

Orchestrate BigQuery Workloads with Dataform

Dataform is a service for data analysts to develop, test, version control, and schedule complex SQL workflows for data transformation in BigQuery. In this course you will explore the components of Dataform core, learn how to define tables and dependencies in SQLX, document BigQuery tables and views, understand BigQuery security settings and how to manage these with Dataform, write assertions, execute SQL workflows, and explore additional advanced use cases.

**DURATION**

1 day

**LEVEL**

200 - Intermediate

**FORMAT**

Instructor-led

What you'll learn

- Document BigQuery tables and views.
- Understand BigQuery security settings using Dataform.
- Use assertions to validate data in Dataform workflows.
- Execute Dataform SQL workflows in an automated fashion.

Overview	7 modules · 4 labs
Who this course is for	• Customers
Products	• Dataform • Lakehouse • Managed Service for Apache Airflow • BigQuery Studio • SQLX Workflows • Knowledge Catalog • Confidential Cloud Run • BigQuery • Cloud Run functions • Cloud Scheduler
Prerequisites	• Knowledge of SQL data analysis and BigQuery as discussed in BigQuery for Data Analysis.

Module 01

Dataform Core Components

Topics	• SQL workflow • Repositories and workspaces • Default files and folders • Compiled graphs
Objectives	• Understand the components of Dataform core.

Module 02

Table Definitions and Dependencies

Topics

- Declare a data source.
- Create a table.
- Create an incremental table.
- Set partitioning and clustering options.
- Create an empty table.
- Create an external BigLake table.
- Create views and materialized views.
- Define dependencies.

Objectives

- Create tables and views in BigQuery using Dataform.
-

Module 03

Document BigQuery Tables and Views

Topics

- Use column descriptions.
- Use globally defined JavaScript constants.
- Add labels.

Objectives

- Document BigQuery tables and views.
-

Module 04

BigQuery Security Settings

Topics

- IAM dataset and table/view access
- Column-level security
- Row-level security

Objectives

- Understand BigQuery security settings using Dataform.
-

Module 05

Assertions

Topics

- Use built-in assertions.
- Create manual assertions.

Objectives

- Use assertions to validate data in Dataform workflows.
-

Module 06

SQL Workflow Executions

Topics

- Dataform code lifecycle.
- What happens during compilation.
- Customize and schedule compilation results.
- Execute workflows (UI, Cloud Scheduler, Cloud Composer).
- Logging and monitoring.

Objectives

- Execute Dataform SQL workflows in an automated fashion.
-

Module 07

Advanced Use Cases

Topics

- Create a BigLake table after file upload using Cloud Run functions.
- Build a Machine Learning pipeline with BigQuery ML.
- Work with Slowly Changing Dimensions Type 2.

Objectives

- Explore additional use cases for Dataform.