rudderstack™

The Definitive Guide

Replace Google Analytics with warehouse analytics

Replace Google Analytics with warehouse analytics

Table of Contents

BREAKING FREE FROM GOOGLE ANALYTICS

It's hard to believe, but Google will soon celebrate the 20th anniversary of Google Analytics. Over the past two decades, Google has steered the service to nearly ubiquitous usage, sitting on a whopping 87% market share[1]. It's not hard to see why: along with defining standards for categories of traffic, Google Analytics is largely plug-and-play, offers very robust out-of-the-box reporting, and perhaps most importantly, has remained completely free.

But users have had a love/hate relationship with the tool for a long time.

With the messy introduction of Google Analytics 4 (GA4) and a looming deadline to migrate off of Universal Analytics (UA), those long-held frustrations are coming to the surface. This has resulted in data and marketing teams rethinking whether they should continue to use Google Analytics at all.

TJ
@teej_m

Google has been screwing up GA for years now and people keep using it.

---

[1] Slintel - Google Analytics Market Share

These teams are increasingly turning to a solution that overcomes all of the limitations of Google Analytics, from data quality to flexibility and security: **their cloud data warehouse**.

Warehouses have long been used for creating reporting for key business metrics, but when you combine their storage and compute power with a dedicated, robust first-party data collection tool like RudderStack, you can open up powerful possibilities for all types of analytics, including traffic and behavioral analysis for websites and mobile apps.

Why companies are leaving Google Analytics

It's worth stepping back to ask why data and marketing teams are ditching a tool with such broad adoption, especially when GA4 seems to be a step forward. While GA4 is a better tool than UA in many ways, it's still plagued by the same old problems and, what's worse, it introduces several new ones.

GA4 is designed to create vendor lock-in

GA4's new integrations with BigQuery, Looker Studio and other tools seem exciting until you realize that Google will make you pay for them.

GA4 promises more flexibility, most notably the ability to export data from GA directly into BigQuery, Google's cloud data warehouse. That sounds great until you realize that

1. GA4's exciting new integrations are only for Google products, and
2. those products have limited free tiers and are designed to make you pay as you scale. For example, Looker Studio (formerly Data Studio) has a query limit, so if you want to leverage more advanced reporting functionality, you'll have to pay.

This is a shift in strategy by Google: instead of monetizing Google Analytics primarily from the top down with GA360, which cost ~$150,000 per year and was purchased only by large enterprises, Google is forcing all users onto GA4 and leveraging their market share to drive broad monetization of their cloud infrastructure and advanced analytics products.

For small companies or companies already running on Google Cloud, the pain of vendor lock-in might not be severe at first, but data infrastructure must be dynamic and evolve to meet the needs of your business over time. Entrusting your most valuable asset—customer data—and all of your infrastructure to a system designed for vendor lock-in is a precarious road indeed.

> **Christina LeVasseur (Brodzky)**
> **@cbrodzky**
>
> ICYMI: About a month ago, Google began enforcing a quota limit in Looker Studio for the GA4 connector. This means that if you reach your limit, which most of us prob will, your dashboard will look like this. There are a few solutions - none of which I like at this time.

GA4 doesn't capture all of your data

Losing visibility into a third of your traffic means you'll never develop a complete picture of your customer journey.

> *"A regular GA deployment, be it directly with gtag.js or Google Tag Manager will lose, depending on industry, anywhere between 20% - 50% of visitor data. There are ways to prevent this kind of blocking but it generally involves a non-insignificant amount of DevOps work behind the scenes."*
>
> *Max Werner, CEO of Obsessive Analytics*

Because Google Analytics is the most popular analytics tool in the world, it's the first target for both ad blockers and privacy-conscious browsers. As a result, GA4 does not capture all of the visitor data from your website or app.

Studies have shown that 40% of internet users leverage some sort of ad blocker[2] and multiple independent tests have shown that Google Analytics can lose 15-30% of traffic[3]!

---

[2] [Share of internet users who used an adblocker in the United States as of the 2nd quarter of 2020, by age group](link)
[3] [How much data is missing from your Google Analytics dashboard?](link)

GA4 still suffers from lack of data fidelity

Google Analytics payloads are designed specifically for Google Analytics, not for multi-purpose first-party analytics.

Along with capturing less data overall, the individual payloads sent by Google Analytics (including GA4) are limited in their detail. The included data is even encoded using proprietary keys, making individual data points more difficult to work with. GA4's new event-based payloads give you more flexibility in instrumentation, which is useful when data is processed by Google Analytics itself, but under the hood the payloads are extremely similar to what was sent by GA Classic.

GA Classic page payload	GA4 page payload
v: 1 _v: j98 a: 104871803 t: pageview _s: 3 dl: https://rudderstack.com/pricing/ dr: $direct dp: /pricing/ ul: en-us de: UTF-8 dt: RudderStack Pricing sd: 24-bit sr: 2560x1440 vp: 1280x787 je: 0 _u: aCDAAEIRAAAAACgDIAB~ jid: gjid: cid: 679471488.1651608029 tid: UA-176997588-1 _gid: 1790365349.1671218530 z: 942919277	v: 2 tid: G-252450485 gtm: 2oebu0 _p: 309274267 cid: 679471488.1651608029 ul: en-us sr: 2560x1440 uaa: arm uab: 64 uafvl: Google%20Chrome;107.0.5304.121|Chromium;107.0.5304.121|Not%3DA%3FBrand;24.0.0.0 uamb: 0 uam: uap: macOS uapv: 12.6.0 uaw: 0 _s: 2 dl: /pricing/ dr: $direct dt: RudderStack Pricing sid: 1671219066 sct: 1 seg: 1 en: page_view _ee: 1

GA4 operates on box data processing & features

GA4's black box features for identity resolution and predictions compromise your most important decisions.

If capturing less data and detail weren't bad enough, Google Analytics still processes key data functions behind the scenes in a black box—and GA4 doesn't solve this problem. UA was notorious for sampling data, and it's still unclear how much sampling GA4 data will be subject to.

Beyond sampling, though, there are much deeper concerns: GA4 provides one-size-fits-all features around identity resolution and predictions that aren't transparent to end users, meaning some of your most important decisions are based on significant unknowns that might not accurately reflect your specific business model or customer journey.

Matteo Zambon
@zambros_it

#GA4 #alert "Partial data shown for date range"
Ok, I've never seen this alert before. GA4
surprise me every day Did you find other alerts?

GA4 faces increasing security and regulatory scrutiny

Google Analytics stores a copy of your user data, creating both security and legal challenges.

Google Analytics is increasingly coming under legal scrutiny, especially in the European Union. In fact, in many countries that have passed GDPR legislation, Google Analytics is illegal to use out of the box[4]. Beyond specific regulatory concerns, Google Analytics also stores a copy of your data on their servers, meaning yet another data silo and security concern when it comes to your users' data.

GA4 still kills site performance

Degraded site performance is one of the most hated aspects of Google Analytics, especially for technical teams focused on site speed and SEO performance. In our own studies, we've seen key site

---

[4] NOYB Data protection complaint & decision - Austria

performance metrics improve by up to 40% simply by removing the Google Analytics and Google Tag Manager scripts.

Google Analytics means slow site load times

■ With Google Analytics   ■ Without Google Analytics
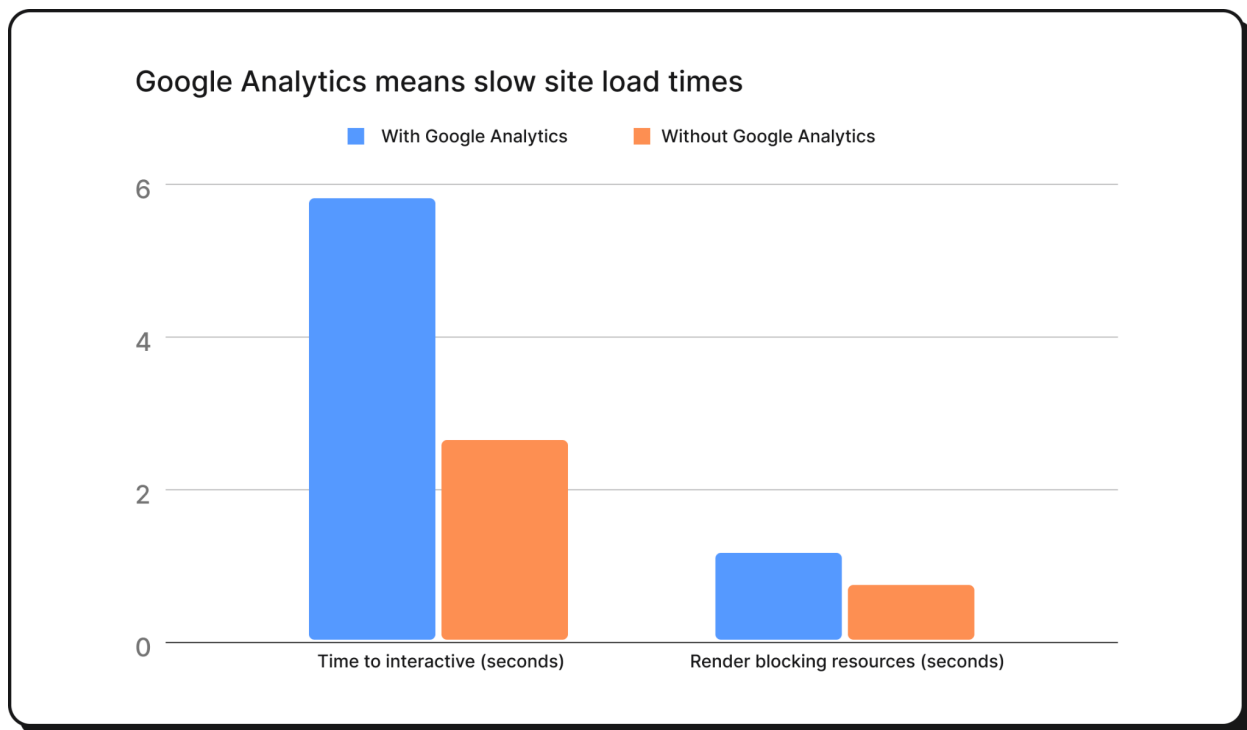
Max Chadwick
@maxpchadwick

After putting this off for months I put GA4 on my site... And it tanked my Lighthouse score by 14 points. Was previously at 100 and now down to 86. Only change was adding the GA4 tag

The benefits of owning your analytics infrastructure

Thankfully, you don't have to subject your data team, or any other team, to the limitations of Google Analytics. Using RudderStack for first-party data collection and your warehouse for storing, modeling

and serving data, you can build a future-proof analytics stack that will deliver data you can trust and easily scale with your business.

> *"Data has changed, and for those still depending on Google Analytics, this is the perfect moment to modernize your data stack. Your customer data is the most valuable asset your business has, and you need to own all of the data and stop being dependent on tools like GA to host or report it. With a solid CDP you can now track all your customer data and pipe it to a warehouse, and own your data. The businesses that own their data in the future will be the ones who win"*
>
> *Dan McGaw, Founder & CEO of McGaw.io, tech stack and analytics experts*

Analytics with infinite optionality

Abstracting your data capture infrastructure away from your data storage and transformation layers, as well as the analytics tools themselves, gives you the ability to modify any individual component of your stack as the analytics needs of your business change.

Do your A/B testing or product teams want to try a new analytics tool? No problem—just point the event stream at a new destination without opening any dev tickets (more on this later). Do your analysts need to modify metrics to reflect a new pricing plan? You can move fast by updating the queries running on your warehouse without having to completely overhaul event instrumentation and your entire visualization layer. What if your data science team is building a new recommendations model on a fresh data lake partition? You can avoid painful and slow batch jobs by simply adding a new data store as a destination for your existing events.

Owning your analytics infrastructure means giving the data team full control and flexibility to quickly adapt to the needs of the business without facing the painful limitations of vendor lock-in.

Capture every site visit

Capturing all of your data requires using an analytics-agnostic tool specifically designed to collect raw, first-party data. RudderStack's SDKs are purpose-built for capturing customer events and can be proxied behind your site or app URL, giving you full data capture.

Visits captured
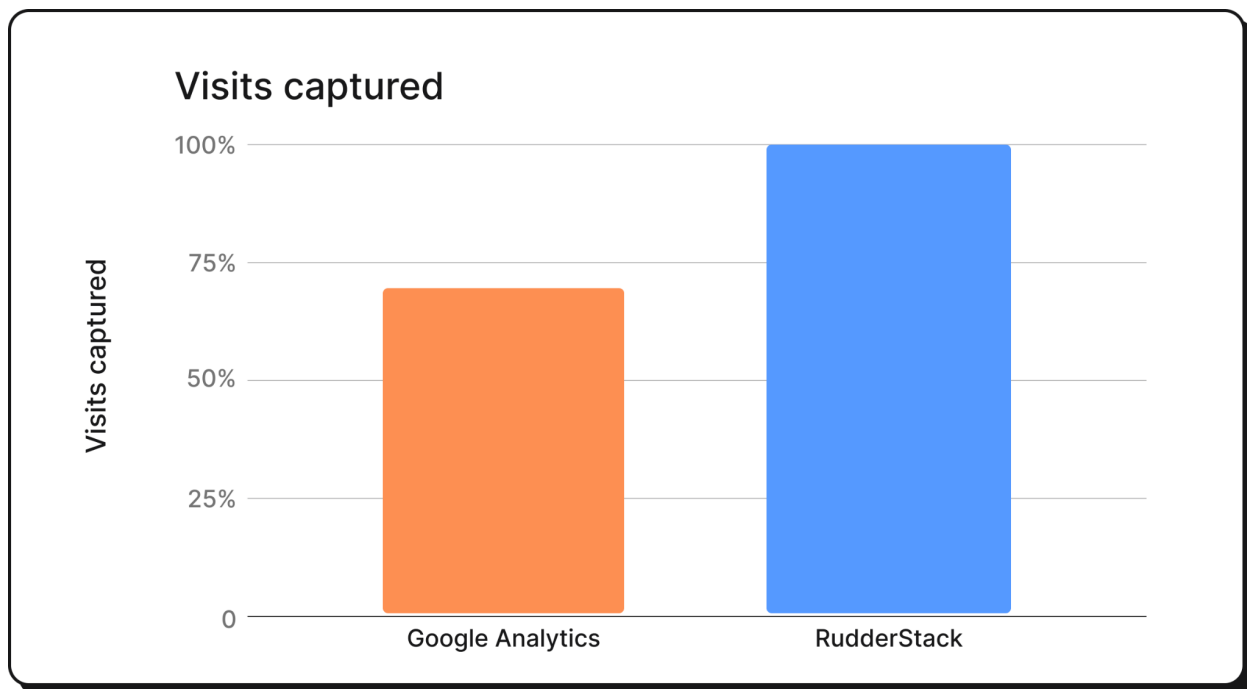
See the whole picture with rich, configurable, transformable payloads

RudderStack's out-of-the-box payloads provide far richer data than GA4. They also feature a standardized, open-source JSON schema with objects you can customize with both event properties and user traits that meet the specific needs of your business. If analytics needs change down the line, which they always do, you can use RudderStack's transformations to rename keys or reshape payloads without having to touch the code in your app or website.

GA4 page payload	RudderStack page payload

```
v: 2
tid: G-252450485
gtm: 2oebu0
_p: 309274267
cid: 679471488.1651608029
ul: en-us
sr: 2560x1440
uaa:
arm
uab: 64
uafvl:
Google%20Chrome;107.0.5304.121|Chromium;107.0.5304.121|Not%3DA%3FBran
d;24.0.0.0
uamb: 0
uam:
uap: macOS
uapv: 12.6.0
uaw: 0
_s: 2
dl: /pricing/
dr: $direct
dt: RudderStack Pricing
sid: 1671219066
sct: 1
seg: 1
en: page_view
_ee: 1
```

```
{
    "channel": "web",
    "context": {
        "app": {
            "build": "1.0.0",
            "name": "RudderLabs JavaScript SDK",
            "namespace": "com.rudderlabs.javascript",
            "version": "1.21.0"
        },
        "traits": {
            "name": "First Name"
        },
        "library": {
            "name": "RudderLabs JavaScript SDK",
            "version": "1.21.0"
        },
        "userAgent": "Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7)
AppleWebKit/537.36 (KHTML, like Gecko) Chrome/107.0.0.0 Safari/537.36",
        "locale": "en-US",
        "os": {
            "name": "",
            "version": ""
        },
        "screen": {
            "density": 2,
            "width": 2560,
            "height": 1440,
            "innerWidth": 1280,
            "innerHeight": 787
        },
        "sessionId": 1671218225671,
        "campaign": {},
        "page": {
            "path": "/pricing/",
            "referrer": "$direct",
            "referring_domain": "",
            "search": "",
            "title": "RudderStack pricing",
            "url": "https://rudderstack.com/pricing/",
            "tab_url": "https://rudderstack.com/pricing/",
            "initial_referrer": "$direct",
            "initial_referring_domain": ""
        }
    },
    "type": "page",
    "messageId": "1671218575125100-c1271cc4-5b0c-4b99-92a7-58cb69310802",
    "originalTimestamp": "2022-12-16T19:22:55.125Z",
    "anonymousId": "8bd1d069-0318-4d97-bfca-6b4a39ab0874",
    "userId": "eric@rudderstack.com",
    "properties": {
        "path": "/pricing/",
        "referrer": "$direct",
        "referring_domain": "",
        "search": "",
        "title": "RudderStack Pricing",
        "url": "https://rudderstack.com/pricing/",
        "tab_url": "https://rudderstack.com/pricing/",
        "initial_referrer": "$direct",
        "initial_referring_domain": ""
    },
    "integrations": {
        "All": true
    },
    "sentAt": "2022-12-16T19:22:55.135Z"
}
```

Build confidence with full transparency in your warehouse

When you store and model all of your data in your warehouse, you never have to wonder what GA4 is doing inside of the black box and whether their algorithms are making the right choice for your business.

> *"We were using the free version of Google Analytics. Everything was anonymized, and we couldn't see what our users were doing on our website or how they were using our product. RudderStack has increased visibility into user behavior and user journeys, given us deeper insight into our funnel and we can run A/B tests that let us customize the user experience."*
>
> *Mona Sami, Director of Data Analytics, InfluxData*

```
1  select * from "RUDDER_WEBAPP_DATA"."RUDDER_WEBAPP"."IDENTIFIES"
2
3
```

Results   Data Preview

✔ Query ID   SQL   789ms   100 rows

Filter result...   Copy

Row	NAME	ANONYMOUS_ID	ID	CHANNEL	CONTEXT_IP
1	Steve Jobs	5d2fd9c5-5231-4b36-96da-...	dd5288ec-667a-4b78-9590-cef69...	web	96.232.85.15
2	Elon Musk	0eb93ac7-9046-404a-a2ce-...	d7a91de0-d510-44f8-bc4f-93b60a...	web	170.10.252.83
3	Walt Disney	81e84cf3-5848-4836-86d9...	7d11407a-bef6-4516-80d3-f7eb72b...	web	134.41.188.254

Building analytics on your warehouse gives you full transparency and flexibility

Deploy and modify flexible data models that match actual behavior

Using tools like dbt, you can build your own models for key use cases identity resolution and sessionization, then modify them as you learn more about how your users interact with your website and apps.

Break free from Google's compliance chaos

Perhaps most importantly, owning your analytics infrastructure means you can say goodbye to Google's compliance chaos once and for all, meaning your infosec, legal and marketing teams can rest easy.

HOW-TO GUIDE: REPLACING GA4 WITH WAREHOUSE-FIRST ANALYTICS

In this section of the guide, we'll cover the definition of warehouse-first analytics and walk through the steps of replacing GA4 with your own data infrastructure.

What are "warehouse-first" analytics?

A warehouse-first data analytics stack is an analytics stack with a single data repository (otherwise known as the data warehouse) for all customer data. As we mentioned above, analytics tools like GA4 create data silos, but so do other customer data repositories like Salesforce. This makes it challenging for organizations to access all the data they need to make complex decisions.

By gathering all of your company's data into a single location, including traffic data from your websites and apps, you're able to build efficient analytics on large, diverse, high-quality datasets to answer questions that analytics tools like GA4 are not able to answer on their own.

For example, suppose your company is a startup in its early stages. In that case, you're probably just trying to understand how customers perceive or interact with your product, so you likely have questions like:

- How many active users did we have this month?
- What's our overall retention rate?
- What are our top sources of referral signups?

For those basic questions, GA4 can get the job done (albeit with limited, siloed data), but as your company grows, you'll need answers to more complex questions like:

- What behavior in a user's first day indicates they're likely to sign up for a paid plan?
- How does lifetime value vary by paid advertising channel?
- What patterns lead to an increase in purchases of our products?
- What products should we recommend to specific users to increase revenue?
- How should we price our product based on the cost of acquisition or cost of resources per customer?

You can't answer those questions with GA4. But you can with warehouse-first analytics. If you have all of the data in your warehouse, your data and analytics teams can not only model the raw data to find the answers, but understand how those answers change over time—all with fully transparent, version-controllable queries.

The components of a warehouse-first data analytics stack

A warehouse-first data analytics stack is made up of a few crucial components to facilitate the collection, storage, and movement of your data to make it easier to analyze:

1. Cloud data warehouse
2. Behavioral data ingestion

3. Data modeling
4. Visualization

For each element, let's look at a few tools fit to get the job done.

1. Centralizing the data: your data warehouse

With raw traffic and behavioral data—along with all of your other customer data—in a single place you get complete ownership of your data, plus you can perform much deeper analyses. It also means you can easily move between and onboard different analytics tools for different teams because each tool has access to all of your historical data.

Numerous data warehouse options are available, but Redshift, Snowflake, and BigQuery have the most adoption. While each of these warehouses has their own unique advantages, they all work well for ingesting event payloads from a tool like RudderStack to drive web analytics. Warehouses are used for more than web analytics, though, so data teams will want to consider other questions like the additional types of data being stored (i.e., semi-structured data), how much performance management is required by engineers and how the cost scales depending on your company's use case.

See our [Appendix](link) for more guidance on how to choose the best warehouse solution for your needs.

2. Collecting the data: event ingestion

Once you have a centralized store for your data, you need to start collecting it. Ingestion is a massive topic when it comes to getting data into your warehouse, so in this guide we're going to focus on what you need in order to replace GA4 for web and mobile analytics.

To build web and mobile analytics, you need to capture user behavior data, which comes in the form of both *events* and *user traits*. **Events** describe a user action, like making a purchase, while **traits** describe the individual user who performed that action.

As modern data teams know all too well, Google Analytics is a notorious silo for this kind of behavioral data and, even though you can technically export the data through ETL jobs or, with GA4 to Google products, it still requires a ton of modeling to be usable...and that's before you even try to figure out how to join user identities with other kinds of customer data from other sources.

All of these problems are made worse when additional scripts from other analytics tools are added to the equation, creating multiple different dedicated ingestion pipelines that feed their own silos (yikes!).

A single source of truth for data ingestion

Data teams building analytics on their warehouse leverage a dedicated tool like RudderStack for behavioral data ingestion. With RudderStack, you can instrument your site with one single SDK, capturing event and user identification payloads one time, then syndicating that single version of the truth both to your warehouse as well as other downstream SaaS analytics tools (more on SaaS integrations below).

Data teams rejoice: a single SDK means no more bloated scripts or sequencing or dealing with API changes and broken pipelines. RudderStack's approach also helps when a team wants to switch tools and add a new tool to the stack. You won't need to touch the codebase to delete or add any new SDKs. Instead, you can simply remove the tool or integrate the new tool via the event streaming tool's dashboard.

When it comes to analytics on your warehouse, RudderStack payloads are automatically flattened and loaded into specific tables according to a standardized schema, meaning your analytics team already has a running start when it comes to creating joins and modeling data.

Here are a few specific features of RudderStack's event streaming pipeline that our users find particularly helpful for building rich behavioral analytics:

- Configurable session tracking data in each payload
- Configurable event properties and user traits
- Ad blocker resilience
- Automatic tracking of anonymous users with a unique id
- Multiple methods for identifying users and assigning traits to users and groups, even across domains

3. Prepping the data for analysis: data modeling

Once your behavioral data ingestion pipeline has pulled in events from your websites and apps, the next step is to make sense of all that data.

Data modeling is all about organizing, transforming, and grouping the events to answer the questions teams formerly tried to answer with GA4. For example, say your marketing team wants to understand which channels drive the most content views (i.e., organic search vs paid social). You'll need to count distinct pageviews for each relevant URL, then group them by referring domain.

It's worth noting that RudderStack's event schemas are specifically designed to make this kind of modeling much easier. In order to perform the above analysis, you'd count pageviews by distinct `anonymousId` (which represents a single user, anonymous or known) to get a unique pageview number, then use the `referrer` value to group pageviews by the site that referred them. To maintain continuity with Google Analytics' default channel grouping, you would assign a channel name to a group of URLs (i.e., Twitter, LinkedIn and Facebook could be grouped into a channel called 'Social').

Once the data is modeled, it's easier to extract value from it, whether in the form of dashboards,reports or as a base for predictive or prescriptive analytics.

Many data teams use SQL or Python to model their data, but in the last few years tools like dbt have made the modeling workflow much better for data teams.

What is dbt?

One of the most popular data modeling tools is [dbt](dbt) built by dbtlabs. dbt (data build tool) enables data analytics engineers to transform data in their warehouses by simply writing SQL select statements. dbt takes those SQL codes and runs them against your data warehouse to create tables and views.

dbt enables data engineers to work like software engineers with version control, continuous integration, and testing built-in. Because dbt is SQL-based, it is straightforward to get started with. Plus, dbt is open-source and has a very active community on Slack.

RudderStack offers multiple dbt models that, when run on event data from RudderStack, can scaffold nearly complete data models for key web analytics use cases:

- Sessionization
- Customer journey mapping
- Identity resolution
- Web analytics

Check out the [knowledge base article](knowledge-base-article) to learn more about the dbt models and how to use them to build analytics

4. Making use of the data: visualization tools

Data modeling is not complete without visualization. Now that you've transformed the data in the warehouse into rich datasets, the next step is to feed it into a visualization tool that will allow you to visually represent those datasets in the form of charts, maps, graphs, or images in order to draw valuable insights from them.

There are a lot of warehouse-focused visualization tools in the market. Some of the popular ones are Tableau, Metabase, Sigma and Hex. However, before you pick a tool, you should consider things like ease of use, learning curve, support for your choice of data warehouse, level of flexibility, customization options, and if it fits your use case.

Tableau
Tableau is often recognized as the undisputed king of data visualization software and with good cause. Due to its ease of use and capacity to generate interactive visuals well beyond standard BI solutions. It is especially well suited to dealing with large and rapidly changing datasets.

Hex
[Hex](Hex) is a modern platform for data science and analytics. From ad-hoc analyses, to beautiful dashboards, to complex apps, Hex streamlines the entire analytics workflow, so your team can focus on generating insight, driving decisions, and moving things forward.

Metabase
Metabase is an open-source business intelligence tool that lets you create charts and dashboards using datasets from your data warehouse. Although SQL is not required to produce visualizations,

Metabase does allow SQL for sophisticated customization. Its simplicity and ease of use are the top reasons why many users love it.

Sigma
Sigma is a cloud analytics platform that uses a familiar spreadsheet interface to give business users instant access to explore and get insights from their cloud data warehouse.

Bonus: powering SaaS analytics (and the rest of your stack) with the same event stream

While this guide focuses on how to migrate off of GA4 and use your data warehouse as the central source of truth for analytics, the reality for various teams like product and marketing is that they still benefit from the use of SaaS analytics tools for specific use cases.

One of the significant benefits of a tool like RudderStack is that, along with the ability to send events and user profiles directly to your data warehouse, **the deep integration library also allows you to send the data to over 180 additional tools, including analytics platforms like Mixpanel, Matomo and...yes, GA4!**

User identification events can also be sent to a variety of other downstream tools to create and update leads, which is extremely useful for sales and customer support teams.

APPENDIX

Choosing a cloud data warehouse

Here are key considerations for teams choosing a data warehouse:

- **What type of data you will be storing**: Will it be structured data (data that fits well into rows and columns) or semi-structured data (such as emails, social media posts, or geographical data)?
- **The quantity of data you plan to store in the data warehouse**: For most use cases, there's no need to worry about this because any data warehouse is typically able to store massive amounts of data without additional costs.
- **How easily can the data warehouse scale**: Are there enough storage and compute resources to process data in times of demand or peak season without affecting performance?
- **Management requirements**: Are you willing to dedicate engineering time to vacuuming, resizing, and monitoring the cluster to ensure performance remains strong? For smaller teams, it would make sense to have a fully managed, self-optimized data warehouse, so your engineers can focus on building your products. However, manually managing a warehouse means you have more flexibility and control and can optimize it precisely for your company's needs.

- **How much it costs**: A data warehouse's pricing structure is based on a mix of storage capacity, run time, and queries. If you frequently execute queries on your data, you should opt for a solution with a reduced compute cost.

Here's a brief overview of the unique characteristics to keep in mind when choosing a cloud data warehouse solution:

Redshift
Redshift is a data warehouse owned by Amazon Web Services (AWS). Redshift is a relational data warehouse and, therefore, accepts only structured data types. Redshift requires some sort of management in the sense that, in times of high demand, if you need to scale, then you need to handle that manually by adding new nodes. Usage costs a minimum of $0.25 based on the type and number of nodes in your cluster. In general, Redshift may be an ideal choice for organizations that have already invested in AWS tooling and deployment for seamless integrations with other AWS offerings.

Snowflake
The Snowflake Data Cloud has a columnar database engine capability, which means that it can handle both structured and semi-structured data, such as JSON and XML. It automatically scales up or down depending on demand and is fully managed with automated administration and maintenance. In addition, it has a decoupled architecture that allows for computing and storage to scale separately, with data storage provided on the user's cloud provider of choice (Google Cloud, AWS, or Azure). Snowflake's pricing is based on the volume of data you store in Snowflake and the compute time you use. So, you can turn off compute resources when they're not in use. If you've already been making use of any of the three cloud storages mentioned and wouldn't want to switch, then Snowflake may be an ideal choice for you.

BigQuery
BigQuery is a serverless, highly scalable, and cost-effective multi-cloud data warehouse built by Google Cloud. It is based on Google's internal column-based data processing technology, "Dremel", and therefore is able to handle both structured and semi-structured data. It is also a fully managed data warehouse and can automatically allocate computing and storage resources as you need them.

BigQuery has two pricing models: Pay per number of bytes processed per query or pay a flat rate by purchasing dedicated virtual CPUs for a certain period of time. One key differentiator of BigQuery is its integration with BigQuery ML. So, if you're looking to build machine learning models on top of your data warehouse to aid predictions, that may be a good reason to go for BigQuery.

rudderstack™

RudderStack is the warehouse-first, customer data platform built for developers. We take a new approach to building and operating your customer data infrastructure, making it easy to collect, unify, transform, and store customer data as well as securely route it to a wide range of marketing, analytics, sales, and product tools. Over 18,000 sites and apps run RudderStack including Crate & Barrel, Acorns, Hinge, Stripe, Allbirds, and more.

🔗 www.rudderstack.com

🐦 @rudderstack