Nov-DEC-2025

Institute of Management & Entrepreneur Development LIBRARY

Total No. of Questions : 5] SEAT No. :

PE-4194 [Total No. of Pages : 3

[6554]-1007

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

EC-11-3-FDS512MJ : FUNDAMENTAL OF DATA SCIENCE

(2024 Pattern) (Semester - I)

Time : 2½ Hours] *[Max. Marks : 50*

Instructions to the candidates :

1) *All questions are compulsory.*
2) *Each question carry equal marks.*

Q1) **Solve any 2 (two) :** **[10]**

a) Explain the evolution of Data Science as a discipline.

b) Describe the various methods of Data Collection for a Data Science Project.

c) Discuss the social impact of Data Science in the field of Public Policy and Governance.

d) Explain the ethical considerations in Data Science.

Q2) **Solve any 1 (one) :** **[10]**

a) Perform LU Decomposition for the matrix $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$.

b) Compute the Eigenvalues of the Matrix $A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Are the Eigenvectors Linearly Independent?

Q3) **Solve any 2 (two) :** **[10]**

a) Describe the 'GROUP BY' function in Pandas with a suitable example.

b) Explain the process of Data Integration. What are the potential issues that can arise?

c) Discuss Exploratory Data Analysis (EDA)? Why is it a crucial step in a Data Science project?

d) Propose a data processing model for **Amazon Inventory Management**, detailing how data is collected, cleaned and processed.

P.T.O.

SPPU Question Papers.com

[10]

Q4) Solve any 1 (one) :

a) Consider the following Dataset and code and Answer the below Questions

Dataset : Monthly Temperature Data (°C)

```
temperature = np.array ([
  [12, 15, 18, 22, 25, 28, 30, 29, 26, 21, 16, 13], # City A
  [08, 10, 15, 20, 25, 30, 33, 32, 28, 22, 15, 10], # City B
  [15, 17, 19, 21, 23, 25, 26, 26, 24, 21, 18, 16] # City C
])
```

Questions:

i) Calculate the annual average temperature for each city

ii) Find the hottest month for each city

iii) Calculate the temperature range (max-min) for each city

iv) Identify cities with average temperature > 20°C

v) Find the month with highest average temperature across all cities

b) Consider the following code and Answer the below Questions

```
import pandas as pd
employee_data = {
'Emp_ID': ['E001', 'E002', 'E003', 'E004', 'E005'],
'Name': ['Alice', 'Bob', 'Charlie', 'Diana', 'Eve'],
'Department': ['IT', 'HR', 'IT', 'Finance', 'HR'],
'Salary': [75000, 65000, 80000, 90000, 60000],
'Experience': [5, 3, 7, 10, 2],
'Age': [30, 28, 35, 42, 25]
}
df = pd.DataFrame(employee_data)
```

Questions :

i) Find the average salary by department

ii) Filter employees with salary > 70000

iii) Add a new column. 'Bonus' as 10% of salary

iv) Find the employee with maximum experience

v) Group by department and calculate average age and salary

D.Y. Patil Institute of Management & Entrepreneurship Development LIBRARY

SPPUQuestionPapers.com CMMP020351 103.243.141.178 17/01/2026 13:35:12 178

Q5) Solve any 1 (one) : **[10]**

a) Consider the following code and Answer the below Question

E-commerce Customer Data

age_groups = ['18-25', '26-35', '36-45', '46-55', '56+']

purchase_frequency = [15, 25, 18, 12,8]

avg_spending = [1200, 2500, 3200, 2800, 2000]

customer_satisfaction = [3.8, 4.2, 4.5, 4.3, 4.0]

product_categories = ['Electronics', 'Clothing', 'Books', 'Home', 'Beauty']

category_sales = [35000, 28000, 15000, 22000, 18000]

Questions :

Write a python code to plot dual-**axis plot** showing purchase frequency and average spending by age group

b) Consider the following code and Answer the below Question **Company Department Data**

departments=['IT', 'HR', 'Finance,' 'Marketing,' 'Operations']

male_employees = [45, 15, 20, 25, 35]

female_employees = [30, 25, 18, 22, 28]

salary_ranges = ['20-30K', '30-40K', '40-50K', '50-60K', '60K+']

employee_count = [25, 40, 35, 20, 15]

Questions :

Write a python code to plot **stacked bar chart** showing gender distribution across departments



SPPUQuestionPapers.com
CMMP02035_1
103.243.114.178 17/01/2026 13:35:12