**How Can an Aspiring Data Scientist Find and Work on Real-World Projects?**

**I’ve done some courses online; what next?**



Image by Author via [Canva](http://canva.com)

As an aspiring data scientist, one of the best ways to gain experience and improve your skills is to work on real-world projects. This will not only allow you to apply what you have learned in a practical setting, but it will also allow you to learn new things and develop a…

The truth is almost all of us know the value of real-world projects. You’ll always work on these when you get a job, but sadly most data science positions ask for experience. This puts us in a frustrating loop: you need real-world experience for the first job and your first job for real-world experience. So the question is, how do you break out of this loop?

Before I start giving suggestions, I want to be honest with you.

None of what I’m about to outline is easy. If it’s easy, we would have all [4.8M learners](https://web.archive.org/web/20220531124205/https:/www.coursera.org/learn/machine-learning) of the most popular machine learning course become data scientists. It need not be for everyone. But if you follow the suggestions and put in the work to go beyond the online courses, you can grow as a data scientist.

With that honest conversation, let’s approach this in a step-by-step fashion, assuming you have just finished a [bunch of online courses](https://datagrads.com/start-here/) and wondering what’s next for you.

**Step 0: Join a data science community or meetup group**

You can start by searching for data science meetup groups in your area on [meetup.com](https://www.meetup.com/) or [eventbrite.com](https://www.eventbrite.com/). Alternatively, you can follow the social media page of the nearest local auditorium or co-working space where generally these meetups are hosted. That’s what I did.

I remember a new co-working space called [Hatch Works](https://hatch.lk/) came up in my early days. I would attend every data science meetup hosted there since it was convenient. 99% of the time, the senior professionals in the AI industry would present their work, distil an evolving topic, and even discuss job opportunities in their organization.

I leave most meetup events with fresh project ideas that I want to add to my personal projects, try at work, or at least write about.

Being an active member of your local community/meetup groups:

* puts you into the heart of your local data science industry.
* give you multiple projects ideas that are relevant to your future job
* creates a support system of peers with whom you can collaborate later
* open networking opportunities that may lead to a job

It’s called step zero because you can start on it today…well..*now!*

**Step 1: Try this at the online data science competition websites**

If you have just completed a few online courses, you are not ready for real-world projects. Nobody is.

The data is insufficient to train a machine learning model. Even if you have the data, it is either messy, imbalanced, missing, or has privacy concerns. This has been the AI industry’s [most pressing bottleneck](https://www.oreilly.com/radar/the-quest-for-high-quality-data/) for ages now.

Why not settle for something more beginner-friendly and transition to real-world projects slowly? Check out online data science competition websites like [Kaggle](http://kaggle.com) and [DrivenData](https://www.drivendata.org/).

Pick one of the past competitions based on your existing knowledge. The dataset is generally from actual companies but cleaned and masked to hide private information. They have solutions and guided tutorials, which can help you get used to data pre-processing, analysis, training and tuning machine learning models.

Most beginners directly attempt the existing competitions and are disheartened when they’re nowhere in the competition leaderboard. They start doubting their skills and go back to do a yet-another online course.

Common mistake.

Our purpose in using Kaggle or DrivenData is only to get used to solving machine learning problems. Once you’ve had a go at one or two past competitions, **only then** attempt the live competitions. You never know; the results may surprise you.

**Step 2: Volunteering at AI for good projects**

I’m not a fan of unpaid internships, but contributing to open-source or volunteering for AI for Good projects is something I continue to do.

Many organizations, such as non-profits and government agencies, need data scientists to help them analyze their data. Volunteering your time to work on a project for one of these organizations can be a great way to gain experience and make a positive impact at the same time.

The choice of organization and the cause you want to support entirely depends on you. I have personally worked with [Omdena](https://omdena.com/projects/), where beginner-to-expert AI practitioners work towards solving end-to-end real-world problems. Volunteering was straightforward, with an [application](https://omdena.com/projects/) and a short interview with the Founder.

Apart from AI for good, contributing to open source has several other benefits. They:

* allow you to collaborate with other experienced data scientists and learn from them. This can help you learn new techniques and stay up-to-date with the latest developments in the field.
* are a valuable addition to your resume. Many employers value open-source contributions as testimonials of your abilities and commitment to the field.

You asked for real-world projects — you’ve got it here in its best form.

**Step 3: Build a real-world data portfolio while getting certified**

Learning doesn’t have to stop while you seek real-world projects.

Udacity is my go-to platform when I want to acquire skills and build a portfolio simultaneously. I’ve completed three nano degrees so far, and the absolute best for data scientists is, unsurprisingly, [the data scientist nano degree](https://bit.ly/udacityDS) (thanks for using my Udacity affiliate link, meaning I get a small commission if you choose to purchase at no extra cost to you, but I paid for mine in full.) The projects you’ll work on are from partnered companies sharing their data which is why they require you to sign data usage agreements.

If you’re still a student or find their nano degrees expensive, you might want to check out their [scholarships page](https://bit.ly/udacityscholarships). They regularly run scholarship programs, and chances are you’ll be able to apply for one right now.

Most are surprised when I reveal this: all of their projects included in the nano degrees are open source on [their GitHub](https://github.com/udacity), waiting for you to be completed and add to your data portfolio.

Don’t believe me? Here’s the one for [the data scientist nano degree.](https://github.com/udacity/DSND_Term2)

**Step 4: Stepping into the AI industry with an internship**

End of the day, nothing can replace the experience you get at a proper job. That’s a fact.

Since you’re an absolute beginner, consider starting with a data science internship or taking on freelance or contract work that involves data analysis and visualization. Almost all my freelancing work came from inbound leads who saw my work online and asked if I could help them. These experiences can demonstrate your skills without yet having a job.

Eventually, to convince an employer to hire you, you need to demonstrate your data science skills. It’s as simple as that. Steps 0–3 are mainly for this purpose. I have written an [extensive guide](https://towardsdatascience.com/how-to-secure-your-first-data-science-internship-7bbfd8b87bdc) to help you get your first internship. To keep it short for this article:

* Acquire skills (you’ve done this)
* create at least one demonstrable project (you’ve multiple of this)
* stop applying online ([read why](https://pub.towardsai.net/you-will-never-succeed-if-you-keep-applying-for-jobs-online-1107ccbc9ac9))
* get referred internally (go back to step 0)
* prepare and practice for interviews (I need to write about this in future, but for now, here’s [my favourite channel](https://www.youtube.com/@DataInterviewPro) by

[Emma Ding](https://medium.com/u/1b25d5393c4f?source=post_page-----48bed91af008--------------------------------)

)

* ask for feedback from the recruiters after the interview (they want you to succeed too)
* play the long game (more on this in a bit)

See, I understand it’s overwhelming at the start. The hardest part is getting stuck without a direction. By reading this, you’ve overcome that. And if you genuinely enjoy data science, you’ll be up and running in no time.

**Concluding Thoughts: Play the long game**

Ultimately, being proactive and persistent is the key to finding and working on real-world projects as an aspiring data scientist.

By joining data science meetup groups, participating in online competitions and forums, volunteering your time on impactful problems, and working on personal projects, you can gain valuable experience and build a strong portfolio of work that will help you stand out as a data scientist.

The bottom line of getting a job comes down to understanding that all recruiters are looking for candidates with proven experience, irrelevant of the level.

Even when I interview now, after 4+ years of experience, I have to showcase my ability to lead teams that solve business problems by delivering scalable machine learning systems. This is a journey in which you constantly improve and prove yourself.

Play the long game.

Offering you the job you want only becomes a no-brainer for the recruiter.