

How Firefly Health built a self-service data function

As a former software engineer, Jacob Mulligan thought about data pipelines the same way he thought about any other product development lifecycle—through the lens of speed and reliability. The advent of cloud warehousing and off-the-shelf data loaders improved analytics velocity, but didn't provide a way to ensure quality and transparency—two things critical for his embedded team to stay aligned. This is the story of how Firefly Health expanded their data stack to fuel product growth, support business operations, and foster a culture of data trust.

As developer #1 at PillPack, a healthcare startup, Jacob Mulligan was accustomed to using data to create better patient experiences. He stepped into the role of data analyst #1 at Firefly Health with the intention to build an analytics team that was not only rewarding to be a part of, but also have a significant impact on the business.

But it wouldn't be easy. Jacob's previous work on data teams was heavily reliant on a team of engineers responsible for pulling, cleaning, and transforming data for analysts. At Firefly Health, Jacob would be a team of one. This felt foreign, but empowering. *"In my previous role an analyst would make a request, the request would drop into an engineer's backlog, and three weeks later a new dimension would appear, often lacking any understanding of the original business context."*

Jacob knew if he had the tools to access and transform data on his terms, he could really make a difference. The advent of the modern data stack, and cloud warehousing specifically, presented that opportunity. As he describes it, *"Cloud warehousing and the data standardization provided by modern data loaders re-defined the role of the analyst. Tasks only accessible to engineers could be performed by analysts without leaving the warehouse."* This would lay the foundation for the high-leverage team he envisioned.

Why Firefly Health needed a new way to manage data transformations

With Snowflake, Fivetran, and Looker in place, Jacob found he could operate as a team of one.

The tradeoff?

A dramatic increase in the volume of low-leverage tasks he faced on a daily basis.

Cleaning was tedious, testing was ad-hoc, and "quick questions" from stakeholders could consume entire afternoons. The new stack worked, but the "bubble gum and tape" method he crafted wouldn't scale. *"To expect anyone new to understand the pipeline I had created, would have been tough. It was just this mysterious analytics server managed by me. I really needed a way to ensure consistency as the team grew."*

1. Data must inform product direction, not just report on it

"One of our core differentiators is how we leverage data to optimize and personalize an otherwise intensely complex and siloed healthcare system. Rapid, reliable, and accessible data is at the center of our ability to deliver value." To execute on this vision, Jacob needed to automate low-leverage tasks that would free his team to focus on more complex business questions. *"Modeling in Looker alone wasn't cutting it; it was hard to maintain, hard to test, and hard to extend models."* If the analytics team was ever going to become a true business partner – informing actions vs. just reporting on results – Jacob needed to be able to automate massive amounts of low-leverage tasks.

2. Data engineering must be self-serve for analysts. Data insights must be self-serve for the entire organization.

Jacob wanted to remove data-access barriers for analysts, enabling them to double as engineers. But he also wanted data consumers to double as analysts – or at least, self-serve answers to questions like, "What does this number mean?" and "How are we defining a day?" before they filed a ticket. *"I needed something that would enable us to spend more time delivering business insights than unwinding broken scripts, or investigating irregular dashboards."* He continued, "My goal was for anyone with a data question to self-serve answers through Looker, or use documentation to answer the question themselves."

3. Data should be consistent and trustworthy across all views

Jacob understood the value of testing, but the best practices he knew didn't quite translate to data use cases. The methods available were tedious, restrictive, and as a result, ad-hoc. *"I really wasn't able to test with the same rigor I was used to as a developer."* As his team grew into an embedded model, consistency across use cases would be critical. Every time data is wrong, analytics teams lose trust. If Jacob was going to build confidence in the data products his team was creating, he needed a reliable testing framework.

Introducing dbt

As a former developer, Jacob immediately recognized the value of dbt Cloud's analytics engineering workflow. Job scheduling, reusable code, and continuous delivery help him increase analytic velocity. Regimented testing, code review, and integrated documentation provide the "reliability" element reminiscent of his developer days. *"dbt Cloud solves all these problems for me. It makes me feel like I'm spending more time delivering business insights than managing infrastructure, trying to pinpoint errors in formatting, or deciphering why a job failed."*

dbt Cloud helped Jacob create a high-leverage data team that excelled in several areas:

Increased speed and productivity with reusable, modular elements: *"In the 12 months before dbt, when we were using our old analytics server/Crontab/Python setup, I created 4 models. In the ~2 months since launching dbt, we've created 15."*

dbt helped the team increase analytic velocity without sacrificing quality. As Jacob describes it, *"Something that got me very excited about dbt was the concept of packages. Leveraging work from the open source community to work more efficiently with data sources like Salesforce and Segment helps me answer questions that are notoriously difficult for almost every organization. We used the Segment dbt package to create an attribution model that combined with session data with Google Ads and Skaffed call logs. For the first time, we had a remarkably granular understanding of the true cost of customer acquisition."*



Reduced risk of error through routine testing: *"We ship every model with tests now. Prior to dbt... we didn't even test our python transform scripts."*

Routine testing had been a struggle for Jacob, but if he was going to earn the trust of business users, it was a problem that needed to be solved. *"With dbt we get tests built in which is very exciting. Before dbt I had a Crontab, a bunch of Jupyter notebooks running on a scheduler, Cronitor for monitoring, and a task in Salesforce to schedule SQL queries. It worked, but I really wasn't able to test the way I needed to. Observability was near impossible without SSHing onto a server, and development was slow."*

Increased cross-org collaboration and transparency: *"Our Looker dashboards serve insights to more than 50 business users who access the system on a daily basis."*

dbt enables a culture of cross-organization data literacy so data teams can focus on high-impact work. This was important to Jacob, so he created a way for self-serve analytics to be more inviting and accessible. *"Most dbt datasets ship with a short video to help data consumers understand how to perform and expand their analysis."* Automated documentation also speeds time insight for users curious about data lineage.



Driving organizational alignment by matching metrics to company values

Firefly Health has three core value drivers which map to several business KPIs. dbt increases accessibility of the data needed to track progress against each. *"dbt allows us to set, and track metrics-driven goals for almost everything we do. Every metric we track aligns to our core values, and is presented in that context to all business users."*

Pillar 1: Crazy convenient care and effective virtual-first medical care across the spectrum

Firefly removes barriers to care by making telehealth onboarding and activation easier. *"Our goal is to make the patient experience as easy and accessible that 100% of new members complete an appointment or health assessment with the team of professionals we match them to."*

How the data team helps: *"We strive to reply to 100% of new patients requesting care in under 20 minutes, and all clinically urgent cases in under 15 minutes. dbt helps us create models we use to track and analyze these response times."*

Pillar 2: Intelligent and effective navigation through every corner of care

Firefly Health empowers patients to be more proactive about their health. But they've also created tools to help employers prioritize workplace health and safety. *"Our Return to Work service enables employees to request daily health assessments directly through the platform. Employees reporting illness receive custom care plans within 30 minutes. Rapid response is crucial here, but quality of care can't be overlooked. We use external calibration data to quickly source the right specialist, and match by proximity, price, availability, and quality of care."*

How the data team helps: As providers reach capacity, searching care plans becomes more difficult. *"Routine testing in dbt given us confidence that the provider we've selected and the care plan assigned is correct."* dbt also plays a role in how employers access that information. *"Employees can check who is and isn't ready to resume work, so they can plan backfill, or reach out to check that they're ok."*

Pillar 3: New alignment among members, sponsor, and Firefly

Firefly Health creates a weekly report that shows the ER admission rate of their patient population. This benchmark is widely used by healthcare professionals to measure effectiveness of preventative care, though real-time access to these insights was usually difficult to achieve.

How the data team helps: *"dbt powers the dashboards we use to quickly compare our rates to national averages. Today, we can see we're out-performing the market, but our goal is to reduce that metric by a further 20%. dbt will help us track progress."*



Location
Woburn, MA

Industry
Telehealth

Founded
2016

Results

20x

Power model development

In the 12 months prior to dbt, the Firefly team created 4 new data models. Two months into their use of dbt, they've created 15.

50+

Business users served

The data team ships "how-to" videos with every new view to empower all business users to optimize their data.

100%

Models shipped with tests

dbt enables this three-person data team to maintain high-velocity analytics without sacrificing consistent testing.