

Data Science for Business Analytics

Regression, visualisation and selected predictive case studies

FA-0725 · 3 days

Data & Analytics

Introduction

Follow a small business-data question through discovery, preparation, model planning, modelling and communication. Use selected Python and R examples to explore regression, classification and visualisation, then review a bounded case study.

Practise with selected examples and supplied templates, then assess results against the business question. Three days provide an introduction to the workflow, not complete mastery of two programming languages, production recommendation engines or every predictive method.

Learning outcomes

- Frame a business question, acquire sample data and document preparation choices.
- Use supplied Python or R examples to manipulate and visualise a small dataset.
- Distinguish linear-regression predictions from logistic-regression classification.
- Explore a churn or other supplied business case with trees and XGBoost.
- Evaluate a result against a suitable baseline and explain its limitations.
- Recognise basic recommender approaches and operationalisation considerations.

Prerequisites

- Comfort with tabular data and basic mathematical/statistical ideas.

- Some familiarity with programming is helpful; supplied templates support the short Python/R orientation, which is not two complete beginner language courses.
- Access to compatible Python, R and RStudio environments and the required packages, or an approved supplied lab.

Training outline

1. Day 1 — Data, tools and regression foundations

- Discover the business question, decision, sample data and success criteria.
- A focused Python refresher and R/RStudio orientation; RStudio is an IDE, while R packages perform the analysis and plotting.
- Acquire approved sample data and inspect structure, types and missingness.
- Introduce linear and logistic regression with suitable business examples and assumptions.

2. Day 2 — Preparation, visualisation and evaluation

- Clean and preprocess the sample data; fit transformations on training data only.
- Create selected Python and R visualisations and compare the observations with the question.
- Plan a small model, retain appropriate evaluation data and inspect basic metrics and a baseline.
- Review a churn-style classification example and communicate uncertainty rather than treat prediction as a guaranteed outcome.

3. Day 3 — Case study, trees and boosting

- Complete a bounded supplied business case with a selected decision-tree or XGBoost model.
- Explain gradient boosting and key settings at an introductory level; compare results using the same evaluation assumptions.
- Introduce recommendation-engine concepts through a supplied example or discussion, not a promised full production engine.
- Review operationalisation requirements, data limitations, human judgement and communication of findings.