DSC 90

Winter 2020

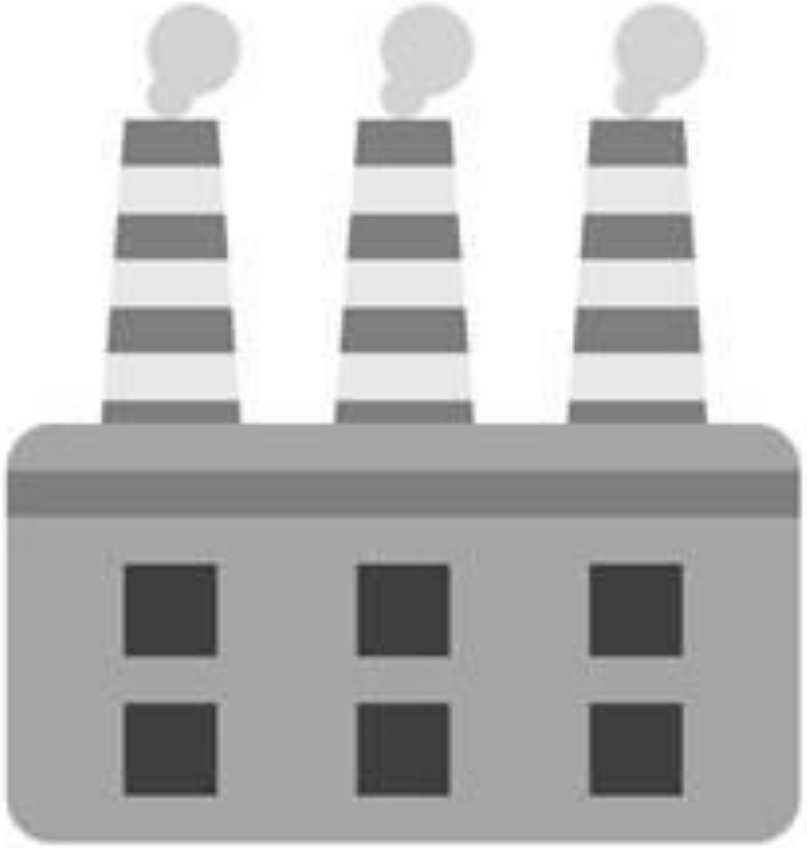
Your First Months on the Job

Your First Few Months on the Job

- What to expect in your first few weeks as a data scientist
- Becoming productive by building relationships and asking questions
- What to do if you're in a bad work environment

When you start working, you will instinctively want to get as much done as possible immediately. Fight that instinct. You need to be sure that you are not just accomplishing tasks but doing them in the right way.

Onboarding at a large organization



Onboarding at a small company

Understanding and setting expectations

Have a meeting with your manager to discuss priorities.

- gives you the knowledge of what you are supposed to be working towards in your job
  - Sometimes: provide analyses to a specific set of stakeholders to help grow a particular part of the business.
  - Other times: make high performing models that help run the business.

Understanding and setting expectations

- It may feel like you should already know what the job expectations are by the job posting and interview process.
  - But things change!

Together, you want to define what success means in your job

- Generally, that is tied to making your team and/or manager successful
  - So you need to understand what problems the team is trying to solve and how performance is evaluated

Understanding and setting expectations

Tools are not Goals!

- Performance usually does not mean "make a machine learning model with 99% accuracy" or "use the newest statistical model in your analysis."
- Common misconception for early career people, because this *is* the goal of most academic research and classes!
- Things like the model's usefulness, level of insight, and maintainability are often more important.

Understanding and setting expectations

Job responsibilities vary, and aren't always communicated clearly.

You might be expected to:

- help with several different projects at once
- drop your work on a moment's notice to help a colleague
- have deliverables on a regular basis
- ignore emails or slack messages in order to achieve your project

Get explicit, frequent feedback.

Knowing your data

You do need to learn about the data science part as well!

Read reports (if they exist) to learn about:

- types of data your company keeps
- some key insights
- tone and style of how you communicate your results

Get access to data.

- what data system?
- what table?

Knowing your data

Start by reading documentation

- Keep an eye out for quirks or know issues
- Learn how the data got to you

Then take a look at a few rows and summary statistics

When you find surprises that aren't documented:

- Talk to the expert on that table.
- Figure out if it is a true issue that needs fixing.

Keep a log of any "gotchas" in the data and a map of where everything lives

Becoming productive

It is totally normal to feel like you are making work harder for your supervisor and coworkers. You are. And that is ok!

- Early on, you're going to be asking more of the company ("can I get access to this, why is this query so slow") than you're getting back (in the form of reports and analyses).
- It will take you longer than you think to get fully productive
- Your employer and boss should expect some ramp-up time
- Focus on positioning yourself to deliver value in the longer term (months, weeks).

Becoming productive

You can still deliver some value early by focusing on simple and entirely descriptive questions

- familiarize yourself with the company's data
- find some of the snags and traps that are waiting
- check-in with your manager and get regular in-process feedback

There are two strategies that can help you become productive more quickly:

- asking questions
- building relationships

Asking questions

One of the biggest things that can hold you back in your career is being afraid to ask questions or say I don’t know.

Data science is such a big field that no one knows even a fraction

But there's also no way to know all the intricacies of your company’s data

Don't spend time stuck!

Asking questions

All questions are not created equal.

Pick the right medium: in person, on a slack channel, in forums, etc.

Doing some research yourself first helps you ask better questions

For lots of questions (or deeper ones) schedule a meeting, and keep a running list

Avoid voicing criticisms veiled as questions, rather genuinely try to understand why things are done the way they are

Building relationships

Engage in some non-technical talk with people

- Set up meetings with people you've never talked to before to get to know them and their work
- It's also normal to message people you'll be working with to ask if you can set up a meeting to get to know each other
- If you work in a large office, ask your manager to help make a list of people you should get to know.
- This is not wasted time - it allows both you and your coworkers to feel more comfortable relying on each other
- Questions are a good way to start a conversation - people like feeling helpful

Once you know a few people, even the largest offices feel less intimidating.

Building relationships

Find out who to go to for specific questions

- Different people for different problems
- It usually isn’t your manager

Meet with all the stakeholders

- If the data science team is under 10 people, try to meet with them all individually
- If there are data engineers or other data people you’ll work with, talk with them

- Do a lot of listening, both in official meetings and in social opportunities like over lunch.
- Meet people who work in data-adjacent areas (which could be everything from engineering to finance to sales ops to marketing analytics) and hear about how they currently do their jobs.
- Don't rush in with "I could do that better" or make commitments early like "we'll build you a machine learning platform to do that;" simply focus on collecting information and thoughts.

Prioritizing Work

Prioritizing Work

Quick tasks directly from stakeholders

- small requests like “make a graph of sales over time.”
- They are often urgent and since they don’t take much time it’s hard to say no.
- But each one is a distraction from more important work and as they build up it’s harder and harder to do productive work.

Long term projects for the business.

- These projects are the core of a data scientist’s job.
- Building dashboards, making long-form analyses, and creating models to put in production all fall into this bucket.
- These tend to be highly important but given that they can take weeks or months they are not always urgent.

Ideas that you personally think have a long-term benefit.

- Might be technical solutions to business problems
- Also includes work that makes you more productive
- No one is asking for this work, but it feels important.