**Don’t Be “Data-Driven.” Be Analytics-Driven.**

Here are some practical steps you can take to get there.



Look at this poor, confused executive trying to make decisions without analytics support. Mountains? Clocks? Chess? What? 🤔 [image from freepik]

“We are a data-driven company.”

I hear you groaning at your screen already. Let’s talk about what this phrase means these days, what we as data scientists and analysts really want it to mean, and how to bridge that gap. Read this before you respond to your next ad-hoc request.

**😤 What does “data-driven” mean these days, and why is it so dangerous?**

It’s been sexy in the last couple years to be a self-proclaimed “data-driven” org. And by this, organizations typically mean: **we make our decisions using data**. Now don’t get me wrong — it’s a great idea to support decision-making processes with data. But there’s something insidious in this statement:**it doesn’t recognize the hard work of data scientists/analysts.** The focus is on the data, not the analyses, which is to say:

We have so much data, we just need to see it and use it to make decisions. Analysts and data scientists just need to get the data for us.

*Data-driven orgs: analysts fetch data.  
Analytics-driven orgs: analysts find answers.*

*Data-driven orgs: hire more and more analysts/DS as “hands”.  
Analytics-driven orgs: invest in infrastructure, tooling, and education.*

*Data-driven: “Can you pull these numbers…?”  
Analytics-driven: “Can you help me think about…?”*

*Data-driven: Finds data that justifies management decisions  
Analytics-driven: generate insights that tell a compelling story management can act on*

I’d say it’s about time to organizations to make the shift from data-driven to analytics-driven.

**🤔 What to do about it**

Unfortunately, there’s no silver bullet here — successfully becoming “analytics-driven” will hinge on how your organization views its analytics org. Are they thought partners or subordinates? SQL monkeys or scientists?

At the end of the day, stakeholders need to get in the habit of asking business-level questions of analytics, not having data requests.

[For a great read on this subject, check out Pedram Navid’s article, [Building the Modern Data Stack](https://pedram.substack.com/p/modern-data-team?r=12stp&utm_campaign=post&utm_medium=web&utm_source=twitter" \t "_blank).]

That said, short of having a long talk with cross-functional leadership, I’ve found that a few procedural changes can help support this shift:

1. **📝 Document your ad-hoc work.**
2. **🍽 Document your data.**
3. **📚 Set up a central place to find your documentation.**

While these won’t directly change corporate culture, they’ll enable you to reduce the load of your ad-hoc requests by making your past work searchable and, dare I say it, self-serviceable.

**📝 Document [and own] your ad-hoc work**

As a data scientist or analyst, you are the person most familiar with data, so naturally data-related requests will come to you. My default response has always been to drop a query/dashboard on stakeholders ASAP, which I’d promptly forget about. But then I’d have to repeat the work the next time I was asked. Plus, this sort of ad-hoc response just reinforces the perception that your relationship with decision-makers is purely transactional.



Give your ad-hoc work a little more weight, and keep it alongside your more robust analyses. [Image from freepik]

**Instead, write up a doc**. Start with the question:

“What question are you trying to answer?”

Then, do your work masterfully and reproducibly — working with data is a science, after all, and should be treated as one.

This’ll have two benefits:

1. **It’ll reduce your load.** You’ll have this for later in case you get asked the same question again (you will).
2. **It’ll put more weight behind your work.** You’ll be forced to give the problem more serious thought. This will produce a better analysis, reaffirm your value as a thought partner, and improve the quality of the decision-making. Analytics-driven here we come. 🚀

**🚗 Document your data [tables + transformations]**



A file cabinet. Drawers are kind of like schemas, right? 🙄 [Image from freepik]

Documenting your work is a fantastic first step, but your analysis is only good as the data you use to make it. If you want to have confidence in your analytics work (and if you want others to have confidence in your work as well), you should deeply understand and document the tables and transformations used in your final analysis. There are two parts to this:

1. **Document your data.**At the least, keeping a central doc somewhere for your team to access can be a great start. If you’re looking for a lightweight hosted solution, [Hyperquery](https://hyperquery.ai/?utm_source=medium&utm_medium=organic-content&utm_campaign=2021-06-07-data-driven" \t "_blank) provides a query-first note-taking tool with a data catalog built in, so you can put both your SQL work and your table docs in one place. Data cataloging companies (like Alation, Collibra, or data.world) or open-source “data discovery” tools (Datahub, Amundsen, metacat, metamapper, Magda) could fit the bill here, but they tend to be a bit heavy to set up/maintain.
2. **For your transformations, use a tool like dbt or dataform to document and version-control your transformations.**Transformations deserve to be visible, and version-controlled, not shoved into views (though this can work stably for quite a while). [dbt](https://www.getdbt.com/" \t "_blank) is a fantastic tool to build, document, and version-control this sort of work.

**📚 Set up a central place to find your docs**



[Image from freepik]

Now that I’ve convinced you to document things, there’s a glaring question left:

Where do we put these docs?

You’ll need to consolidate all this documentation somewhere. I’ve worked with/at companies who have used Confluence, Notion, or even Github to manage this, which can work in a pinch. At Airbnb, we used [Knowledge Repo](https://github.com/airbnb/knowledge-repo" \t "_blank) in conjunction with [Dataportal](https://medium.com/airbnb-engineering/democratizing-data-at-airbnb-852d76c51770) to store these as git-tracked publications, but this always felt a bit heavy for ad-hoc work.

[Hyperquery](https://hyperquery.ai/?utm_source=medium&utm_medium=organic-content&utm_campaign=2021-06-07-data-driven) works really well for organizing work in this way — Hyperquery combines a Notion-/Confluence-style note-taking environment with executable query blocks, so you can start your SQL work there, share it easily with the rest of your org, and later, use your work as a starting point to build up a full-fledged set of docs.

But at the end of the day, use tools that work for you. The important thing is that your work needs to be visible, accessible, and searchable by the rest of your team and your company.

**Final thoughts**

Analytics organizations don’t exist to write SQL. We are here to problem solve. While it’s going to be a herculean feat to try to force stakeholders to view it this way, let’s at least document our data work so the problem-solving aspects of our work are more visible. Let’s show the world why analytics-driven orgs can be so powerful! 🙌