Watch the Data Cloud Summit
(https://cloudonair.withgoogle.com/events/summit-data-cloud-2022) on demand and learn about
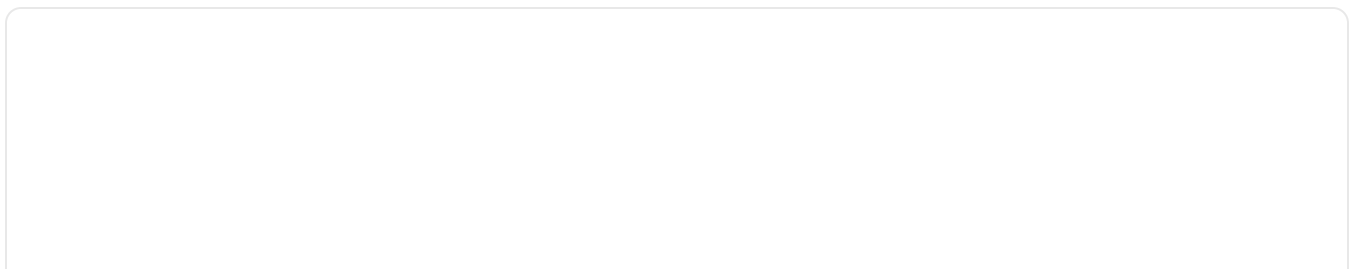the latest innovations in analytics, AI, BI, and databases.

Datasets

Enhance your analytics and AI initiatives with pre-built data solutions and valuable
datasets powered by BigQuery (https://cloud.google.com/bigquery), Cloud Storage
(https://cloud.google.com/storage), Earth Engine (https://earthengine.google.com), and other
Google Cloud services.

Explore datasets (https://console.cloud.google.com/marketplace/browse?filter=solution-type:dataset)

Start Analytics Hub free trial (https://cloud.google.com/analytics-hub)

Expand your data ecosystem

Increase the value of your data assets when you augment your analytics or AI
initiatives with external data. Discover and access unique and valuable datasets and
pre-built solutions from Google, public, or commercial providers. With fully managed
data pipelines, you can stay focused on what matters most: delivering insights and
business value.

(https://www.youtube.com/watch?v=CmRfB3h5H2Q)

Learn more about our public datasets

International Google Search terms now in BigQuery

Enhance your analysis with Google Trends from over 50 countries using proprietary Google Search data in BigQuery and Analytics Hub.

→

Inside the public dataset onboarding pipeline

Check out our release notes for the latest updates and datasets onboarded using our public datasets pipeline repository.

→

Announcing Analytics Hub in Preview

Read our blog about how Analytics Hub enables efficient and secure exchange of valuable data and analytics assets across organizational boundaries.

→

Featured datasets

CATEGORY

Google datasets ⌃

📈 Google Trends

(https://console.cloud.google.com/marketplace/product/bigquery-public-datasets/google-search-trends)

View the Top 25 and Top 25 rising queries from Google Trends from the past 30-days with this dataset. Each term includes 5 years of historical data across the 210 Designated Market Areas (DMAs) in the US and now over 50 countries across the globe.

✅ What are the most popular retail items people have searched for across the area?

📍

Community Mobility Reports
(https://console.cloud.google.com/marketplace/product/bigquery-public-datasets/covid19_google_mobility)

This dataset reports movement trends over time by geography, across different categories of places such as retail and recreation, groceries and pharmacies, parks, transit stations, workplaces, and residential.

✅ Use case: Identifying the difference in retail traffic on weekends

▪

Google Analytics (Sample)
(https://console.cloud.google.com/marketplace/product/obfuscated-ga360-data/obfuscated-ga360-data)

The dataset provides 12 months (August 2016 to August 2017) of obfuscated Google Analytics 360 data from the Google Merchandise Store
(https://www.googlemerchandisestore.com/shop.axd/Home?
utm_source=Partners&utm_medium=affiliate&utm_campaign=Data%20Share%20Promo)
to show what an ecommerce website would see, including traffic source, content, and transactional data.

✅ What is the total number of transactions generated per device browser?

💡

Google Patents Research
(https://console.cloud.google.com/marketplace/product/google_patents_public_datasets/google-patents-research-data)

Google Patents Research Data contains the output of much of the data analysis work used in Google Patents (patents.google.com), including machine translations of titles and abstracts from Google Translate, embedding vectors, extracted top terms, similar documents, and forward references.

✅ What are the 20 most recent patents filed?

Commercial datasets ⌄

CRUX
Crux Informatics
 (https://console.cloud.google.com/marketplace/product/crux-marketplace/crux-deliver)

Crux Deliver is a managed service for data engineering and operations. Crux wires up all of the traditional and alternative data providers on behalf of its clients and manages all aspects of onboarding, data engineering, and operations. Every dataset is validated so that we only deliver clean and actionable data.

✅ What are the datasets Crux can help me onboard into my data ecosystem?

❮❯EDI
Exchange Data International (https://www.exchange-data.com/)

Exchange Data International (EDI) helps the global financial and investment community make informed decisions. EDI's extensive content database includes worldwide equity and fixed income corporate actions, dividends, static reference data, closing prices, and shares outstanding.

✅ Understand historical events that affect Equity Shares and ETFs.

FACTSET
Factset (https://www.factset.com/data)

FactSet is a global provider of integrated financial information, analytical applications, and industry-leading service that delivers superior content, analytics, and flexible technology.

✅ Track multiple versions of merger deals to enhance your investment process.

HouseCanary (https://console.cloud.google.com/marketplace/product/housecanary-com/housecanary)

Instant access to reliable property, loan and valuation information for 100M homes. ML algorithms process hundreds of data sources to provide Home Price Indices for 381 Metros, 18,300 ZIP codes and 4M blocks covering >95% of the US residential market.

✅ Make investment decisions from 40-year historical volatility or 3-year forecast.

LinkUp (https://pages.linkup.com/General-Data-Landing-Capital-Markets.html)

LinkUp, the global leader in accurate, real-time, and predictive job market data and analytics offers proprietary data solutions that give customers the ability to derive valuable insights into the global labor market and help investors generate alpha at the macro, sector, geographic, and individual company level.

✅ Create models and signals to assess and predict job growth at the sector level.

London Energy Brokers Association (https://www.lebaltd.com/)

LEBA's solution gives customers the ability to access a unique, consolidated view of the Energy markets from across the main energy brokers. Energy, Oil and Gas producers, wholesale users, utilities, and financial traders benefit from independent market information based on traded activity rather than price assessments.

✅ Understand the energy prices across countries in Europe

Neustar (https://www.home.neustar/)

Neustar, Inc., a TransUnion company, is a leader in identity resolution providing the data and technology that enable trusted connections between companies and people at the moments that matter most. Neustar offers industry-leading solutions in marketing, risk and communications.

✅ Improve customer data assets and build privacy-focused consumer databases

RS Metrics (https://rsmetrics.com/)

RS Metrics, the leading company for asset-level, real-time, objective and verifiable ESG data, gives customers the ability to access accurate insights into EV manufacturers' factory inventory levels.

✅ Create independent, verifiable, and objective benchmarks of EV car production.

URSA

Ursa Space Systems_ (https://ursaspace.com/)

Ursa Space Systems, a global satellite intelligence infrastructure provider, gives customers the ability to monitor global economic trends with data derived from satellite imagery, updated on a weekly basis.

✅ What is the likely direction of oil price benchmarks and regional spreads?

Public datasets ⌄

Severe Storm Event Details
(https://console.cloud.google.com/marketplace/product/noaa-public/severe-storm-events)

The Storm Events Database is an integrated database of severe weather events across the United States from 1950 to this year, with information about a storm event's location, azimuth, distance, impact, and severity, including the cost of damages to property and crops.

✅ Use case: home improvement retailer understanding impact of storms on inventory

✅ Technical Reference Pattern: <u>Dynamic insurance pricing model using this dataset</u>
   (https://cloudonair.withgoogle.com/events/summit-finserv?
   talk=fireside_insurance_dynamic_pricing_noaa)

Census Bureau US Boundaries
(https://console.cloud.google.com/marketplace/product/united-states-census-bureau/us-geographic-boundaries)

These are full-resolution boundary files, derived from TIGER/Line Shapefiles, the fully supported, core geographic products from the US Census Bureau.These include information for the 50 states, the District of Columbia, Puerto Rico, and the outlying island areas.

✅ Use case: Developing an urbanization index for retailers

American Community Survey
 (https://console.cloud.google.com/marketplace/product/united-states-census-bureau/acs)

The American Community Survey (ACS) is an ongoing survey that provides vital information on a yearly basis about our nation and its people by contacting over 3.5 million households across the country. The resulting data provides incredibly detailed demographic information across the US aggregated at various geographic levels.

✅ Use case: Population growth trends as inputs to facility/site selection analysis

🛒

All public datasets (https://console.cloud.google.com/marketplace/browse?filter=solution-type:dataset)

Search for and access over 200 datasets listed in Google Cloud Marketplace.

✅ What datasets can help provide deeper context for our analytics or ai workflows?

Earth Engine datasets ⌄

🌐

Earth Engine (https://developers.google.com/earth-engine/datasets)

Earth Engine's public data archive includes more than forty years of historical imagery and scientific datasets, updated daily and available for online analysis.

✅ How has surface temperature changed over the past 30 years?

✅ What did this area look like before year 2000?

Kaggle datasets ⌄

k

Kaggle Datasets (https://www.kaggle.com/datasets)

Inside Kaggle you'll find all the code and data you need to do your data science work. Use over 80,000 public datasets and 400,000 public notebooks to conquer any analysis in no time.

✅ Can you tackle some of the most vexing and provocative problems in data science?

Synthetic datasets ⌄

Cymbal

Cymbal Investments
(https://console.cloud.google.com/marketplace/product/cymbal/cymbal_investments)

The synthetic data represents transactions from automated trading bots operated by the fictional Cymbal Investments group, each using a single algorithm to guide its trading decisions. The records are derived from FIX protocol (version 4.4) <u>Trade Capture Reports</u> (https://fiximate.fixtrading.org/legacy/en/FIX.4.4/body_54526569.html) loaded into BigQuery.

✅ How much did traders make from each individual trade?

Research datasets ⌄

G

Dataset Search_ (https://datasetsearch.research.google.com/)

Google's Dataset Search program has indexed almost 25 million datasets from across the web, giving you a single place to search for datasets and find links to where the data is. Filter by recency, format, topic, and more.

✅ What datasets exist for < keyword you're interested in >?

✅ Which sustainability datasets exist from last year are free for commercial use?

Feeling inspired? Let's solve your challenges together.

Learn how Google Cloud datasets transform the way your business operates with data and pre-built solutions.

Contact sales (https://cloud.google.com/contact)

If there is a public dataset you would like to see onboarded, please contact public-data-help@google.com.

With **BigQuery sandbox** (https://cloud.google.com/bigquery/docs/sandbox), you can try the full BigQuery experience without a billing account or credit card.

Data partners and customer stories

Learn more from both sides of the dataset ecosystem: data providers and data consumers.

BLOG POST

NOAA and Google Cloud: A data match made in the cloud

5-min read

See all customers (https://cloud.google.com/customers)

Take the next step

Start building on Google Cloud with $300 in free credits and 20+ always free products.

Get started for free (https://console.cloud.google.com/freetrial)

Need help getting started?

Contact sales (https://cloud.google.com/contact/)

Work with a trusted partner

Find a partner (https://cloud.withgoogle.com/partners/)

Continue browsing

See all products (https://cloud.google.com/products/)