**What Stops You From Getting Your First Data Science Job?**

**Walkthrough of the data science interview funnel and steps to get your first offer**

Even though there are many blogs, videos and tutorials on how to become a data scientist, a lot of these resources suffer from what I like to call “the curse of knowledge”. What is that? Well, most of them are created by individuals that have been data scientists for quite some time and so have forgotten the bottlenecks and struggles of getting your first job as a data scientist.

I remember when I started applying for data science positions in London (at that point, I had a bit over two years of experience as a data analyst), there were a lot of rejections in almost all of the stages. I probably made hundreds of applications in order to have only a couple move into the next stage. Even for those applications that passed the screening process, there were a lot of rejections before I managed to get my first offer. In fact, there were times that I decided to abandon the process altogether (sometimes even for months), being convinced that my dream of becoming a data scientist was not attainable. Sometimes I wish I could go back in time and help myself navigate through the complexities and unknowns of the data science interview funnel and save some of that lost time.

With that in mind, I wanted to construct this article a little differently. Instead of focusing on the technical requirements like statistics, programming and machine learning (don’t worry, we won’t skip these entirely), I wanted to create a guide that focuses on the stages of the interview process and where to focus in order to get your first offer. More specifically, the article will be structured as follows:

1. **Mapping the data science interview funnel**
2. **Open the gates with your CV**
3. **Technical (take-home) assessment**
4. **Live interviews with hiring data scientists**
5. **Learn from your mistakes**

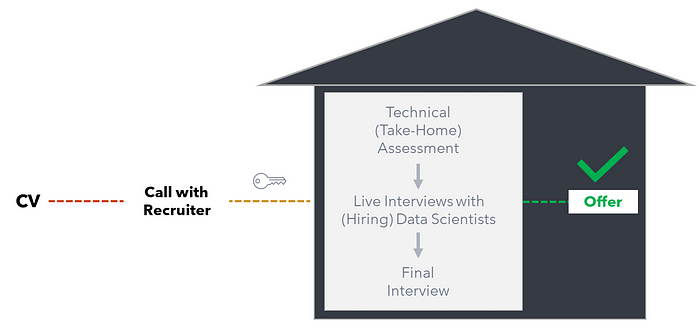
**1. Mapping the Data Science Interview Funnel**

The interview process for most entry data science roles usually involves three to five stages. Although these will be slightly different based on the type of data scientist you want to be and the company you apply for, the building blocks will be the same.

The first two stages are part of the screening process. These are your **resume (CV)** and the **initial call with the recruiter**. At this point, you are still knocking on the company’s front door. The idea is to narrow down the candidate pool to a smaller group of individuals that will be invited “inside the house”.

Once inside the house, you still need to prove that you are adequately qualified for the position. To put it differently, you need to demonstrate that you have the right attitude and set of skills so that they would want to hire you but also be able to justify it. That translates into the three stages below (note that not all stages will be present in every interview process):

* **Technical (Take-Home) Assessment:** This stage is designed to test your logical thinking and competence level in statistics and programming (SQL, R or Python). In essence, they want to assess whether your knowledge of the fundamentals is adequate to move further into the house.
* **Live Interviews with (Hiring) Data Scientists**: This stage involves a video call (after Covid hit), where senior or lead data scientists will ask you to solve problems and answer questions based on the day-to-day tasks you will be expected to complete if hired. At this point, you are well inside the house.
* **Final Interview:** This stage is not always present, but in established data science teams, as the last barrier, you will have a short call with a senior leader from the data science team (Head or Director). The idea is for them to know you and give the ‘green light’ or pick up any red flags.



Data Science Interview Funnel [Image by the author]

So now that we have a better understanding of the stages let’s dive deeper into each one of them so that you can better understand how they are structured and how to approach them in order to land your first data science role.

**2. Open the Gates With Your CV**

When creating your CV, think carefully about the type of data scientist you want to be. If you want to forge the right key, you should first consider the type of door you want it to open. There are two main types of (entry) data scientist positions (i.e. doors):

* **Analytics Data Scientist**: main responsibilities are AB testing, ad-hoc analytics to improve campaigns, offers or products, and building simple (mostly exploratory) machine learning models, such as clustering.
* **Core Data Scientist:** main responsibilities are building predictive models (with the highest accuracy) that can be pushed to a production environment.

If you want to learn more about the different types of data scientist roles and their respective day-to-day tasks, you can also check the video below.

While a typical key will be in the form of a small piece of shaped metal with incisions cut to fit the wards of a particular lock, your CV will have **keywords** and **work experience**.

Unfortunately, even if you forge the right key, the success rate at this stage will be the lowest in the entire interview process. This is primarily due to the inherent randomness and hastiness of this stage (you have probably heard that most CVs are looked at for 2–3 seconds before a decision is made).

While applying without getting responses can be dishearting on the flip side, you are given all the ammunition you need to forge the right key. Most job descriptions will already list these keywords, so what is left for you is to a) identify which roles are right for you, b) pick up the common keywords and required skills, and c) make sure they are clearly stated in your CV. Trust that simple process and with the right volume of applications you will pass onto the next stage (a simple case of convergence due to the law of large numbers).

As mentioned earlier, the second part of the screening process is a phone call with the recruiter. This usually lasts 15–20 minutes, and the (non-technical) recruiter starts by giving you some background on the company and the role. Afterwards, the recruiter will ask you about your work experience, notice period, visa requirements, compensation expectations and some questions regarding the position. Common questions to prepare for are:

* Can you quickly walk me through your CV/work experience?
* What interests you about this role in our company?
* What are you looking for in a new role or framed differently what tasks or responsibilities do you look for?
* What tools and techniques have you used in your current or previous roles?

Even though these might seem simple, I recommend practising them a couple of times. As the quote by **Abraham Lincoln** goes,​​ “give me six hours to chop down a tree, and I will spend the first four sharpening the axe”.

**3. Technical (Take-Home) Assessment**

The technical take-home assignment will usually be in the form of a series of questions. These will usually include the following:

* Write a function to tackle a simple problem (like calculating the first 10 Fibonacci numbers that are bigger than 1000) using R or Python.
* Given one or more tables (usually no more than two), you will be asked to write code in SQL to extract or manipulate the data. These questions usually start simple, like counting the total number of rows that fit certain criteria, with the last questions requiring the usage of different types of joins and window functions.
* Given a chart or a small table, you will be asked to use statistics to draw the right conclusions. The most frequent questions are around skewed distributions, averages and the Central Limit Theorem.
* Given a certain business problem, you will be asked to write a quick proposal on the KPIs and analysis you would select to achieve the desired result.

I cannot emphasise enough the importance of double (or triple) checking your answers and validating your code. I also recommend adding these checks, when possible, into your solutions or code. You can find more information on SQL statements and testing your code in the article below.

Alternatively, you might be given a data challenge. More specifically, you will be given a dataset and an overview of the assignment or problem. For example, in my take-home assignment for Deliveroo, the task was to analyse the performance of their Rider Gets Rider program compared to other marketing channels. The last step of the task involved presenting a summary of the information to senior non-technical stakeholders and suggesting possible next steps. If you need more help on this step, I recommend reading the guide below.

**4. Live Interviews With Hiring Data Scientists**

This will be the most challenging part of the interview process. Not only is it a live interview, but the difficulty level is also the highest in the funnel. Having said that, you can expect to be given a fair evaluation as you will have the interviewer’s full attention for the duration of the video call.

Before COVID-19, most technical interviews were hosted at the company site (literally inside the ‘house’). However, nowadays, the vast majority is done via zoom or other video conference tools. These interviews are conducted by one or two data scientists and usually consist of the following three parts:

**4.1 Behavioural / Attitude Assessment**

In this part, the questions are mostly non-technical and aim to assess your soft skills and character. Common questions are:

* Can you describe a situation where you encountered a problem and explain the steps you took to address it?
* Walk me through a project you are proud of and its impact on the business.

While preparing your answers (following the advice of our friend Abe Lincoln), it is important to always have in mind that there is no right answer to these types of questions. Instead, what they are looking for is the ability to understand that in any environment, you will have obstacles, and what is truly important is the ability to showcase perseverance and the skills to provide value. The **STAR** method is a good framework to follow by discussing the specific Situation, Task, Action, and Result.

**4.2 Experience Assessment**

In this part, you will be given a specific problem from the interviewer, usually based on the task(s) you will be expected to complete if hired. Typical scenarios are the following:

* We recently launched a campaign or offer and want to assess its performance. Describe how you would approach the problem from start to finish (liaising with the business stakeholders, understanding the objectives, getting the required data and analysing and presenting the results).
* Describe how you would design an AB test for a particular feature or initiative.
* If the AB test shows that the desired metric is going down while another metric is going up, how would you make a decision on whether we should launch or discard the feature?

Do not be afraid to ask the interviewer clarifying questions and walk them through your train of thought. The interviewer will judge you both on your ability to understand better a vague description (as they mostly appear in a real environment) and on your approach to tackling it. Also, do not hesitate to ask the interviewer for a few minutes to structure the problem before answering. Finally, remember that what they are looking for in a new hire is the ability to provide value to the business, so always structure your answers with that north star in mind.

**4.3 Technical Skills Assessment**

In this part, you will be judged on your technical skills in statistics, AB testing and machine learning. Example questions are:

* How would you describe a p-value and confidence intervals to a 5-year-old?
* Can you explain what bootstrapping is and how it is used?
* How would you estimate the sample size for an AB test before launching it?
* How would you analyse the results of an AB test with respect to outliers and pre-test imbalances?

If you need a refresher or introduction to the above questions, you can find a full guide on AB testing in a business environment in the link below.

I also highly recommend the free book [An Introduction to Statistical Learning: With Applications in R](https://www.statlearning.com/) for key topics in statistical learning. The book also covers questions such as:

* Can you explain the bias-variance trade-off in machine learning?
* Can you describe what overfitting is and how to tackle it when building predictive models?
* Can you describe the difference between supervised and unsupervised modelling?
* Can you describe what cross-validation is and how it is used in model evaluation?

**5. Learn From Your Mistakes**

One of the most overlooked parts of the interview process is asking for feedback. This is understandable as the majority of times, you will not get an answer, or it will be an automated (not useful) response.

My advice is to think of getting feedback like a focus group. You need ten or so good responses in order to pinpoint the common areas that you need to improve but, equally important, what worked well and what you should keep unchanged.

Another point I want to emphasise is the importance of doing your own assessment of the company based on the quality of their interview process. Was sufficient effort put into the different stages, or did it look like a hasty effort? Did they evaluate you on what you would expect? Did they take the time and effort to give you feedback? Even if you don’t get hired today in that role, you can always put an asterisk on the company for future applications, indicating a place to avoid or a place to look for.

Finally, do not think of a rejection as a failure. The more interviews you take, the better you will understand what is required for the position you aim for and equally important, the more prepared you will be for it! As with everything else, **the only way to improve is by practising**. Similarly to the first time you rode a bike, you weren’t interested in making it on the first try, nor were you so embarrassed by failing that you never tried it again.

**Summary**

🚀🚀 And with that last step, we have reached the end of the guide. Below, you can also find a quick summary of the steps:

✅ Make sure you have a clear understanding of the task and responsibilities you want in your first role, and use keywords from the relevant job descriptions to forge the right key (CV).

✅ Always structure your answers with the overarching goal of providing value to the business. Remember, the company is not looking per se for statistics or programming practitioners but rather for individuals that can provide value from big data.

✅ Practice may seem boring or redundant, but it will make all the difference in the battle for your first data science offer.

✅ Ask for feedback when possible to pinpoint what is working and where you need to improve.