

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA
(An Autonomous Institute Affiliated to AKTU, Lucknow)

B.Tech

SEM: IV - 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) | Common programming tools used in Data Science include languages (CO1, K1) | 1 |
| | (a) Photoshop | |
| | (b) AutoCAD | |
| | (c) CorelDRAW | |
| | (d) Python | |
| (1.2) | Which of the following is effective way of checking validity of data analysis? (CO1, K2) | 1 |
| | (a) Re-run the analysis | |
| | (b) Review the code | |
| | (c) Check the sensitivity | |
| | (d) All of the mentioned | |
| (1.3) | The data that doesn't fit neatly into structured databases and tables but has some structure- (CO2, K1) | 1 |
| | (a) Structured data | |
| | (b) Semi-structured data | |
| | (c) Unstructured data | |
| | (d) Categorical data | |
| (1.4) | The type of data can be represented using charts, graphs, or plots- (CO2, K1) | 1 |
| | (a) Numeric data | |
| | (b) Categorical data | |
| | (c) Graphical data | |
| | (d) High Dimensional Data | |

- (1.5) Concept hierarchy generation is primarily concerned with: (CO3, K1) 1
- (a) Data reduction
 - (b) Data transformation
 - (c) Aggregation
 - (d) Organizing data into hierarchies
- (1.6) Adjusted R-Square takes into account the number of independent variables in a regression model to: (CO3, K2) 1
- (a) Increase its value
 - (b) Decrease its value
 - (c) Not affect its value
 - (d) Make it negative
- (1.7) In time series analysis, what is the purpose of trend and seasonality analysis? (CO4, K2) 1
- (a) Identifying spatial patterns
 - (b) Understanding patterns in data collected at irregular intervals
 - (c) Decomposing time series data into trend and seasonal components
 - (d) Analyzing cross-sectional data
- (1.8) How can APIs be used in data wrangling for data acquisition? (CO4, K2) 1
- (a) To improve data security
 - (b) To increase data dimensionality
 - (c) To make data more complex
 - (d) To interact
- (1.9) Sorting data is important to: (CO5, K2) 1
- (a) Randomize values
 - (b) Arrange logically
 - (c) Delete values
 - (d) Merge datasets
- (1.10) Which chart type is best for showing proportions? (CO5, K1) 1
- (a) Line
 - (b) Pie
 - (c) Bar
 - (d) Scatter

2. Attempt all parts:-

- (2.1) What is the need for Data Science? (CO1, K2) 2
- (2.2) Differentiate between numeric and categorical data. (CO2, K4) 2
- (2.3) How can a line chart be utilized to reveal trends or patterns in data? (CO3, K3) 2
- (2.4) What is the main focus of time series analysis? (CO4, K1) 2
- (2.5) Define the primary purpose of data visualization. (CO5, K1) 2

SECTION-B

30

Answer any one of the following:-

(3.1)	Compare and contrast Data Science and Business Analytics by identifying their key differences in purpose, methods, and applications.(CO1, K4)	6
(3.2)	Explain the importance of Data Science in solving real-world problems and decision-making processes.(CO1, K2)	6
Answer any one of the following:-		
(3.3)	Analyze the differences between structured, semi-structured, and unstructured data, and examine how each type is utilized in Data Science with relevant real-world examples.(CO2, K4)	6
(3.4)	Apply the concepts of mean, median, and mode to different datasets and explain when each measure of central tendency should be used with suitable examples.(CO2, K3)	6
Answer any one of the following:-		
(3.5)	Apply data discretization techniques in data preprocessing. (CO3, K3)	6
(3.6)	Explain the Knowledge Discovery in Databases (KDD) process and describe its stages. (CO3, K2)	6
Answer any one of the following:-		
(3.7)	Explain the concepts of data transformation and dimensionality reduction in the context of data analysis. (CO4, K2)	6
(3.8)	Analyze the concept of class separation in LDA and examine its significance in achieving effective classification performance.(CO4, K4)	6
Answer any one of the following:-		
(3.9)	Apply the concepts of filtering and pivoting in Tableau to prepare datasets for visualization, and demonstrate how these processes improve data analysis.(CO5, K3)	6
(3.10)	Analyze the role of data mapping in data visualization and examine why it is particularly useful in improving interpretation and decision-making with a relevant example.(CO5, K4)	6
<u>SECTION-C</u>		50
4. Answer any <u>one</u> of the following:-		
(4.1)	Analyze the different types of data analysis and examine how each type contributes to solving problems and supporting decision-making in real-world contexts.(CO1, K4)	10
(4.2)	Explain how different skills contribute to the role of a Data Scientist in real-world scenarios? (CO1, K2)	10
5. Answer any <u>one</u> of the following:-		
(5.1)	Apply the concepts of covariance and correlation to analyze the relationship between two variables, and demonstrate how they differ using suitable examples.(CO2, K3)	10
(5.2)	Examine the importance of data manipulation in data analysis and explain how it improves the quality of insights, using an example in Python or R.(CO2, K4)	10
6. Answer any <u>one</u> of the following:-		
(6.1)	Explain the key attributes of a histogram and describe its uses in data analysis. (CO3, K2)	10

(6.2) Apply appropriate data reduction techniques to different types of datasets and demonstrate when and why each technique should be used in data analysis.(CO3, K3) 10

7. Answer any one of the following:-

(7.1) Apply the concept of correlation analysis to a dataset and demonstrate how it can be used to identify the relationship between variables using an example.(CO4, K3) 10

(7.2) Discuss the methods for identifying outliers and the considerations when deciding whether to remove or transform them in a data analysis. (CO4, K2) 10

8. Answer any one of the following:-

(8.1) A company wants to create a dashboard to track key performance indicators (KPIs) such as website traffic, sales, and social media engagement. Apply appropriate data visualization techniques to design an effective dashboard in Tableau or any other visualization tool, and justify the choice of visual elements used for each KPI. (CO5, K3) 10

(8.2) Compare and contrast calculated fields and table calculations in Tableau. When would you use one over the other? (CO5, K4) 10