# [What are the Best Practices for Reducing Cost?](#_bunpy6t7t22v)

# BigQuery Best Practices: Controlling Costs

## 1. Avoid SELECT \*

**Best practice:** Query only the columns that you need.

Using SELECT \* is the most expensive way to query data. When you use SELECT \*, BigQuery does a full scan of every column in the table.

If you are experimenting with data or exploring data, use one of the [data preview options](https://cloud.google.com/bigquery/docs/best-practices-costs#preview-data) instead of SELECT \*.

Applying a LIMIT clause to a SELECT \* query does not affect the amount of data read. You are billed for reading all bytes in the entire table, and the query counts against your free tier quota.

Instead, query only the columns you need. For example, use SELECT \* EXCEPT to exclude one or more columns from the results.

If you do require queries against every column in a table, but only against a subset of data, consider:

* Materializing results in a destination table and querying that table instead
* [Partitioning your tables by date](https://cloud.google.com/bigquery/docs/creating-partitioned-tables) and querying the relevant partition; for example, WHERE \_PARTITIONDATE="2017-01-01" only scans the January 1, 2017 partition

## 2. Sample data using preview options

**Best practice:** Don't run queries to explore or preview table data.

If you are experimenting with or exploring your data, you can use table preview options to view data for free and without affecting quotas.

BigQuery supports the following data preview options:

* In the web UI, on the **Table Details** page, click **Preview** to sample the data.
* In the CLI, use the [bq head](https://cloud.google.com/bigquery/docs/managing-table-data#browse-table) command and specify the number of rows to preview.
* In the API, use [tabledata.list](https://cloud.google.com/bigquery/docs/reference/rest/v2/tabledata/list) to retrieve table data from a specified set of rows.

## 3. Price your queries before running them

**Best practice:** Before running queries, preview them to estimate costs.

Queries are billed according to the number of bytes read. To estimate costs before running a query use:

* The query validator in the web UI
* The --dry\_run flag in the CLI
* The dry Run parameter when submitting a query job using the API
* The [Google Cloud Platform Pricing Calculator](https://cloud.google.com/products/calculator/)

## 4.Limit query costs by restricting the number of bytes billed

**Best practice:** Use the maximum bytes billed setting to limit query costs.

You can limit the number of bytes billed for a query using the maximum bytes billed setting. When you set maximum bytes billed, if the query will read bytes beyond the limit, the query fails without incurring a charge.

If a query fails because of the maximum bytes billed setting, an error like the following is returned:

Error: Query exceeded limit for bytes billed: 1000000. 10485760 or higher required.

## 5.LIMIT doesn’t affect cost

**Best practice:** Do not use a LIMIT clause as a method of cost control.

Applying a LIMIT clause to a query does not affect the amount of data that is read. It merely limits the results set output. You are billed for reading all bytes in the entire table as indicated by the query.

The amount of data read by the query counts against your free tier quota despite the presence of a LIMIT clause.

## 6.View costs using a dashboard and query your audit logs

**Best practice:** Create a dashboard to view your billing data so you can make adjustments to your BigQuery usage. Also consider streaming your audit logs to BigQuery so you can analyze usage patterns.

You can [export your billing data](https://support.google.com/cloud/answer/7233314?hl=en) to BigQuery and visualize it in a tool such as Google Data Studio. For a tutorial on creating a billing dasboard, see [Visualize GCP Billing using BigQuery and Data Studio](https://medium.com/google-cloud/visualize-gcp-billing-using-bigquery-and-data-studio-d3e695f90c08).

You can also stream your [audit logs](https://cloud.google.com/bigquery/audit-logs) to BigQuery and analyze the logs for usage patterns such as query costs by user.

## 7.Partition data by date

**Best practice:** Partition your tables by date.

If possible, [partition](https://cloud.google.com/bigquery/docs/creating-partitioned-tables) your BigQuery tables by date. Partitioning your tables allows you to query relevant subsets of data which improves performance and reduces costs.

For example, when you query partitioned tables, use the \_PARTITIONTIME pseudo column to filter for a date or a range of dates. The query processes data only in the partitions that are specified by the date or range.

## 8.Materialize query results in stages

**Best practice:** If possible, materialize your query results in stages.

If you create a large, multi-stage query, each time you run it, BigQuery reads all the data that is required by the query. You are billed for all the data that is read each time the query is run.

Instead, break your query into stages where each stage materializes the query results by writing them to a [destination table](https://cloud.google.com/bigquery/querying-data#permanent-table). Querying the smaller destination table reduces the amount of data that is read and lowers costs. The cost of storing the materialized results is much less than the cost of processing large amounts of data.

## 9.Consider the cost of large result sets

**Best practice:** If you are writing large query results to a destination table, use the default table expiration time to remove the data when it's no longer needed.

Keeping large result sets in BigQuery storage has a cost. If you don't need permanent access to the results, use the [default table expiration](https://cloud.google.com/bigquery/docs/managing-datasets#table-expiration) to automatically delete the data for you.

For more information, see [storage pricing](https://cloud.google.com/bigquery/pricing#storage).

## 10. Use streaming inserts with caution

**Best practice:** Use streaming inserts only if your data must be immediately available.

There is no charge for [loading data](https://cloud.google.com/bigquery/docs/loading-data) into BigQuery. There is a charge, however, for [streaming data](https://cloud.google.com/bigquery/streaming-data-into-bigquery) into BigQuery. Unless your data must be immediately available, load your data rather than streaming it.

# Link for cost controlling best practices for bigquery)

(https://cloud.google.com/bigquery/docs/best-practices-costs)