

Data Science Pathway

The guide to understanding data science pathway

✓ Step 1

Introduction to Data Science —

Objective : Understand the basics of data science and its applications.

Topics Covered :

- What is Data Science?
- Overview of the Data Science Process
- Introduction to Key Programming Languages (Python, R, Julia)
- Basic Statistics and Probability

✓ Step 2

Programming Foundations —

Objective : Gain proficiency in one or more programming languages essential for data science.

Topics Covered :

- Python Programming for Data Science
- R for Statistical Analysis
- Julia for High-Performance Scientific Computing
- Data Manipulation and Management Techniques across Languages
- Data Structures and Algorithms
- Working with Data Libraries (Pandas, NumPy)
- Introduction to SQL for Data Handling

✓ Step 3

Data Analysis and Visualization —

Objective : Learn to analyse and visualise data to extract meaningful insights using different tools.

Topics Covered :

- Exploratory Data Analysis (EDA) in Python, R, or Julia
- Statistical Analysis Techniques across Languages
- Data Visualization (Matplotlib and Seaborn for Python; ggplot2 for R; Plots.jl for Julia)
- Advanced Data Handling (SQL and NoSQL Databases)

✔ Step 4

Machine Learning Basics

Objective : Understand the fundamentals of machine learning and build simple models with different languages.

Topics Covered :

- Supervised vs Unsupervised Learning in Python, R, and Julia
- Regression and Classification Models across Languages
- Decision Trees and Random Forests
- Model Evaluation and Validation Techniques

✔ Step 5

Advanced Machine Learning

Objective : Deepen knowledge in advanced machine learning techniques and algorithms using multiple languages.

Topics Covered :

- Ensemble Methods and Support Vector Machines in Python and R
- Introduction to Neural Networks and Deep Learning in Python and Julia
- Natural Language Processing (NLP) Techniques
- Time Series Analysis in R

✔ Step 6

Big Data and AI Technologies

Objective : Explore technologies used for handling big data and building AI systems.

Topics Covered :

- Big Data Tools and Frameworks (Hadoop and Spark – APIs available in Python, Scala, Java)
- Cloud Computing for Data Science (platform-agnostic principles; AWS, Azure, Google Cloud)

- Advanced Machine Learning and AI Concepts (using TensorFlow or PyTorch in Python; MXNet in R)

✔ Step 7

Real-World Applications and Projects —

Objective : Apply knowledge to real-world data science problems and projects using various languages.

Topics Covered :

- Capstone Projects (choice of language)
- Industry-Specific Applications (healthcare analysis in R, financial modelling in Python, or engineering computations in Julia)
- Ethical Considerations in Data Science
- Data Security and Privacy Compliance

✔ Step 8

Career Preparation and Advancement —

Objective : Prepare for a career in data science with knowledge of multiple programming environments.

Topics Covered :

- Building a Data Science Portfolio (showcasing projects)
- Effective Resume and Cover Letter Preparation
- Interviewing Skills for Data Science (technical and problem-solving skills across languages)
- Networking and Professional Development