

## Introduction

A STEM career such as data science can appeal to some because it combines skills from complex systems, math, computer science, and statistics. As said by NYU professor Vasant Dhar, "Data science is the study of the generalizable extraction of knowledge from data."\*(1) In other words, it can analyze and make predictions based on both quantitative and qualitative data.

## Prerequisites for Success in Data Science

As said earlier, data science combines skills from complex systems, math, computer science, and statistics. For the math portion, that usually requires standard calculus. The rest would be learned from earning a bachelor's degree in the subject. Considering that it's a job high in demand, many people also get a masters degree or even a PhD.



Students Graduating from the University of California, Berkely, One of America's Top Data Science Programs\*(2)

More specifically, there's certain skills and attributes that will be exercised and later become vital to a fruitful career in data science. Master's in Data Science, a division of 2U Inc, has condensed this down to a simple list\*(3), including:

- Programming
- Machine Learning Techniques
- Data Visualization and Reporting
  - Risk Analysis
- Statistical Analysis and Math
  - Effective Communication
  - Software Engineering Skills
- Data Mining, Cleaning, and Munging
  - Research
  - Big Data Platforms
  - Cloud Tools
- Data Warehousing and Structures

People in data science must be curious, self-assured, and inventive. To prove this, Lisa Qian, a Data Scientist for Airbnb, says, "Successful data scientists have a strong technical background, but the best data scientists also have great intuition about data. Are the features

meaningful, and do they reflect what you think they should mean? Given the way your data is distributed, which model should you be using? What does it mean if a value is missing, and what should you do with it? The best data scientists are also great at communicating, both to other data scientists and non-technical people. In order to be effective at Airbnb, our analyses have to be both technically rigorous and presented in a clear and actionable way to other members of the company.”\*(3)

## Relationship with Competitive Robotics

There are many ways to get a good foothold into the world of STEM at a young age. Competitive robotics can help people learn skills that will help them have successful careers in the world of STEM. As said earlier, data scientists must be curious, self-assured, and inventive. This can be improved upon with teamwork skills from competitive robotics. The Vex IQ Challenge directly claims it puts “Classroom STEM concepts are put to the test as students learn lifelong skills in teamwork, leadership, communications, and more.”\*(4)



Students Cooperating at a Robotics Competition.\*(5)

This also puts in the mindset of using methods similar to that of an engineer, by dissecting the problem and forming intuitive solutions. This will later be shown by some of the most well-known data scientists that this technique is surprisingly useful.

## Scholars in Data Science

Many people have done great and wonderful things with data science. Here are some examples of those people.

Benjamin Fry is an MIT graduate known for being the sole founder of Fathom Information Design, who claims that it “...works with clients to understand complex data through interactive tools and software for mobile devices, the web, and large format installations.”\*(6) In other words, they help companies better analyze data so they can then use it to improve said establishment. They help institutions like General Electric, Warner Brothers, the Bill and Melinda Gates Foundation, Nike, National Geographic, Volkswagen, Mayo Clinic, Samsung, J.P. Morgan, and even Google. They also cooperate with some of the most well known universities, such as Harvard and Stanford. With this, they also dedicate some of their time spreading awareness about data science and inspiring others to join the field. As for Ben Fry more specifically, he traces his expertise from facing problems similar to how an engineer would, which leads to the way that Fathom and his other works have such a good reputation among the data science world.

After receiving a PhD from the University of California, Berkeley in 1976, Chien-Fu Jeff Wu now teaches at the Georgia Institute of Technology. He is also a faculty member for other

schools such as the University of Michigan. In 1997, Wu gave a lecture there about what differentiates it from statistics. Even now, there is still a debate over what *exactly* data science is. Some like Wu and later DJ Patil argued that it is different from statistics. A lecture by Wu, known as "Statistics=Data Science?", helped reinstate what separates it from statistics. He concludes that statistics is more about analyzing a small set of quantitative data while data science is more about making conclusions from large sets of both quantitative *and* qualitative data. He also says that "The current state of statistical work can be described by a **Statistical Trilogy**:

1. Data Collection (experimental design, sample surveys)
2. Data Modeling and Analysis
3. Problem Understanding/ Solving, Decision Making"(7).

Now, not all people in data science have to be hardcore mathematicians or college professors. A good example of this would be Nate Silver, a "BA" student from the University of Chicago the The Telegraph said that "emerged as the other major winner during the US election for predicting all major results and Barack Obama's victory."\*(8) He's most well known for using data science to accurately predict what every state will pick as well as the District of Columbia for the 2012 presidential election. He also predicted forty-nine states during the 2008 presidential election. In fact, he even became one of Time's "The World's 100 Most Influential People" in 2009 shortly after the election. He is also one of the main creators of PECOTA, which predicts the likely career path of incoming MLB players using data science. Lastly, unlike Chien-Fu Jeff Wu or DJ Patil, Nate Silver believes that data science is really just an example of statistics.

The last example here will be DJ Patil. Like Chien-Fu Jeff Wu and unlike Nate Silver, he believes that data science is different from statistics. What sets him different is that he was in fact the first ever U.S. Chief Data Scientist from 2015-2017, as in, he worked as a data scientist, but for the Federal Government. It was a role created by President Obama to give the president extra info about what is more accurately in the national interest. He would also try to keep America as a global superpower in the technology industry. The last part of the job was to try to convince some of the best scholars in data science to help the U.S. Government as well as inspire people to join the field.

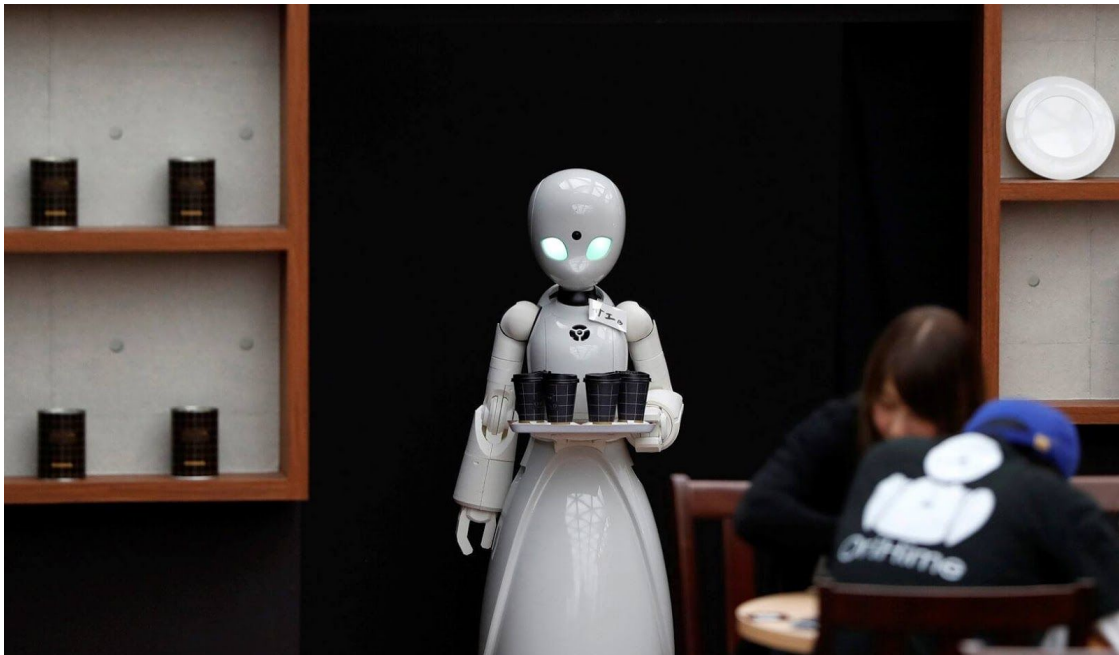


DJ Patil\*(9)

Before then, he also worked for the Department of Defense to help detect potential dangers to the country. After gaining a PhD from the University of Maryland, College Park, he also worked for companies like Salesforce, LinkedIn, Skype, PayPal, and eBay.

## What is the Future of Data Science?

It's no secret that data science is becoming more and more of an essential part of Western society. As the job becomes higher in demand and more PhDs in data science are given out, people might be under pressure to end the debate of what "data science" is and set the definition in stone. With this, there is also a threat considering that computers used by large workspaces are becoming more and more advanced. Some of the more simple parts could soon be replaced by a computer while the more complex material is reserved for us, the humans.



Some Blue-Collar Jobs are Beginning to be Replaced by Computers.\*(10)

This would raise the bar and add more competition to the field, which can be a good thing. Competition has caused some of the greatest discoveries in human history. For example, Christopher Columbus might've never actually contacted the New World if it weren't for competing against the other European powers during the spice trade. The same might be true for data science. As of 2020, we're in a window of time where many jobs will be from in the hands of people to in control of computers. Higher stakes could evolve data science and make it much more valuable and prosperous to those who take the opportunity to use it wisely. New technologies might be invented, and more people should be excited for the beautiful tomorrow that will soon enough be upon us.

---

Written By: Brady White

Edited By: Sarah Land and Sam Snow

Team 46929C

## Career Readiness- Data Science

### References

\*(1)

Dhar, V., 2013. Data Science And Prediction. [online] Communications of the ACM. Available at: <<https://cacm.acm.org/magazines/2013/12/169933-data-science-and-prediction/fulltext>> [Accessed 17 October 2020]

\*(2)

University of California, Berkeley, n.d. Data Science Grads Shine At Commencement. [image] Available at: <<https://data.berkeley.edu/news/data-science-grads-shine-commencement>> [Accessed 18 October 2020].

\*(3)

Master's in Data Science. 2020. *What Is A Data Scientist?*. [online] Available at: <<https://www.mastersindatascience.org/careers/data-scientist/>> [Accessed 17 October 2020].

\*(4)

VEX Robotics. 2002. *VEX IQ*. [online] Available at: <<https://www.vexrobotics.com/iq>> [Accessed 17 October 2020].

\*(5)

The Quad-Cities Daily, n.d. Students Cooperating At Robotics Competition.. [image] Available at: <<http://quadcitiesdaily.com/?p=561208>> [Accessed 18 October 2020].

\*(6)

Fathom Information Design. 2010. *Fathom*. [online] Available at: <<https://fathom.info/>> [Accessed 18 October 2020].

\*(7)

Wu, C., 1997. *Statistics=Data Science?*.

\*(8)

Hugh, A., 2012. *Nate Silver: Politics 'Geek' Hailed For Barack Obama Wins US Election Forecast*. [online] The Telegraph. Available at: <<https://www.telegraph.co.uk/news/worldnews/us-election/9662363/Nate-Silver-politics-geek-hailed-for-Barack-Obama-wins-US-election-forecast.html>> [Accessed 18 October 2020].

\*(9)

Vox, n.d. DJ Patil. [image] Available at: <<https://www.vox.com/platform/amp/2016/12/28/14106602/full-transcript-chief-data-scientist-dj-patil-science-technology-policy>> [Accessed 18 October 2020].

\*(10)

Stanford Graduate School of Business, n.d. Our Mismatched Fear Of Job-Stealing Robots. [image] Available at: <<https://www.gsb.stanford.edu/insights/mismatched-fear-job-stealing-robots>> [Accessed 18 October 2020].