BigQuery overview

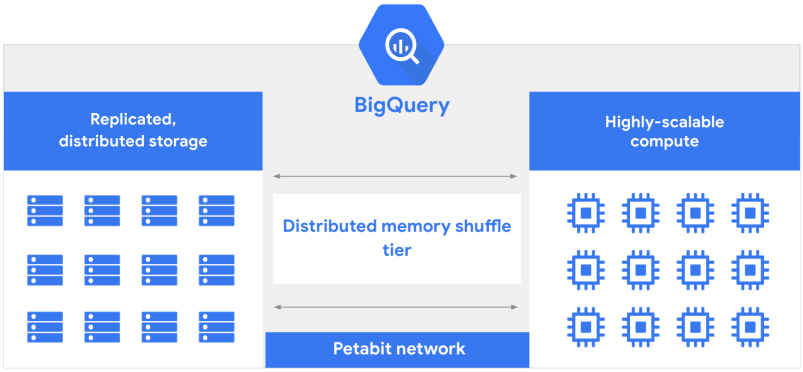
BigQuery is a fully managed, AI-ready data platform that helps you manage and analyze your data with built-in features like machine learning, search, geospatial analysis, and business intelligence. BigQuery's serverless architecture lets you use languages like SQL and Python to answer your organization's biggest questions with zero infrastructure management.

BigQuery in a minute

BigQuery provides a uniform way to work with both structured and unstructured data and supports open table formats like Apache Iceberg, Delta, and Hudi. BigQuery streaming supports continuous data ingestion and analysis while BigQuery's scalable, distributed analysis engine lets you query terabytes in seconds and petabytes in minutes.

BigQuery's architecture consists of two parts: a storage layer that ingests, stores, and optimizes data and a compute layer that provides analytics capabilities. These compute and storage layers efficiently operate independently of each other thanks to Google's petabit-scale network that enables the necessary communication between them.

Legacy databases usually have to share resources between read and write operations and analytical operations. This can result in resource conflicts and can slow queries while data is written to or read from storage. Shared resource pools can become further strained when resources are required for database management tasks such as assigning or revoking permissions. BigQuery's separation of compute and storage layers lets each layer dynamically allocate resources without impacting the performance or availability of the other.



This separation principle lets BigQuery innovate faster because storage and compute improvements can be deployed independently, without downtime or negative impact on system performance. It is also essential to offering a fully managed serverless data warehouse in which the BigQuery engineering team handles updates and maintenance. The result is that you don't need to provision or manually scale resources, leaving you free to focus on delivering value instead of traditional database management tasks.

BigQuery interfaces include Google Cloud console interface and the BigQuery command-line tool. Developers and data scientists can use client libraries with familiar programming including Python, Java, JavaScript, and Go, as well as BigQuery's REST API and RPC API to transform and manage data. ODBC and JDBC drivers provide interaction with existing applications including third-party tools and utilities.

As a data analyst, data engineer, data warehouse administrator, or data scientist, BigQuery helps you load, process, and analyze data to inform critical business decisions.

Get started with BigQuery

You can start exploring BigQuery in minutes. Take advantage of BigQuery's free usage tier or no-cost sandbox to start loading and querying data.

1. BigQuery's sandbox (/bigquery/docs/sandbox): Get started in the BigQuery sandbox, risk-free and at no cost.
2. Google Cloud console quickstart (/bigquery/docs/quickstarts/quickstart-web-ui): Familiarize yourself with the power of the BigQuery Console.
3. Public datasets (/bigquery/public-data): Experience BigQuery's performance by exploring large, real-world data from the Public Datasets Program.

Explore BigQuery

BigQuery's serverless infrastructure lets you focus on your data instead of resource management. BigQuery combines a cloud-based data warehouse and powerful analytic tools.

BigQuery storage

BigQuery stores data using a columnar storage format that is optimized for analytical queries. BigQuery presents data in tables, rows, and columns and provides full support for database transaction semantics (ACID (https://en.wikipedia.org/wiki/ACID)). BigQuery storage is automatically replicated across multiple locations to provide high availability.

- Learn about common patterns to organize BigQuery resources (/bigquery/docs/resource-hierarchy#patterns) in the data warehouse and data marts.
- Learn about datasets (/bigquery/docs/datasets-intro), BigQuery's top-level container of tables and views.
- Load data into BigQuery (/bigquery/docs/loading-data) using:
  - Stream data (/bigquery/docs/streaming-data-into-bigquery) with the Storage Write API (/bigquery/docs/write-api).
  - Batch-load data (/bigquery/docs/batch-loading-data) from local files or Cloud Storage using formats that include: Avro (/bigquery/docs/loading-data-cloud-storage-avro), Parquet (/bigquery/docs/loading-data-cloud-storage-parquet), ORC (/bigquery/docs/loading-data-cloud-storage-orc), CSV (/bigquery/docs/loading-data-cloud-storage-csv), JSON (/bigquery/docs/loading-data-cloud-storage-json), Datastore (/bigquery/docs/loading-data-cloud-datastore), and Firestore (/bigquery/docs/loading-data-cloud-firestore) formats.
- BigQuery Data Transfer Service (/bigquery/docs/dts-introduction) automates data ingestion.

For more information, see Overview of BigQuery storage (/bigquery/docs/storage_overview).

BigQuery analytics

Descriptive and prescriptive analysis uses include business intelligence, ad hoc analysis, geospatial analytics, and machine learning. You can query data stored in BigQuery or run queries on data where it lives using external tables or federated queries including Cloud Storage, Bigtable, Spanner, or Google Sheets stored in Google Drive.

- ANSI-standard SQL queries (SQL:2011 support (https://www.iso.org/standard/53681.html)) including support for joins, nested and repeated fields, analytic and aggregation functions, multi-statement queries, and a variety of spatial functions with geospatial analytics - Geographic Information Systems.

- Create views (/bigquery/docs/views-intro) to share your analysis.
- Business intelligence tool support including BI Engine (/bigquery/docs/bi-engine-intro) with Looker Studio (/bigquery/docs/visualize-looker-studio), Looker (/bigquery/docs/looker), Google Sheets (/bigquery/docs/connected-sheets), and 3rd party tools like Tableau and Power BI.
- BigQuery ML (/bigquery/docs/bqml-introduction) provides machine learning and predictive analytics.
- BigQuery Studio (/bigquery/docs/query-overview#bigquery-studio) offers features such as Python notebooks, and version control for both notebooks and saved queries. These features make it easier for you to complete your data analysis and machine learning (ML) workflows in BigQuery.
- Query data outside of BigQuery (/bigquery/external-data-sources) with external tables (/bigquery/docs/external-tables) and federated queries (/bigquery/docs/federated-queries-intro).

For more information, see Overview of BigQuery analytics (/bigquery/docs/query-overview).

BigQuery administration

BigQuery provides centralized management of data and compute resources while Identity and Access Management (IAM) (/iam/docs) helps you secure those resources with the access model that's used throughout Google Cloud. Google Cloud security best practices (/security/best-practices) provide a solid yet flexible approach that can include traditional perimeter security or more complex and granular defense-in-depth approach (/security/overview/whitepaper#technology_with_security_at_its_core).

- Intro to data security and governance (/bigquery/docs/data-governance) helps you understand data governance, and what controls you might need to secure BigQuery resources.
- Jobs (/bigquery/docs/jobs-overview) are actions that BigQuery runs on your behalf to load, export, query, or copy data.
- Reservations (/bigquery/docs/reservations-intro) let you switch between on-demand pricing and capacity-based pricing.

For more information, see Introduction to BigQuery administration (/bigquery/docs/admin-intro).

BigQuery resources

Explore BigQuery resources:

- Release notes (/bigquery/docs/release-notes) provide change logs of features, changes, and deprecations.
- Pricing (/bigquery/pricing) for analysis and storage. See also: BigQuery ML (/bigquery/pricing#bqml), BI Engine (/bi-engine/pricing), and Data Transfer Service (/bigquery-transfer/pricing) pricing.
- Locations (/bigquery/docs/locations) define where you create and store datasets (regional and multi-region locations).
- Stack Overflow (https://stackoverflow.com/questions/tagged/google-bigquery) hosts an engaged community of developers and analysts working with BigQuery.
- BigQuery Support (/bigquery/docs/getting-support) provides help with BigQuery.
- Google BigQuery: The Definitive Guide: Data Warehousing, Analytics, and Machine Learning at Scale (https://www.google.com/books/edition/Google_BigQuery_The_Definitive_Guide/-Jq4DwAAQBAJ) by Valliappa Lakshmanan and Jordan Tigani, explains how BigQuery works and provides an end-to-end walkthrough on how to use the service.

APIs, tools, and references

Reference materials for BigQuery developers and analysts:

- SQL query syntax (/bigquery/docs/reference/standard-sql/query-syntax) for details about using GoogleSQL.
- BigQuery API (/bigquery/docs/reference/libraries-overview) and client libraries (/bigquery/docs/reference/libraries) present overviews of BigQuery's features and their use.
- BigQuery code samples (/bigquery/docs/samples) provide hundreds of snippets for client libraries in C# (/docs/samples?l=csharp&p=bigquery), Go (/docs/samples?l=go&p=bigquery), Java (/docs/samples?l=java&p=bigquery), Node.js (/docs/samples?l=nodejs&p=bigquery), Python (/docs/samples?l=python&p=bigquery), Ruby (/docs/samples?l=ruby&p=bigquery). Or view the sample browser (/docs/samples?p=bigquery).

- DML (/bigquery/docs/data-manipulation-language), DDL (/bigquery/docs/reference/standard-sql/data-definition-language), and user-defined functions (UDF) (/bigquery/docs/reference/standard-sql/data-definition-language#create_function_statement) syntax lets you manage and transform your BigQuery data.
- bq command-line tool reference (/bigquery/docs/reference/bq-cli-reference) documents the syntax, commands, flags, and arguments for the `bq` CLI interface.
- ODBC / JDBC integration (/bigquery/docs/reference/odbc-jdbc-drivers) connect BigQuery to your existing tooling and infrastructure.

BigQuery roles and resources

BigQuery addresses the needs of data professionals across the following roles and responsibilities.

Data Analyst (#data-analyst) | Data Administrator ... | Data Scientist (#data-scientist) | Data Developer ...

Task guidance to help if you need to do the following:

- Query BigQuery data (/bigquery/docs/query-overview) using interactive or batch queries using SQL query syntax (/bigquery/docs/reference/standard-sql/query-syntax)
- Reference SQL expressions, functions, and operators (/bigquery/docs/reference/standard-sql/functions-and-operators) to query data
- Use tools to analyze and visualize BigQuery data including: Looker (/bigquery/docs/looker), Looker Studio (/bigquery/docs/visualize-looker-studio), and Google Sheets (/bigquery/docs/connected-sheets).
- Use geospatial analytics (/bigquery/docs/gis-intro) to analyze and visualize geospatial data with BigQuery's Geographic Information Systems
- Optimize query performance (/bigquery/docs/best-practices-performance-overview) using:
  - Partitioned tables (/bigquery/docs/partitioned-tables): Prune large tables based on time or integer ranges.

- Materialized views (/bigquery/docs/materialized-views-intro): Define cached views to optimize queries or provide persistent results.
- BI Engine (/bi-engine/docs): BigQuery's fast, in-memory analysis service.

To take a tour of BigQuery's data analytics features directly in the Google Cloud console, click **Take the tour**.

Take the tour (https://console.cloud.google.com/?walkthrough_id=bigquery--ui-tour-data-analyst)

BigQuery video tutorials

The following series of video tutorials get you started with BigQuery:

Title	Description
How to get started with BigQuery (https://www.youtube.com/watch?v=BH_7_zVk5oM) (17:18)	An overview that summarizes what is BigQuery and how to use it. Segments include: ETL pipelines, pricing and optimization, BigQuery ML and BI Engine, and wrapping up with a demo of BigQuery in Google Cloud console.
What is BigQuery? (https://www.youtube.com/watch?v=d3MDxC_iuaw) (4:39)	An overview of BigQuery of how BigQuery is designed to ingest and store large amounts of data to help analysts and developers alike
Using the BigQuery sandbox (https://www.youtube.com/watch?v=StEuT-pntZQ) (3:05)	How to set up a BigQuery sandbox, letting you run queries without needing a credit card
Asking questions, running queries (https://www.youtube.com/watch?v=Davhwj_8b8Q) (5:11)	How to write and run SQL queries in the BigQuery UI - plus picking a winning jersey number
Loading data into BigQuery (https://www.youtube.com/watch?v=Abzj-Vyhi74&t=60s) (5:31)	How to ingest and analyze data in real time, or just a one-time batch analysis of data - plus cats v. dogs
Visualizing query results (https://www.youtube.com/watch?v=AlKKZuULIxM) (5:38)	How data visualization is useful for making complex datasets easier to

Title	Description
	understand and internalize
Managing access with IAM (https://www.youtube.com/watch?v=ZTpOAmJWhAk) (5:23)	How to allow other users to query your datasets in BigQuery with IAM permissions and access control
Saving and sharing queries (https://www.youtube.com/watch?v=AYmoUQ_Iyeg&list=PLIivdWyY5sqLAbIdmcMwsxWg-w8Px34MS&index=8&t=8s) (6:17)	How to save and share your queries in BigQuery hassle-free
Protecting sensitive data with authorized views (https://www.youtube.com/watch?v=8fNnqshJ2Nw) (7:12)	How to easily share datasets with different users by setting customized access controls
Querying external data with BigQuery (https://www.youtube.com/watch?v=jBfx1K3-97k) (5:49)	How to set up an external data source in BigQuery and query data from Cloud Storage, Cloud SQL, Google Drive, and more
What are user-defined functions? (https://www.youtube.com/watch?v=c3dtgLWRycs) (4:59)	How to create user-defined functions (UDFs) for analyzing datasets in BigQuery

What's next

- For an overview of BigQuery storage, see Overview of BigQuery storage (/bigquery/docs/storage_overview).
- For an overview of BigQuery queries, see Overview of BigQuery analytics (/bigquery/docs/query-overview).
- For an overview of BigQuery administration, see Introduction to BigQuery administration (/bigquery/docs/admin-intro).
- For an overview of BigQuery security, see Overview of data security and governance (/bigquery/docs/data-governance).

Last updated 2024-08-21 UTC.