✨ Member-only story

# The Portfolio that Got Me a Data Scientist Job

Spoiler alert: It was surprisingly easy (and free) to make

Matt Chapman · Follow
Published in Towards Data Science
10 min read · Mar 24, 2023



Photo by John Schnobrich on Unsplash

Getting a Data Scientist job is hard.

This isn't 2015 anymore: it's not enough to know a few pandas functions and put the words "Big Data" on your résumé. Competition for the top jobs is fierce. On a recent

trawl through the LinkedIn jobs board, I struggled to find a London-based Data Scientist role with less than 100 applicants.

The good news is that this competition is not due to a lack of jobs. Even in 2023, Data Science remains a fast-growing field, and the U.S. Bureau of Labour Statistics has estimated that the number of Data Scientist jobs will grow by 36% between 2021 and 2031 [1].

The challenge, then, is not that there are no Data Science jobs — there are tons! Rather, it's that a huge number of people are trying to break into the industry, making it increasingly difficult to stand out from the crowd and land that lucrative first role.

## The solution: Build a data science portfolio

In the current job market, I am convinced that one of the best ways to differentiate yourself from the competition is through building a personal portfolio of personal Data Science projects to showcase your skills and experience. This is especially important if you don't have loads of commercial experience working in Data Science roles. To borrow the words of Data Scientist Will Stanton [2]:

> *If you don't have any experience as a data scientist, then you absolutely* have *to do independent projects.*

In this article, I'll walk you through the process I followed to create my portfolio of personal projects, and attempt to give you a bit of inspiration for how you can create one for yourself.

If the idea of creating a portfolio feels a bit daunting, don't worry: it is surprisingly easy (and free) to get something up and running, and you don't need any knowledge of web development to start.

## My background (AKA why I created my portfolio)

My journey into Data Science only began relatively recently. My first jobs fresh out of school/university were in fields like Project Management, Sales and Marketing. When I decided to make the switch to the Data Science in 2020, I had very little hands-on experience working with data and didn't have a technical undergraduate degree in a field like Computer Science or Statistics.

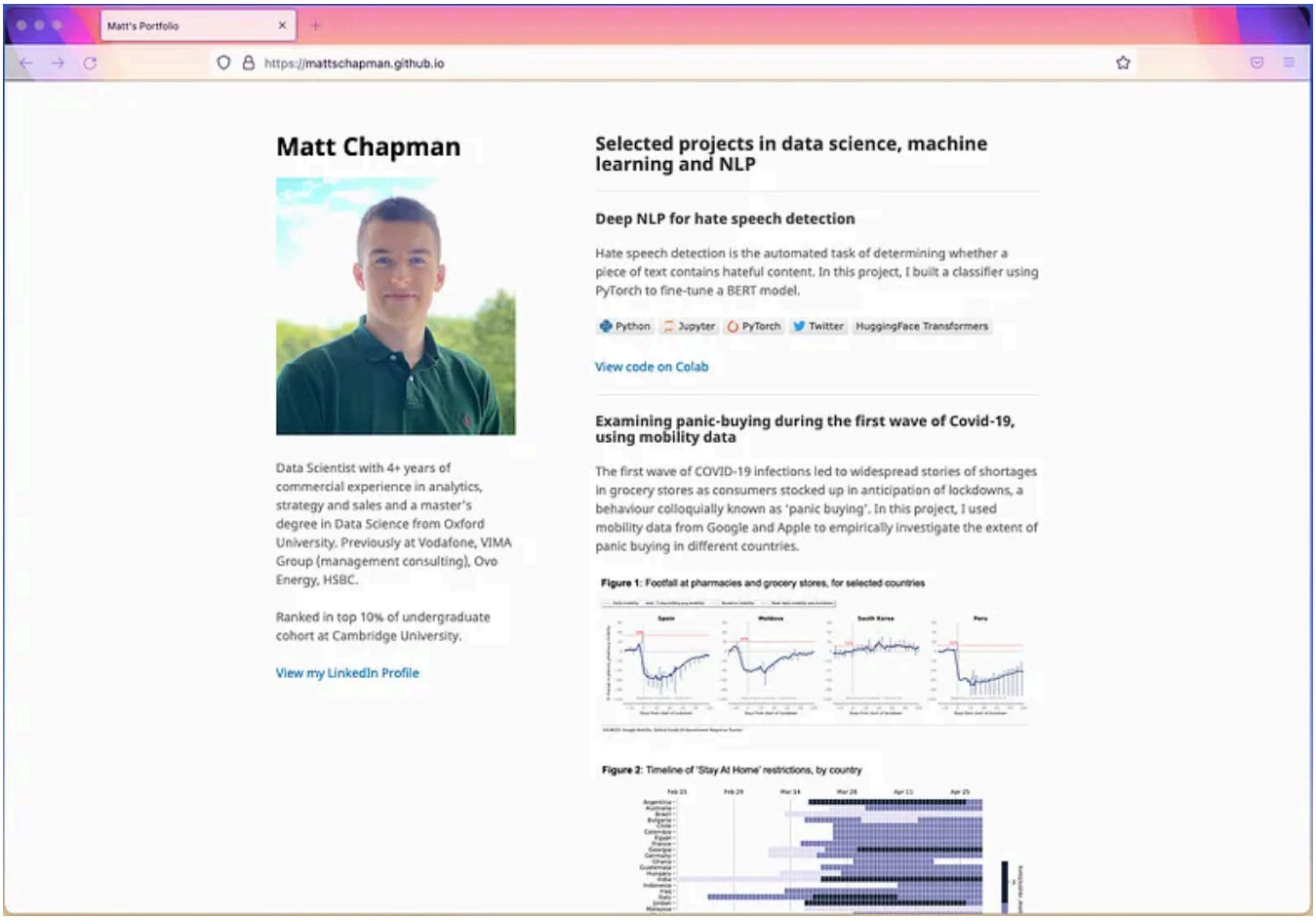
Yikes. That's not a very strong place to start.

In my previous post, I shared how I addressed this by taking a bunch of online courses, studying for a Master's degree in Data Science, and doing a couple of Data Science internships. Along the way, I worked on many different projects and built a broad repertoire of data-related skills.

However, for the first few jobs I applied to, I found that I was struggling to get my foot in the door: it felt like recruiters weren't even taking a second look at my application, **despite me meeting all of the minimum requirements.** At the time, this was incredibly frustrating. I believed that I had a lot to offer, but it seemed like the conversations were being shut down before they even began.

To make matters worse, I learned that recruiters often spend as little as 6–8 seconds reviewing a CV for an over-subscribed job posting [3]. That convinced me that I needed a quicker way of communicating my skills and standing out, especially given that I didn't have reams of experience as a Data Scientist.

To tackle the this problem, I decided to create an online portfolio as a way of showcasing my experience in a visually compelling and quickly digestible format.




My Data Science portfolio: https://mattschapman.github.io/. Image source: author's own.

Then, once I'd built my portfolio, I included a link at the top of my résumé, to make it super easy for recruiters to find:


The heading of my résumé / CV. I liked the idea of having a clickable link or QR code that would take employers straight to my portfolio and help me stand out from the crowd. Image source: author's own.

## How I created my portfolio

When creating my portfolio, there were a few basic principles that I followed:

1. Include projects that have something genuinely unique and interesting about them
2. Keep it short and simple
3. Make it pretty
4. Don't waste time on the "web development" side of things

In the rest of this article I'll briefly discuss each of these and explain how they helped me create a winning portfolio.

### Include projects that have something genuinely unique and interesting about them

This might sound obvious, but it is such an important point that it bears repeating:

*Don't include content that everybody else includes.*

If the whole point of your portfolio is to differentiate yourself, why would you bother including generic projects that have already been done to death by others? This is a point echoed by Jeremie Harris in The 4 fastest ways not to get hired as a data scientist [4], where he said:

> *It's hard to think of a faster way to have your resume thrown into the 'definite no' pile than featuring work you did on trivial proof-of-concept datasets among your highlighted personal projects.*
>
> *When in doubt, here are some projects that hurt you more than they help you:*
>
> ** Survival classification on the Titanic dataset.*

> ** Hand-written digit classification on the MNIST dataset.*
>
> ** Flower species classification using the iris dataset.*

What this means is that if you're really interested in Computer Vision, include a couple of projects relating to that. Or, if you just can't get enough of Time Series problems, chuck in something on that theme. Honestly, it doesn't matter whether you're interested in Bayesian Optimisation or Basket Weaving: the point is, just make sure it's something that you genuinely find interesting and think will add value to others. Personally, I'm really interested in problems related to Natural Language Processing and the social sciences, so I tried to include projects related to those. But honestly, it doesn't matter so much what you include as long as you can convincingly show *why* it's adding value.

#### Keep it short and simple

No one wants to spend time on a boring or lengthy webpage, and this is especially true when it comes to online portfolios. You're giving a snapshot, not writing a PhD thesis, and I honestly believe that you'll be wasting your time if you spend too long writing out lengthy project descriptions and results sections.

To appreciate this point, we have to go back to the original purpose of a portfolio, and at the heart of that is one simple idea:

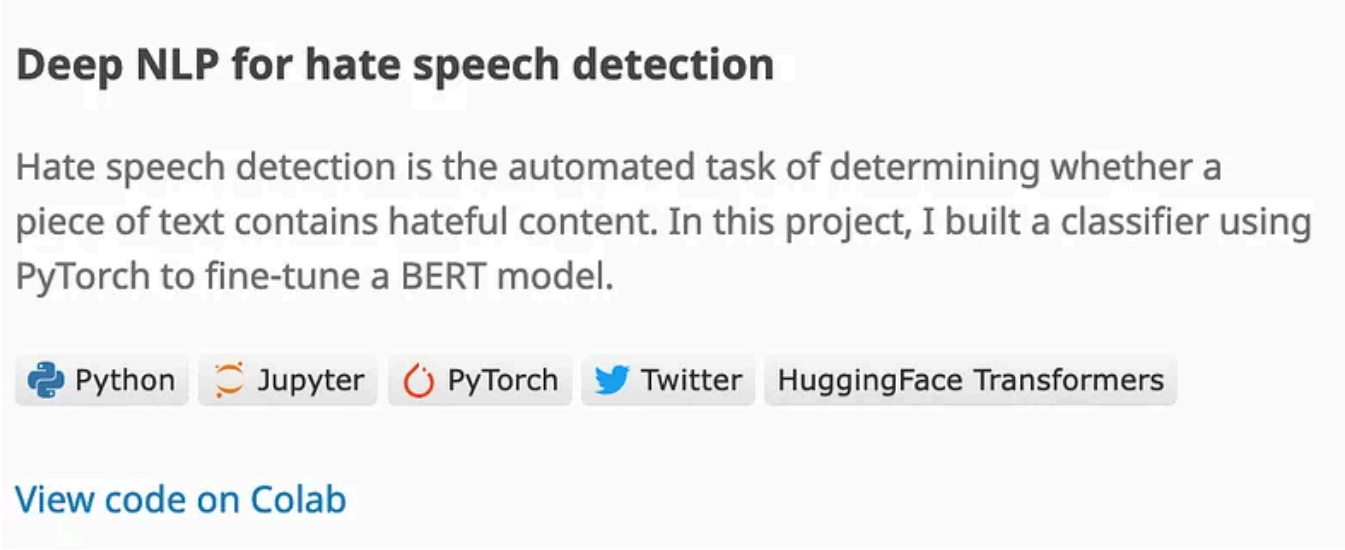
*The purpose of your portfolio is to get your foot in the door, not to get you the job.*

In other words, you're never going to get a job offer solely off the back of your portfolio. The reality is that, even if recruiters like the look of your portfolio, you're still going to have to go through their interviews and assessments. The purpose of the portfolio is just to provide a quick snapshot of your skills and show 'em what you can do.

In my portfolio, for instance, I mentioned seven projects that I'd previously worked on. For six out of the seven projects, I wrote no more than two short sentences of description.

Yep, you read that right: **two** sentences.

Here's an example of a project description in my portfolio:




An example project in my portfolio. Image source: author's own.

As you can see, I didn't waste time setting the context and explaining the nuances of the results. Sure, that stuff is super important. But if the prospective employer is interested, they'll ask you in the interview. At this stage, you just need a short description which can pique their interest and prompt further discussions later on down the line. If you like, you can even include a link to more description elsewhere; that's kind of what I've done by including a link to the code underneath the project description. But on the main page of your portfolio, you just need the high-level details. Anything else is wasting time.

**Make it pretty**

My next tip is to make your portfolio visually appealing.

Put yourself in the shoes of a prospective recruiter: you've read hundreds (or even thousands) of bland résumés and are struggling to remember what the colour green looks like because it's been so long since you saw something that wasn't black-and-white. Then, along comes a résumé/portfolio which — hold the front door — is actually quite nice to look at.

You don't have to be a rocket scientist (or even a Data Scientist) to work out that the recruiter is going to enjoy looking at your portfolio.

To make your portfolio pretty, my top tip is to include lots of interesting graphs/plots. For example, here is the description of another project I included in my portfolio, which has some graphs about Covid-19 restrictions and population movements during lockdowns:

# Examining panic-buying during the first wave of Covid-19, using mobility data

The first wave of COVID-19 infections led to widespread stories of shortages in grocery stores as consumers stocked up in anticipation of lockdowns, a behaviour colloquially known as 'panic buying'. In this project, I used mobility data from Google and Apple to empirically investigate the extent of panic buying in different countries.

**Figure 1**: Footfall at pharmacies and grocery stores, for selected countries

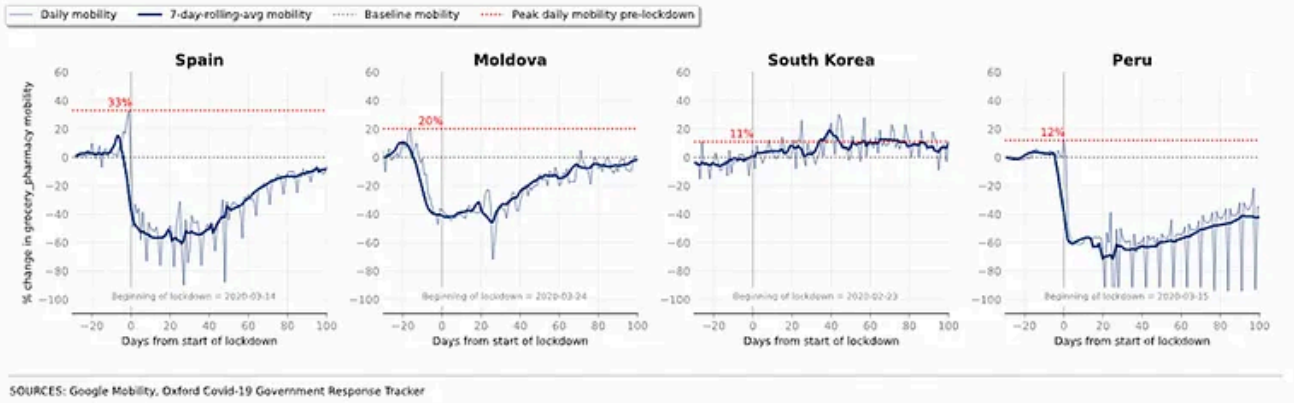
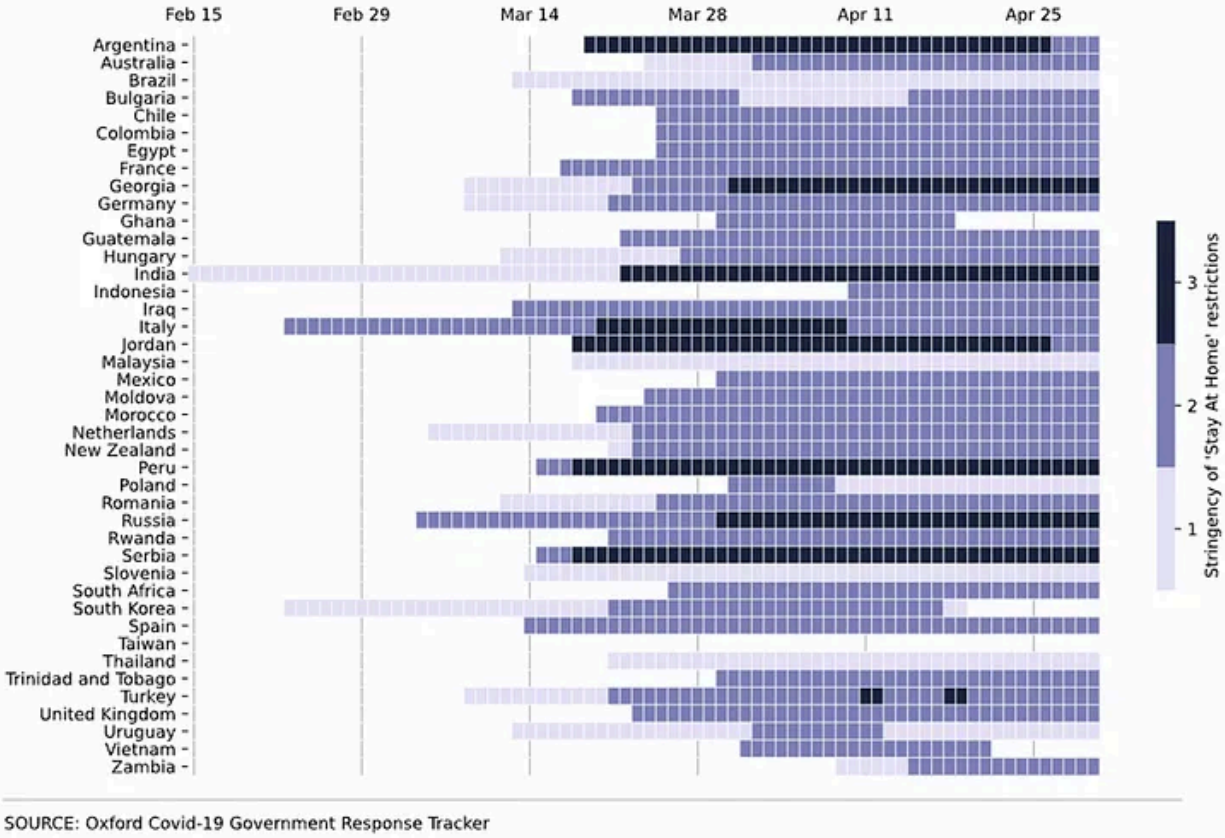
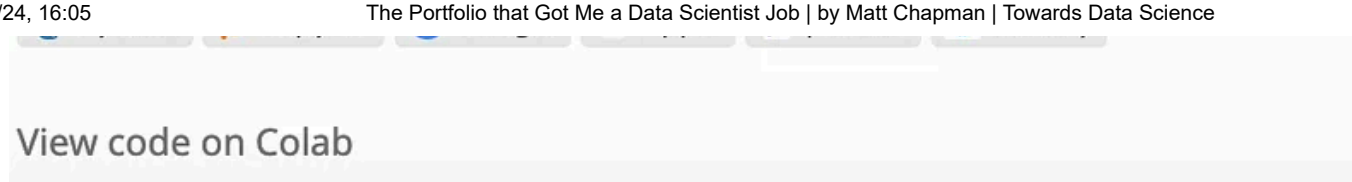



**Figure 2**: Timeline of 'Stay At Home' restrictions, by country




Python Jupyter Google Apple pandas NumPy



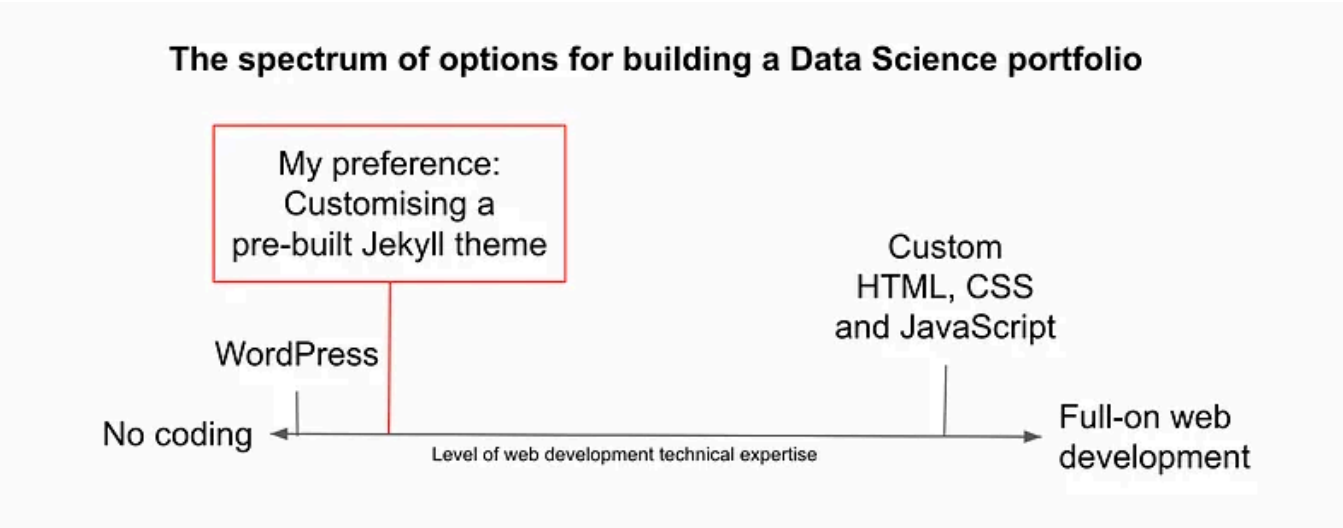

One of the project descriptions in my portfolio. Image source: author's own.

The graphs show prospective employers that I know how to communicate ideas visually and make pretty pictures. This might seem trivial, but it's really not: as a Data Scientist, you will frequently have to communicate your findings to non-technical stakeholders. By including some graphs in your portfolio, you're showing your readers that you recognise the importance of communication and have the skills needed to put it into practice. In other words, it's a fantastic way to show off your skills and differentiate yourself in a crowded field.

**Don't spend unnecessary time on web development**

My final tip is to avoid spending too much time actually "building" the website which will contain your portfolio. Remember: you're applying for a job as a Data Scientist, not a Web Developer. You won't be evaluated on your HTML and CSS skills; you'll be evaluated on your Data Science skills.

Consequently, unless you have extensive prior skills in web development, I'd recommend going for low-code/no-code option to build your website, for example using a tool like WordPress or Webflow which allows you to "drag-and-drop" sections of your portfolio and choose from a range of pre-made templates.




The spectrum of options for building a Data Science portfolio. Image source: author's own.

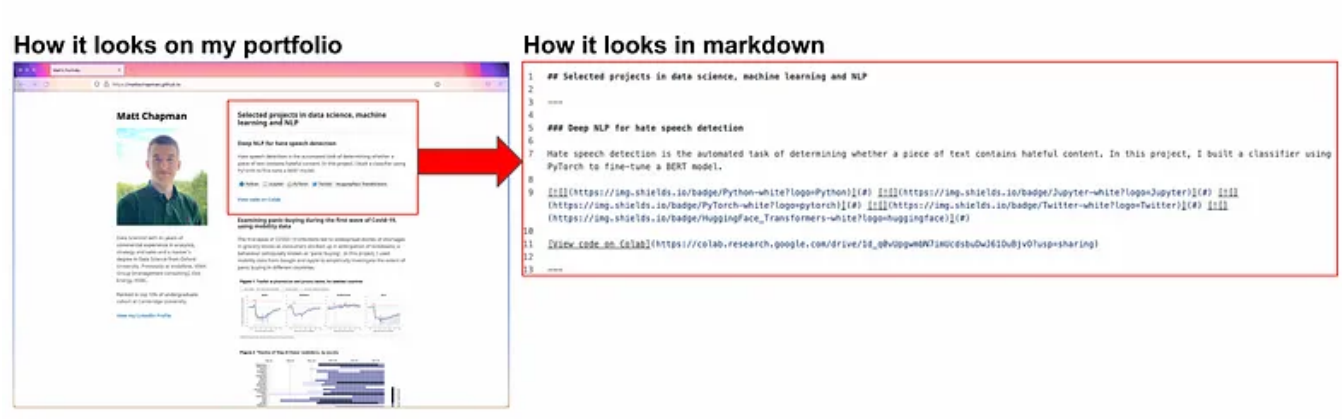
In my case, I opted for something a little different, and followed this excellent guide written by Ivanna Kacewica, who used the Minimal Jekyll theme for GitHub Pages to construct the barebones of the website [5]. The main reasons I went for this option

were (a) it's free, and (b) it doesn't require you to include ads or the word "wordpress" in the URL.

An added bonus of Ivanna's approach is that it allows you to host your portfolio for free on GitHub Pages at the address *username*.github.io, where *username* is your username (or organization name) on GitHub. Since your GitHub username is included in the URL, it's super easy for people to find your GitHub account, which is useful if you're got additional projects stored on GitHub.

I won't go into the details of how to use Ivanna's theme (as she does an excellent job of explaining things), but the basic idea of this approach is that you add text and images using Jekyll and markdown. If you're not familiar with Jekyll, it's a simple framework for building websites that requires no coding or knowledge of HTML/web development.

Using Jekyll, you build the basic structure of the website, and then using markdown you populate the site with your text and images. If you've not used markdown before, don't worry — markdown is just a really simple way of formatting text. And when I say "simple", I mean *really simple*: to write a heading, for example, you just put a hashtag in front of the text you want to be in your heading. To make a bullet point, you just put an asterisk in front of the text you want to be in your bulleted list. It's that easy.




How to add content using markdown. Image source: author's own.

If you want to see the full details of how I customised the theme, take a look at the `index.md` file in my GitHub repository.

## Conclusion

Building an online portfolio is a great way to stand out in the crowded Data Science job market. In this article, I've given some insight into how I went about building my personal portfolio and my top tips for building your own.

Let me know how you get on, and if you'd like any feedback or advice on your portfolio, drop a link in the comments and I'll be sure to take a look.

If you'd like some more advice on how to come up with portfolio ideas, you might want to check out one of my other articles:

**How to Find Unique Data Science Project Ideas That Make Your Portfolio Stand Out**

Forget Titanic and MNIST: Pick a unique project that builds your skills and helps you stand out from the crowd

towardsdatascience.com

## Oh, one more thing —

I've started a free newsletter called AI in Five where I share 5 bullet points each week on the latest AI news, coding tips and career stories for Data Scientists/Analysts. There's no hype, no "data is the new oil" rubbish and no tweets from Elon — just practical tips and insights to help you develop in your career. Subscribe here if that sounds up your street!

**AI in Five | Matt Chapman | Substack**

The latest news, career stories and coding tips from the world of Data Science and AI, summarised in 5 bullet points...

aiinfive.substack.com

## Sources

[1] U.S. Bureau of Labor Statistics. Occupational Outlooks Handbook: Data Scientists. (September 8 2022). https://www.bls.gov/ooh/math/data-scientists.htm

[2] Will Stanton. Creating a great data science blog. (July 15 2015). https://will-stanton.com/creating-a-great-data-science-resume/

[3] Standout CV. How long do recruiters spend looking at your CV? (November 2022). https://standout-cv.com/how-long-recruiters-spend-looking-at-cv#:~:text=Research%20shows%20that%20recruiters%20spend,15%20minutes%20reviewing%20a%20CV.

[4] Jeremie Harris. The 4 fastest ways not to get hired as a data scientist. (June 12 2018). https://towardsdatascience.com/the-4-fastest-ways-not-to-get-hired-as-a-data-scientist-565b42bd011e

[5] Ivanna Kacewica. Set Up Your Portfolio Website in Less Than 10 Minutes with Github Pages. (March 1 2019). https://medium.com/@evanca/set-up-your-portfolio-website-in-less-than-10-minutes-with-github-pages-d0efa8ff56fd

Data Science | Portfolio | Data Science Careers | Careers | Data Analysis


Follow

## Written by Matt Chapman

6K Followers · Writer for Towards Data Science

Data Scientist based in Oxford, UK. Learn SQL at the-sql-gym.com or get my portfolio template at shop.mattchapman.co

### More from Matt Chapman and Towards Data Science

Matt Chapman in Towards Data Science

## Stop Overusing Scikit-Learn and Try OR-Tools Instead

Many Data Scientists overuse ML and neglect Mathematical Optimisation, even though it's great for your career and easy to learn

✦ · 9 min read · Jan 26, 2024

409 6

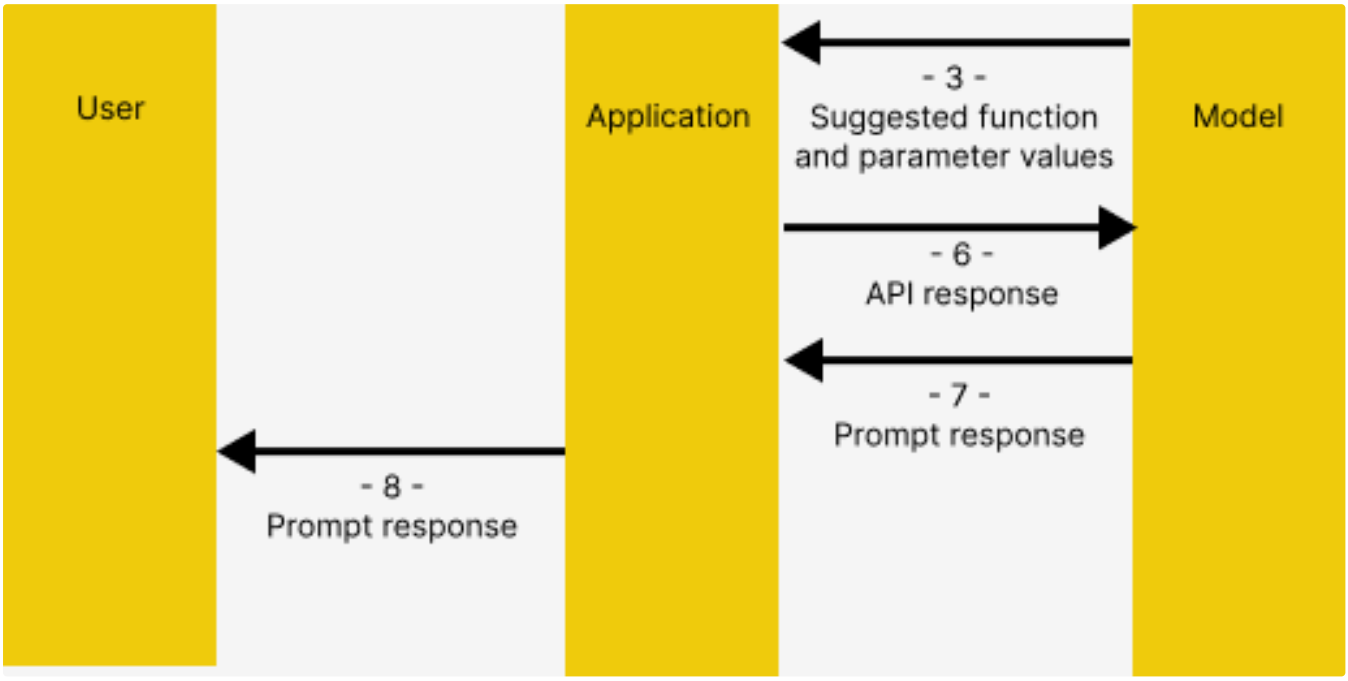
Rami Krispin in Towards Data Science

## Setting A Dockerized Python Environment—The Elegant Way

This post provides a step-by-step guide for setting up a Python dockerized development environment with VScode and the Dev Containers...

9 min read · Apr 2, 2024

787 1




Julian Yip in Towards Data Science

## Build Autonomous AI Agents with Function Calling

Transform your chatbot into an agent that can interact with external APIs

11 min read · Apr 2, 2024

1K 8

Matt Chapman in Towards Data Science

## How I Stay Up to Date With the Latest AI Trends as a Full-Time Data Scientist

No, I don't just ask ChatGPT to tell me

· 8 min read · May 1, 2023

2.8K 31

See all from Matt Chapman

See all from Towards Data Science

## Recommended from Medium

```
*__, a, b, *_ = [1, 2, 3, 4, 5, 6]
print(__, _)
```

What does this print?

- A) Syntax error
- B) `[1] [4, 5, 6]`
- C) `[1, 2] [5, 6]`
- D) `[1, 2, 3] [6]`
- E) `<generator object <genexpr> at 0x1003847c0>`

Liu Zuo Lin

## You're Decent At Python If You Can Answer These 7 Questions Correctly

# No cheating pls!!

✨ · 6 min read · Mar 6, 2024

2.7K 19

Natassha Selvaraj in Towards Data Science

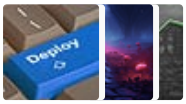
## How I Got a Data Analyst Job in 6 Months

And how you can do the same in 2023, by leveraging ChatGPT and generative AI.

· 10 min read · Nov 18, 2023

509 7

## Lists



**Predictive Modeling w/ Python**

20 stories · 1117 saves

**How to Career Plan When You've Already Started a Career**

10 stories · 327 saves

**Practical Guides to Machine Learning**

10 stories · 1339 saves

**ChatGPT prompts**

47 stories · 1456 saves

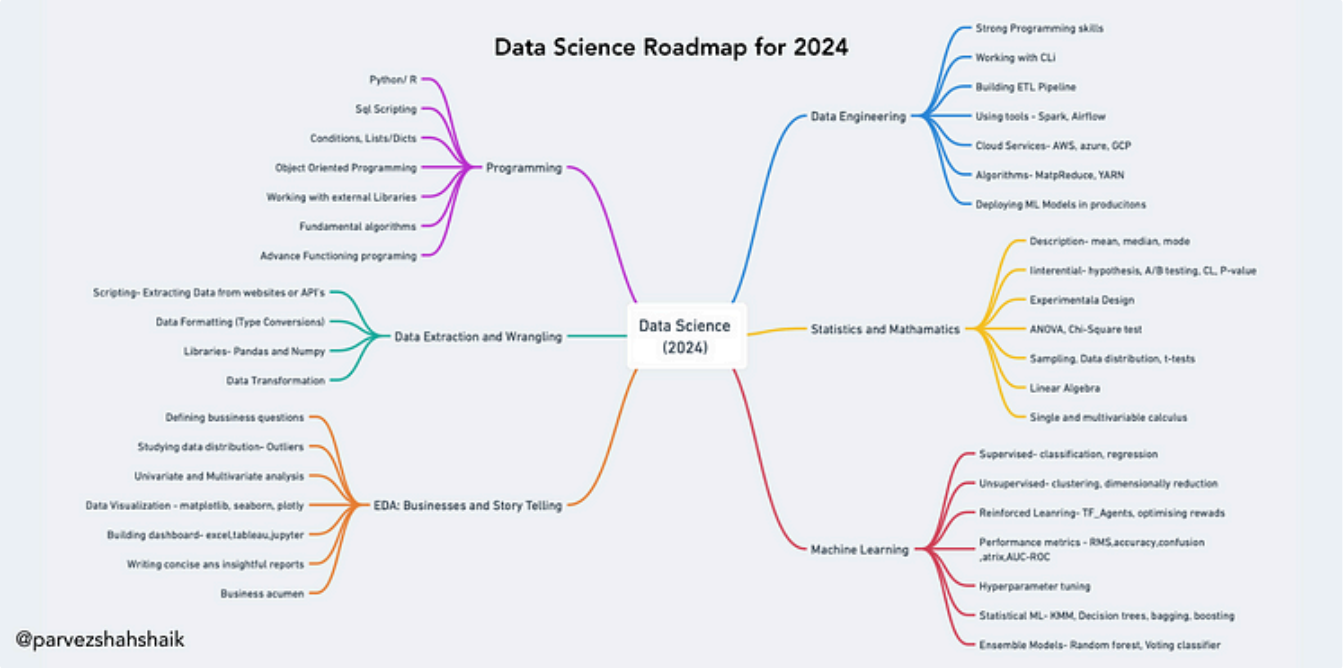
Somnath Singh in Level Up Coding

# The Era of High-Paying Tech Jobs is Over

The Death of Tech Jobs.

· 14 min read · Apr 1, 2024

8.5K 247



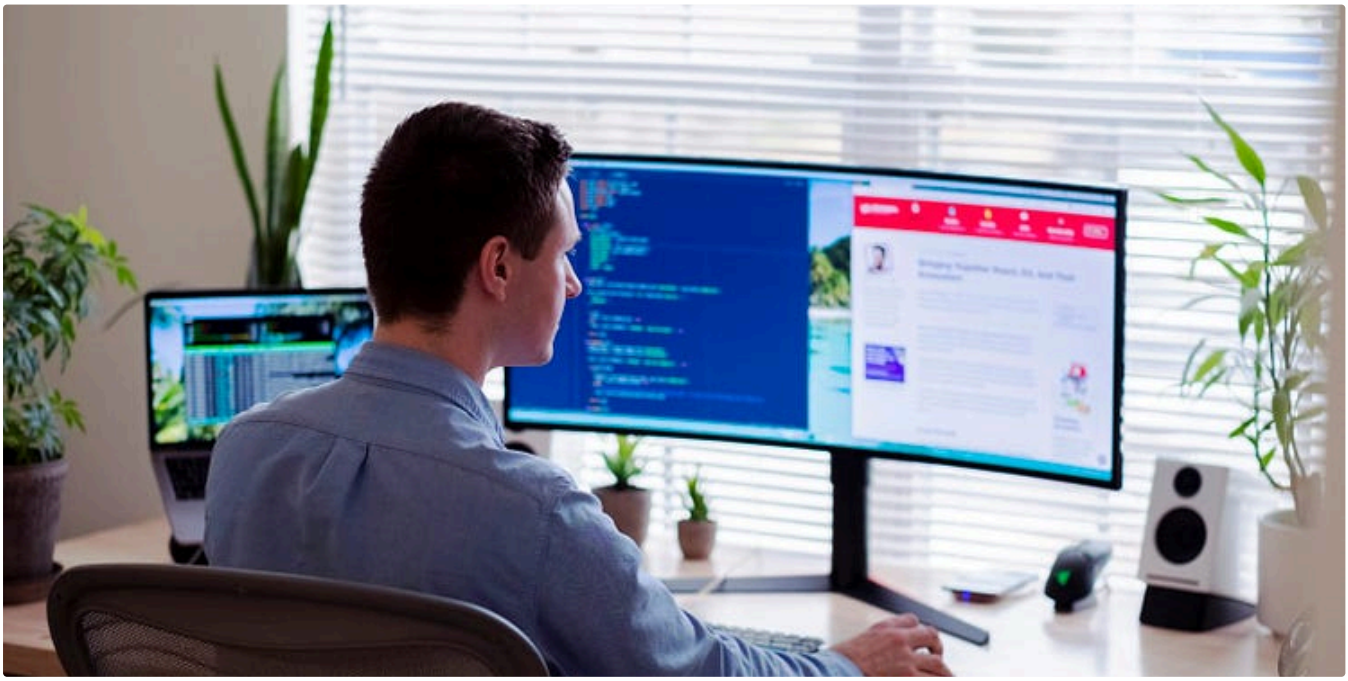

Parvez Shah Shaik in Bootcamp

# DataScience RoadMap for 2024

This article aims to guide you in building a strong portfolio showcasing your data science skills. It provides a framework, resources, and...

8 min read · Dec 27, 2023

150



Hammad Hassan

## I Tried Upwork For 30 Days

I challenged myself to earn $100 in the first month, and it turned out earning more than that.

· 4 min read · Feb 17, 2024

6.1K 125

Hazel Paradise

## How I Create Passive Income With No Money

many ways to start a passive income today

5 min read · Mar 27, 2024

7.3K 172

See more recommendations