

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

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

B.Tech

SEM: VI - 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) | The process of evaluating data through analytical and statistical tools. (CO1 K1) | 1 |
| | (a) Data Mining | |
| | (b) Data Exploration | |
| | (c) Data Analysis | |
| | (d) Data Visualisation | |
| (1.2) | Data science is the process of diverse set of data through? (CO1 K1) | 1 |
| | (a) Organizing data | |
| | (b) Processing data | |
| | (c) Analysing data | |
| | (d) All of the above | |
| (1.3) | Define lifestyle observation technique for personality study. (CO2, K1) | 1 |
| | (a) Inventory | |
| | (b) Case History | |
| | (c) Rating scale | |
| | (d) Projective | |
| (1.4) | Pictorial representation of data using symbols is known as: (CO2, K1) | 1 |
| | (a) Bar Graph | |
| | (b) Pictograph | |
| | (c) Pie chart | |
| | (d) None of the above | |
| (1.5) | Examples of Ordinal can be: (CO3, K1) | 1 |

- (a) ID Numbers, eye color, zip codes
 (b) Rankings, grades, height
 (c) Calender dates, phone numbers
 (d) length, time, counts
- (1.6) Examples of Interval can be: (CO3, K1) 1
- (a) ID Numbers, eye color, zip codes
 (b) Rankings, grades, height
 (c) Calender dates, phone numbers
 (d) length, time, counts
- (1.7) is a dimensionality reduction technique which is commonly used for the supervised classification problems. (CO4, K1) 1
- (a) Value analysis
 (b) Function analysis
 (c) Pure analysis
 (d) None of the above
- (1.8) The primary use of data cleaning is: (CO4, K1) 1
- (a) Removing the noisy data
 (b) Correction of the data inconsistencies
 (c) Transformation for correcting the wrong data
 (d) All of the above
- (1.9) Mention dashboard components typically used in Tableau. (CO5 K2) 1
- (a) Vertical
 (b) Horizontal
 (c) Image Extract
 (d) All the above
- (1.10) The most popular data visualization library in python is ____ (CO5, K1) 1
- (a) matinfolib
 (b) matplotlib
 (c) matpiplib
 (d) pip

2. Attempt all parts:-

- (2.1) Define semi-structured data (CO1 K1) 2
- (2.2) Explain how to identify data sources. (CO2, K2) 2
- (2.3) Justify the term Regression (CO3, K3) 2
- (2.4) Identify dimensionality reduction technique used for supervised classification. (CO4 K2) 2
- (2.5) Define Tree Map? (CO5, K1) 2

SECTION-B

30

Answer any one of the following:-

- (3.1) Explain the security challenges of Data Science? (CO1, K2) 6

(3.2)	Differentiate between Data science and Big Data. (CO1 K2)	6
Answer any one of the following:-		
(3.3)	Explain the advantages of data manipulation. (CO2, K2)	6
(3.4)	How to handle high dimensional data? (CO2, K3)	6
Answer any one of the following:-		
(3.5)	Explain data cleaning and why is it important?(CO3,K2)	6
(3.6)	Explain data transformation with various examples. (CO3, K2)	6
Answer any one of the following:-		
(3.7)	Explain in detail how can we use exploratory data analysis in real life situations? (CO4, K2)	6
(3.8)	Differentiate between Linear Discriminant analysis and Principal Component analysis. (CO4, K4)	6
Answer any one of the following:-		
(3.9)	Explain different types of functions commonly used in Tableau (CO5, K2)	6
(3.10)	Design a view in a map such that if a user selects any state the cities under that state have to show profit and sales. (CO5, K3)	6
SECTION-C		50
4. Answer any <u>one</u> of the following:-		
(4.1)	Discuss the evolution of Data Science and explain the concept of Datafication with suitable examples. (CO1 K2)	10
(4.2)	If you are a data scientist, how will you collect the data? What will be your data acquisition and retention strategy? (CO1 K3)	10
5. Answer any <u>one</u> of the following:-		
(5.1)	Compare different data formats used in Data Science and discuss their advantages and limitations. (CO2, K4)	10
(5.2)	Discuss data manipulation techniques for different file formats such as CSV, PDF, XML, HTML, Text, JSON, and Image files. (CO2, K3)	10
6. Answer any <u>one</u> of the following:-		
(6.1)	When using Python to clean dataset, what are some of the common issues that arise and how do you deal with them? (CO3, K1)	10
(6.2)	Describe ways to address problems arising from data inflow from a variety of sources (CO3, K2).	10
7. Answer any <u>one</u> of the following:-		
(7.1)	Can PCA be used to reduce the dimensionality of a highly nonlinear dataset? Explain various limitations of PCA. (CO4, K2)	10
(7.2)	Explain the concepts of Data Munging and Data Wrangling in Data Science applications.. (CO4, K2)	10
8. Answer any <u>one</u> of the following:-		
(8.1)	Explain the uses of calculated fields? What is the main difference between calculated fields and table calculations in Tableau? (CO5 , K2)	10
(8.2)	Suppose I want to design a view, without using a line or bar chart, to show the	10

region-wise profit and sales. How should I go about doing it? Explain. (CO5, K2)

REG_JAN_JUNE_2026