



Managing a Data Mesh with Dataplex

Course description

In this comprehensive course, you will delve into the principles of Data Mesh architecture and learn how to effectively implement and manage it using Google Cloud's Dataplex. You will gain a deep understanding of how Dataplex serves as a foundational platform for building a self-service data ecosystem, enabling domain-oriented data product creation, federated computational governance, and enhanced data discoverability and quality. Through theoretical concepts and practical applications, you will acquire the skills to design, deploy, and operationalize a Data Mesh strategy within your organization, transforming raw data into valuable, governed, and easily accessible data products.

- Course level: Advanced
- Duration: 2 days

Activities

This course includes presentations, architectural discussions, hands-on labs, demonstrations, and group exercises focusing on real-world scenarios.

Course objectives

In this course, you will learn to:

- Explain the core principles and benefits of a Data Mesh architecture.

- Understand the role and capabilities of Google Cloud Dataplex in supporting Data Mesh.
- Design and structure data domains and data products using Dataplex Lakes, Zones, and Assets.
- Implement robust data discoverability and cataloging mechanisms for data products.
- Establish federated computational governance and data quality rules with Dataplex.
- Enable self-serve data infrastructure capabilities for data domain teams.
- Integrate Dataplex with other Google Cloud services to build comprehensive data pipelines and data products.
- Develop strategies for operationalizing and scaling a Data Mesh environment.

Intended audience

This course is intended for professionals involved in data architecture, engineering, and governance:

- Data Architects
- Data Engineers
- Data Product Owners
- Data Governance Leads
- Cloud Architects
- Technical Leads
- Data Platform Engineers
- Analytics Engineers

Prerequisites

We recommend that attendees of this course have:

- Familiarity with Google Cloud Platform (GCP) fundamentals (e.g., IAM, Cloud Storage, BigQuery).
- Basic understanding of data warehousing, data lakes, and data pipelines concepts.
- Conceptual knowledge of modern data architectures.

Course outline

Module 1: Introducing Data Mesh and Dataplex Fundamentals

- Understanding the challenges of traditional data architectures
- Core principles of Data Mesh (Domain Ownership, Data as a Product, Federated Computational Governance, Self-Serve Data Platform)
- Introduction to Google Cloud Dataplex: capabilities and value proposition
- Dataplex's role as an intelligent data fabric for Data Mesh
- Demo: Overview of the Dataplex UI and key features

Module 2: Designing Data Domains and Data Products

- Defining data domains and their boundaries within an organization
- Conceptualizing data products: input ports, output ports, API contracts
- Mapping Data Mesh constructs to Dataplex Lakes, Zones, and Assets
- Strategies for onboarding existing data sources into Dataplex
- Hands-on Lab: Structuring a Dataplex Lake with Raw, Curated, and Business Zones

Module 3: Enhancing Data Discoverability and Cataloging

- Automated metadata harvesting and ingestion in Dataplex
- Enriching the data catalog with business metadata, tags, and custom properties
- Leveraging Dataplex Data Catalog for data product search and discovery
- Data lineage and impact analysis within Dataplex
- Demo: Populating and searching the Dataplex Data Catalog for data products

Module 4: Implementing Data Quality and Observability

- Defining data quality dimensions and metrics for data products
- Configuring and executing data quality rules using Dataplex data quality tasks
- Monitoring data quality trends and alerts within Dataplex
- Integrating data quality checks into data product pipelines
- Hands-on Lab: Configuring and running Dataplex data quality tasks for a data product

Module 5: Establishing Federated Computational Governance

- Principles of federated data governance in a Data Mesh
- Implementing access control and security for data products with IAM
- Leveraging Dataplex Policy Tags for fine-grained access control and data classification

- Data masking and anonymization strategies for sensitive data
- Hands-on Lab: Applying IAM and Policy Tags for data product access control

Module 6: Building a Self-Serve Data Platform with Dataplex

- Dataplex as the control plane for data ingestion, processing, and consumption
- Integrating data processing services (BigQuery, Dataflow, Dataproc) with Dataplex
- Exposing data products via APIs and managed services
- Best practices for enabling self-service for data domain teams
- Demo: Creating a data product lifecycle from ingestion to consumption via Dataplex

Module 7: Operationalizing and Scaling Your Data Mesh

- Monitoring Dataplex resources and data product health
- Automation strategies for Dataplex setup and management (e.g., Infrastructure as Code)
- Handling schema evolution and data product versioning
- Best practices for scaling a Data Mesh across multiple domains and regions
- Hands-on Lab: Capstone - Designing and implementing a simple data product end-to-end using Dataplex

Module 8: Course Wrap-up

- Review of Data Mesh principles and Dataplex capabilities
- Future trends in data architecture and governance
- Next steps and additional resources for advanced topics
- Course summary and Q&A

Your course overview ends here.
But your cloud journey doesn't!

Unlock advanced training and certification pathways with **CloudThat**.



About CloudThat

Established in 2012, CloudThat is an award-winning company and the first in India to offer cloud training and consulting services for individuals and enterprises worldwide. Recently, it won Google Cloud's New Training Partner of the Year Award for 2025, becoming the first company in the world in 2025 to hold awards from all three major cloud giants: AWS, Microsoft, and Google. CloudThat notably won consecutive AWS Training Partner of the Year (APJ) awards in 2023 and 2024 and the Microsoft Training Services Partner of the Year Award in 2024, bringing its total award count to an impressive 12 awards in the last 8 years.

CloudThat specializes in Cloud Migration, Data Platforms, DevOps, IoT, and cutting-edge technologies like Gen AI & AI/ML. As a Microsoft Solutions Partner, AWS Premier Tier Training Partner, Google Cloud Platform Partner, and collaborator with leading organizations like HPE, Oracle, Nvidia and Databricks, CloudThat has trained over 850,000 professionals across 600+ cloud certifications, empowering students and professionals worldwide to advance their skills and careers.

To know more about our cloud certification training, email at sales@cloudthat.com or call us at [+91-8880002200](tel:+91-8880002200) or [+1-8555588830](tel:+1-8555588830) for those from outside India.