

Google Cloud Big Data and Machine Learning Fundamentals

Code:	GCP-BD-ML
Length:	1 days
URL:	View Online

This one-day instructor-led course introduces participants to the big data capabilities of Google Cloud. Through a combination of presentations, demos, and hands-on labs, participants get an overview of the Google Cloud and a detailed view of the data processing and machine learning capabilities. This course showcases the ease, flexibility, and power of big data solutions on Google Cloud.

Skills Gained

This course teaches participants the following skills:

- Identify the purpose and value of the key Big Data and Machine Learning products on Google Cloud.
- Use Cloud SQL and Cloud Dataproc to migrate existing MySQL and Hadoop/Pig/Spark/Hive workloads to Google Cloud.
- Employ BigQuery and Cloud Datalab to carry out interactive data analysis.
- Train and use a neural network using TensorFlow.
- Employ ML APIs.
- Choose between different data processing products on Google Cloud.

Who Can Benefit

This class is intended for the following:

- Data analysts, Data scientists, Business analysts 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.

Prerequisites

To get the most of out of this course, participants should have:

- Basic proficiency with common query language such as SQL.
- Experience with data modeling, extract, transform, load activities.
- Developing applications using a common programming language such Python.
- Familiarity with machine learning and/or statistics.

Course Details

The course includes presentations, demonstrations, and hands-on labs.

Course Outline

Module 1: Introducing Google Cloud Platform

- Google Platform Fundamentals Overview.
- Google Cloud Platform Big Data Products.

Module 2: Compute and Storage Fundamentals

- CPUs on demand (Compute Engine).
- A global filesystem (Cloud Storage).
- CloudShell.
- Lab: Set up a Ingest-Transform-Publish data processing pipeline.

Module 3: Data Analytics on the Cloud

- Stepping-stones to the cloud.
- Cloud SQL: your SQL database on the cloud.
- Lab: Importing data into CloudSQL and running queries.
- Spark on Dataproc.
- Lab: Machine Learning Recommendations with Spark on Dataproc.

Module 4: Scaling Data Analysis

- Fast random access.
- Datalab.
- BigQuery.
- Lab: Build machine learning dataset.

Module 5: Machine Learning

- Machine Learning with TensorFlow.
- Lab: Carry out ML with TensorFlow
- Pre-built models for common needs.
- Lab: Employ ML APIs.

Module 6: Data Processing Architectures

- Message-oriented architectures with Pub/Sub.
- Creating pipelines with Dataflow.

- Reference architecture for real-time and batch data processing.

Module 7: Summary

- Why GCP?
- Where to go from here
- Additional Resources

Refer a friend or colleague and get up to \$100 Amazon gift card* — when they book training!

[Learn More](#)

ExitCertified® Corporation and iMVP® are registered trademarks of ExitCertified ULC and ExitCertified Corporation and Tech Data Corporation, respectively
Copyright ©2021 Tech Data Corporation and ExitCertified ULC & ExitCertified Corporation.
All Rights Reserved.

Generated 12