

Updated on 12/04/2024

Sign up

# Google Cloud Dataform training

2 days (14 hours)

## Presentation

Since its emergence, Dataform has established itself as a pillar in the field of cloud-based data management. Positioned as an innovative alternative, the tool enables analytics teams to create efficient data pipelines based on software engineering best practices.

In this Google Cloud Dataform training course, you'll learn how to set up a Dataform-specific development environment, enabling you to develop, test and deploy your first data models.

You'll also discover how [Dataform](#) integrates seamlessly into today's data architectures, and how it addresses the most pressing data management challenges facing businesses.

By the end of this training course, you'll have mastered the main Dataform functionalities. We'll also introduce you to the latest advances in Dataform with [Core v2.8](#).

## Objectives

- Understanding the importance of data management in business
- Set up a Dataform project and establish connections to various data sources
- Developing data models with SQL
- Automate data workflows
- Collaborate effectively on data projects

## Target audience

- Data Scientists
- Data Analysts
- Project managers

- Data engineers
- Business Analysts

## PREREQUISITES

- An active Google Cloud Platform (GCP) account
- Master the basic concepts of data management
- Have a GitHub account linked to your Gmail account

## OUR DATAFORM TRAINING PROGRAM

### INTRODUCTION TO DATAFORM AND DATA MANAGEMENT

- Introducing Dataform and its role in modern data management
- The importance of data management in companies
- Key concepts: data, databases, data warehouses
- Introduction to the Dataform user interface and its main components
- Account creation and initial configuration of work environment

### DATAFORM PROJECT CONFIGURATION

- Creating and configuring a dataform.json project file
- Establish connection to various data sources
- Defining dependencies between tables and modeling files
- Management of development, test and production environments
- Best practices for structuring data projects

### DATA MODEL DEVELOPMENT WITH SQL

- Introduction to relational data modeling: entities, relationships, attributes
- Creating tables, views and functions with Dataform and SQL
- Data transformation, filtering and aggregation techniques
- Practical exercises based on real-life use cases
- Tips for writing clean, maintainable SQL code

### AUTOMATION OF DATA WORKFLOWS

- Overview of Dataform tasks and their configuration
- Task planning and execution: triggers and schedulers
- Dependency management and workflow orchestration
- Real-time task execution monitoring and debugging
- Practical examples of data workflow automation

### COLLABORATION AND SHARING IN DATA PROJECTS

- Data team collaboration techniques
- Using Dataform's code review and sharing features
- Access and permissions management for secure projects
- Set up notifications and comments to improve communication
- Conflict resolution and code review strategies

## MONITORING, OPTIMIZATION AND TROUBLESHOOTING

- Introduction to the principles of data pipeline monitoring
- Techniques for optimizing SQL query performance
- Using indexes, partitions and aggregations in Dataform
- Strategies for detecting and resolving performance problems
- Case study: troubleshooting a data workflow with Dataform

## 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, with 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.

