

# Statistics Are Useful, But Not In The Way Popular Media Suggest

By ACSH Staff — October 10, 2018



Credit: Shutterstock [1]

“Data science” is hot right now. The number of undergraduate degrees in statistics has [tripled](#) [2] in the past decade, and as a statistics professor, I can tell you that it isn’t because freshmen love statistics.

Way back in 2009, economist Hal Varian of Google dubbed statistician the “[next sexy job](#) [3].” Since then, statistician, data scientist and actuary have topped various “[best jobs](#) [4]” lists. Not to mention the enthusiastic press coverage of industry applications: [Machine learning!](#) [5] [Big data](#) [6]! [AI](#) [7]! [Deep learning](#) [8]!

But is it good advice? I’m going to voice an unpopular opinion for the sake of starting a conversation. Stats is indeed useful, but not in the way that the popular media – and all those online data science degree programs – seem to suggest.

## **Super-employees**

While all the press tends to go to the sensationalist applications – [computers that watch cat videos](#) [9], anyone? – the data science boom reflects a broad increase in demand for data literacy, as a baseline requirement for modern jobs.

The “big data era” doesn’t just mean large amounts of data; it also means increased ease and ability to collect data of all types, in all walks of life. Although the [big five](#) <sup>[10]</sup> tech companies – Google, Apple, Amazon, Facebook and Microsoft – represent about 10 percent of the U.S. market cap and dominate the public imagination, they employ only one-half of one percent of all employees.

Therefore, to be a true revolution, data science will need to infiltrate non-tech industries. And it is. The U.S. has seen its impact on political campaigns. I myself have consulted in the medical devices sector. A few years back, [Walmart held a data analysis competition](#) <sup>[11]</sup> as a recruiting tool. The need for people that can dig into the data and parse it is everywhere.

[In a speech](#) <sup>[12]</sup> at the National Academy of Sciences in 2015, Steven “Freakonomics” Levitt related his insights about the need for data-savvy workers, based on his experience as a sought-after consultant in fields ranging from the airline industry to fast food. He concluded that the next-generation super-employee is someone with a bit of business sense, a bit of computing know-how and a bit of statistics under his or her belt.

Data is increasingly being called on to inform all our decisions. But this broad utility means that it isn’t sexy. The sexy jobs – working on [self-driving cars](#) <sup>[13]</sup> or Go-playing computers – are going to require more than an undergrad major in statistics or a week-long bootcamp on prediction using Python. In fact, I was once told by an industry colleague that the term “data scientist” was coined to placate Ph.D. physicists who were tasked with running linear regressions all day long.

So, the way I see it, there will be egghead types off at the edge of the field, and there will some folks doing the necessary drudge work, and there will be a lot of people in between, looking carefully at the data and trying to glean useful insights. But – and this is the big point – everyone had better know how to make basic graphs and poke around a database.

### **So where do I sign up?**

Five years ago there was no such thing as a data science degree, and now [the list](#) <sup>[14]</sup> runs for pages and pages. And that’s not counting the traditional statistics programs, or programs in related subjects like computer science or operations research. LinkedIn’s sidebar strongly feels I should consider an online master’s degree in data analytics, from several different places.

The proliferation of these programs speaks to the inadequacy of many people’s undergraduate educations in terms of statistics and data competency. Although stats majors have tripled, there were only 3,000 last year, compared to [370,000 business degrees and 117,000 psych degrees](#) <sup>[15]</sup>. More of these students should certainly give statistics (or one of the newer data science degrees) a hard look, given that a bachelor’s degree is borderline [compulsory](#) <sup>[16]</sup> these days.

But I worry that the premise behind the appeal of these degrees – especially at the master’s level – is the idea that the technology alone can solve problems. Nothing could be farther from the truth. Statistics is a tool for understanding data, but cannot by itself understand anything. Probably the biggest mistake people make when applying statistical or machine learning methods is not recognizing that the data being analyzed is insufficient to answer the relevant question. A degree that teaches you only about the hottest predictive analytics technology, like deep learning, is a bit

like learning how to drive without knowing the first thing about how to navigate.

Setting realistic expectations for the added value of a statistics education is important to me because I'm a true believer. I feel that more people should learn statistics and how to analyze data because it is a powerful way to understand modern life. In addition to boosting one's job prospects, a statistics education can teach you when to [ignore your doctor's bad advice](#) [17], help you understand important financial ideas and, in general, help you be [wrong less often](#) [18]. These real virtues are undermined by big data hype.

So yes, lots more folks are studying statistics at the college level than in the past and, absolutely, even more people should be. But I think focusing on the surge in data science specialists is misinterpreting the nature of the demand. Everyone should have more of these skills, even if it isn't their primary job title.

By [P. Richard Hahn](#) [19], Associate Professor of Statistics, Arizona State University. This article is republished from The Conversation under a Creative Commons license. Read the [original article](#) [20]

---

COPYRIGHT © 1978-2016 BY THE AMERICAN COUNCIL ON SCIENCE AND HEALTH

---

**Source URL:** <https://www.acsh.org/news/2018/10/10/statistics-are-useful-not-way-popular-media-suggest-13489>

#### Links

[1] <https://www.shutterstock.com/image-photo/business-documents-on-office-table-smart-414153883>

[2] <http://magazine.amstat.org/blog/2018/08/01/2017-degree-report/>

[3] <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>

[4] <http://time.com/money/5114734/the-50-best-jobs-in-america-and-how-much-they-pay/>

[5] <https://www.engadget.com/2018/03/16/new-yorker-machine-learning-blocked-bike-lanes/>

[6] <https://www.theguardian.com/commentisfree/2018/sep/27/the-guardian-view-on-big-data-and-insurance-knowing-too-much>

[7] <https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html>

[8] <https://blogs.wsj.com/cio/2018/09/07/deep-learning-is-it-approaching-a-wall/>

[9] <https://www.wired.com/2012/06/google-x-neural-network/>

[10] <https://www.bloomberg.com/view/articles/2017-11-15/the-big-five-could-destroy-the-tech-ecosystem>

[11] <https://www.kaggle.com/c/walmart-recruiting-store-sales-forecasting>

[12] <https://www.youtube.com/watch?v=r5jATFtKtI8>

[13] <https://medium.com/@andrewng/self-driving-cars-are-here-aea1752b1ad0>

[14] <https://www.discoverdatascience.org/programs/bachelors-in-data-science/>

[15] [https://nces.ed.gov/programs/digest/d17/tables/dt17\\_322.10.asp](https://nces.ed.gov/programs/digest/d17/tables/dt17_322.10.asp)

[16] <https://press.princeton.edu/titles/11225.html>

[17] <https://medium.com/@mijordan3/artificial-intelligence-the-revolution-hasnt-happened-yet-5e1d5812e1e7>

[18] [https://books.google.com/books/about/How\\_Not\\_to\\_Be\\_Wrong.html?id=2k9KAgAAQBAJ](https://books.google.com/books/about/How_Not_to_Be_Wrong.html?id=2k9KAgAAQBAJ)

[19] <https://theconversation.com/profiles/p-richard-hahn-562205>

[20] <https://theconversation.com/statistics-and-data-science-degrees-overhyped-or-the-real-deal-102958>