Szkolenie: Google Cloud

Managing a Data Mesh with Dataplex

Google Cloud

Cel szkolenia:

Dataplex is an intelligent data fabric that enables organizations to centrally discover, manage, monitor, and govern their data across data lakes, data warehouses, and data marts. You can use Dataplex to build a data mesh architecture to decentralize data ownership among domain data owners.

In this course, you will learn how to discover, manage, monitor, and govern your data across data lakes, data warehouses, and data marts through guided lectures and independent exercises using sample data.

What you'll learn

- Identify the importance of a modern data platform
- Configure and set up Dataplex
- Secure data lakes, zones, and assets
- Implement tagging for resources and use tags to search for assets
- Process data using Dataplex tasks
- Design, execute and report on data quality processes

Audience

This course is primarily intended for data engineers, architects, and analytics professionals who want to design, govern, and manage data mesh architectures using Google Cloud Dataplex.

Products

- Dataplex
- Cloud Storage
- BigQuery
- Dataflow

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70

BNP Paribas Bank Polska S.A.
Oddział w Warszawie
Nr konta: PL 20 1600 1462 1853 6451 3000 0001

Plan szkolenia:

- Introduction to Dataplex
  - Topics
    - Modern Data Platforms and Data-Oriented Design
    - Pillars of Data Governance
    - What is Dataplex?
    - Dataplex Capabilities
    - Dataplex compared with other products on Google Cloud
  - Objectives
    - Identify the importance of a modern data platform
    - Explain the role of Dataplex on Google Cloud
- Creating a Data Mesh on Dataplex
  - Topics
    - What is a data mesh?
    - Dataplex concepts
    - Creating data lakes and zones
    - Assets in Dataplex
  - Objectives
    - Define key Dataplex concepts
    - Configure and set up Dataplex
  - Activities
    - Lab: Provision a Data Mesh using Dataplex
- Processing Data on Dataplex
  - Topics
    - Processing data on Dataplex
    - Data preparation tasks
    - Ingestion jobs
    - Dataflow and Spark tasks
  - Objectives
    - Understand different data processing options in Dataplex
    - Configure and run data preparation tasks on Dataplex
  - Activities
    - Lab: Standardize Data using Dataplex Tasks
- Managing Data Security through Dataplex
  - Topics

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70
BNP Paribas Bank Polska S.A.
Oddział w Warszawie
Nr konta: PL 20 1600 1462 1853 6451 3000 0001

        - IAM permissions and roles
        - Securing your data lake
        - Policy management
        - Metadata security
    - Objectives
        - Secure data lakes, zones, and assets in Dataplex
    - Activities
        - Lab: Manage Data Security using Dataplex
- Data Tagging and Data Catalog
    - Topics
        - Introduction to Data Catalog
        - Technical metadata vs. business metadata
        - Tags and tag templates
        - Entries and entry groups
        - Data lineage
    - Objectives
        - Implement tagging for resources and use tags to search for assets
    - Activities
        - Lab: Data Catalog and Data Lineage
- Data Quality and Profiling
    - Topics
        - Data quality tasks and AutoDQ
        - Reporting on data quality
        - Data profiling
    - Objectives
        - Design, execute and report on data quality processes
    - Activities
        - Lab: Data Quality and Profiling your Data in BigQuery
- Dataplex Best Practices
    - Topics
        - Best practices
        - End-to-end demo
    - Objectives
        - Implement best practices for Dataplex
    - Activities
        - Challenge Lab: Managing a Data Mesh with Dataplex

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70
BNP Paribas Bank Polska S.A.
Oddział w Warszawie
Nr konta: PL 20 1600 1462 1853 6451 3000 0001

Wymagania:

To get the most out of this course, participants are encouraged to have completed the "Modernizing Data Lakes and Data Warehouses with Google Cloud" and "Building Batch Data Pipelines on Google Cloud" courses in the "Data Engineer" learning path or equivalent experience using Google Cloud.

Poziom trudności



Certyfikaty:

The participants will obtain certificates signed by Google Cloud (course completion).

Prowadzący:

Authorized Google Cloud Trainer

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70

BNP Paribas Bank Polska S.A.
Oddział w Warszawie
Nr konta: PL 20 1600 1462 1853 6451 3000 0001