Training: Google Cloud
Build Streaming Data Pipelines on Google Cloud

Google Cloud

TRAINING GOALS:

In this course you will get hands-on in order to work through real-world challenges faced when building streaming data pipelines. The primary focus is on managing continuous, unbounded data with Google Cloud products.

What you'll learn

- Ingest and manage streaming data using Pub/Sub and Managed Service for Apache Kafka
- Build and deploy streaming data pipelines with Dataflow
- Implement streaming data solutions for real-time analytics and application serving with BigQuery and Bigtable

Audience

This course is designed for Data Engineers, Data Analysts and Data Architects.

CONSPECT:

- Module 01
  - Topics
    - This module introduces the fundamentals of building streaming data pipelines on Google Cloud, providing a foundation for the entire course. It begins by outlining the course's overall learning objectives and introducing a practical, hands-on scenario that will be used throughout the content and labs to make the concepts tangible.
  - Objectives
    - Introduce the course learning objectives, and the scenario that will be used to bring hands on learning to building streaming data pipelines.
    - Describe the concept of streaming data pipelines, challenges associated with it, and the role of these pipelines within the data engineering process.
- Module 02

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.
EUR: PL 79 1600 1462 1853 6451 3000 0006
USD: PL 52 1600 1462 1853 6451 3000 0007

- Topics
  - This module provides an introduction to streaming data use cases and architectures. You will learn about the applications and common architectural patterns for real-time data processing across four key scenarios: Streaming ETL, Streaming AI/ML, Streaming Application, and Reverse ETL.
- Objectives
  - Learn about the various streaming use cases and their applications, including Streaming ETL, Streaming AI/ML, Streaming Application, and Reverse ETL.
  - Identify and describe common sampe architectures for streaming data including Streaming ETL, Streaming AI/ML, Streaming Application, and Reverse ETL.

- Module 03
  - Topics
    - This module provides a comprehensive overview of building streaming data pipelines on Google Cloud, covering the core services for messaging, processing, and analysis. It's designed to give you a hands-on understanding of how these components work together in a cohesive, real-time architecture.
  - Objectives
    - Define messaging concepts
    - Use the console to create various PS and Kafka elements
    - Know when to use Pub/Sub or Managed Service for Apache Kafka
    - Describe the DF service and challenges with streaming data
    - Build and deploy a streaming pipeline
    - Explore various data ingestion methods into BQ
    - Learn about BigQuery continuous queries and using BigQuery ETL and reverse ETL
    - Configure Pub/Sub to BigQuery streaming
    - Architecting BigQuery into your streaming pipelines
    - Describe the big picture of data movement and interaction
    - Establish a streaming pipeline from Dataflow to Bigtable
    - Analyze the BT continuous data stream for trends using BQ
    - Synchronize the trends analysis back into the user-facing application

- Module 04
  - Topics
    - This module provides a comprehensive wrap-up of the course, summarizing the key concepts you've learned for building resilient and robust streaming data pipelines on Google Cloud.
  - Objectives
    - Summarize the course and what you learned about the various Google products, what you achieved throughout the course, and what you're enabled to do next as a result of completing the course.

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.
EUR: PL 79 1600 1462 1853 6451 3000 0006
USD: PL 52 1600 1462 1853 6451 3000 0007

REQUIREMENTS:

- Proficiency in a common programming language like Python
- A strong understanding of SQL
- Data fundamentals like data modeling, formats, and ETL/ELT processes
- Familiarity with the Google Cloud Platform (GCP)

Difficulty level



CERTIFICATE:

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

TRAINER:

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.
EUR: PL 79 1600 1462 1853 6451 3000 0006
USD: PL 52 1600 1462 1853 6451 3000 0007