

Google Cloud Big Query

G-Cloud 15 – Service Definition

Service Overview	2
Key Features	2
Key Benefits	3
Target Use Cases	3
Contact Information	4

Service Overview

Google BigQuery is an autonomous, serverless, and highly scalable data-to-AI platform designed to help organizations manage and analyze data at petabyte scale. Positioned as a leader in the 2025 Gartner® Magic Quadrant™ for Cloud Database Management Systems, BigQuery automates the entire data life cycle—from ingestion and preparation to AI-driven insights. By decoupling storage and compute, it provides the flexibility to run complex analytical queries and machine learning (ML) models without the overhead of managing infrastructure.

Key Features

- **AI-Powered Experiences:** Features Gemini in BigQuery, providing conversational and agentic assistance for data engineering, data science, and analytics.
- **BigQuery ML:** Allows users to create, evaluate, and execute machine learning models (e.g., linear regression, k-means, time series) directly within BigQuery using familiar SQL.
- **Multimodal Data Support:** A unified platform for analyzing structured, semi-structured, and unstructured data, supporting open formats like Apache Iceberg.
- **Real-Time Analytics:** Built-in streaming capabilities and continuous queries enable event-driven analysis, making data available for querying the moment it is ingested.
- **Integrated Governance:** Managed through Dataplex Universal Catalog, offering automatic metadata harvesting, data profiling, lineage, and quality controls.
- **Data Clean Rooms:** Provides a secure environment for organizations to collaborate and share data insights with partners without moving or exposing raw underlying data.
- **Serverless Execution:** Utilizes Google's advanced infrastructure (Borg, Colossus, Dremel) to scale automatically to meet workload demands.

Key Benefits

- **Lower Total Cost of Ownership (TCO):** Offers up to 54% lower TCO compared to cloud-based alternatives through compressed storage, compute autoscaling, and flexible pricing models.
- **Faster Time to Insight:** Eliminates the need for complex ETL pipelines by allowing AI and ML to run where the data resides.
- **Enterprise-Grade Resilience:** Includes managed disaster recovery with cross-region replication and failover capabilities for mission-critical workloads.
- **Democratized AI:** Simplifies the path to AI for data analysts by making advanced generative AI functions (like text summarization and sentiment analysis) available via SQL.
- **Open and Interoperable:** Seamlessly integrates with open-source frameworks like Apache Spark and Python (via BigQuery DataFrames) while maintaining unified security and governance.

Target Use Cases

- **Modern Data Warehousing & Migration:** Consolidating siloed data and migrating legacy on-premises or cloud data warehouses (e.g., Redshift, Snowflake, Teradata) to a managed cloud environment.
- **Predictive Analytics & Data Science:** Building end-to-end ML workflows, from exploratory data analysis in Colab Enterprise notebooks to deploying models for real-time predictions.
- **Real-Time Business Intelligence:** Monitoring business operations in real-time, such as tracking supply chain logistics or detecting fraudulent transactions as they happen.
- **Geospatial Analysis:** Unlocking planetary-scale insights for site selection, sustainable sourcing, and infrastructure planning using built-in geospatial functions and Google Maps Platform integration.
- **Privacy-Centric Data Sharing:** Collaborating with external partners in Data Clean Rooms to analyze joint datasets (e.g., retail and marketing data) while maintaining strict data privacy and compliance.

Contact Information

Netpremacy Public Sector Team

publicsector@netpremacy.com

Call: +44 1420 481450