happy horizon // VARIFY.IO®

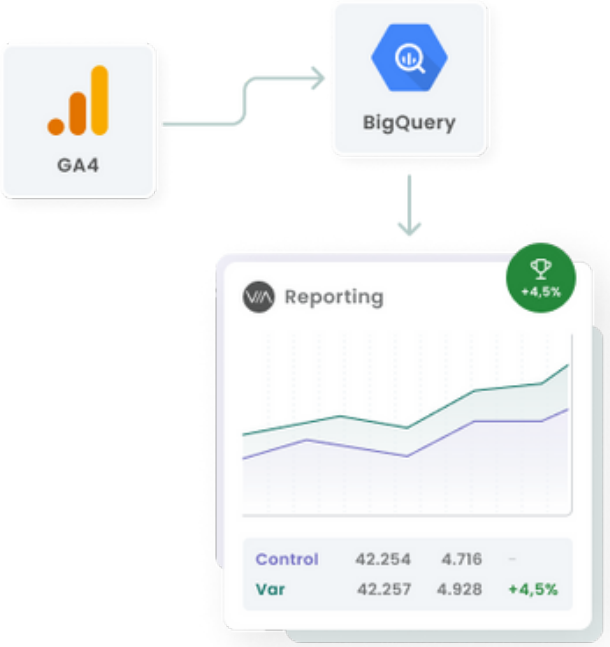
BigQuery for Experiment Analysis

Everything you need to know to get started

A Practical Guide to Reliable A/B Test Analysis with BigQuery

TABLE OF CONTENTS

Executive Summary



Running trustworthy A/B tests means having data you can trust. But many teams still rely on GA4 reports to decide which variation wins – and that can be risky.

GA4 sometimes produces misleading results due to data sampling, aggregation, or a technique called **HyperLogLog++**.

In some cases, this can make a test result look more confident than it really is.

This whitepaper, based on the webinar by Bas Linders (Happy Horizon) and Steffen Schulz (Varify), explains how BigQuery helps reduce that risk. You'll learn what makes BQ different from GA4, how to avoid common data pitfalls, and **how to get started**.

With simple **SQL templates, real-life testing examples, and hands-on tips**, you can begin using BigQuery right away – even without a full data team.

This is not just about tools. It's about creating a more reliable foundation for decision-making.

Introduction

Many teams use GA4 to run experiments. It's fast to set up, works with Looker Studio, and feels like a good default.

But tools like GA4 weren't designed for detailed test evaluation. They focus on speed and usability – not necessarily accuracy.
These visualizations are often based on modeled or sampled data. While this helps load dashboards faster, it can make the real effects of a test harder to judge.

In the webinar, we explained how this can lead to misleading conclusions. When the underlying data is pre-processed and simplified, **important details might be lost**. This makes it harder to understand what really happened during an experiment.

BigQuery, by contrast, provides raw event-level data. It allows you to explore the data without filters, modeling, or aggregation. You can group it how you want, compare what matters, and see the actual differences.

If you're running experiments, the way you access and analyze data matters. GA4 and similar tools may be great for reporting – but not for deep statistical analysis.

That's why many teams shift their analysis layer to BigQuery. It's not about replacing GA4. It's about **taking ownership** of the interpretation layer – and making sure it reflects what truly happened.

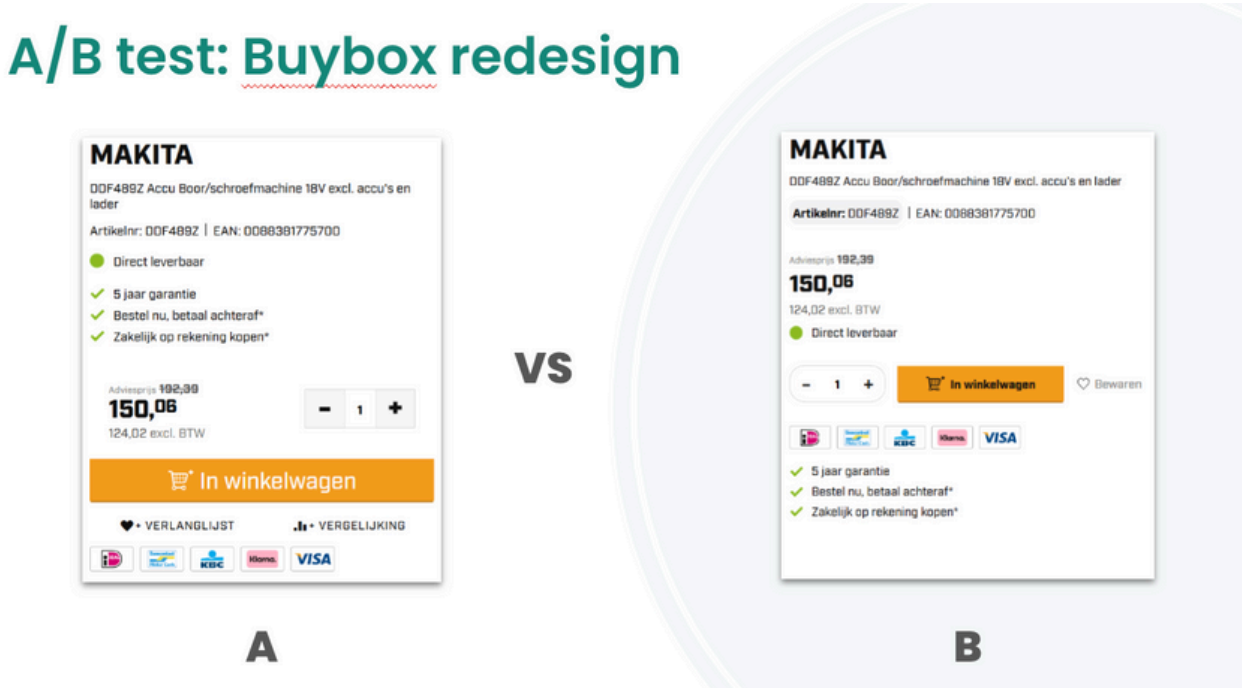
GA4 vs. BigQuery – Key Differences Explained

GA4 and BigQuery may use the same underlying data, but the way they present and access that data is fundamentally different.

Google Analytics 4 uses probabilistic methods (HLL++) to estimate unique users, especially when your experiment has a large number of visitors (over 12,000)[1]. As a result, metrics like users and conversions may be reported higher than they actually are. This is important to keep in mind when evaluating the results of high-traffic experiments.

Even small percentage shifts **(like +2% users, +3% conversions)** caused by modeling can create the illusion of a clear winner. But when the same data is analyzed in BigQuery, using raw numbers, the uplift may turn out smaller and statistical certainty lower. To demonstrate this Bas brought an A/B test with him, showcasing exactly this.

The A/B Test



[1]**Source:** https://blog.analytics-toolkit.com/

GA4 vs. BigQuery – Key Differences Explained

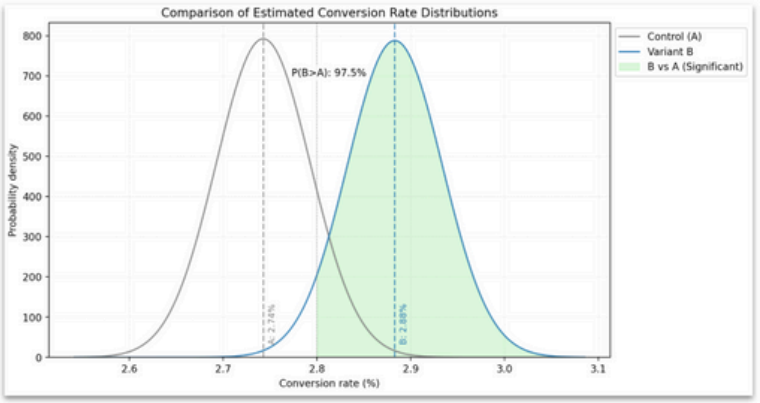
Results with GA4 Reporting

Experience variant	Users	Transactions
A	107742	2955
B	106554	3072

Test results for B vs A

- **Statistically significant result** for B with p-value: 0.0247!
- **Observed power:** 62.56%
- **Conversion rate change** for B: 5.12%

Variant **B** is a **winner**, congratulations!



Chance to win: 97,5%

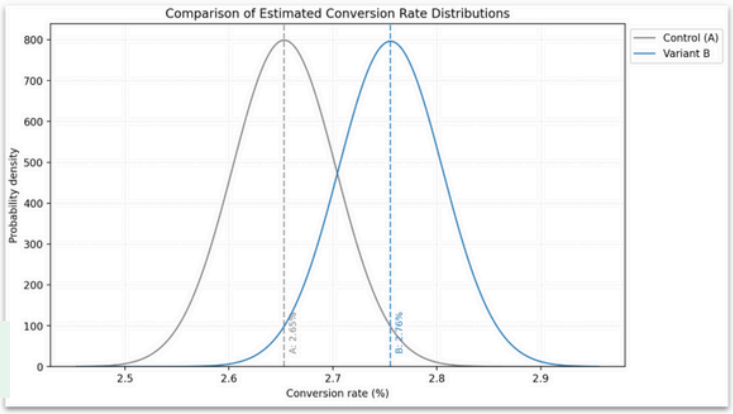
Results with BQ Reporting

Experience variant	Users	Transactions
A	105495	2799
B	104809	2888

Test results for B vs A

- **Confidence interval for difference in conversion rates:** (-0.0004, 0.0024)
- **Observed power:** 42.11%
- **p-value:** 0.0741
- **Conversion rate change for B:** 3.86%
- **P-value (non-inferiority test):** 0.0000

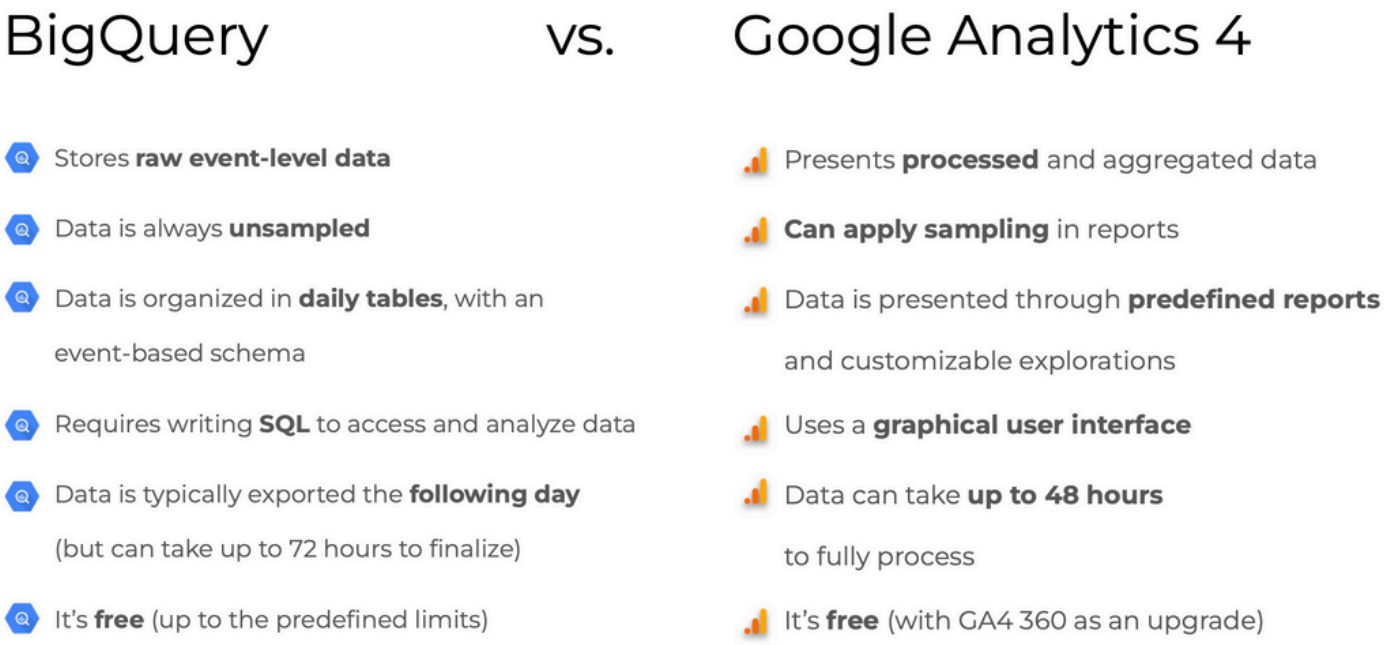
Although the Z-test is not statistically significant (p = 0.0741), the non-inferiority test suggests that B is **not significantly worse** than A within the predefined margin. This suggests that the variant may perform similarly to the control.



Chance to win: 92,49%

GA4 vs. BigQuery – Key Differences Explained

That doesn't mean GA4 is not-good. It's excellent for **reporting trends, funnel visualizations, or campaign monitoring**. But if you need precise answers for experiment evaluation, BigQuery is a safer bet. To summarize and properly showcase the differences, please refer to the attached comparison:



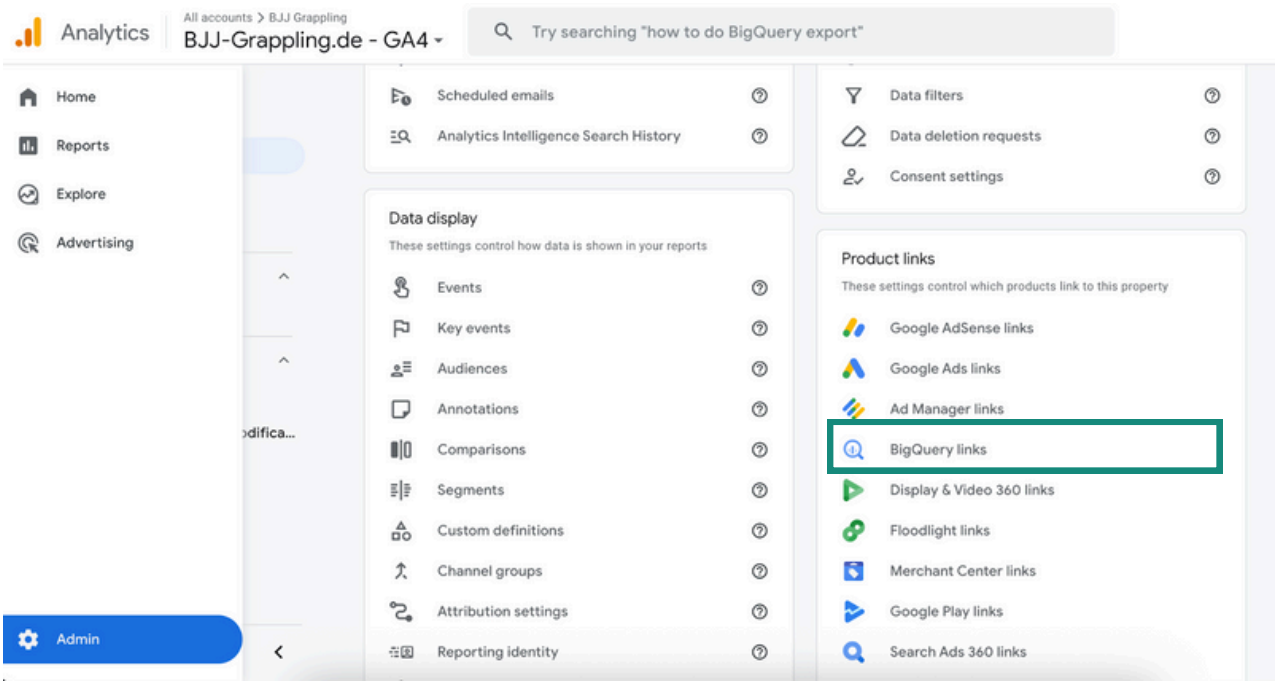
Importantly, this isn't just about GA4. Many A/B testing tools also model and smooth their data. That's why the principle applies more broadly: **trust the raw data** – and question the polished interfaces.

The first steps in BigQuery

Connecting GA4 to BigQuery

Connecting GA4 to BigQuery is easier than you might think – and it doesn't require a data engineering team. GA4 already collects event data in a structured and reliable way. The challenge starts when that data is processed for dashboards or reporting. That's where BigQuery comes in.

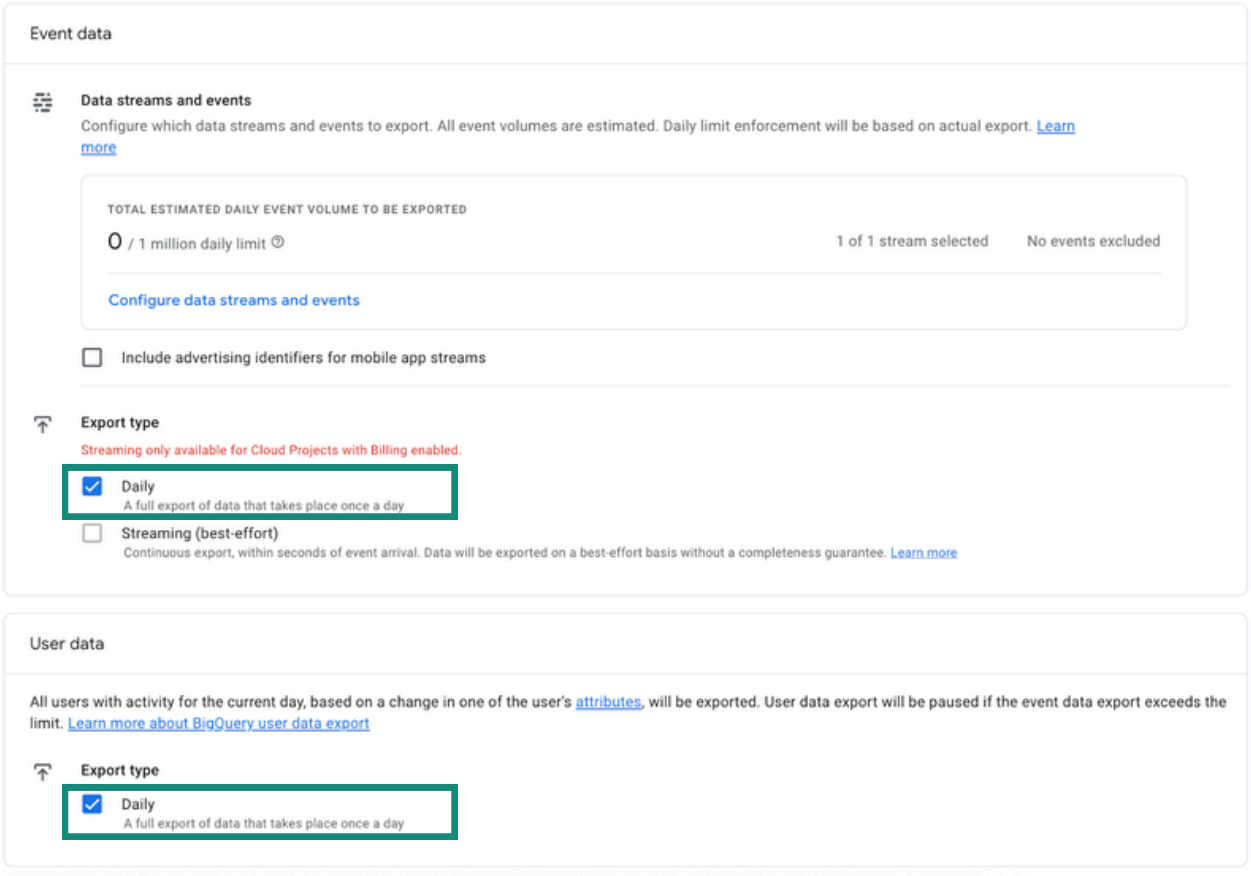
To connect your existing GA4 to BigQuery you want to simply turn it on. To do so you will have to find the "BigQuery links" section in your settings:



The first steps in BigQuery

Connecting GA4 to BigQuery

After clicking on “Link” at the top right corner you have to connect the corresponding data source of your previously created BigQuery project, (if not automatically set up by Varify.io) and adjust the settings the following:



These settings are recommended by us to ensure a robust data export.

The first steps in BigQuery

Ready-made SQL Queries to start with

After successfully connecting GA4 to BigQuery you can start testing it out via the Job explorer and your queries. To make this even easier, we shared two practical SQL templates during the webinar:

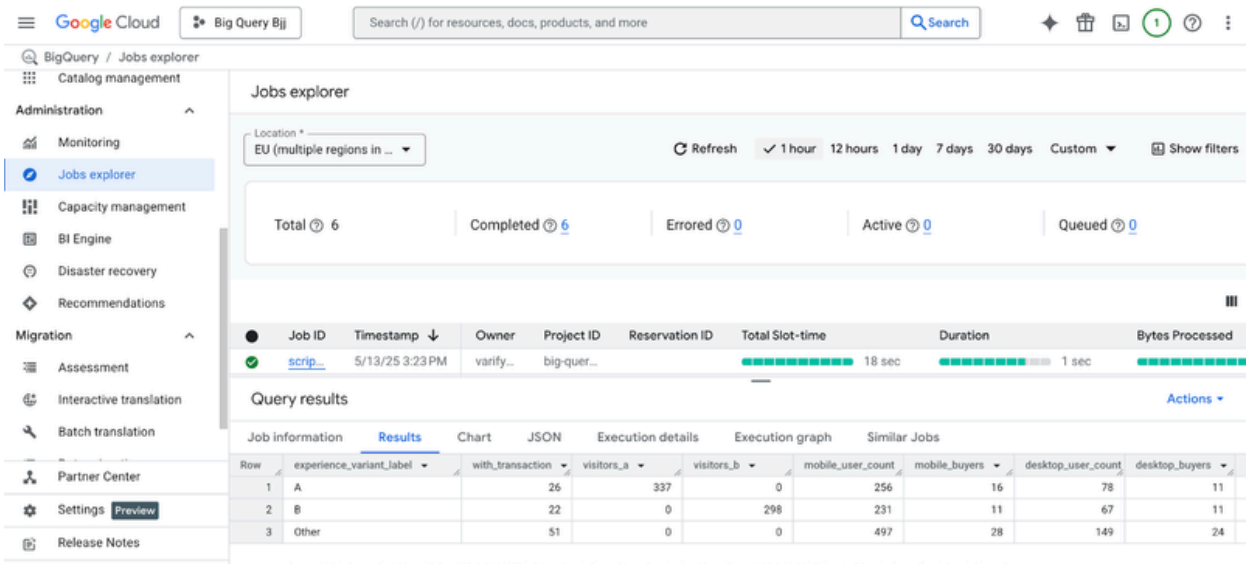
Transactional test evaluation – comparing the number of purchases between two variations

LINK

LINK

Revenue-based test evaluation – comparing average order value or total revenue uplift

After inserting the queries into the Job explorer, you should see something like this:



Practical Tips for Successfully Using BigQuery

Once your data is in BigQuery, the next challenge is working with it efficiently. During the webinar, we shared a few techniques that help you write better queries and avoid unnecessary costs.

Here's what we recommend:

- Use **_TABLE_SUFFIX** to filter by date. This lets you limit the amount of data scanned and keeps queries fast and cheap.
- Select only the fields you need. Avoid **SELECT *** – it's tempting, but it loads too much data and slows things down.
- Filter early. Apply conditions like **event_name = 'purchase'** or experiment-specific filters at the start of your queries.
- Use **SAFE_DIVIDE** when calculating rates. It protects you from division-by-zero errors.
- Start with a LIMIT. When testing your query, use **LIMIT 1000** to preview results without scanning your full dataset.

These tips make it easier to get started – especially if you're new to SQL. And if you're unsure how to build a query from scratch, the templates shared earlier are a great starting point.

Recommendations & Next Steps

Here's what we recommend if you want to use BigQuery to improve how you analyze experiments:

- **Compare your data sources**. Before making any decisions, take a step back and compare the results from different analytics platforms – not just GA4 and BigQuery, but also your A/B testing tools. Do the conversion numbers match? Are user counts consistent? If not, understand why. Each tool processes data differently, and these differences can shape your interpretation. Establishing clarity about what you treat as your 'source of truth' is the foundation for reliable analysis.

- **Work with SQL templates.** Use tested queries from the webinar or your team as a starting point. They help you avoid mistakes and make results easier to share.

- **Understand your GA4 schema.** If you're analyzing GA4 data in BigQuery, know your structure: event_name, event_params, user_pseudo_id, and so on. That's the backbone of any query.

- **Optimize before scaling.** Don't just run queries – improve them. Filter by date, avoid joins where possible, and document what each step is doing.

Recommendations & Next Steps

- **Use "safe spaces"** If you're using Varify, you already have one. Our visual analysis board lets you safely explore your BigQuery data without writing a single line of SQL. It's a guided interface built on top of your own experiment data – including variation exposure, conversions, and custom metrics. That means you can start analyzing confidently before ever opening the query editor.

- For teams writing their own queries, we recommend building **reusable base templates**. These help reduce errors and save time – especially when you're working with complex experiment setups.

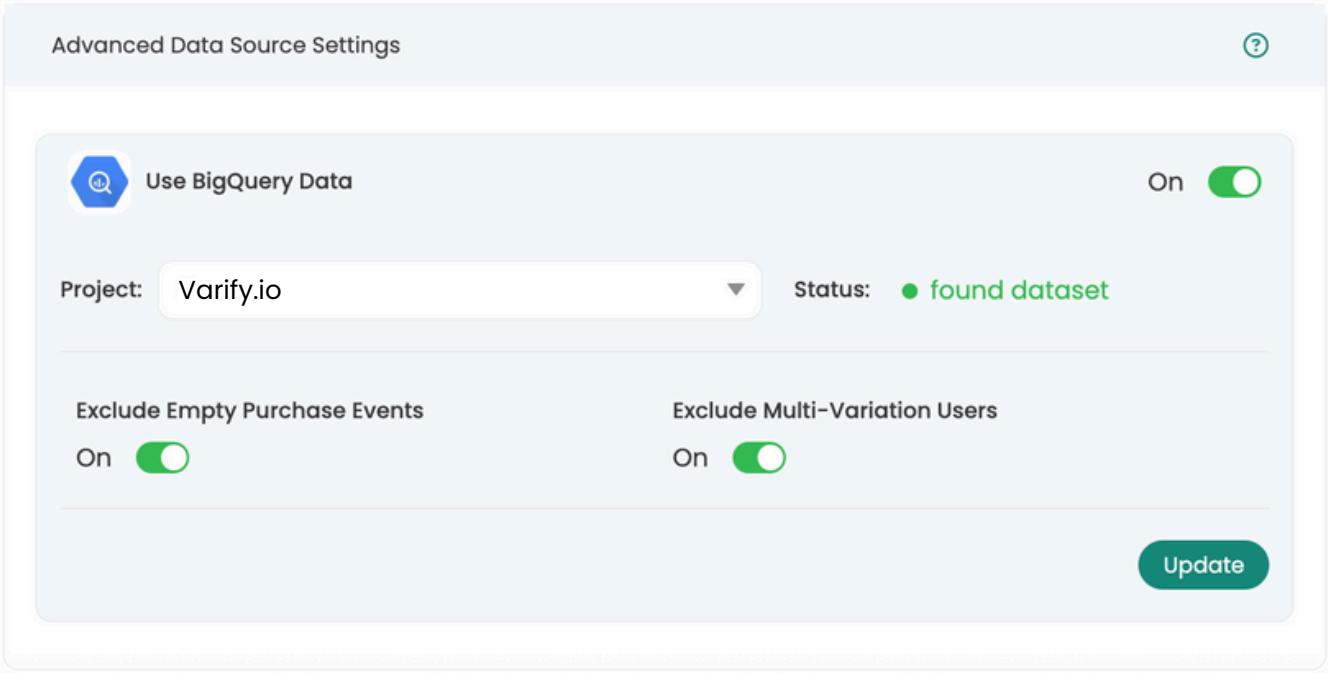
- **Ask for help early**. If something feels off, ask your data analyst or reach out to someone with SQL experience, like Bas. It's better to be slow and correct than fast and wrong.

BigQuery gives us a powerful way to **own** our experiment data – but only if we use it well. The earlier we build clean habits and simple workflows, the more confident we can be in every result we share.

Test BigQuery through Varify.io

Free and with no SQL

If you're using Varify, you don't need to work with SQL yourself. After connecting Varify with BigQuery via our auto setup or by selecting the right data source, the experiment data is synced automatically – including exposure events, conversions, and variation assignments.

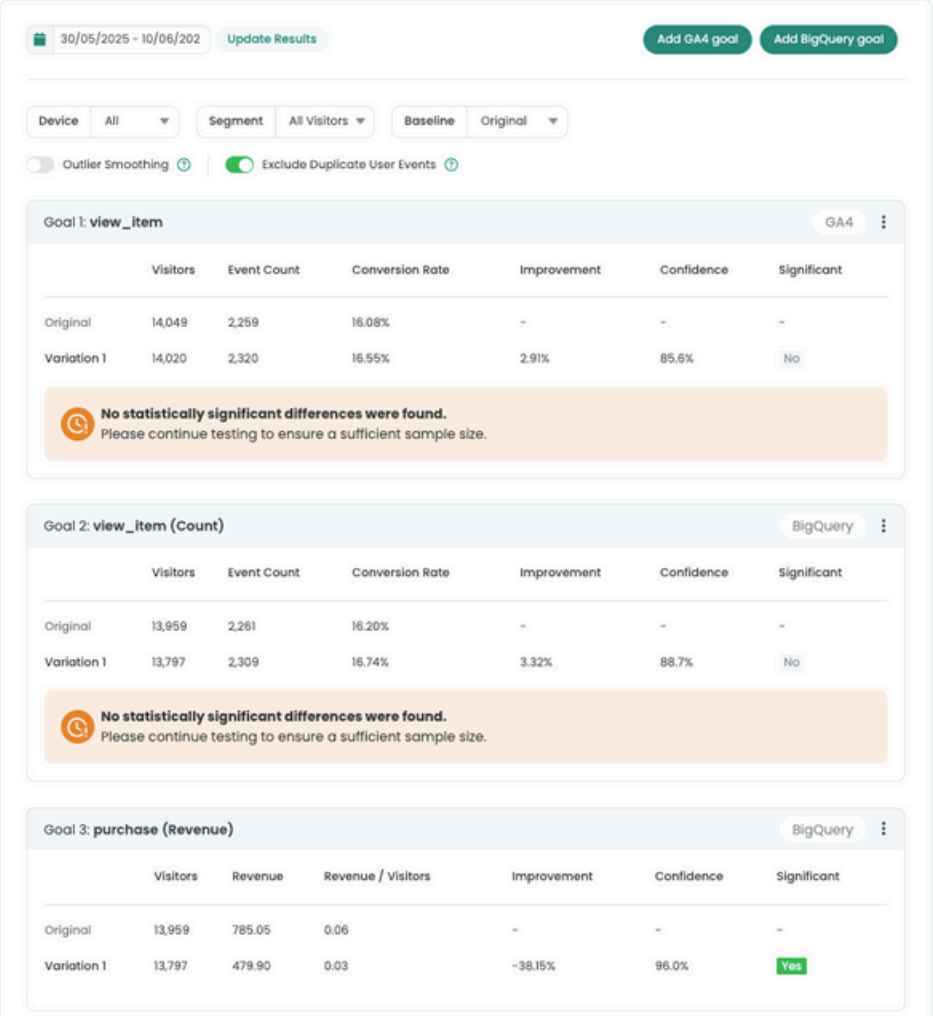


Curious to try it out? Join Varify's BQ-community, link on the last slide!

Test BigQuery through Varify.io

Free and with no SQL

After successfully connecting BigQuery in Varify.io you will start seeing your GA4 events as you are used to - but this time without them being extrapolated or you needing to create SQL queries.



Curious to try it out? Join Varify's BQ-community, link on the last slide!

Additionally you are now able to directly compare your GA4 and BQ data - **Are you surprised?**

References

- Based on the webinar of Bas Linders & Steffen Schulz: BigQuery for Experiment Analysis
- View the full recording of the webinar **here**

Bas Linders
Happy Horizon
Product Owner CRO | Data scientist
baslinders@happyhorizon.com

Reach out on Linkedin

Join our educational BigQuery-Community!

Whether you have open questions or simply want to try BigQuery for free through our plattform - join our Slack Community full of beginners and experts!