

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

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

**B.Tech**

**SEM: V - 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-a. A collection of information about a related topic is referred to as \_\_\_\_\_(CO1,K1) 1
- (a) Visualisation
- (b) Analysis
- (c) Conclusion
- (d) Data
- 1-b. Activity that is not an application of Data Science.(CO1,K1) 1
- (a) Recommendation Systems
- (b) Image & Speech Recognition
- (c) Online Price Comparison
- (d) Privacy Checker
- 1-c. Structured data is usually stored in:(CO2,K1) 1
- (a) Tables
- (b) Images
- (c) Audio
- (d) Video
- 1-d. Categorical data contains:(CO2,K1) 1
- (a) Labels
- (b) Heights
- (c) Salaries
- (d) Temperatures

- 1-e. 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-f. Data Selection is.....(CO3,K1) 1
- (a) The actual discovery phase of a knowledge discovery process
  - (b) The stage of selecting the right data for a KDD process
  - (c) A subject -oriented integrated time-variant non-volatile collection of data in support of management
  - (d) Record oriented classes finding
- 1-g. Find the outlier s in the given dataset: 15, 11, 9, 15, 63 (CO4,K1) 1
- (a) 11
  - (b) 9
  - (c) 63
  - (d) 15
- 1-h. PCA reduces the dimension by finding a few.....(CO4,K1) 1
- (a) Hexagonal linear combination
  - (b) Orthogonal linear combination
  - (c) Octagonal linear combination
  - (d) Pentagonal linear combination
- 1-i. The "Show Me" panel in Tableau is used for:(CO5,K1) 1
- (a) Formatting text
  - (b) Selecting chart types
  - (c) Importing data
  - (d) Saving workbooks
- 1-j. A collection of multiple visualizations on a single screen in Tableau is called:(CO5,K1) 1
- (a) Worksheet
  - (b) Dashboard
  - (c) Story
  - (d) Filter
2. Attempt all parts:-
- 2.a. Differentiate between Structured and Unstructured Data.(CO1,K1) 2
- 2.b. Define numeric data with examples.(CO2,K1) 2
- 2.c. Highlight the importance of Data Cleaning as a preliminary step in data preprocessing.(CO3,K1) 2
- 2.d. State the key motivations for performing Exploratory Data Analysis before model development.(CO4,K1) 2
- 2.e. Mention one use of Tableau Data Interpreter.(CO5,K1) 2

**SECTION-B**

30

3. Attempt all parts:-	
3.a. Answer any <u>one</u> of the following:-	
3.a.(i) Describe the 5 V's in Big Data.(CO1,K2)	6
3.a.(ii) How to deploy a Big Data Model? Mention the key steps involved.(CO1,K2)	6
3.b. Answer any one of the following:-	
3.b.(i) Describe numeric and categorical data with examples.(CO2,K2)	6
3.b.(ii) Explain graphical data with examples.(CO2,K2)	6
3.c. Answer any one of the following:-	
3.c.(i) Enumerate various strategies commonly adopted for Data Discretization.(CO3,K2)	6
3.c.(ii) Demonstrate the working of Binning as a technique for reducing noise.(CO3,K3)	6
3.d. Answer any one of the following:-	
3.d.(i) Identify core elements or components that constitute Dimensionality Reduction.(CO4,K2)	6
3.d.(ii) Justify the role of Exploratory Data Analysis in supporting business decision-making.(CO4,K3)	6
3.e. Answer any one of the following:-	
3.e.(i) Compare line charts and bar charts in terms of usability.(CO5,K4)	6
3.e.(ii) Examine the effect of applying multiple filters in a dashboard.(CO5,K4)	6
<b><u>SECTION-C</u></b>	<b>50</b>
4. Answer any <u>one</u> of the following:-	
4-a. Tabulate the differences between analysis and reporting.(CO1,K3)	10
4-b. Explain the steps for a Data analytics project.(CO1,K2)	10
5. Answer any <u>one</u> of the following:-	
5-a. Write in detail about social network data.(CO2,K2)	10
5-b. Differentiate between primary data source and secondary data source with examples.(CO2,K3)	10
6. Answer any <u>one</u> of the following:-	
6-a. Differentiate between Qualitative and Quantitative Data Attributes based on characteristics.(CO3,K3)	10
6-b. Justify the role of Dimensionality Reduction in enhancing processing efficiency.(CO3,K3)	10
7. Answer any <u>one</u> of the following:-	
7-a. Examine PCA from the viewpoint of assumptions, mathematical structure, and model formulation.(CO4,K4)	10
7-b. Interpret the characteristics of Outliers and describe systematic approaches for their management.(CO4,K4)	10
8. Answer any <u>one</u> of the following:-	
8-a. Evaluate the strengths and weaknesses of Tableau when used for analyzing large business data. Write with real-life examples.(CO5,K5)	10
8-b. Write a plan for cleaning and structuring data before using it in Tableau. Explain how errors in raw data can affect the visualization.(CO5,K3)	10