+33 (0)1 48 10 09 46
formation@ambient-it.net
140 Avenue Jean Lolive 93500 Pantin - France

Updated 03/21/2025

Sign up

Dataiku Advanced Training

2 days (14 hours)

Presentation

Master your entire data workflow with our advanced Dataiku training. Whether you're a Data Engineer, Data Analyst or Data Scientist, our advanced training program will enable you to design complex, scalable data pipelines.

We'll explore advanced data pipelines. You'll learn how to create, optimize and automate them, based on complex scenarios and precise rules.This first step is essential for anyone wishing to certify as an advanced designer on Dataiku.

You'll discover how to create, evaluate and refine models directly in the Dataiku environment. From applying your models to understanding them in depth using the tools of responsible AI.

Finally, whether you want to integrate your own code into your projects, manage that code efficiently, or even create plugins for your colleagues, this training you the keys to becoming a certified Dataiku developer.

For this course, we'll be using the latest version of Dataiku: DSS 12.

Objectives

- Understanding the benefits of Dataiku for data management and Machine Learning
- Know how to use the DSS interface to prepare and analyze data
- Build and automate data transformation and analysis workflows
- Developing and training Machine Learning models with DSS
- Automate and deploy templates and workflows in production

Target audience

- Companies migrating from Saas to dataikus
- Data Analysts
- Data Scientists
- Data Engineer

Prerequisites

- Basic data manipulation skills (SQL, Python, Excel or BI tools)
- Basic knowledge of statistics and Machine Learning desirable, but not mandatory
- Initial experience with a data analysis (SAS, R, Power BI, Tableau, etc.) is a plus.

Software requirements

Licensing Dataiku.

Our Dataiku Advanced training program

Day 1

Introduction

- Introducing Dataiku's advanced features
- Overview of the tool's capabilities

Dataiku's advanced visual tools

- Advanced visual recipes
- Using variables
- Data Pipelines
- Automation
- Collaboration
- Quiz
- Summary

Introduction Machine Learning

- Concepts
- Predictive modeling
- Prediction: Regression and classification
- Clustering
- Quiz
- Summary

Analyze your data with Interactive Stats

- Interface
- Univariate and bivariate analysis
- Curve and distribution fitting
- Correlation matrix
- Principal component analysis
- Statistical tests
- Quiz
- Summary

Machine Learning

- Setting up a model
- Model adjustment
- Understanding model predictions
- Quiz
- Summary

Scoring

- Deploying your model
- Scoring Data
- Evaluating your model
- Quiz
- Summary

Advanced Machine Learning

- Model development strategies
- Model diagnosis
- Actionability
- Quiz
- Summary

Stratified or partitioned models

- Setting up a stratified model
- Quiz
- Summary

TALN - Visually

- Introduction to TALN

- Text data preparation
- Text feature management for ML
- Quiz
- Summary

Time series analysis and forecasting

- Time series fundamentals
- Time series analysis
- Time series forecasting
- Quiz
- Summary

Temporal data preparation

- Resampling recipe
- Interval extraction recipe
- Windowing recipe
- Exterma extraction recipe
- Quiz
- Summary

Day 2

Code in Dataiku

- Code notebooks
- Code recipes
- Code environments
- External IDE integration
- Quiz
- Summary

Share your code

- Introduction
- Libraries
- Import from Git
- Code samples
- Practice
- Quiz
- Summary

Custom ML templates

- Custom pre-processing in visual ML
- Custom modeling in visual ML
- Quiz
- Summary

Variables for coders

- Variable definition
- Using variables in a code recipe
- Changing variable values
- Practice
- Quiz
- Summary

Visualization

- Webapps
- Static Insights
- Quiz
- Summary

Managed folders

- How to use Managed folders
- Practice
- Quiz
- Summary

Dataiku APIs

- Introduction
- The dataiku package and the public API
- Practice
- Quiz
- Summary

Deployment and monitoring

- Dataiku Govern
- Model follow-up
- Batch deployment
- Real-time API
- Quiz
- Summary

Advanced Designer Certificate exam (Optional)

ML Practitioner Certificate exam (Optional)

Developer Certificate exam (Optional)

Companies concerned

This course is aimed at both individuals and companies, large or small, wishing to train their teams in a new advanced computer technology, or to acquire specific business knowledge or modern methods.

Positioning on entry to training

Positioning at the start of training complies with Qualiopi quality criteria. As soon as registration is finalized, the learner receives a self-assessment questionnaire enabling us to assess his or her estimated level of proficiency in different types of technology, as well as his or her expectations and personal objectives with regard to the training to come, within the limits imposed by the selected format. This questionnaire also enables us to anticipate any connection or security difficulties within the company (intra-company or virtual classroom) which could be problematic for the follow-up and smooth running of the training session.

Teaching methods

Practical course: 60% Practical, 40% Theory. Training material distributed in digital format to all participants.

Organization

The course alternates theoretical input from the trainer, supported by examples, brainstorming sessions and group work.

Validation

At the end of the session, a multiple-choice questionnaire verifies the correct acquisition of skills.

Sanction

A certificate will be issued to each trainee who completes the course.