

Data, ML, and AI in **Google Cloud**

Ground Rules

Observe the following rules to ensure a supportive, inclusive, and engaging classes



Give full attention
in class



Mute your microphone
when you're not talking



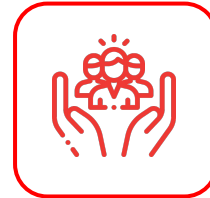
Keep your
camera on



Turn on the CC Feature
on Meet



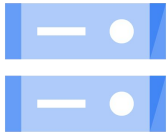
Use raise hand or chat
to ask questions



Make this room a safe place
to learn and share

Quiz

GCP Data Storage Options



Cloud
Storage



Cloud SQL



Cloud
Spanner



Cloud
Firestore



Cloud
Bigtable



BigQuery

















Cloud
Memorystore



Cloud Storage

Object storage for companies of all sizes. Store any amount of data.
Retrieve it as often as you'd like.

Cloud Storage - Deployment Use Case

	Standard (Multi-regional)	Standard (Regional)	Nearline	Coldline	Archive
Common usage	For highest availability of frequently accessed data	For data accessed frequently within a region	For data access less than once a month	For data accessed roughly less than once a quarter	For long term retention, data accessed less than once a year
Customer use case examples	<ul style="list-style-type: none"> Streaming videos Images Websites Documents	<ul style="list-style-type: none"> Video transcoding Genomics General data analytics & compute	<ul style="list-style-type: none"> Serving rarely accessed docs Backup	<ul style="list-style-type: none"> Serve rarely used data Movie archive Disaster recovery	<ul style="list-style-type: none"> Regulatory archives Tape replacement



Cloud SQL

Fully managed relational database service for MySQL,
PostgreSQL, and SQL Server.

Cloud SQL Features

- Fully managed
 - No software installation.
 - Automated backups, replication, patches, and updates.
- Performance and scalability
 - Scales to 96 cores and 624 GB of RAM.
 - Up to 64 TB of storage.
 - Automatic Storage Increase feature.
- Reliability and security
 - High availability with automatic failovers.
 - Easily configure replication and backups.
 - Data is encrypted.
- Compatibility
 - Accessible from almost any app.
 - Easy to move and migrate data.



Cloud Spanner

The one and only enterprise-grade, globally-distributed, and strongly-consistent managed database service built for the cloud specifically to combine the benefits of relational database structure with non-relational horizontal scale.

Cloud Spanner **Features**

- Scale + SQL
 - Scales horizontally.
 - Low latency, transactional consistency, and high availability.
 - Future-proofs database backends.
- Fully managed
 - Create or scale a globally replicated database in a few clicks.
 - Synchronous replication and maintenance built in.
- Launch faster
 - Relational semantics.
 - ACID transactions.
 - Schemas.
- Enterprise grade security
 - Data-layer encryption.
 - IAM integration.
 - Audit logging.



Cloud Firestore

A fast, fully managed serverless NoSQL document database that built for automatic scaling, high performance, and ease of application development - in real time at global scale.

Available in 2 modes



Datastore

*“Scales to **over tens of millions writes per second**, for server or container workloads.”*

- Compatible Cloud Datastore API
- No DB limit on write requests per second
- Server-side SDKs only



Native

*“Scales to **over a million concurrently connected** mobile, web and IoT devices.”*

- Document Model and API
- DB limit of 10k write requests per second
- Mobile & Server-side SDKs
- Security Rules using Firebase Auth or Cloud Identity Platform
- Live Data Synchronization & Offline Mode



Cloud Bigtable

A fully managed, scalable NoSQL database service for large analytical and operational workloads and use cases where low latency random data access, scalability and reliability are critical.

Cloud Bigtable **Features**

- Fast and performant
 - High performance under high loads.
 - Faster, more reliable, and more efficient.
 - Low latency.
- Seamless scaling and replication
 - Billions of rows and thousands of columns.
 - No downtime during reconfiguration.
 - Replication adds high availability.
- Fully managed
 - Database configuration and tuning handled by Google.
 - Data backups created for disaster recovery.
- Integrated and secure
 - Integrated with open-source big data tools for powerful data analysis, i.e. Hadoop and Apache HBase



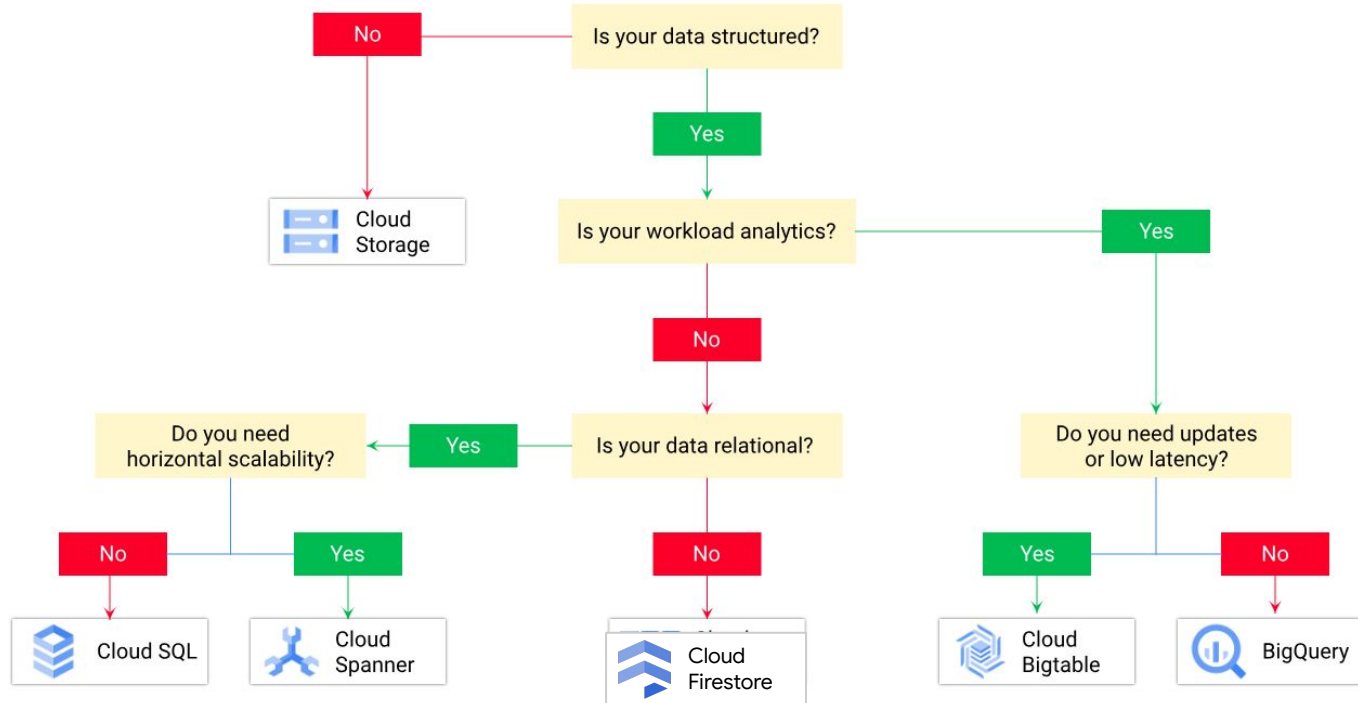
BigQuery

Serverless, highly scalable, and cost-effective multi-cloud data warehouse designed for business agility.

BigQuery **Features**

- Enterprise Data Warehouse
 - For analytical workload.
 - Queries using standard SQL syntax.
- Serverless
 - Fully managed.
 - No need to deploy any resources, such as disks and virtual machines.
 - No cluster maintenance is required.
- Runs on Google's high-performance infrastructure
 - Compute and storage are separated with a terabit network in between.
 - You only pay for storage and processing used.
 - Automatic discount for long-term data storage.

What storage type is **best for me?**



But, what if I need **in-memory data storage**?



Cloud Memorystore

A fully managed in-memory data store service for Redis and Memcached.
Build application caches that provide sub-millisecond data access.

<https://cloud.google.com/memorystore/>

Cloud Memorystore **Features**

- Choice of engines
 - Support Redis and Memcached.
 - Choose the right engine that fits your cost and availability requirements.
- Fully managed
 - Provisioning, replication, failover, and patching are all automated.
- Security
 - Protected from the internet using VPC networks and private IP.
 - IAM integration.
 - Systems are monitored 24/7/365.
 - Memorystore for Redis provides in-transit encryption and Redis Auth to secure your sensitive data.

How about **Big Data in Google Cloud?**

Traditional Big Data Platform

- Storage
- Server
- Network
- Assembly required
- Maintenance
- Technicians



“Don’t manage infrastructure, focus on **finding insight!**”

Scalable Big Data Tools to Overcome Data Challenges

Ingest



Bigquery
Storage
(import)

Transform



Bigquery
Analysis
(SQL)



Cloud
Dataprep
(preparation)

Store



Cloud
Storage
(buckets)



Bigquery
Storage
(tables)

Analyze



Bigquery
Analysis
(SQL)

Visualize



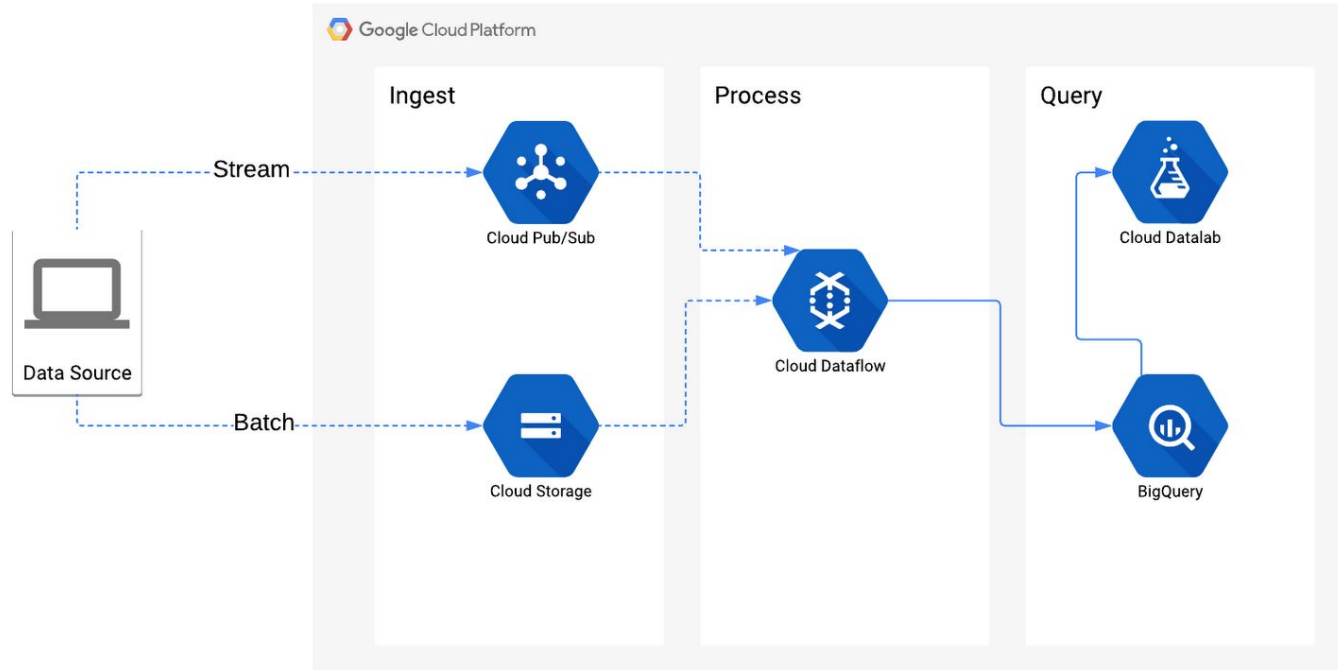
Google
Data Studio

On-Demand **Cloud**

- Ad Hoc Querying
- Scalable Storage
- Scalability on query processing time
- Separation of storage and computing power enables efficient resource allocation
- Pay for only the resource you are using.



Example Architecture



Introduction to **Machine Learning**

The GCP Machine Learning **Tool Spectrum**

Advanced Models	Modeling for Analytics	Pretrained Models	Minimal Efforts
TensorFlow <ul style="list-style-type: none">• Data Scientist• Data Engineers	ML on BigQuery <ul style="list-style-type: none">• Data Analysts	Pretrained ML APIs <ul style="list-style-type: none">• Data Analysts• Data Scientists• Data Engineers	AutoML <ul style="list-style-type: none">• Everyone

Google Cloud **Pre-Trained ML Models**

Pre-Trained ML Models

Sight



Vision



Video
Intelligence

Language



Translation



Natural Language

Conversation



Dialogflow



Text-to-Speech



Speech-to-Text

Structured Data

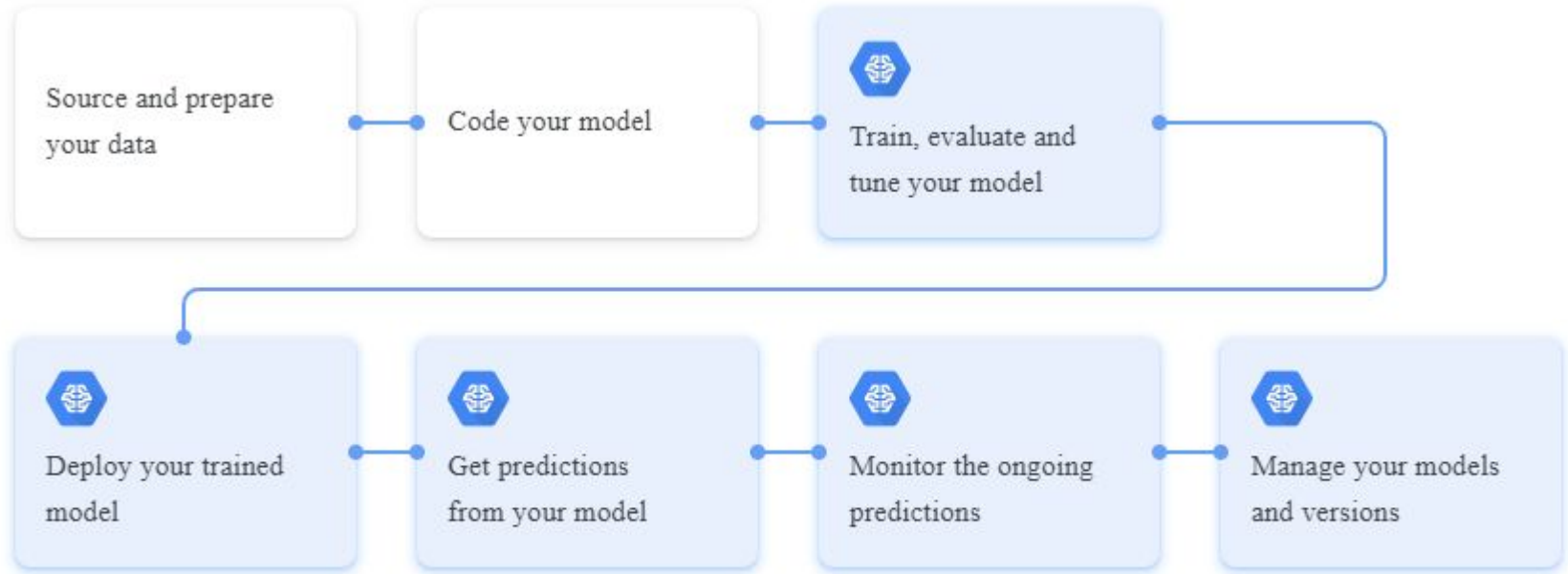


Recommendation AI

AI Platform

Use AI Platform to train your machine learning models at scale, to host your trained model in the cloud, and to use your model to make predictions about new data.

Where AI Platform fits in the ML workflow?



ML workflow

Sharing Session

Demo Link

<https://www.qwiklabs.com/focuses/1760?parent=catalog>

<https://www.qwiklabs.com/video/36225>

Discussion

Thank You