



Building Batch Data Pipelines on Google Cloud

Damon Runion
Technical Curriculum Developer

Welcome to Building Batch Data Pipelines on Google Cloud.
I'm Damon and I'm a Technical Curriculum Developer at Google.

Data Engineering on Google Cloud course series



Modernizing Data Lakes and Data
Warehouses with Google Cloud



**Building Batch Data Pipelines on
Google Cloud**



Building Resilient Streaming
Analytics Systems on Google Cloud



Smart Analytics, Machine Learning
and AI on Google Cloud

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

Course agenda

- 01 Introduction to Building Batch Data Pipelines
- 02 Executing Spark on Dataproc
- 03 Serverless Data Processing with Dataflow
- 04 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 routines.

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.