

Summary of GCP DCA Module 4 Video 3

Emergence of Cloud Data Tools

Cloud data tools are evolving rapidly and are tailored to specific stages of the data lifecycle.

Tool Selection

Choosing the right tool depends on the task and stage of the data lifecycle.

BigQuery

A serverless data warehouse that supports SQL queries, machine learning, and business intelligence.

Looker

A data visualization and reporting tool that integrates various file types and publishes dashboards.

Dataproc

A managed service for running open-source tools like Hadoop and Spark, useful for ETL and data lake modernization.

Dataflow

Enables streaming and batch data processing in a serverless environment, ideal for building data pipelines.

Cloud Data Fusion

A GUI-based tool for integrating datasets and managing data pipelines without coding.

Dataplex

Central hub for managing and monitoring data across lakes, warehouses, and marts, accessible via BigQuery.

BigLake

A storage engine that unifies data warehouses and lakes, offering access control and multi-cloud storage.

Future of Cloud Tools

These tools represent standard options for data analysts, with more innovations expected.