

Subject Code	Subject Name	Credits
26CS012	FOUNDATION OF DATA SCIENCE	4

Course Objectives:

1. To introduce core concepts and scope of data science.
2. To develop basic analytical and problem-solving skills.
3. To understand statistical and mathematical foundations.
4. To gain programming skills for data handling.
5. To learn data preprocessing, visualization, and interpretation techniques.

Learning Objectives:

On successful completion of the course, students will be able to:

1. Explain the data science lifecycle and applications.
2. Apply statistical methods to analyze datasets.
3. Write programs for data manipulation using Python/R.
4. Perform data cleaning and visualization effectively.
5. Understand basic machine learning concepts.

Unit 1 - Introduction to data science (12 hrs.):

Introduction to data science: Definition and importance of Data Science - Data Science lifecycle - Types of data: structured, unstructured, semi-structured- Sources of data - Applications of data science in business, healthcare, finance - Overview of tools: Python, R, Excel

Unit 2 – Mathematical & statistical foundations (12 hrs.):

Mathematical & statistical foundations: Basics of statistics: mean, median, mode, variance, standard deviation - Probability concepts and distributions - Correlation and regression - Introduction to linear algebra - Data sampling and hypothesis testing

Unit 3 - Programming for data science (12 hrs.):

Programming for data science: Introduction to Python/R programming - Data types, variables, operators - Control structures - Functions and modular programming - Working with arrays and lists - Libraries: NumPy, Pandas .

Unit 4- Data preprocessing and visualization (12 hrs.):

Data preprocessing and visualization: Data collection methods and sources - Data cleaning - Outlier detection and treatment - Data transformation and normalization - Exploratory Data Analysis - Data visualization tools: Matplotlib, Seaborn, charts and graphs

Unit 5 - Data analytics & machine learning Coding (12 hrs.):

Data analytics & machine learning: Basic concepts of data analytics - Types of analytics: descriptive, predictive, prescriptive - Introduction to machine learning - Supervised vs unsupervised learning - Simple algorithms: linear regression, classification techniques - Ethics in data science

Reference Books:

1. Joel Grus - Data Science from Scratch (2019)
2. Wes McKinney - Python for Data Analysis (2022)
3. Jiawei Han, Micheline Kamber – Data Mining: Concepts and Techniques (2011)
4. Cathy O'Neil, Rachel Schutt - Doing Data Science (2013)
5. James D. Miller - Statistics for Data Science (2017)