

DATA SCIENCE

ADVANCED

5 DAYS

Data Science with Python

Build a strong foundation in Python for data analysis and machine learning by learning core programming concepts and visualization techniques. Gain hands on experience with real world examples and popular Python frameworks to build and apply machine learning models effectively.

Why Data Science with Python?

Faster Insights

Quickly analyze large volumes of data, identify trends, and make data-driven decisions.

Cost-effective

Python, being an open-source programming language, provides a cost-effective solution for you to utilise data analytics without significant expenditures.

Scalable

Manage large datasets and adapt to the evolving requirements of a growing business, create tailored data science solutions that seamlessly integrate with other tools and systems.

What you'll be able to do

- ✓ Gain a solid understanding of the basics of Python programming
- ✓ Gain an understanding of data analysis and visualisation
- ✓ Get introduced to machine learning and discover popular Python frameworks for building machine learning models
- ✓ Acquire hands-on experience with real-world examples and projects

Who this is for

- Data Analysts
- Data Scientists
- Business analysts
- IT professionals

Curriculum

01 Python Foundations for Data Science

- Use Python syntax, types, functions, and control flow
- Organize scripts and notebooks for analysis work
- Manage environments and dependencies for reproducibility
- **Hands-on:** Build a small data processing script

02 NumPy for Numerical Computing

- Work with arrays and vectorized operations
- Apply numerical routines common in data science
- Understand broadcasting and performance basics
- Prepare numeric data for downstream libraries

03 Data Manipulation with Pandas

- Load, filter, join, and aggregate tabular data
- Reshape data for analysis and modeling
- Handle missing values and dtypes deliberately
- Build reusable pandas transformation steps

04 Visualization with Matplotlib

- Create core chart types for exploratory analysis
- Customize plots for clarity and stakeholder use
- Combine multiple views for comparison
- Export visuals for reports and notebooks

05 Data Cleaning and Preprocessing

- Apply cleaning patterns for ML-ready datasets
- Encode, scale, and split data correctly
- Detect outliers and data quality issues
- Document preprocessing assumptions

06 Machine Learning Fundamentals

- Frame supervised and unsupervised problem types
- Understand train/validation/test discipline
- Choose baseline models before complexity
- Connect ML outputs to business decisions

07 Models with Scikit-learn

- Build classic ML models with scikit-learn
- Compose preprocessing + model pipelines
- Compare algorithms on the same evaluation setup
- **Hands-on:** Train and score a baseline pipeline

08 Model Evaluation and Selection

- Select metrics aligned to the problem type
- Use cross-validation and model selection patterns
- Diagnose bias/variance issues
- Pick a model with a clear trade-off rationale

09 Deep Learning with TensorFlow

- Introduce neural network building blocks
- Train a simple TensorFlow model end to end
- Monitor training behavior and overfitting
- Know when deep learning is and is not warranted

10 Text Processing and Web Scraping

- Process text with NLTK-style workflows
- Scrape and parse web data responsibly
- Structure scraped data for analysis
- Respect rate limits, robots, and legal boundaries

11 Model Deployment and Project Customization

- Package models for scoring environments
- Choose simple deployment patterns for first releases
- Customize project structure for team workflows
- Define monitoring and retraining next steps