.

Printed Page:-3 **Subject Code:- AMIBABA0712**

Roll. No:

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA
(An Autonomous Institute Affiliated to AKTU, Lucknow)
MBA (Integrated)
SEM: VII - THEORY EXAMINATION (2025 - 2026)
Subject: Introduction to Data Science

Time: 2.5 Hours **Max. Marks: 60**

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 15

1. Attempt all parts:-

1-a. Skill sets linked to Data Science usually revolve around this combination of abilities (CO1, K2) 1

(a) Painting & music
(b) Cooking & design
(c) Statistics, coding & domain knowledge
(d) Gymnastics

1-b. Meaningful insights appear once data groups are arranged based on shared characteristics. (CO2, K4) 1

(a) Characteristics are irrelevant
(b) Characteristics reduce accuracy
(c) Characteristics apply only in maths
(d) Characteristics drive classification

1-c. Purpose of preprocessing aligns with the goal of boosting data quality for analysis (CO3, K3) 1

(a) Adding noise
(b) Improving quality
(c) Reducing accuracy
(d) Hiding patterns

1-d. Techniques that uncover hidden structure in high-dimensional data come under PCA. (CO4, K5) 1

(a) PCA
(b) Histogram

(c) Pie chart
(d) Box plot

1-e. A chart suited for showing parts of a whole dataset in Data Visualization. (CO5, K3) 1

(a) Pie chart
(b) Line plot
(c) Heat map
(d) Box plot

2. Attempt all parts:-

2.a. Introduce the idea of data science and its role in modern decision-making. (CO1, K2) 2

2.b. Give an overview of key data categories found in industry datasets. (CO2, K4) 2

2.c. Indicate reasons analysts rely on preprocessing steps. (CO3, K3) 2

2.d. Provide a short evaluation of factor analysis. (CO4, K5) 2

2.e. Mention reasons visualisation helps in data understanding. (CO5, K3) 2

SECTION-B 15

3. Answer any three of the following:-

3-a. Comparison between analysis, analytics, and reporting based on purpose and output. (CO1, K2) 5

3-b. Examination of structured and unstructured data through relevant examples. (CO2, K4) 5

3.c. Illustration of typical preprocessing workflow in a machine learning project. (CO3, K3) 5

3.d. Critical examination of exploratory data analysis as a decision-support practice. (CO4, K5) 5

3.e. Description of pie chart and 3D pie chart usage. (CO5, K3) 5

SECTION-C 30

4. Answer any one of the following:-

4-a. A startup wants to build an automated recommendation engine but lacks clarity on required data skills. Outline the skill set pathway for their team (CO1, K2). 6

4-b. A company receives large social-media data but struggles to decide which analytics type to apply. Build a reasoning flow for selection (CO1, K2). 6

5. Answer any one of the following:-

5-a. A factory collects sensor readings and employee logs. (a) Categorise the two forms of data. (b) Propose a manipulation step needed before analysis. (CO2, K4) 6

5-b. A research lab maintains unstructured experiment notes. (a) Recommend a data classification method. (b) State one advantage of classification. (CO2, K4) 6

6. Answer any one of the following:-

6-a. A medical research lab collected noisy patient records. Prepare a preprocessing plan identifying needed steps (CO3, K3). 6

6-b. A marketing agency wants to extract valuable variables from survey data. Construct 6

.

a variable-identification workflow (CO3, K3).

7. Answer any one of the following:-

7-a.	A company wants to reduce dimensionality in a large dataset. Recommend a PCA-based reasoning sequence (CO4, K5).	6
7-b.	A psychologist analyses behavioural factors in students. Evaluate how Factor Analysis can support insights (CO4, K5).	6

8. Answer any one of the following:-

8-a.	A business wants to present regional sales differences on a single dashboard. Create a geospatial visualization plan (CO5, K3).	6
8-b.	A teacher wants to show distribution of exam scores to students. Propose an appropriate visualisation method (CO5, K3).	6