Does Data Science require coding?  
Well, it does. The question is how much and what can you do about it.



**Anish Mahapatra**

Contents

[Introduction 3](#_Toc106233299)

[Does Data Science require coding? 3](#_Toc106233300)

[Why is Coding Required in Data Science? 4](#_Toc106233301)

[How Much Coding is Needed for Data Science? 5](#_Toc106233302)

[What Programming Languages are Used in Data Science? 6](#_Toc106233303)

[Can You Become a Data Scientist Without Coding? 8](#_Toc106233304)

[How Can You Start Learning Coding for Data Science 8](#_Toc106233305)

[What Jobs in Data Science Require Coding? 9](#_Toc106233306)

[Conclusion 9](#_Toc106233307)

[Author 10](#_Toc106233308)

# Introduction

“I have been researching, and Data Science is a good field to be in.”

Whether a fresher or an experienced professional, people looking to move into Data Science always want to know, “Does data science need coding?”.   
  
Let’s understand a little more first. Data Science is a field that is a combination of *mathematics, business and technology*. In a constantly evolving field, the **mathematical** understanding of Data Science remains *consistent*. Now, it is a question of the rest. Let’s understand more:

* **Business**: Data Science is a business-agnostic field. Whatever domain you come from, you can leverage your business knowledge to do better data science.   
    
  For instance, if you are from a CA background, you can help Fintech companies. In addition, given your strong understanding of financial data, you can understand more than most. However, based on your interest, it is possible to work in any domain of Data Science.
* **Technology**: Technology is a field that keeps evolving every day. A lifelong learning mindset must be applied to keep up with the pace of technology.   
    
  Once we understand the foundational elements of technology, it becomes vital that we keep upgrading ourselves based on the latest in technology, like by doing the [insight data science Bootcamp](https://www.knowledgehut.com/data-science/data-science-bootcamp-training).

Now that we have established that we need to be up to date with technology let’s understand a little more about our original question.

## Does Data Science require coding?

**The short answer**   
*Yes*

**The long answer**

* **For freshers**: Coding in Data Science is not how you did it back in school or college. It takes a different form in the real world. However, much of what we learn in Data Science already exists in easily usable functions. Data Science is practical. Google will be your saviour and have all the answers, but to ask the right questions, you need to understand how to code. A little controversial, but in my opinion, the fact of the matter is,

No Data Scientist is a Data Scientist without Google.

Looking at the data, you need to figure out what will be the input and the possible output. We have the input and available functions you can use as code, and then we get the result. A significant chunk of the work is to interpret the output. As we discussed, coding is required, and a [data science certificate](https://www.knowledgehut.com/data-science-courses) can help.   
  


* **For working professionals who do not code**: I have been asked often – do we need to know coding for data science? I will assume that you have some fear about whether you can learn code or not. That is probably the wrong thing to wonder – because the answer is yes. The question is, are you willing to learn to code.   
    
  If you do not code at your job right now, you likely don’t like to code. However, let’s say you are in a management position in Data Science going ahead. To accurately guide your team, you need hands-on coding experience to know what they are talking about. Coding is required.
* **For working professionals who code:** Coding is required in Data Science, and you can pick it up. There is a learning curve in Data Science because, along with code, you will also need to unlearn and relearn mathematics and business. The [knowledgehut insight data science bootcamp](https://www.knowledgehut.com/data-science/data-science-bootcamp-training) can help here.   
    
  Instead of just deliverables, being a Data Scientist is challenging and fun. You will need to think about how you can add value to the bottom line of the company you work with. This will include business and thought leadership elements you have not considered before.

# Why is Coding Required in Data Science?

Data Science is a field where **experiments** are carried out on data to help improve the quality or bottom line of the enterprise. We just use Excel to analyze data. Large volumes of data are generally present on a cloud platform, and a Data Scientist must perform analytics.

To do this, a Data Scientist needs to have a robust toolkit where they are free to experiment. Any experimentation, data manipulation and visualization should be possible to strive to achieve the end result. It’s not engineering; it’s actual science that consists of performing experiments, where some succeed, and most fail.

The freedom to do anything comes with code.

Coding is required in Data Science because:

* **Sourcing Data:** Regardless of the cloud platform or source, code can help get the data from wherever it is stored. Code enables us to manipulate data while pulling it right from the start.
* **Data Transformation:** Knowing how to code can help to manipulate, fix and transform the data as required – this can be done via multiple platforms. For instance, Python code can be applied on almost any cloud platform or tool.
* **Exploratory Data Analysis:** The patterns in data can be deciphered with the help of code; it is vital to explore large datasets to understand the visible and hidden patterns.
* **Experimenting with Data:** Working on different hypotheses to see if there is backing for a data-driven decision can be done with the help of code
* **Machine Learning & Modelling:** Having the freedom to make models and perform machine learning on data can be done with the help of code
* **Visualization:** Giving a Data Scientist the ability to visualize data in multiple ways is a powerful tool. It can transform how we go about solving a problem, as visualizing data can help business stakeholders make data-driven decisions better

The freedom to do anything is the main reason coding is required in Data Science. So, in the next section, let’s go over how much coding is needed for Data Science.

# How Much Coding is Needed for Data Science?

Depending on your selected role, varying degrees of coding are required for each position. However, a good start would be understanding the fundamentals of one coding language and a querying language. Remember, when you code in the real world, Google is your best friend. All of us are Data Scientists because Google is there to help us.

Data Scientists would not be Data Scientists if Google did not exist.

I have another article – Is Data Science for all, explaining the roles in detail. In this section, let’s go over some of the roles and the amount of coding that is required:

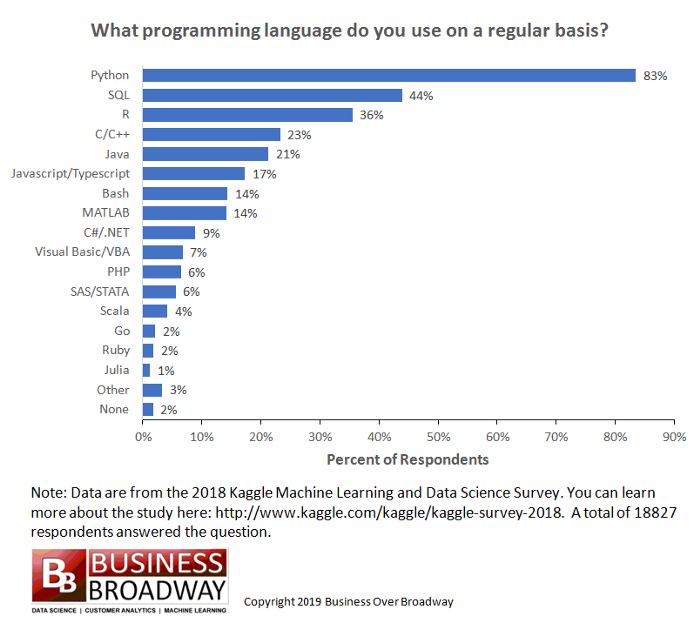
* **Data Engineer:** A data engineer would need to be an expert in SQL or a data query language and understand the fundamentals of python/R to manipulate data as required. A knack for attention to detail can help you become a better Data Engineer.   
     
  Over time, a Data Engineer will gain expertise in a Cloud platform such as Amazon Web Service (AWS), Google Cloud Platform (GCP) or Microsoft Azure. Doing certifications on these cloud platforms can help aid your entry and expedite your career path in Data Engineering.
* **Machine Learning Engineer:** A Machine Learning Engineer needs expertise in a coding language such as Python/R and understands the fundamentals of a querying language such as SQL. Value addition for this role is the fundamentals of Software Engineering, like basic Data Structures.
* **Business Analyst:** Depending on the company you are applying for, this is a role that requires less coding. Understanding the fundamentals of SQL and a visualization tool such as Power BI and Tableau can help you become a better Business Analyst.
* **Data Scientist:** A Data Scientist needs to know everything mentioned above. There must be a keen interest to learn, irrespective of the technology stack or problem. Data scientists must keep learning throughout their career, irrespective of platform, coding language, tools and technologies.   
    
  This can be daunting if you are trying to enter Data Science. However, knowing the fundamentals of a language and an eagerness to learn is what companies are looking for.

Now that we know how much coding is required in Data Science, let’s briefly discuss the programming languages used.

# What Programming Languages are Used in Data Science?

If you are setting out to learn a new language specifically for Data Science, the best language to learn is Python. Some blogs highlight a whole host of languages, tools and technologies. I’ll tell you about the ones to learn to crack 90% of interviews.

Before that, let’s look at a survey of the top programming languages used in the world.

  
[Source](https://towardsdatascience.com/programming-languages-for-data-scientists-afde2eaf5cc5)

If you want to dive in, you can find multiple statistics regarding the world in Kaggle’s recent survey of Data Scientists on Kaggle [here](https://www.kaggle.com/kaggle-survey-2021). The two coding languages to focus on are:

* **Python**: Data scientists worldwide primarily use Python as their language of choice. It is a highly diverse language and fits nicely into multiple technology stacks companies use. Python also has excellent support from the developer community.   
    
  Without fail, it is asked in all Data Science technical interviews. The focus should be on mastering concepts and general logic rather than trying to become an expert in the syntax of Python. Language(s) simply enable you to implement logic.
* **SQL**: Companies test SQL as a fundamental querying language skill. SQL enables us to query databases in a simple language. SQL is a reasonably intuitive language to learn and can be one of the first languages to pick up to give you the initial boost of confidence.

# Can You Become a Data Scientist Without Coding?

This is a subjective question, so everything here will be strictly my opinion and not representative of any organizations I may be associated with.

Yes, it is possible to enter as a Data Scientist with no coding. But, it is impossible to survive as a Data Scientist unless you are open to lifelong learning and coding

There are no two ways about it. A sales/marketing person can sell you whatever you want to hear. So an excellent way to know more about the involvement of coding as a Data Scientist is to talk to one! So, to reduce personal bias, I spoke to my friends in the industry to get their opinion on the subject – it is the same. You need coding to survive and grow in Data Science.

# How Can You Start Learning Coding for Data Science

Sitting on the fence about if Data Science is for you? Have you been trying to understand how you can get a head start to get your foot into the door with Data Science? As your research might be pointing to gradually, learning to code is the best way to get into Data Science. So my two cents on learning how to code:

The best way to learn to code is to simply do it.

Great job, Anish, solid advice! I kid you not – it might seem like I am joking with this profound statement. From personal experience, I was definitely not into coding. There is an inherent fear of the unknown. Coding can seem daunting, like doing advanced algebra as a kid. However, it is simply a barrier you must overcome to be a good Data Scientist.

To put it away as much as possible, I would involve myself in theory from YouTube, Udemy, Textbooks, friends, courses etc. I would learn as much about coding as possible but just not do it. Remember, Data Scientists are in demand because there aren’t many great ones that can fulfil the need. And not everyone becomes a Data Scientist because of a difficult barrier to cross.

Once you can cross the code barrier, Data Science is within reach.

Having said that, here are some resources to start learning to code.

* **YouTube:** The best resource out there to answer all your questions. The one tricky thing to get past is that knowledge is scattered and is difficult to collate, so you really need to know what you are looking for.
* **Udemy:** A decent place where industry practitioners teach courses
* [**Insight Data Science Bootcamp**](https://www.knowledgehut.com/data-science/data-science-bootcamp-training)**:** An end to end Bootcamp structured to get you started with Data Science
* **Books on Data Science:** Books on Data Science can be of great help when it comes to upskilling yourself
* **Coding with a friend:** A great exercise once you pick up the fundamentals is to sit down online or face to face with a friend and code together! There is a surprising amount of learning that can come from this exercise.

# What Jobs in Data Science Require Coding?

**All jobs in Data Science** require some degree of coding and experience with technical tools and technologies. We discussed more on this in our section on “What Programming Languages are used in Data Science?” Let’s summarize:

* **Data Engineer**: Moderate amount of Python, more knowledge of SQL and optional but preferrable is knowledge on a Cloud Platform
* **Machine Learning Engineer**: More amount of Python, a moderate amount of SQL and a keen interest in experimenting with data
* **Business Analyst**: Strong understanding of business, knowledge of a visualization tool, minimal coding (depending on company profile for Business Analyst)
* **Data Scientist**: End-to-end understanding of the data pipeline. Needs coding.

Gradually, as the industry matures, we may have more roles requiring less coding. You may have read about various “No-code” platforms. Although it would be ideal, many companies aren’t using these platforms. This is because they are not mature enough to offer as much flexibility as just coding it out and cannot handle all tasks.

The only job that comes to mind where it might be possible to do *less* coding is a Business Analyst. Still, even that would depend on the company.

# Conclusion

Now you understand whether coding is required for Data Science and the answer is a resounding yes! Many of these opinions have been formed, having spoken to over 2000+ people in Data Science. So, as you read the entire blog, there are parts where it will feel like I am talking to you. You can do it! Depending on your nature and the role that you are going for, there are multiple ways you can and will pick up on coding!

# FAQs

### Q. Can I become a Data Scientist without coding?

No, it is not advisable to become a Data Scientist without coding. However, you somehow may be able to get a job as a Data Scientist. Growth in the industry will be almost impossible unless you are willing to learn and code!

### How much coding do you need for data science?

It would depend on the role, project, position and company. Just like the cycle of every other project, there might be phases. For example, in the initial stages of your career, you may need to code a lot for Data Science. Then, as you learn and grow to a more senior position, your hands-on coding time will reduce over time.

### What coding does data science use?

Data Science primarily works on leveraging pre-defined packages for the task at hand. Therefore, almost everything we would like to do already exists as modules on the internet in various packages.

In the context of Python, there are widely-used packages such as pandas, NumPy, sklearn etc., that simply need to be called and used in the code. However, data Science does not use the traditional Data Structures and Algorithms principles very often.

### Is C++ required for data science?

No, C++ is not required for Data Science. However, knowing the fundamentals of C++ or Java could help you understand some of the basics of Python. In addition, having any experience with code, however rudimentary, would put you in a stronger position to do Data Science. Having said that, even if you have no coding experience, it is possible to do it.

### Is Python sufficient for data science?

Yes, Python as a primary language is sufficient to start with Data Science. However, as we learn more and technology evolves, a Data Scientist will learn more, depending on the particular project or company. A core question is what I need to know to become a Data Scientist.

Python would be able to satisfy about 70% of what you would need to crack the Data Science interview, and knowing SQL will give you an additional 10-15% edge. The rest would depend on the kind of projects and certifications you can showcase to the recruiter.

# Author

**Anish Mahapatra**   
*Senior Data Scientist*   
LinkedIn: [https://www.linkedin.com/in/anishmahapatra](about:blank) ([LinkedIn](about:blank))   
Email: anishmahapatra01@gmail.com