APril-May-2026

Dr. D. Y. Patil Institute of Management & Entrepreneur Development ★ LIBRARY

Total No. of Questions : 5]

SEAT No. :

PF12193

[Total No. of Pages : 3

[6651]-57

F.Y.M.C.A. (Management)

FDS512MJ : FUNDAMENTAL OF DATA SCIENCE

(2025 Pattern) (Semester - I)

Time : 2½ Hours] *[Max. Marks : 45*

Instructions to the candidates:

1) *All questions are compulsory.*
2) *Use of non-programmable calculator is allowed.*
3) *Figures to right indicate full marks.*

Q1) Solve the following (Any Three). **[9]**

a) Explain various Data Reduction Techniques with suitable examples.

b) Discuss the different types of data such as Structured, Semi-Structured, and Unstructured Data with suitable examples.

c) Describe the various stages involved in the Data Science Lifecycle.

d) Explain the roles and responsibilities of a Data Scientist and a Data Analyst in an organization.

Q2) Solve the following (Any One). **[9]**

a) Solve the following system of linear equations using the Gaussian Elimination Method:

$2x + y - z = 8$

$-3x - y + 2z = -11$

$-2x + y + 2z = -3$

Also verify the solution using Matrix Inversion Method.

b) Using the Bisection Method, find the root of the equation:

$x^3 - 4x - 9 = 0$ in the interval [2, 3] up to three iterations.

Q3) Solve the following (Any Three). **[9]**

a) Explain the stages involved in data processing methods.

b) Discuss the importance of Data Cleaning and Pre-processing in data analytics.

c) Explain the process of Handling Missing Data and Dealing with Outliers with suitable examples.

d) Write a short note on Normalization and Scaling in Data Wrangling Process.

P.T.O.

SPPU Question Papers.com

Q4) Solve the following (Any One). **[9]**

a) Consider the following NumPy arrays:

```
import numpy as np
A = np.array ([[2, 4, 6], [1, 3, 5]])
B = np.array ([[10, 20, 30], [40, 50, 60]])
```

Perform the following operations:

i) Describe the shape and datatype of arrays A and B.

ii) Perform element-wise addition and multiplication of arrays A and B.

iii) Apply array slicing to extract: First row of A and Last column of B.

iv) Split array B into two equal sub-arrays column-wise.

v) Generate a 2 × 3 random integer array using NumPy random module and calculate its maximum and minimum values.

b) A university maintains student records in the following dataset:

Roll No	Name	Department	Marks	City
101	Rahul	CS	78	Pune
102	Priya	IT	85	Mumbai
103	Amit	CS	69	Nashik
104	Sneha	IT	91	Pune
105	Karan	CS	74	Mumbai

Using Pandas and EDA techniques, perform the following tasks:

i) Create a Pandas DataFrame for the above dataset.

ii) Display only the Name and Marks columns using selection and slicing techniques.

iii) Sort the DataFrame based on Marks in ascending order.

iv) Apply GROUP BY operation on Department and calculate average marks department-wise.

v) Write the Pandas command to export the DataFrame into a CSV file.

Q5) Solve the following (Any One). **[9]**

a) A company recorded monthly sales data for the year 2025 as follows:

Month	Sales (₹ in Lakhs)
Jan	12
Feb	15
Mar	18
Apr	20
May	22
Jun	25

Using Matplotlib, write a Python program to perform the following tasks:

i) Create a Line Plot for the above sales data.

ii) Create a Bar Chart for the same dataset.

iii) Customize the plots by adding:

1) Title
2) X-axis and Y-axis labels
3) Different colors and line styles

b) Consider the following dataset showing marks obtained by students:

Student	Marks
A	65
B	72
C	88
D	91
E	76
F	84
G	69

Using Seaborn and Matplotlib libraries, perform the following tasks:

i) Write Python code to create a Histogram for the marks data.

ii) Create a Box Plot to identify the spread and outliers in the dataset.

iii) Customize the plots using titles, labels, colors, and styles.

✪ ✪ ✪

SPPUQuestionPapers.com
CMMP020351 103.68.11.122 01/06/2026 09:47:52 103