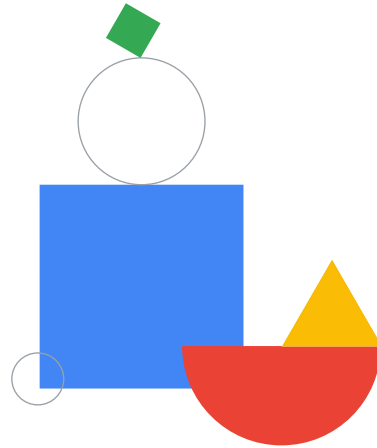


Building Batch Data Pipelines on Google Cloud



Data Engineering learning path

- 1 Modernizing Data Lakes and Data Warehouses with Google Cloud
- 2 [Building Batch Data Pipelines on Google Cloud](#)
- 3 Building Resilient Streaming Analytics Systems on Google Cloud
- 4 Smart Analytics, Machine Learning and AI on Google Cloud

Building Batch Data Pipelines on Google Cloud is the second course of the Data Engineering learning path.

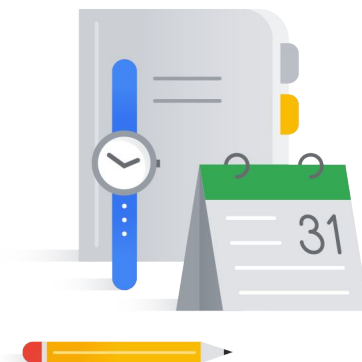
Course agenda

Introduction to Building Batch Data Pipelines

Executing Spark on Dataproc

Serverless Data Processing with Dataflow

Manage Data Pipelines with Cloud Data Fusion and Cloud Composer



In the previous course, we talked about data lakes and warehouses as storage options for your data.

In this course, we'll concentrate on how to build batch data pipelines to get data into storage using extract load, extract transform load, and extract load transform. We'll cover several technologies for data transformation on Google Cloud, including BigQuery, executing Spark on Dataproc, serverless data processing with Dataflow, and leveraging Google Cloud in pipelines with Cloud Data Fusion and Cloud Composer.