

HR and People Analytics Foundations

Statistics, preparation and selected workforce-data examples in Python and R

FA-0731 · 3 days

Data & Analytics

Introduction

Explore how statistics and analytical workflows can inform a workforce question using approved sample data. Practise acquisition, cleaning, charts and reporting through selected Python and R examples, then review introductory modelling and text-analysis concepts.

No prior programming is required, but the short course uses supplied templates rather than promising independent expertise in two languages. Findings support human review and must be interpreted with data quality, privacy and potential bias in mind.

Learning outcomes

- Frame a workforce question and identify suitable, permitted data and measures.
- Use supplied Python and R examples for small data-preparation and charting tasks.
- Explain basic statistical summaries and model/evaluation concepts.
- Recognise regression, trees and recommendation approaches and their limits.
- Inspect a bounded text-analysis example and distinguish surface language patterns from validated meaning.
- Communicate findings, uncertainty and data-handling decisions in a short report.

Prerequisites

- Ability to use the internet.

- Experience with either Windows, MacOS or Ubuntu (either one).
- Any programming experience will help, but not required.
- Fluency in English.

A compatible supplied Python/R environment or approved local installation with required packages; sample or appropriately de-identified data is used rather than live personnel records.

Training outline

1. Day 1 – Questions, statistics and data preparation

- Define a workforce question and distinguish a descriptive measure from a causal or predictive conclusion.
- Review basic statistics and document data purpose, ownership and permitted access.
- Acquire a supplied sample and inspect missingness, types and quality; retain appropriate data-handling decisions.
- Use a focused Python orientation with NumPy, pandas and selected scikit-learn concepts; prepare a small table using a supplied template.

2. Day 2 – R, visualisation and model awareness

- Orient to R in RStudio and inspect a comparable prepared data/chart example; distinguish the IDE from R packages.
- Create selected views and consider whether aggregate results fairly represent the available data.
- Training/evaluation separation, linear or polynomial regression and a decision-tree demonstration.
- Introduce feature engineering, prediction and recommendation concepts; compare a baseline and review error or bias limitations.

3. Day 3 – Text, case study and reporting

- Inspect a small prepared text-analytics example and introduce semantic interpretation limits.
- Use a bounded workforce-data case to connect preparation, visualisation and selected analysis results.
- Document privacy, data limitations, potential bias and appropriate human interpretation; an exploratory model is not a validated basis for individual employment decisions.
- Prepare a concise report with evidence, uncertainty and next questions, and identify further language/statistics practice needed.