



Google Cloud Big Data and Machine Learning Fundamentals

This course introduces the Google Cloud big data and machine learning products and services that support the data-to-AI lifecycle. It explores the processes, challenges, and benefits of building a big data pipeline and machine learning models with Vertex AI on Google Cloud.

 **DURATION**
1 day

 **LEVEL**
Beginner

 **FORMAT**
Instructor led
On-demand

What you'll learn

- Recognize the data-to-AI lifecycle on Google Cloud and the major products of big data and machine learning.
- Design streaming pipelines with Dataflow and Pub/Sub.
- Analyze big data at scale with BigQuery.
- Identify different options to build machine learning solutions on Google Cloud.
- Describe a machine learning workflow and the key steps with Vertex AI.
- Build a machine learning pipeline using AutoML.



Overview	7 Modules · 4 Labs · 45 Videos
Who this course is for	<ul style="list-style-type: none">• Data analysts, data scientists, and business analysts who are getting started with Google Cloud• Individuals responsible for designing pipelines and architectures for data processing, creating and maintaining machine learning and statistical models, querying datasets, visualizing query results, and creating reports• Executives and IT decision makers evaluating Google Cloud for use by data scientists
Products	BigQuery, BigQuery ML, Dataflow, Pub/Sub, Apache Beam, Looker, Data Studio, Vertex AI, AutoML, Vertex Workbench, Document AI, Contact Center AI (CCAI), TPU (Tensor Processing Unit), Kubernetes Engine, Compute Engine
Prerequisite	Basic understanding of one or more of the following: <ul style="list-style-type: none">• Database query language such as SQL• Data engineering workflow from extract, transform, load, to analysis, modeling, and deployment• Machine learning models such as supervised versus unsupervised models

Module 0 Course Introduction

Topics	This section welcomes learners to the Big Data and Machine Learning Fundamentals course and provides an overview of the course structure and goals.
Objectives	<ul style="list-style-type: none">• Recognize the data-to-AI lifecycle on Google Cloud• Identify the connection between data engineering and machine learning

Module 1 Big Data and Machine Learning on Google Cloud

Topics	This section explores the key components of Google Cloud's infrastructure. We introduce many of the big data and machine learning products and services that support the data-to AI lifecycle on Google Cloud.
Objectives	<ul style="list-style-type: none">• Identify the different aspects of Google Cloud's infrastructure.• Identify the big data and machine learning products on Google Cloud.
Activities	<ul style="list-style-type: none">• Lab: Exploring a BigQuery Public Dataset• Quiz



Module 2 Data Engineering for Streaming Data

Topics	This section introduces Google Cloud's solution to managing streaming data. It examines an end-to-end pipeline, including data ingestion with Pub/Sub, data processing with Dataflow, and data visualization with Looker and Data Studio.
Objectives	<ul style="list-style-type: none">• Describe an end-to-end streaming data workflow from ingestion to data visualization.• Identify modern data pipeline challenges and how to solve them at scale with Dataflow.• Build collaborative real-time dashboards with data visualization tools.
Activities	<ul style="list-style-type: none">• Lab: Creating a Streaming Data Pipeline for a Real-Time Dashboard with Dataflow• Quiz

Module 3 Big Data with BigQuery

Topics	This section introduces learners to BigQuery, Google's fully managed, serverless data warehouse. It also explores BigQuery ML and the processes and key commands that are used to build custom machine learning models.
Objectives	<ul style="list-style-type: none">• Describe the essentials of BigQuery as a data warehouse.• Explain how BigQuery processes queries and stores data.• Define BigQuery ML project phases.• Build a custom machine learning model with BigQuery ML.
Activities	<ul style="list-style-type: none">• Lab: Predicting Visitor Purchases Using BigQuery ML• Quiz

Module 4 Machine Learning Options on Google Cloud

Topics	This section explores four different options to build machine learning models on Google Cloud. It also introduces Vertex AI, Google's unified platform for building and managing the lifecycle of ML projects.
Objectives	<ul style="list-style-type: none">• Identify different options to build ML models on Google Cloud.• Define Vertex AI and its major features and benefits.• Describe AI solutions in both horizontal and vertical markets.
Activities	Quiz



Module 5 The Machine Learning Workflow with Vertex AI

Topics This section focuses on the three key phases—data preparation, model training, and model preparation—of the machine learning workflow in Vertex AI. Learners can practice building a machine learning model with AutoML.

- Objectives**
- Describe a ML workflow and the key steps.
 - Identify the tools and products to support each stage.
 - Build an end-to-end ML workflow using AutoML.
 - Lab: Vertex AI: Predicting Loan Risk with AutoML
 - Quiz
-

Module 6 Course Summary

Topics This section reviews the topics covered in the course and provides additional resources for further learning.

Objectives Describe the data-to-AI lifecycle on Google Cloud and identify the major products of big data and machine learning.

