

Data Analysis and Governance with Python

An introductory analytics workflow for business professionals

FA-0724 · 3 days

Data & Analytics

Introduction

Explore how raw data becomes a small, interpretable analysis while keeping ownership, access and quality decisions visible. This introductory course for business professionals combines foundational Python, pandas, charts and a guided analytics workflow.

Three days provide a high-level practical foundation, not a complete governance-program implementation or mastery of all predictive methods. Selected models and optional recommendation examples are reviewed with appropriate evaluation and limitations.

Learning outcomes

- Write simple Python code and use selected libraries for an analysis task.
- Acquire and clean sample data with pandas and create selected Matplotlib or Seaborn charts.
- Describe the analytics lifecycle and distinguish exploratory findings from validated predictions.
- Recognise basic governance responsibilities for data purpose, ownership, access and quality.
- Explain training/evaluation separation and inspect selected clustering or classification examples.
- Document assumptions and identify further learning or governance decisions needed before operational use.

Prerequisites

- Basic computer, file-management and spreadsheet/data familiarity; no prior Python expertise is assumed.

- A supported Python/notebook environment with compatible packages and approved sample data.
- A Google account and Drive-sharing permissions are needed only if the chosen delivery uses those services; follow the approved sharing policy.

Training outline

1. Day 1 – Python and governed data inputs

- Syntax.
- Strings.
- Numbers.
- Lists.
- Dictionaries.
- Conditions.
- Loops.
- Functions.
- Lambda.
- Libraries.

Define the analytical question, permitted data use, data owner and access boundaries before collecting inputs.

Introduce pandas sources and connections using supplied files or approved endpoints. Scraping is a selected permitted-source example, not an assumption of access to every website.

2. Day 2 – Preparation and visual exploration

- Inspect types, missing values, duplicates and inconsistent entries; document cleaning choices.
- Use pandas to prepare a small dataset and verify the resulting rows and columns.
- Create selected Matplotlib and Seaborn views and check interpretation against the source data.
- Describe a lightweight governance process: ownership/stewardship, definitions, quality checks, access decisions and issue review.
- Maintain a record of inputs, transformations, assumptions and sharing/retention requirements appropriate to the organisation.

3. Day 3 – Model awareness and workflow review

- Separate exploratory analysis from prediction and choose an appropriate target or clustering question.
- Split evaluation data before fitting preprocessing; inspect a small pipeline and baseline.

- Introduce k-means, k-nearest neighbours, decision trees and dimensionality reduction through selected worked examples, not four promised deep implementation labs.
- Evaluate results and explain limitations, leakage risks and required human judgement.
- Review the complete workflow against its analytical question and governance responsibilities.
- If time permits, compare user-based and item-based collaborative recommendation concepts using a supplied example.