
(3.6)	Demonstrate mode with example. (CO3, K3)	6

Answer any one of the following:-

(3.7)	Differentiate normalization and standardization with example. (CO4, K4)	6
(3.8)	Outline Min-Max normalization technique. (CO4, K2)	6

Answer any one of the following:-

(3.9)	Illustrate primary databases with examples. (CO5, K3)	6
(3.10)	Outline secondary databases with examples. (CO5, K2)	6

SECTION-C 50

4. Answer any one of the following:-

(4.1)	Discuss different types of data with examples. (CO1, K2)	10
(4.2)	Discuss the application of data analytics in biotechnology. (CO1, K3)	10

5. Answer any one of the following:-

(5.1)	Discuss features and advantages of JSON format. (CO2, K2)	10
(5.2)	Classify different types of missing values with examples. (CO2, K3)	10

6. Answer any one of the following:-

(6.1)	Illustrate range, variance, and standard deviation with examples. (CO3, K3)	10
(6.2)	Describe correlation and its types. (CO3, K2)	10

7. Answer any one of the following:-

(7.1)	Apply Min-Max normalization to a dataset using formula and suitable example. (CO4,K3)	10
(7.2)	Outline Z-score standardization with steps and example. (CO4, K3)	10

8. Answer any one of the following:-

(8.1)	Classify bioinformatics tools and their types. (CO5, K3)	10
(8.2)	Describe sequence analysis tools with examples. (CO5, K2)	10