**How to Be an Effective Data Science Manager**

**Lessons and reflections from a Data Science & Analytics leader**

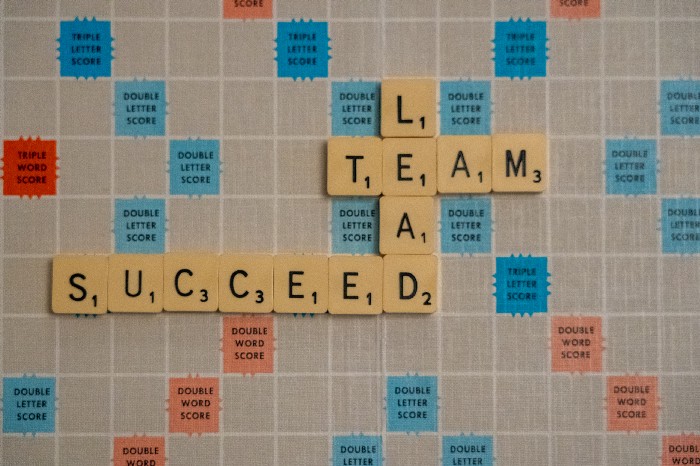


Photo by [Nick Fewings](https://unsplash.com/@jannerboy62?utm_source=medium&utm_medium=referral) on [Unsplash](https://unsplash.com?utm_source=medium&utm_medium=referral)

Being a data science manager is an aspiration for many data scientists. I love watching the Olympics. We see various countries’ sporting heroes battling it out for the prized gold medal. Unfortunately, the chance to compete in the Olympics is for selected and gifted few. For the rest of…

However, data science is at a challenging time. We’re past the hype of data science. We are in the middle of two towering challenges:

* *“How do we productionise it?”*: more than 90% of data science solutions still struggle to be productionised. Out of the productionised model, business adoption is still low relative to the efforts spent.
* *“Show me the money!”*: funding is scarce, and everyone’s competing for the CFO. Executives have seen the benefits of data analytics and reporting to achieve data-driven decisions. But, many are not convinced of the benefits of data science projects.

Wait, I didn’t sign up for this. I thought being a Data Science Manager was about leading a team of Data Scientists. You know, lead, and motivate the team to deliver. This sounds too hard.

The problem is that data science is still a maturing profession. In many companies I have worked in, it’s usually a greenfield or brownfield context. Everyone knows what the well-established marketing team does. However, not everyone knows that data science does. Companies have operated well for years without data science.

It’s not enough to be a “good” data science manager; we need effective data science managers. I have had years of experience practicing this dark art. I have learned my painful lessons, and I hope it will spare some of yours. Being effective means achieving our vision, a data science manager that leads a successful data team that delivers valuable predictive models and is loved by the organisation.

A bit about myself, I’m a qualified actuary (sometimes known as the original data scientist) who has been journeying a long career in data science and analytics. I have held many data leadership roles for nearly a decade for teams of various sizes, geographies, and styles. I’m at a stage where I could contribute my learnings back to the community and help future data leaders.

I like to develop ways for people to remember this content easily. Therefore, we will use the PPT acronym. No, it’s not our favourite PowerPoint. It’s **P**eople, **P**rocess, and **T**echnology.

**People**



You Dont Understand Logan Paul GIF By SHOWTIME Sports. 2021. Retrieved from [Giphy](https://media.giphy.com/media/CCbudQwZBz2DtSuEvT/giphy.gif)

When I first became a data science manager, I knew I had to work with people — but I didn’t realize it was that complex and challenging. With coding, we get the error feedbacks, and we re-run. With people, feedbacks are scarce, and a “re-run” is not straightforward.

If you are like me, we spent decades learning and upskilling our data science skills. This includes a business degree, an actuarial qualification, and technology certifications.

However, a must-have requirement for a Data Science Manager is people skills. Well, where do we learn people skills? Would an MBA degree suffice? It might help, but we really learn people skills through the school of life. Personally, I have tried various ways, some are successful, and some are not. However, all of them are valuable experiences that brought me to where I am in my data leadership journey.

We have been successful data scientists because we have delivered excellent data science solutions. You might be rewarded as a Senior Data Scientist already, or you might already be appointed a Data Science Manager. However, what got you here won’t get you where you want to be.

As a Data Science Manager, you will work with many people. For simplicity, we will categorise them into internal and external people.

*Internal People*

Internal people cover all the people within your team, immediate or neighbouring departments. This includes your direct reports, your data colleagues, the broader technology team, or the other department team members (depending on where the data science team sits).

To deploy data science solutions successfully, you need all of their support. We need to work with the data engineers to source and productionise our data. The other technology team needs to be across of your deployment plans. All the other department team members needs to be aware of your team’s capabilities and deliveries.

Internal people expect collaboration. Therefore, my tip is to start by understanding their existing processes and align them with the Data Science team. This includes broader strategy, project prioritization, and people dependencies. There is no need to create your own world because it leads to further silos. While aspects of a Data Science team are in a start-up context, we need to leverage the existing culture to pave the way.

Another tip is to catch up with them regularly. Sometimes, we’re all about business and only meet people when there is a project. However, it’s worth understanding everything that is happening around you. Sometimes you might be able to help directly, and sometimes you could be a good listening ear.

When I started as a data science manager, I spent too much time with other business leaders and stakeholders. I thought that my high-performing team was self-motivated individual contributors. I was not wrong, but they needed more transparency and guidance from me, particularly as many data science projects have ambiguous scopes. When I received this feedback through the employee survey, I rearranged my calendar and spent more time with the team individually and collectively. Since then, the team is not just happier, but they are more innovative as my partners in solving business problems together.

*External People*

External people cover all the people outside your team. This includes your direct stakeholders and the wider organisation.

For direct stakeholders, you need to keep a good relationship. The team has likely delivered successful work previously, and there’s a good amount of trust. Therefore, my tip is to nurture this relationship and build closer relationships between the teams.

For the wider organisation, this is probably very vague — and it is. As a new (or “unknown”) team on the block, the wider organisations don’t know about your shiny team and what it could do them. Our objective is to identify them. My tip is to ask for referrals from your direct stakeholders and colleagues. Once they have been identified, don’t just sell data science to them. You need to understand what they do, their aspiration, and how you may (or may not) be able to help using data science.

When I started as a data science manager, I spent too much time with the stakeholders I already knew. That was great until my executive asked how I was embedding data science in all the teams in the organisation. Since then, I have made conscious efforts to get to know new stakeholders. Although uncomfortable initially, they were super appreciative that the “cool” tech team reached out. Nowadays, I see my role also covers broader data science and analytics evangelisation.

**Process**



GIF By Lush Interior Design. 2021. Retrieved from [Giphy](https://media.giphy.com/media/quuIo0rCMQK6KHMrJD/giphy.gif)

Let me start with this LinkedIn feed that I came across recently:

Startups can take a decade to build.

Year 1: stay in the game

Year 2: find product market fit

Year 3: find paying customers

Year 4: build a team and delegate

Year 5: build processes to scale

Year 6–10: just stay in the game

Be that 10-year overnight success

Source: Andrew Gazdecki

As mentioned before, Data Science is still an emerging profession. And we need to treat it like a start-up.

When the Data Science Manager position exists, the company is typically three years plus in the journey. If they are already 10 years plus, and you want to grow the team further, then building a process to scale remains the right action.

So, what processes do we need? I will cover three main focus areas:

*People process*

In Australia, coffee is one of our religions. If you want to meet new people, buy them a coffee. If you want to hang out with your team, buy them a coffee. If you want to make customers happy, buy them a coffee.

The point is people don’t just magically work together. In particular, the hybrid working arrangement brings challenges to team collaboration.

An effective Data Science Manager must explore and implement healthy habits to achieve the team’s potential. Lao Tzu, the writer of Art of War, eloquently advises us to:

Watch your thoughts; they become habits. Watch your habits; they become character. Watch your character; it becomes your destiny. “ -**Lao Tzu**”

For example, I am a big fan of OKR (stands for “Objective and Key Results”) for agile goal setting. It starts a quarterly conversation on what outcomes we are trying to achieve. Then, we have weekly catch-up meetings on what’s working or not. At the end of the quarter, we will have a retrospective which will feed the next OKR. In addition, I also encourage ongoing coaching and feedback across my team.

The above is an example of a people process for delivery, but we also need a people process for social. Yes, we need to be intentional about our social life. After all, we spent a lot of hours at work.

When I started as a data science manager, creating a people process felt like micro-managing. As a “macro” people manager, I let meetings and “processes” run spontaneously, which worked fine when the team or workload was small. As we grew bigger, things got more chaotic, and the team struggled to keep up. That’s when the team asked to create a people process to manage our time and people better. And it did make a significant difference where people were less stressed out because they knew the key people and steps needed towards the outcomes.

*Deployment process*

When the early days of the data science team, deployment was quite ad-hoc. On average, I have seen data science teams generate between 1 to 3 productionised models in their first three years. With a small team, it’s usually a simple refresh, recalibrate, or troubleshoot.

Imagine if we want to have 20 or 50 models in production. We either need a massive army of data scientists or a robust process. From my experience, your CFO will choose the latter. Development time also needs to be factored in, which would be time-consuming.

The success of the data science team will be determined by your team’s ability to deliver models fast and reliably. Based on my experience, MLOps is the way to go, which is the practice of productionise machine learning artifacts in a scalable and reliable manner. As the name suggests, it involves applying and adapting Software Engineering and DevOps principles.

However, even MLOps is still a greenfield area. While there are principles and examples, you need to design something that works for the team. As a tip, start simple and see whether you can leverage existing market tools. It is not a competition for building grandeur (also known as complex) solutions, but it is about achieving business value — and how we do it efficiently, consistently, and reliably.

In the past, I drove deployment agility through frequent “war rooms”. Soren Kierkegaard once said that “progress can be understood backwards; but it must be lived forward”. This echoes true when our team fell behind schedule or we just wanted to go above-and-beyond expectations. Fast and collaborative communications worked to iron out practical improvements to our deployment process one at a time. For example, we shaved our data deployment process from 4 weeks to 1 to 2 weeks — and we’re still optimistic about improving it further!

*Sales process*

Whoa, I’m a Data Scientist; I’m not a salesperson. Yes, you aren’t a salesperson. But you are responsible for the team’s pipeline. If you are early in the maturity, the work pipeline looks pretty empty. And we could learn a few things from our sales friends that do this for a living.

From the above people section, we should have identified our direct stakeholders and several potential ones. For your existing client, you need to take a proactive approach. A common mistake is to wait for business requests. The issue is the requests would be minuscule such as data refresh or some sample data. We want data science to drive significant business value — hence you need to drive the data science team as a business partner, not just a service provider.

As a start, you might ask to be involved in your stakeholders’ regular planning and strategy sessions. This brings you up to speed with their contexts, challenges, and aspirations. From there, you would be in a better position to identify where data science might help. Then, start building regular partnering sessions with them to build the requirements.

In the past, I have adopted a Kanban approach to a good degree of success. The board is split into key opportunity stages, such as prospect, connect, qualify, propose, objection, close, and referral. It created a very visual representation, which worked well for my focus and to show others.

**Technology**



Black Panther Magic GIF. 2019. Retrieved from [Giphy](https://media.giphy.com/media/26SSWjcPAKCyXdCXgh/giphy.gif)

As a Data Science Manager, you are responsible for the why, what, and how. For the how part, technology plays a huge enablement role.

There are two main schools of thought: to build or to buy. To build means that your team codes the solutions from scratch, involving various scripts such as Python and SQL deployed through containerisation. To buy means that your team procures a solution that is pre-configured and ready-to-use for the given purpose.

The total cost comparison for both options is mixed. From a dollar cost point of view, build is “cheaper” as there are many open-source options. However, we absorb the build times, maintenance operations, and additional headcounts. The buy option costs more on subscription dollars. However, it is ready and continuously maintained plus improved by the vendors.

In an effective data science team, the technology issue surfaces from a capability rather than costs. Your high-performing team is eager to showcase their coding capability in building customised and well-made data products.

As mentioned above, data science is a business profession that applies our quantitative, technology, and commercial skills. Therefore the team’s success is based on its ability to deliver business outcomes at the right time.

In light of this, time might not be our friend and it’s often wise to choose the buy option. Most of these vendor tools allow customisation so you could steer it manually. It also enables the data science team to address more business problems and create bigger organisational impact.

Once the team has delivered results and scales, the technology options remain. By then, you would be aware of the innovation boundaries that require a customised and well-made data solution.

When I started as a data science manager, I spent too much time on best technology comparisons. I wanted to make accurate and “data”-driven decisions. However, these comparisons took months, and people still disagreed with the “best” technology. I realised that we needed to factor organisation and team culture on top of features analysis. From there, we conducted proof-of-values and gathered real organisational feedback that we didn’t have before. This created a more robust business case and allowed us to deliver business value at the same time.

**Moving forward together**

Being an effective data science manager is not easy, but it’s achievable. We have dedicated most of our learnings to the latest mathematical techniques and technologies. Now, we need to adjust our learnings toward the softer skills of people, processes, and technology.

When you are at an early stage or learning to be a data science manager, the temptation is to learn every management course and boil the ocean. However, they usually cover general management skills such as communication and planning. We also need more specific resources in the emerging profession of data science.

Thank you for reading this article! Please add your favourite tips in the comments section below.