[**Teach Yourself Data Science: the learning path I used to get an analytics job at Jet.com**](http://127.0.0.1/wordpress/wp-admin/post.php?post=73&action=edit)

**by Dan Stern**

How can you go from zero programming skills to a job in technology or analytics?

If you’re interested in learning these skills, whether for fun or for a career change, what’s the best way to go about it?

Countless lists of the best online courses exist, but how can you forge your own learning path with all of the noise?

I personally never thought I’d learn any practical skills around programming, data analysis, machine learning, or technology in general. As a finance major, I always assumed I’d be the “business guy.” Yet somehow, I taught myself **Python** and **SQL,** and found myself working in analytics at [Jet.com,](http://jet.com/) using one of these languages everyday.

Why Python and SQL, you might ask?

Python is the [fastest growing programming](http://www.zdnet.com/article/java-javascript-c-no-python-is-the-fastest-growing-programming-language/)language out there, and for good reason. It has an insane number of libraries that you can use for machine learning applications, data analysis, visualization, web apps, API integrations, and much more. Plus, it’s one of the easier languages to pick up and learn. As for SQL, databases power technology companies, and SQL allows you to better understand, explore, and make use of the troves of collected data.

[Fossbytes.com](https://fossbytes.com/python-fastest-growing-programming-language/)

Below, I outline the path I took in learning these languages that brought me into analytics. To be clear, this path was incredibly challenging; I spent countless evenings feeling frustrated and confused. Many nights I wanted to just throw in the towel and settle for being the business guy.

But your **motivation** remains the key to pushing forward through the obstacles you’ll inevitably face. Whether you want to move into a data analysis or data science type role, or just want to have a better grasp on programming and technology for the fun of it (which it does become fun!), you have to figure out how to stay motivated and disciplined if you want to actually learn these skills.

For me, setting aside specific amounts of time almost every day (about 90 minutes to 2 hours) to learn or practice immediately after I got home from work allowed me to develop consistent habits and hammer home concepts I found confusing.

Here’s the path that I took; hopefully it can help you get started on your own.

### **The Core Foundation**

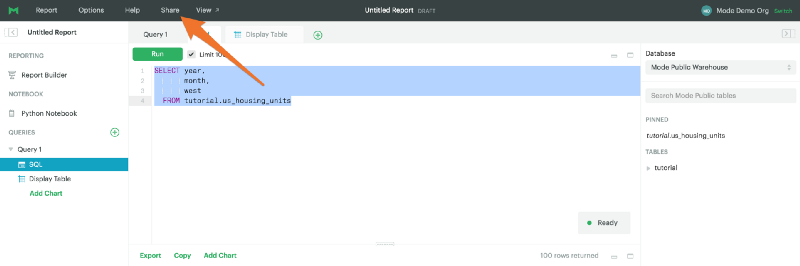
1. [Learn Python the Hard Way](https://learnpythonthehardway.org/)

This is one of the best courses I’ve ever taken, period. It’s self-directed and challenging, but Zed provides you with enough detail and guidance to start to actually begin programming in Python. He makes programming feel accessible, and the material gives you the confidence week after week to actually feel as if you can effectively learn Python.

2. [Mode Analytics: Pandas](https://community.modeanalytics.com/python/tutorial/pandas-dataframe/)

Mode Analytics provides an awesome introduction to Python and includes tutorials on one of its most powerful data structures: the Pandas DataFrame. This is perfect for learning the basics of data analysis once you have the fundamentals of Python down.

3. [Mode Analytics: SQL](https://community.modeanalytics.com/sql/tutorial/introduction-to-sql/)

 [Mode Analytics](https://community.modeanalytics.com/sql/tutorial/sql-select-statement/)

The other Mode Analytics tutorial on SQL is fantastic too. You can learn all of the key concepts and create a strong SQL foundation here. They even have their own SQL editor and data you can play around with.

In conjunction with Mode Analytics, [W3 Schools](https://www.w3schools.com/sql/) can help answer any SQL question you ever have as you go make your way through the tutorials.

### **Diving Right Into Machine Learning**

Before I fully had a strong grasp of Python, I took a shot and applied for Udacity’s self-driving car nanodegree. I knew it was completely over my head, but I thought, why not try?

It’s easier to motivate yourself to learn Python and machine learning when you’re fascinated by the practical applications.

I had about a month before the class began, so I took as many classes around data science and machine learning as possible.

Here were the best **free** **introductory** **courses** I found that were incredibly helpful:

* [Udacity Machine Learning Intro](https://www.udacity.com/course/intro-to-machine-learning--ud120)
* [Udacity Intro to Statistics](https://www.udacity.com/course/intro-to-statistics--st101)
* [Udacity Intro to Data Science](https://www.udacity.com/course/intro-to-data-science--ud359)

Yes, you can see I think quite highly of Udacity.

While not free, I’d also highly recommend checking out the [Grokking Deep Learning](https://www.manning.com/books/grokking-deep-learning) book. It provides extremely clear and relatable examples on the fundamentals of machine learning.

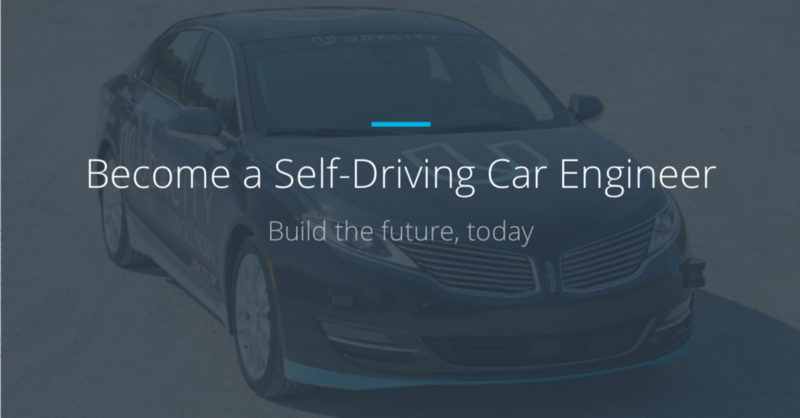
 [KDNuggets](https://www.kdnuggets.com/2016/03/distributed-tensorflow-arrived.html)

[TensorFlow](https://www.tensorflow.org/), developed by Google, is an open source library for machine learning that can be written in Python. It’s incredibly powerful, and absolutely worth becoming familiar with.

Check out the [MNIST exercise](https://www.tensorflow.org/get_started/mnist/beginners) for a fantastic introduction to the framework.

I found the [Stanford CS231](https://www.youtube.com/playlist?list=PLkt2uSq6rBVctENoVBg1TpCC7OQi31AlC) class to be a useful resource too; it covers convolutional neural networks (what we use for image or facial recognition software) extensively, which I read would be incredibly helpful for the self-driving car Nanodegree. If you’re interested at all in using machine learning with images or video, you won’t find much better than this course.

Finally, after using these resources to build a solid foundation, I began the [Udacity Self Driving Car Nanodegree](https://www.udacity.com/course/self-driving-car-engineer-nanodegree--nd013).

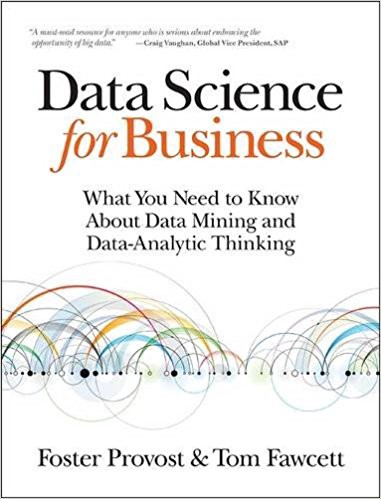
 [Udacity](http://udacity.com/)

I’m not going to talk about it too much since there are already great write ups of the course [here](https://techcrunch.com/2017/09/19/udacity-adds-a-new-intro-to-self-driving-cars-nanodegree/) and [here](https://medium.com/@dhruvp/self-driving-car-students-in-their-own-words-ace90d66f369). What I will say is that, to my own shock, despite being the most challenging course I’ve ever taken, I was able to understand most of the content. Armed with the right base knowledge, you’d be surprised at how deep your understanding of a complex topic can be.

### **Continued Analytics and Data Science Learning**

After diving intensely into machine learning for a few months, it was helpful to take a step back and reinforce my understanding of practical analytics and data science principles.

I started with [Data Science, Deep Learning, & Machine Learning with Python](https://www.udemy.com/data-science-deep-learning-in-python/), a fantastic course on **Udemy**. While touching upon machine learning, it completely covers principles in analytics, data science, and statistics, particularly around different data mining techniques and practical scenarios to deploy them.

 [Amazon.com](https://www.amazon.com/Data-Science-Business-Data-Analytic-Thinking/dp/1449361323)

The book [***Data Science For Business***](http://shop.oreilly.com/product/0636920028918.do), also explains incredibly well the HOW and WHY certain models work when solving problems in a specific context; it hammers into you an analytical framework and mindset that can be applied to any situation revolving around data problems. It’s the **best** resource I found that connects different analytical approaches to specific business situations and problems.

Of course, if you’re interested in pursuing a career in analytics or data science, you should always be honing old skills or adding new skills into your toolkit. [**FreeCodeCamp**](https://medium.freecodecamp.org/) and [**Hackernoon**](https://hackernoon.com/) publish informative articles and tutorials on all things data science and software engineering. My favorite article recently was a well-written tutorial on writing your own [blockchain](https://hackernoon.com/learn-blockchains-by-building-one-117428612f46).

You want to know the **best** way to continue learning though?

Build something. Anything. Explore a dataset. Find a practical problem that you or your company faces, and try to solve it.

Even if you don’t have access to high-quality data at your company, there are plenty of [open source datasets](https://www.kaggle.com/datasets) that you can play around and practice with. I bet you’ll learn just as much, if not more, working on your own data projects than taking any course or reading any book.

 [Unsplash.com](http://unsplash.com/)

Finally, meeting and learning from people who have the skills you want to acquire is hugely beneficial. I highly recommend using [**Meetup**](https://www.meetup.com/) to find groups of analytics or software professionals in your area. Many of these groups have free tutorial or study sessions, and you’ll meet plenty of insanely smart people who can provide tips and tricks to accelerate your learnings.

In New York City, some of the groups that have helped me tremendously are:

* [Machine Learning Society](https://www.meetup.com/Machine-Learning-Society-NY/)
* [Google Developer Group](https://www.meetup.com/gdgnyc/)
* [NYAI](https://www.meetup.com/NYAImeetup/)
* [New York Data Science](https://www.meetup.com/NYDataScientists/)

### **Have fun learning, and let me know how your own journey goes!**