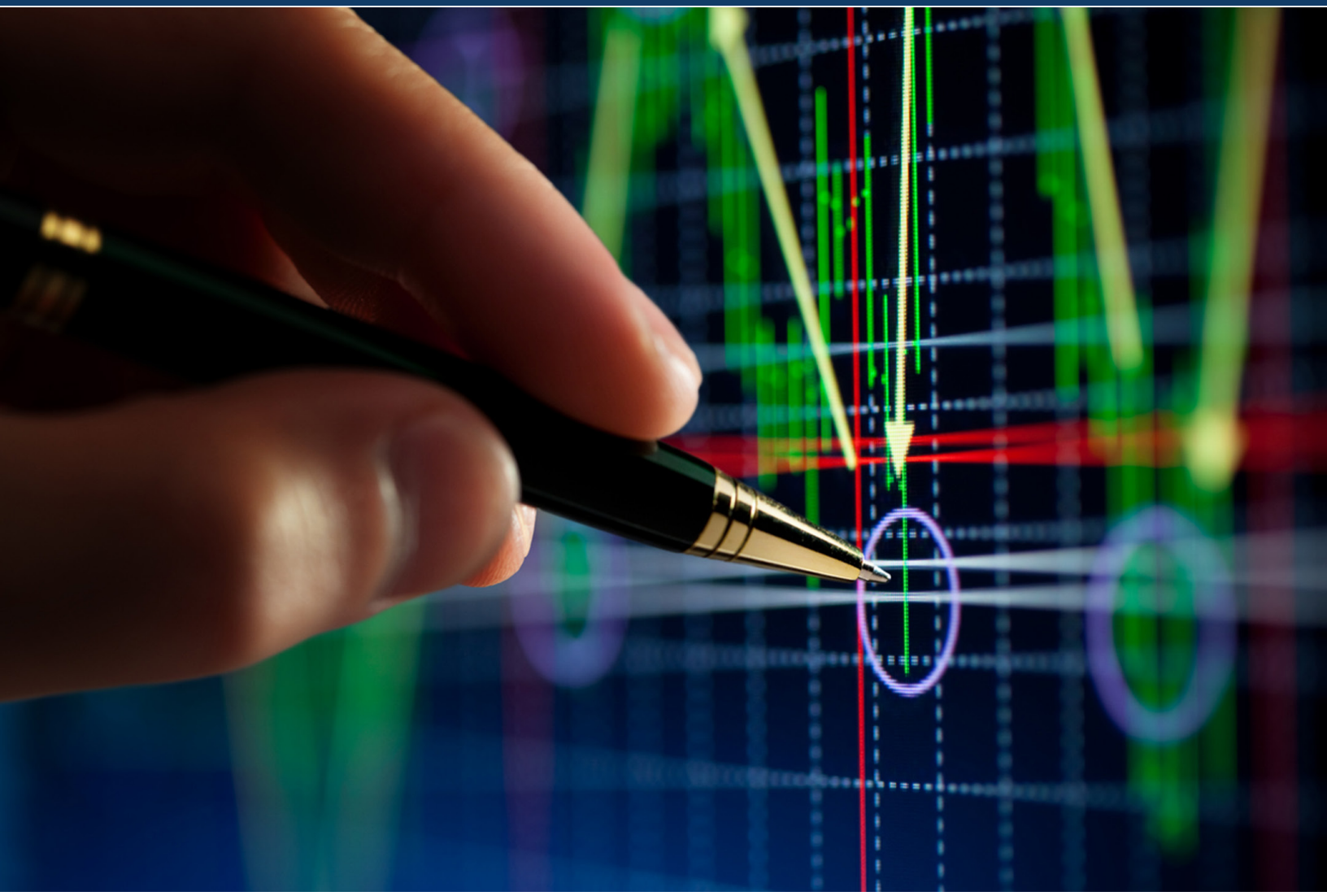
How to build a successful data scientist career



TechRepublic

How to build a successful data scientist career

Published by TechRepublic
March 2017

Disclaimer

The information contained herein has been obtained from sources believed to be reliable. CBS Interactive Inc. disclaims all warranties as to the accuracy, completeness, or adequacy of such information. CBS Interactive Inc. shall have no liability for errors, omissions, or inadequacies in the information contained herein or for the interpretations thereof. The reader assumes sole responsibility for the selection of these materials to achieve its intended results. The opinions expressed herein are subject to change without notice.

TechRepublic
9920 Corporate Campus Dr.
Suite 1000
Louisville, KY 40223
Online Customer Support:
http://techrepublic.custhelp.com/

Credits

Editor In Chief
Jason Hiner

Managing Editor
Bill Detwiler

Feature Editors
Jody Gilbert
Mary Weilage

Editorial Assistant
Amy Talbott

Graphic Designer
Kimberly Kalisik

Cover Image

iStockphoto.com/Irochka_T

Contents

Want a big data job? First decide whether to follow a tech or management career path

By Mary Shacklett

Careers in big data are beginning to solidify, as companies ramp up their analytics teams. Big data professionals usually have to make a decision early in their careers regarding whether to follow the highly technical path or pursue management and people-centric roles. Here's a look at some of the positions and career paths you should be aware of if you want to work in analytics.

Technical big data and analytics career paths

Database management: Data analysts, data modelers, DBAs, data architects

Those choosing the database path begin as data analysts and data modelers, where they work with data, prepare data, and develop data organizational schemes that use databases and data marts. They help with technical definitions of the data and data relationships so data can be aggregated appropriately for the different business functions using the data. They might also be responsible for keeping data current and secure.

Those who work in this area typically begin as junior data analysts and progress into data modeling; ultimately, they can become corporate database administrators (DBAs) or data architects, and both of these senior positions can command six-figure salaries in large enterprises.

Systems management: Systems programmers and systems engineers

Those choosing this path enjoy implementing, maintaining, and fine-tuning system hardware, operating systems, and systems software that underlie the big data applications the company runs. These people often have engineering and/or programming backgrounds, and they spend a lot of time programming systems.

The typical career path of a person working in technical systems begins as a junior systems programmer and progresses into senior levels of system programming and engineering. Those progressing to a distinguished engineer designation in a large enterprise can earn six-figure salaries.

Application development: Application developers, big data managers, data scientists

Those opting to develop end-business applications for big data and analytics use programming languages and tools designed to extract meaning from unstructured data. They often begin as junior application developers who use third-party tools for extracting data and developing simple reports and dashboards. They progress into senior positions where they develop applications from scratch, using advanced programming languages.

The ultimate position in big data application development can be a manager or a data scientist who blends concepts from engineering, arithmetic algorithm development, and statistical analysis to develop complex data queries that get answers to questions companies haven't been able to answer. Fields that use data scientists include pharmaceuticals, medicine, and other complex research and development industries. Data scientists can earn six-figure salaries.

People and management big data and analytics career paths

Business analysts and chief data officers

Those who want a nontechnical IT career path in big data and analytics often begin as junior business analysts. These people are responsible for working with end-business users in the company, coordinating with IT, and identifying the business cases and needs for analytics applications.

These individuals develop models in plain English that describe how data is to be used. These descriptions are then submitted to database analysts, application developers, and others on the technical side. Business analysts also study the business processes within the company and determine where, when, and how analytics reports should be delivered.

Junior business analysts progress into senior business analyst roles, analytics middle managers, and ultimately senior directors and even chief data officers (CDOs), which are often well compensated six-figure positions in large enterprises.

Managers of analytics-transformed user departments

The explosion of big data/analytics is giving managers of established non-IT disciplines new opportunities to grow their careers. For example: A market research analyst might be asked for market data that affects other company functions; or a GIS manager could be called upon to provide GIS information to new business areas. These managers must step out of their offices, meet with end users, and help to identify new uses for the information they have worked with for a long time. As they do this, they get new visibility and can be promoted into director and senior manager roles.

Chief data officers

Senior managers/directors on the business side of analytics can continue to build their careers into a CDO position in a major company. The CDO coordinates corporate-wide analytics and information usage, is a C-level officer with executive compensation, and often reports directly to the CEO or the board of directors.

The hottest new big data analytics jobs

By Mary Shacklett

Big data and analytics are transforming the way we work. Many HR managers are even shortening the review and revision process on job responsibilities and the task lists on position descriptions to reflect the active role that analytics plays.

Those already in the workforce should take steps to improve their skills in analytics, as more employees will be expected to work in this area. For individuals seeking their first jobs, taking an analytics course or two are strong entries to add to your resume.

The big data and analytics boom is also leading to the creation of new jobs. Here's a list of new analytics roles that some employees are filling.

Citizen data scientist

Data scientists are hard to find, and many companies can't afford to hire one. The result is the birth of the "citizen" data scientist—an analytically talented individual from the organization who does not have a formal degree in data science or engineering but who takes on the mantle of developing complex algorithms and queries of data that can yield breakthrough information for the company.

In April 2015, Gartner research analyst Alexander Linden described citizen data scientists as "people on the business side that may have some data skills, possibly from a math or even social science degree—and putting them to work exploring and analyzing data." Shawn Rogers, chief research officer at Dell Statistica, predicted a rise in the number of citizen data scientists "because users throughout the business want a more democratized approach to big data and analytics. Not every company can afford a data scientist, which is a big reason why citizen data scientists will become a big part of the data ecosystem as it evolves."

IoT specialist

Manufacturing engineers and other machine operators and technicians at the edges of enterprises where machines are being used may soon take on roles that involve harvesting information from sensors in these machines and moving the sensor-based input into software and systems that are running machines, coordinating machine-based operations and handoffs, and checking on machine health.

Other IoT roles will include engaging and programming advanced robotics and even stationary field-based sensors that report on environmentals at remote locations. IoT will transform field-based technician and engineering jobs beyond the current boundaries of machine maintenance and operation and into data gathering and software-based analytics that feed into onsite and remote, centralized systems.

Data hygienist

Think about your electronic toothbrush for a second—it does a pretty good job, but there are always those hard-to-reach areas you have to go after yourself. Big data is no different. There are tools for big data cleaning, sorting, scrubbing, deduplication, etc., but to really get data to the level of business precision organizations want, employees with first-hand knowledge of the business must refine the data by hand.

These "by hand" employees who refine and clean up data come from administrative and clerical functions and will increasingly assume the role of internal data hygienists. Along the way, they will pick up new skills in data preparation and classification.

Data orchestrator

Most organizations have data architects who come from DBA ranks and create an overall architecture of the data throughout the enterprise. What's needed now is someone who can orchestrate the *movements* of this data. How much data from the edges of the enterprise will need to be locally stored there, and how much will need to be instantaneously transferred to different points throughout the enterprise?

For data that requires real-time or near-real-time velocity, the questions that need to be answered are: How much data is there, and are the enterprise's data pipelines and systems adequate to transport and store it? There is no formal job title for this function, but increasingly, application developers and systems analysts will assume key roles in determining the different speeds and resting places of data throughout the enterprise.

Paraprofessional analytics

Analytics machines like IBM Watson are already assisting with medical diagnoses and legal research. In law, medicine, and other fields, these analytics are generating new forms of research work that paraprofessionals (e.g., physician assistants, paralegals, etc.) in organizations will most likely perform.

Machine teachers

As analytics machines like IBM Watson play greater roles in organizations, they will have to be continuously "taught" so they can do their work better. In many cases, machine learning can be done by the machines, but in some cases, this learning will have be guided by a new generation of "teachers" culled from the ranks of business analysts and subject matter experts within the enterprise.

Six big data trends to watch in 2017

By Mary Shacklett

Major strides were made in big data and analytics last year, and companies will expect even more from those types of projects in 2017. If you're seeking a big data job role, it's a good idea to keep a close eye on the emerging trends. Here are six big data and analytics trends you should be aware of.

1: Movement to the cloud

Small and midsize companies and even large enterprises are mapping strategies that take more of their applications to the cloud and out of the data center—and this holds true for big data and analytics as much as it does for traditional transaction processing systems. Companies want to see reduced spend in their data centers and greater flexibility in terms of plugging into and out of solutions. The ability to do this comes with subscriptions to services and not having to lock-in for multiple years to on-premises equipment.

An additional factor for big data and analytics is the difficulty that even large organizations have in finding the talent to run in-house Hadoop clusters and processing. This is forcing many organizations to go to the cloud and to cloud services providers that offer the big data processing platform as well as the expertise.

2: Aggregation of digital unstructured and machine IoT data

For all that's been written about it, using and incorporating Internet of Things (IoT) data is a future endeavor for most organizations. What we *do* know is that everyone is thinking about it.

Organizations' big data aggregation goals will expand to visions where standard digital data originally entered by humans and data issued from machines will be aggregated into composite visualizations that will transform the way work is done. A good example is drone-hosted data that will combine an assortment of sensory and standard IT inputs into a single-pane-of-glass view for an operator of how a drone is functioning. Big data and analytics vendors and consultants will be called upon to assist companies in defining and achieving these new data aggregation goals.

3: The use of more dark data

Companies will begin to troll the wealth of information contained in paper-based documents, photos, videos, and other corporate assets that are lying dormant in vaults and storage closets but that could be put to use in big data aggregation. These assets can give organizations a more comprehensive view of historical performance trends and product cycles that can be useful for planning. The data can also provide supporting evidence for trademark infringement and/or intellectual property violation claims.

4: Stronger administration of data security permissions

As more big data moves to data warehouses and repositories, the goal is a "single version of the truth," where users all use the same data but can't necessarily access all of it. Companies will tighten up data access permissions to ensure that each data user has the correct access permissions in place. This will likely involve creating or revising data access permissions policies and implementing technology that monitors and detects potential data exfiltration by users. Data exfiltration is a process in which users—without authorization—copy, transfer, or retrieve data that exceeds their access clearances.

5: Immediately gratifying analytics

Like it or not, executives and line managers want to see analytics that give them immediately actionable data. They don't want to wait for batch analytics reports, which the majority of big data analytics still is. The pressure will be on IT to deliver actionable analytics results faster and to focus more big data and analytics activities on real-time or near real-time data.

6: "Where's the beef?" business evaluations of big data projects

In 1984, hamburger chain Wendy's rolled out a "Where's the beef?" commercial that focused on the amount of meat customers were actually getting on their sandwiches. Many CIOs and big data project leaders will be encountering their own "Where's the beef?" moments as they get called in by CEOs, board members, and other chief stakeholders to demonstrate how big data and analytics projects that the company faithfully invested in are now delivering tangible value to the end business. If tech leaders can show only projects that were successful in terms of IT execution, without significant returns on investment for the end business, they could find themselves in the line of fire.

Conclusion

Big data and analytics endeavors can be enticing and exciting, but they must still pay off for companies. More analytics systems will be moved into the mission-critical category in 2017, but they will also be expected to perform well operationally, meet governance standards, and fulfill promises of business value to the company.

Five ethics principles big data analysts should follow

By Michael Kassner

Big data is not only big—it is also powerful and error prone, said Susan Etlinger, an industry analyst with Altimeter Group, in her 2014 TED talk. "At this point in our history... we can process exabytes of data at lightning speed, which also means we have the potential to make bad decisions far more quickly, efficiently, and with far greater impact than we did in the past."

Besides the potential for bad decisions, Etlinger believes that humans place too much faith in technology, including, for example, our blind acceptance of charts and graphs developed from big data analysis.

An ethical framework for big data analysis

As to what might be done to improve the situation, Etlinger and Jessica Groopman write in their Altimeter report The Trust Imperative: A Framework for Ethical Data Use (PDF) that businesses and organizations building and/or using big data platforms need to start adhering to ethical principles.

To incorporate ethics, Etlinger and Groopman suggest studying The Information Accountability Foundation's (IAF) paper A Unified Ethical Frame for Big Data Analysis, and paying particular attention to the following principles:

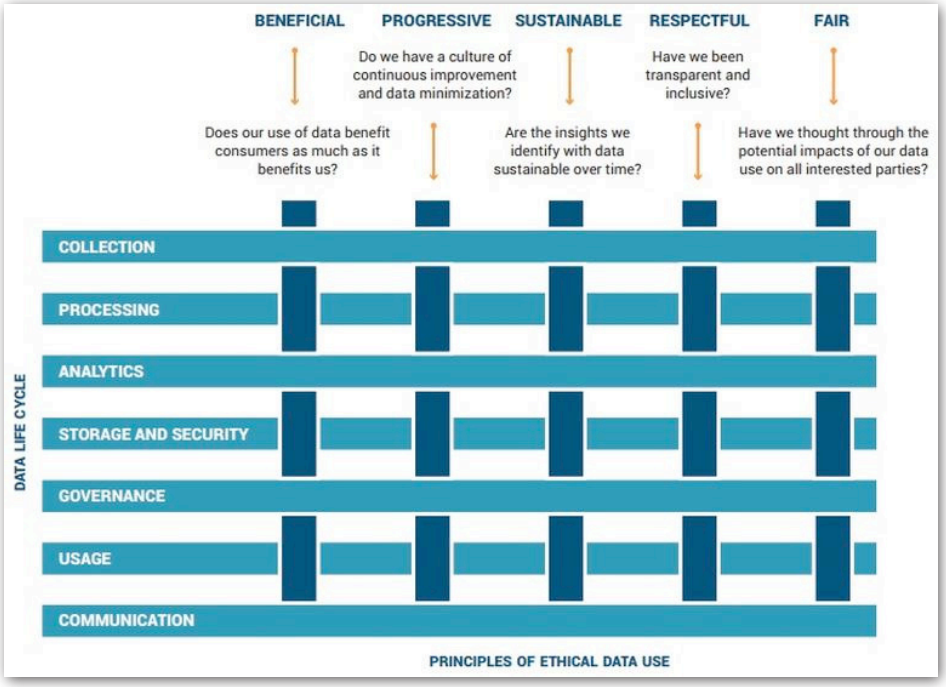


Image: Altimeter, Susan Etlinger, and Jessica Groopman

1. Beneficial

"Data scientists, along with others in an organization, should be able to define the usefulness or merit that comes from solving the problem so it might be evaluated appropriately." (IAF)

"The first principle for ethical data use is that it should be done with an expectation of tangible benefit," Etlinger and Groopman said. "Ideally, it should deliver value to all concerned parties—the individuals who generated the data as well as the organization that collects and analyzes it."

They offered Caesars Entertainment as an example. Joshua Kanter, senior vice president of revenue acceleration at Caesars Entertainment, said, "Before conducting any new analysis, we ask ourselves whether it will bring benefit to customers in addition to the company. If it doesn't, we won't do it."

2. Progressive

"If the anticipated improvements can be achieved in a less data-intensive manner, then less-intensive processing should be pursued." (IAF)

The value of progressiveness, according to Etlinger and Groopman, relies on the following:

The expectation of continuous improvement and innovation: In other words, what organizations learn from applying big data should help deliver better and more valuable results.

Minimizing data usage: Businesses should use the least amount of data necessary to meet the desired objective, with the understanding that minimizing data usage promotes more sustainable and less risky analysis.

The above principles should help eliminate hidden insights or correlations, such as disenfranchising individuals based on race or demographics.

3. Sustainable

"Big-data insights, when placed into production, should provide value that is sustainable over a reasonable time frame." (IAF)

Sustainability, the authors said, is broken down into these categories:

Data sustainability: Sustaining value is closely related to the access organizations have to different social data sets. "While this is a fact of access and economics, it can wreak havoc when sets of data from public and private sources are combined," Etlinger and Groopman said. "The issue of sourcing also comes into play.... Inconsistencies in sample sizes or methodologies affect the integrity of the data and the sustainability of the algorithm."

Algorithmic sustainability: A critical element of sustainability is an algorithm's longevity. The Altimeter report suggests longevity is affected by how the data is collected and analyzed.

Device- and/or manufacturer-based sustainability: A third consideration is the lifespan of the data being collected. "For example, if a company develops a wearable or other networked devices that collect and transmit data, what happens if that product is discontinued, or the company is sold, and the data is auctioned off to a third party?" Etlinger and Groopman asked.

4. Respectful

"Big-data analytics affect individuals to whom the data pertains, organizations that originate the data, organizations that aggregate the data, and those that might regulate the data in different ways." (IAF)

"The advent of social and device-generated data captured in real time decimates the norms for data analytics.... As a result, even seemingly minor decisions can have tremendous downstream implications," Etlinger and Groopman said.

As can be expected, the individual who originated the data will be affected the most by big data analysis, in particular making private, semi-private, or even public information more public.

5. Fair

"In lending and employment, United States law prohibits discrimination based on gender, race, genetics, or age. Yet, big data processes can predict all of those characteristics without actually looking for fields labeled gender, race, or age." (IAF)

Etlinger and Groopman considered the ability to predict characteristics at any level just by asking for what they call "unintended consequences." To counter unintended consequences, they again used Caesars Entertainment as an example:

"Caesars has a simple yet effective litmus test for fairness, which it calls the Sunshine Test—whether the issue can be discussed openly and the final decision disclosed without any sense of misgiving. Before deciding on a course of action that requires customer data, the company's executives imagine how people would react if all of the details were out in the open, in the light of day. Would it strengthen or threaten customer relationships?"

Joshua Kanter added, "If the initiative fails the Sunshine Test, we do not move forward."

Final thoughts

Etlinger and Groopman suggested that applying the five principles of ethical data use is a pragmatic approach for businesses. "At the same time, data complexity, differences in business models, emerging technologies, and most importantly, people, mean that no single approach will address every scenario."

The 20 best schools to study big data analytics

By Conner Forrest

Big data, data analytics, data science—whichever direction you want to take it, there's no denying that the technology fields around data and data services are hot. Data is the currency of the enterprise, and more organizations than ever are trying to get some value out of it.

Many folks in tech careers who want to pursue a job in big data often head back to school to get the proper credentials. Value Colleges, an online publication that compares universities, recently released its list of the Top 50 Best Value Big Data Graduate Programs of 2016, where it details the best schools to learn about data.

Here are the top 20.

1. Carnegie Mellon University

Carnegie Mellon University's Heinz College houses the School of Information Systems and Management, where students can earn a master's degree in information systems management with a focus in data analytics. The school also offers a program for aspiring CISOs.

2. Stanford University

One of the building blocks of Silicon Valley, it's no surprise that Stanford is on this list. Stanford's master of science degree in statistics with a concentration in big data, combined with their ties to the entrepreneurial community, make it a great fit for someone who wants to work in startups.

3. Santa Clara University

Santa Clara University is one of the oldest schools in the state of California. Its Leavey School of Business offers students a master of science degree in business analytics, which is one of the best in the country.

4. University of Michigan - Dearborn

Michigan isn't just the home of the biggest automakers, it's also a good place to learn about big data. The university's Dearborn campus offers a master of science in business analytics, and is fairly well priced compared to the private universities on the list.

5. University of Texas at Dallas

They say everything's bigger in Texas, and that includes the data. The University of Texas at Dallas Jindal School of Management is a great option to pursue a graduate degree in analytics, as the surrounding Dallas economy is home to many major businesses.

6. University of Virginia

The University of Virginia is known, all-round, as a quality school, and its Data Science Institute is a great place to study in the field. Its Master of Science in Data Science (MSDS) is only 10 months long, offering a quick turnaround for professionals.

7. University of Florida

Coming in at no. 7 is the University of Florida, where students can pursue a master of science in information systems and operations management. The university is also routinely ranked as a top business school, making it a great option for those who wish to also pursue management down the line.

8. Purdue University

Considered by some to be the Ivy League school of the Midwest, Purdue is a great place to study data analytics.

9. University of Maryland

One the whole, the University of Maryland is known as a great public university, but it also has a strong graduate program for aspiring data scientists. Also, being close to Washington, DC means that government jobs are a great place to start your career.

10. Georgia Institute of Technology

The Georgia Institute of Technology Master of Science in Analytics is a great interdisciplinary degree, and nearby Atlanta is home to a booming tech job market.

11. University of California - San Diego

Most of the schools in the University of California fold are known for being tech-centric, but UCSD's Rady School of Management offers students an in-depth graduate degree business analytics.

12. Bentley University

Bentley University is known for its history in accounting, but it also is a data powerhouse. The school was also ranked by Value Colleges as a top school for UX design, for students looking to broaden their skills.

13. Oklahoma State University

At Oklahoma State University, students can take two paths to data dominance: a master's degree in management information systems or in business analytics. The low cost of tuition is an added perk.

14. Northwestern University

Northwestern University's McCormick School is a great option for an advanced degree in analytics. The school is also known for business and journalism, which could provide additional skills in management or data reporting.

15. North Carolina State University

The birthplace of SAS, North Carolina State University's Institute for Advanced Analytics has been working with data for a long time. Also, the state's Research Triangle is home to many tech companies that are likely looking for data analysts.

16. Indiana University - Bloomington

Indiana University's main campus in Bloomington hosts the School of Informatics and Computing, where students can get a master's degree in data science in only 30 credit hours. The school is also known for its focus on business, making it a great option for students who are also looking to go into management.

17. Worcester Polytechnic Institute

Though a smaller school, Worcester Polytechnic Institute in Worcester, MA, offers one of the best master's degree in data science and analytics.

18. Univerity of Massachusetts - Amherst

Regarded as one of the best research universities in the nation, University of Massachusetts Amherst is a great place to study data science as well. Its College of Information and Computer Sciences has been highly ranked as a research institution.

19. Seattle University

Like many schools on this list, Seattle University offers a master's degree in business analytics. However, the close proximity to tech giants like Microsoft and Amazon make it a strategic geographic location to start your big data journey.

20. George Mason University

Rounding out the top 20 is George Mason University. Known for its work in computing, business, and economics, George Mason is a great option for students who want to use data analytics in business environment.

Why data science is just grade school math and writing

By Matt Asay

Say the words “data science” and images of number-crunching propellerheads come to mind, or rocket scientists moonlighting as interpreters of an enterprise’s Hadoop cluster. But according to Noah Lorang, a data scientist for Basecamp, “Data scientists mostly just do arithmetic.”

Basic math? Really? Is that what companies are paying outsize salaries to recruit and retain? People who can add and subtract?

Yes there are two paths you can go by...

The answer, of course, is “maybe.” It depends on the target audience. As I’ve outlined before, data science breaks down into two categories: Data science intended for human consumption and data science intended for machine consumption.

For the latter audience, data science “involves complex digital models that ingest large amounts of data and extract insights using machine learning and algorithms, then act autonomously to display certain ads or make stock trades in real time.” As such, machine-oriented data scientists require “exceptionally strong mathematical, statistical, and computational fluency to build models that can quickly make good predictions,” as former Google and Foursquare data scientist Michael Li wrote.

This may be the vision we have of data scientists, generally, but the aforementioned skillset is different from that required to thrive in human-oriented data science.

“[N]umbers have no way of speaking for themselves. We speak for them. We imbue them with meaning,” noted statistician Nate Silver. We bias our data the minute we start collecting it, as we determine what we’ll collect, not to mention the types of questions we’ll ask of it. There is no such thing as unbiased data, be it machine-oriented or human-oriented.

Bias is the natural state of all data.

Once we understand this, the role of the human-oriented data scientist becomes clear: Help the data tell clear stories. In an interview with Zoomdata CEO Justin Langseth, he warned against the facile expectations of machine-driven data science, holding that “algorithmic insight... generates false positives that drive any human reviewer nuts or cause them to not trust the system.”

By contrast, data visualization, with its explicit human involvement, facilitates “exploration [which can] lead... to ‘aha insights.’”

In short, good data science requires good storytelling and data visualization. All of which starts with basic math.

And it makes me wonder...

This brings us back to Lorang, who suggests, “In the last two weeks, the most ‘sophisticated’ math I’ve done has been a few power analyses and significance tests.”

So what does he spend all his well-paid data scientist time doing?

“Mostly what I’ve done is write SQL queries to get data, performed basic arithmetic on that data (computing differences, percentiles, etc.), graphed the results, and wrote paragraphs of explanation or recommendation,” Lorang said.

Really? Really.

“I haven’t coded up any algorithms, built any recommendation engines, deployed a deep learning system, or built a neural net.”

While he leaves room for more “sophisticated” data science down the road, he insists that Basecamp doesn’t need it now and others probably don’t, either:

The dirty little secret of the ongoing “data science” boom is that most of what people talk about as being data science isn’t what businesses actually need. Businesses need accurate and actionable information to help them make decisions about how they spend their time and resources. There is a small subset of business problems that are best solved by machine learning; most of them just need good data and an understanding of what it means that is best gained using simple methods.

In his view, businesses need to better understand their data, which is an inherently human problem. Langseth echoed this sentiment when he told me,”The best [data] visual[ization] is the one that allows a normal human with understanding of a business system to quickly see how the visuals match up with the system.”

Ultimately, Lorang said, “Knowing what matters is the real key to being an effective data scientist.” And that, it would seem, generally comes down to common sense, a bit of math, and the ability to tell a story with data.

Six things to avoid when heading up a big data project

By Mary Shacklett

Most organizations have a set of big data best practices they have formulated from their successful project work. But an equally important list outlines the pitfalls that organizations should stay away from when it comes to big data and analytics. Here are six don'ts to keep in mind during your big data projects.

1: Swinging for the fences

The most successful big data initiatives build a strong foundation for big data and analytics and use them. The best way to do this is by creating a constant path of new big data deliverables that incrementally and continuously improve the organization's ability to tackle strategies and operational issues with richer and better data.

2: Making things unnecessarily complicated

Dashboard and spreadsheet-style data delivery that enables the end business to drill down into data and ask more questions works exceptionally well. A big reason why is that users are already familiar with these types of data capture and manipulation tools.

The more at ease users are with the tools they have to access and manipulate data, the more they will believe in and adopt big data and analytics.

3: Bypassing security as a project consideration

Security is one of the largest missing pieces in big data projects. These are some security questions to consider:

What types of security and risks are in play when big data begins to get captured and moved off machines at the edge of the enterprise—or from outside of it?

How secure is your internal data preparation environment? Do only authorized users have access to it?

How do you vet the many types of unstructured data to ensure that it is tamper-proof?

If you are aggregating data from third-party vendors, what levels of security and governance do they use for their data? There are so many diverse sources and types of big data, so security of this data is still an area that most enterprises are struggling with.

4: Pursuing projects without end user engagement

If you don't know the critical questions that areas of the end business want to solve with big data, you can't deliver the solutions. Engage heavily with end users about the nagging questions in the business and collaborate with them as you strategize about how to obtain and extract information from big data.

5: Working with low-confidence data

If data isn't properly cleaned and vetted for accuracy, the results it could indicate might be erroneous and catastrophic to your company.

Remember the New Coke initiative of the 1980s? A market research team conducted more than 200,000 taste tests that showed participants preferred New Coke over both Classic Coke and Pepsi. Coke soon learned that taste preference wasn't the only factor that went into consumer purchasing decisions—tradition was also a major factor.

In the end, the $4 million that had been sunk into New Coke development was wasted, as was another $30 million of New Coke syrup that sat on the shelf while Coke restored its Classic Coke.

6: Limiting your innovation

Big data projects, like all projects, must return value and show results. Consequently, those leading big data projects tend to focus on areas of low-hanging fruit where they know they can produce results quickly.

It's also important to keep some experimental big data work going. Why? Because experimental big data work—where there is no immediate pressure of timelines or results—has the potential to produce breakthroughs.

To pursue big data innovations, you need buy-in from the CEO and other corporate managers. These projects have high failure rates that everyone must accept. You also need a mechanism within these R&D projects to pull the plug as soon as you see they aren't going to produce results.

About TechRepublic

TechRepublic is a digital publication and online community that empowers the people of business and technology. It provides analysis, tips, best practices, and case studies aimed at helping leaders make better decisions about technology.

Resources

Subscribe to our free newsletters: Stay on top of business technology trends, learn about innovative new products, and hone your skills with our how-to's and tutorials.

Check out the TechRepublic discussion forums: Touch base with your peers and share tips, advice, solutions, and opinions.

Catch the latest videos and photo galleries: Our video library offers interviews with entrepreneurs, IT pros, and CXOs; short clips on the latest tech news; and overviews of emerging technologies. Our galleries offer a look at everything from the hottest mobile devices to autonomous cars to the gadgets, tools, and accessories that are headed your way.