Data Engineering on Google Cloud Platform

Course Overview

A Data Engineer designs and builds systems that collect and transform the data used to inform business decisions. This learning path guides you through a curated collection of concepts and labs that provide you with real-world, hands-on experience using Google Cloud technologies to the Data Engineer role.

Duration: 08 days / 64 hours

Level: Professional

Prerequisites: There is no prerequisite for this learning path.

Course Outcome: Learner can take **Google Cloud Certified Professional Data Engineer** exam

Table of Content

Google Cloud Big Data and Machine Learning Fundamentals

- Big Data and Machine Learning on Google Cloud
- Data Engineering for Streaming Data
- Big Data with BigQuery
- Machine Learning Options on Google Cloud
- The Machine Learning Workflow with Vertex AI

Modernizing Data Lakes and Data Warehouses with Google Cloud

- Introduction to Data Engineering
- Building a Data Lake
- Building a Data Warehouse

Building Batch Data Pipelines on Google Cloud

- 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

Building Resilient Streaming Analytics Systems on Google Cloud

- Introduction to Processing Streaming Data
- Serverless Messaging with Pub/Sub
- Dataflow Streaming Features
- High-Throughput BigQuery and Bigtable Streaming Features
- Advanced BigQuery Functionality and Performance

Serverless Data Processing with Dataflow: Foundations

- Beam Portability
- Separating Compute and Storage with Dataflow
- IAM, Quotas, and Permissions
- Security

Serverless Data Processing with Dataflow: Operations

- Monitoring
- Logging and Error Reporting
- Troubleshooting and Debug
- Performance
- Testing and CI/CD
- Reliability
- Flex Templates

Serverless Data Processing with Dataflow: Develop Pipelines

- Beam Concepts Review
- Windows, Watermarks Triggers
- Sources & Sinks
- Schemas
- State and Timers
- Best Practices
- Dataflow SQL & DataFrames
- Beam Notebooks