Updated 05/13/2025

Sign up

Dataiku training: The DSS platform
2 days (14 hours)

Presentation

Dataiku is a data science platform that enables the creation of models while industrializing the entire processing chain in the background: data collection, preparation...

Thanks to this Dataiku training course, your team will be able to learn more about Dataiku's functionalities. They'll be able interface with Cloud environments and improve your organization by putting data at the heart of your strategy.

This collaborative tool will provide your company with Machine Learning and Data Science functionalities.

But also convert data into predictions, use it to explore and develop your own data products. You can prepare your platform to automate workflow and deploy production.

For this training, we use the latest version: DSS 11.

Objectives

- Identify the fundamental concepts of Dataiku DSS and its interface (Flow, Flowzones, recipes, engines, projects)
- Import, prepare and enrich datasets using visual recipes
- Structuring a complete project with a clear Flow logic and export of results
- Create interactive visualizations in a dashboard
- Automate processing using scenarios
- Implement a complete end-to-end project in Dataiku

Target audience

- Data analysts
- Data scientists
- Data engineers
- SAS Analyst

Prerequisites

- Basic knowledge of SQL
- Basic knowledge of Data Science

Software requirements

- As a minimum, active access to the 14-day trial version of Dataiku
- A Dataiku DSS instance installed and accessible to participants
- Access to databases (SQL, NoSQL) or data files for exercises
- Connection to APIs and cloud services as required

Our Dataiku training program

Introduction to Dataiku DSS

- Presentation of the DSS platform: interface, project logic/Flow
- Understanding Flowzones to organize treatments
- Storage types and runtime engines
- File import (CSV, Excel)
- Case study: creating a DSS project and importing data for the first time

DATA PREPARATION AND CLEANSING

- Explore imported data: column types, formats, quality
- Data cleansing with the "Prepare" recipe:
  - Format standardization
  - Handling missing values
  - Renaming and filtering columns
- Case study: preparing a set of customer data

DATA ENRICHMENT AND TRANSFORMATION

- Enrichment by joining multiple datasets (Join recipe)
- Data grouping and aggregation (Group recipe)
- Structuring a processing flow with dedicated zones
- Case study: enriching a dataset with orders or product information

FLOW STRUCTURING AND EXPORT

- Flowzone creation and logical project organization
- Recipe documentation and dependency management
- Data export (CSV, Excel, database)
- Case study: finalizing a complete Flow, ready for export

DATA VISUALIZATION AND DASHBOARDS

- Best practices in data visualization
- Creating graphs from a dataset
- Layout and filtering in an interactive Dashboard
- Case study: building a sales or customer dashboard

CUSTOMIZED: RECIPES AND TOOLS TAILORED TO BUSINESS NEEDS

- Listening and identifying participants' specific use cases
- Presentation of recipes or plugins relevant to their business contexts
- Targeted demonstrations on appropriate data
- Support for free exploration or in-depth study on a case-by-case basis

AUTOMATED PROCESSING

- Introduction to scenarios in Dataiku
- Creating a simple scenario: cleaning, enriching, exporting
- Adding steps and triggers
- Case study: automating project execution

FINAL CASE STUDY: COMPLETE PROJECT

- Project management from A to Z:
  - Import, cleaning, enrichment
  - Flow structuring
  - Dashboard visualization
  - Scenario-based automation
- Objective: to demonstrate mastery of all the steps covered during training
- Final feedback and individual project feedback

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 which enables 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 for 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.