

Second Semester FYUGP Degree (Reg/Sup/Imp) Examination
April 2026

KU2MDCCAP104 - INTRODUCTION TO DATA SCIENCE
2024 Admission onwards

Time : 1.5 hours

Maximum Marks : 50

Section A

Answer any 6 questions. Each carry 2 marks.

1. What is Python, and why is it popular for Data Science?
2. List any four applications of Data Science.
3. List two real-world applications of Machine Learning.
4. What is underfitting in Machine Learning?
5. What is data normalization, and why is it used in preprocessing?
6. Define hypothesis testing in statistics.
7. List out 2 clustering algorithms.
8. Differentiate between machine learning and deep learning.

Section B

Answer any 4 questions. Each carry 6 marks.

9. Apply descriptive statistics to summarize a dataset containing student exam scores. Provide insights based on mean, median, standard deviation, and interquartile range.
10. Explain data preprocessing and the importance of data preprocessing in data science.
11. Evaluate the effectiveness of different visualizations (bar chart, line graph, pie chart) in representing categorical and numerical data.
12. Analyze how PCA works on a dataset.
13. Analyze the effectiveness of interactive visualizations over static charts.
14. Apply the concept of activation functions in a neural network.

Section C

Answer any 1 questions. Each carry 14 marks.

15. Design a program that demonstrates: a) The creation of different data types in Python (int, float, string, list, tuple, set, dictionary). (7 marks) b) The conversion of one data type into another with examples and built in functions also. (7 marks)
16. Compare and contrast logistic regression and decision tree algorithms

Don Bosco Arts and Science College
Library
Angadikadavu, Kannur
lib.donbosco.ac.in

