COURSE OUTLINE

GC-DICDF - Data Integration with Cloud Data Fusion

DURATION	LEVEL	TECHNOLOGY	DELIVERY METHOD	TRAINING CREDITS
2 days	200 - Intermediate	Google Cloud	VILT / ILT	N/A

INTRODUCTION

This course introduces Google Cloud data integration using Cloud Data Fusion. It covers integration challenges, visual batch and streaming pipelines, metadata, lineage, transformations, connectors, execution environments, monitoring, and troubleshooting.

AUDIENCE PROFILE

Data engineers and data analysts.

PRODUCTS

- Hybrid Data Fusion
- Cloud Data Fusion

PREREQUISITES

Completed Introduction to Data Engineering.

COURSE OBJECTIVES

- Identify Cloud Data Fusion use cases and core components.
- Design and execute batch and real-time pipelines.
- Use Wrangler for transformations.
- Integrate varied data sources with connectors.
- Configure, monitor, and troubleshoot pipeline execution.
- Explain metadata and data lineage.

COURSE OUTLINE

COURSE CONTENT

Module 01 | Introduction

Topics
Course introduction

Objectives
Introduce course objectives.

Module 02 | Introduction to Data Integration and Cloud Data Fusion

Topics
- Data integration challenges
- Industry tools
- User personas
- Cloud Data Fusion
- Critical capabilities
- User interface

Objectives
- Explain data integration needs and challenges.
- Create a Cloud Data Fusion instance.
- Identify core components.

Activities
- Graded lab
- Quiz
- Discussion

Module 03 | Building Pipelines

Topics
- Architecture
- Core concepts
- Pipelines and DAGs
- Pipeline lifecycle
- Pipeline Studio

Objectives
Design, deploy, and execute a pipeline.

Activities
Graded lab & Quiz

COURSE OUTLINE

Module 04 | Designing Complex Pipelines

Topics
- Branching, merging, and joining
- Actions and notifications
- Error handling and macros
- Scheduling, import, and export

Objectives
Build complex pipelines and manage runtime execution.

Activities
- Graded labs
- Quiz

Module 05 | Pipeline Execution Environment

Topics
- Schedules and triggers
- Compute profiles and provisioners
- Monitoring

Objectives
Configure execution, logging, metrics, alerts, and monitoring.

Activities
Quiz

Module 06 | Transformations with Wrangler

Topics
- Wrangler
- Directives
- User-defined directives

Objectives
Transform data with the UI and directives.

Activities
- Graded lab
- Quiz

COURSE OUTLINE

Module 07 | Connectors and Streaming Pipelines

Topics
- Connectors
- DLP
- Streaming architecture
- Streaming pipelines

Objectives
Use connectors and DLP and build a streaming pipeline.

Activities
- Graded lab
- Quiz
- Discussion

Module 08 | Metadata and Data Lineage

Topics
- Metadata
- Data lineage

Objectives
Differentiate metadata types and explain lineage.

Activities
- Graded lab
- Quiz

Module 09 | Summary

Topics
Course summary

Objectives
Review course objectives and concepts.