

BigQuery public datasets

A public dataset is any dataset that is stored in BigQuery and made available to the general public through the [Google Cloud Public Dataset Program](#) (/public-datasets). The public datasets are datasets that BigQuery hosts for you to access and integrate into your applications. Google pays for the storage of these datasets and provides public access to the data via a [project](#) (/docs/overview#projects). You pay only for the queries that you perform on the data. The first 1 TB per month is free, subject to [query pricing details](#) (/bigquery/pricing#analysis_pricing_models).

Public datasets are available for you to analyze using either legacy SQL or [standard SQL](#) (/bigquery/docs/reference/standard-sql/query-syntax) queries. Use a fully qualified table name when querying public datasets, for example `bigquery-public-data.bbc_news.fulltext`.

You can access BigQuery public datasets by using the [Google Cloud console](#) (<https://console.cloud.google.com/marketplace/partners/bigquery-public-data>), by using the [bq command-line tool](#) (/bigquery/docs/cli_tool), or by making calls to the [BigQuery REST API](#) (/bigquery/docs/reference/v2) using a variety of [client libraries](#) (/bigquery/docs/reference/libraries) such as [Java](#) (<https://developers.google.com/api-client-library/java/apis/bigquery/v2>), [.NET](#) (https://developers.google.com/api-client-library/dotnet/get_started), or [Python](#) (<https://developers.google.com/api-client-library/python/>). You can also [view and query public datasets through Analytics Hub](#) (/bigquery/docs/analytics-hub-view-subscribe-listings#view_listings), a data exchange platform in [preview](#) (/products#product-launch-stages) that helps you discover and access data libraries.

[Go to Analytics Hub](#) (<https://console.cloud.google.com/bigquery/analytics-hub>)

You can find more details about each individual dataset by clicking the dataset's name in the Datasets section of Cloud Marketplace.

[Go to Datasets in Cloud Marketplace](#) (<https://console.cloud.google.com/marketplace/browse?filter=so>)

Note: The **Last Updated** date on a Cloud Marketplace dataset page indicates when the dataset page was last updated. To find out when the data itself was last updated, see [Accessing public datasets in the Google Cloud console](#) (#public-ui).

Before you begin

To get started using a BigQuery public dataset, you must create or select a project. The first terabyte of data processed per month is free, so you can start querying public datasets without enabling billing. If you intend to go beyond the [free tier](/bigquery/pricing#free-tier) (/bigquery/pricing#free-tier), you must also enable billing.

1. If you're new to Google Cloud, [create an account](https://console.cloud.google.com/freetrial)

(<https://console.cloud.google.com/freetrial>) to evaluate how our products perform in real-world scenarios. New customers also get \$300 in free credits to run, test, and deploy workloads.

2. In the Google Cloud console, on the project selector page, select or [create a Google Cloud project](/resource-manager/docs/creating-managing-projects) (/resource-manager/docs/creating-managing-projects).

★ **Note:** If you don't plan to keep the resources that you create in this procedure, create a project instead of selecting an existing project. After you finish these steps, you can delete the project, removing all resources associated with the project.

[Go to project selector](https://console.cloud.google.com/projectselector2/home/dashboard) (<https://console.cloud.google.com/projectselector2/home/dashboard>)

3. Make sure that billing is enabled for your Cloud project. Learn how to [check if billing is enabled on a project](/billing/docs/how-to/verify-billing-enabled) (/billing/docs/how-to/verify-billing-enabled).

4. BigQuery is automatically enabled in new projects. To activate BigQuery in a preexisting project,
Enable the BigQuery API.

[Enable the API](https://console.cloud.google.com/flows/enableapi?apiid=bigquery) (<https://console.cloud.google.com/flows/enableapi?apiid=bigquery>)

Public dataset locations

Each public dataset is stored in a specific location like `US` or `EU`. Currently, the BigQuery sample tables are stored in the `US` multi-region [location](/bigquery/docs/locations) (/bigquery/docs/locations). When you query a sample table, supply the `--location=US` flag on the command line, choose `US` as the processing location in the Google Cloud console, or specify the `location` property in the `jobReference` section of the [job resource](/bigquery/docs/reference/rest/v2/jobs) (/bigquery/docs/reference/rest/v2/jobs) when you use the API. Because the sample tables are stored in the `US`, you cannot write sample table query results to a table in another region, and you cannot join sample tables with tables in another region.

Access public datasets in the Google Cloud console

You can access public datasets in the [Google Cloud console](#)

(<https://console.cloud.google.com/bigquery>) through the following methods:

- View the **bigquery-public-data** project in the [Explorer panel](#) (/bigquery/docs/bigquery-web-ui#explorer_panel) of the navigation pane. The public dataset project is pinned to every project. To view the public datasets and tables in this project, see [Displaying resources](#) (/bigquery/docs/bigquery-web-ui#displaying_resources).

If the project is not shown, search for "bigquery-public-data" and click "Broaden search to all projects" to find this project. You can also pin the project to your Explorer panel.

- Use Analytics Hub (Preview) to [view and subscribe to public datasets](#) (/bigquery/docs/analytics-hub-view-subscribe-listings).

To find out when a data table was last updated, go to the table's **Details** section as described in [Getting table information](#) (/bigquery/docs/tables#getting_table_information), and view the **Last modified** field.

Other public datasets

There are many other public datasets available for you to query, some of which are also hosted by Google, but many more that are hosted by third parties. Other datasets include:

- [Cloud Life Sciences public datasets](#) (/life-sciences/docs/resources/public-datasets)
- [NIH chest x-ray dataset](#) (/healthcare/docs/resources/public-datasets/nih-chest)
- [The Cancer Imaging Archive \(TCIA\) dataset](#) (/healthcare/docs/resources/public-datasets/tcia)
- [Datasets publicly available on BigQuery \(reddit.com\)](#). (<https://www.reddit.com/r/bigquery/wiki/datasets>)
- [Dataset of release notes for the majority of generally available Google Cloud products](#). (https://console.cloud.google.com/bigquery?p=bigquery-public-data&d=google_cloud_release_notes&t=release_notes&page=table)

Share a dataset with the public

You can share any of your datasets with the public by changing the dataset's access controls to allow access by "All Authenticated Users". For more information about setting dataset access controls, see [Controlling access to datasets](#) (/bigquery/docs/dataset-access-controls).

When you share a dataset with the public:

- Storage charges are incurred by the billing account attached to the project that contains the publicly-shared dataset.
- Query charges are incurred by the billing account attached to the project where the query jobs are run.

For more information, see [Overview of BigQuery pricing](#) (/bigquery/pricing#overview_of_pricing).

Sample tables

In addition to the [public datasets](#)

(<https://console.cloud.google.com/marketplace/browse?filter=solution-type:dataset&q=public%20data>), BigQuery provides a limited number of sample tables that you can query. These tables are contained in the [bigquery-public-data:samples dataset](#) (<https://console.cloud.google.com/bigquery?p=bigquery-public-data&d=samples&page=dataset>).

The requirements for querying the BigQuery sample tables are the same as the requirements for querying the public datasets.

The [bigquery-public-data:samples](#) dataset includes the following tables:

Name	Description
gsod (https://console.cloud.google.com/bigquery?p=bigquery-public-data&d=samples&t=gsod&page=table)	Contains weather information collected by NOAA, such as precipitation amounts and wind speeds from late 1929 to early 2010.
github_nested (https://console.cloud.google.com/bigquery?p=bigquery-public-data&d=samples&t=github_nested&page=table)	Contains a timeline of actions such as pull requests and comments on GitHub repositories with a nested

schema. Created in September 2012.

github_timeline

(https://console.cloud.google.com/bigquery?p=bigquery-public-data&d=samples&t=github_timeline&page=table)

Contains a timeline of actions such as pull requests and comments on GitHub repositories with a flat schema. Created in May 2012.

nativity

(<https://console.cloud.google.com/bigquery?p=bigquery-public-data&d=samples&t=nativity&page=table>)

Describes all United States births registered in the 50 States, the District of Columbia, and New York City from 1969 to 2008.

shakespeare

(<https://console.cloud.google.com/bigquery?p=bigquery-public-data&d=samples&t=shakespeare&page=table>)

Contains a word index of the works of Shakespeare, giving the number of times each word appears in each corpus.

trigrams

(<https://console.cloud.google.com/bigquery?p=bigquery-public-data&d=samples&t=trigrams&page=table>)

Contains English language trigrams from a sample of works published between 1520 and 2008.

wikipedia

(<https://console.cloud.google.com/bigquery?p=bigquery-public-data&d=samples&t=wikipedia&page=table>)

Contains the complete revision history for all Wikipedia articles up to April 2010.

Contact us

If you have any questions about the BigQuery public dataset program, contact us at bq-public-data@google.com.

What's next

Learn how to query a table in a public dataset at [Quickstart using the Google Cloud console \(/bigquery/docs/quickstarts/quickstart-web-ui\)](https://cloud.google.com/bigquery/docs/quickstarts/quickstart-web-ui).

Except as otherwise noted, the content of this page is licensed under the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) (<https://creativecommons.org/licenses/by/4.0/>), and code samples are licensed under the [Apache 2.0 License](https://www.apache.org/licenses/LICENSE-2.0) (<https://www.apache.org/licenses/LICENSE-2.0>). For details, see the [Google Developers Site](#)

Policies (<https://developers.google.com/site-policies>). Java is a registered trademark of Oracle and/or its affiliates.

Last updated 2022-10-03 UTC.