MOBILE

BIG DATA
R E P U B L I C | Transform Your Business With Data

HOME   BLOGGERS   MESSAGES   POLLS   WEBINARS   RESOURCES   VIDEOS | STRATEGY   ANALYTICS   TECHNOLOGY   FB   TW   LI   G+   RSS

FINANCIAL SERVICES   HEALTHCARE   GOVERNMENT   EDUCATION   RETAIL

DATA SCIENTIST

Data Scientist Wanted: Are You Qualified?

Robert Plant, Associate Professor, School of Business Administration, University of Miami
8/14/2013
Comment
30 comments

Login

👍 👎
100%   0%

With the recent explosion in public awareness of big data, it's only natural that a career in this emergent area is attractive. But what exactly is a data scientist? It's a question I'm frequently asked. My students and professional colleagues are intrigued by the possibilities it may offer them.

According to Tom Davenport and D.J. Patil, a data scientist is "The Sexiest Job of the 21st century," quite a pinnacle for anyone to live up to. Patil and Jeff Hammerbacker, who were the leads of data and analytics at LinkedIn and Facebook respectfully, developed the title "data scientist" with respect to "a high ranking professional with training and curiosity to make discoveries in the world of big data."

The traits of a data scientist
They describe the data scientist as having traits encompassing "a hybrid of hacker, analyst, communicator and trusted advisor." In addition, they also indicate the need to be able to code, and have an intellectual curiosity.

Davenport and Patil's definition is a useful guide for executives, but in practical terms, to understand what a data scientist is, we need to drill down further. One way to see what data scientist means in execution is to examine how leading companies are recruiting and developing talent; beyond employing a singular individual titled Data Scientist.

One company that has led in this talent management area is Netflix. An examination of its openings (as of August 4th, 2013) reveals 22 positions in nine categories within its Data Science and Engineering Department. These encompass a wide range of position descriptions:

- Data Science Production Platform Enginee;
- Director -- Algorithm Experimentation and Analytics
- Director -- Marketing Analytics; Director -- Content Analytics
- Marketing Experimentation Lead
- Senior Business Intelligence Engineer, Customer Service Analytics;
- Senior Data Analyst -- Acquisition & Retention
- Senior Data Analyst -- Device Analytics
- Senior Data Analyst -- Marketing
- Senior Data Engineer -- Acquisition & Retention
- Senior Data Engineer -- Product Analytics
- Senior Data Engineer/Scientist -- Streaming Activity
- Senior Data Engineer -- Experimentation & Testing
- Senior Data Integration Architect -- Content
- Senior Data Scientist -- Experimentation
- Senior Data Tools Engineer
- Senior Engineer/Architect -- Data Platform Architecture
- Senior Software Engineer -- Platform Architecture
- Senior Software Engineer -- Data Pipeline/Hadoop Infrastructure
- Senior Software Engineer -- Big Data Platform Support
- Senior Software Engineer/Architect -- Big Data Platform
- Solutions Architect -- Big Data Platform

Data science team
From this exhaustive list, it becomes clear that the idea of corporate data science being based on a single data scientist entity is misplaced. In reality, it's a team of experts with a common goal: the use of data to drive decision making in the business. This is captured in the definition of the responsibilities Netflix requires of its Director of Marketing Analytics:

The Marketing Analytics team has broad responsibility to investigate and understand the drivers that motivate our customers to join, watch, retain, and leave Netflix. You'll work

MORE BLOGS FROM ROBERT PLANT

A Review: Big Data in 2013
39 comments
Data scientists, NSA, C-suite buy in... Big data changed a lot in the last 12 months.

What Would Seymour Cray Think of Big Data?
13 comments
How would the tech legend respond to the 4 Vs?

The 360 Degree View of a Customer
56 comments
Big data is helping companies like Netflix see what you really want (and even when you go to sleep).

Lessons for Big Data From President Obama's Healthcare Implementation
24 comments
Three big data lessons for CEOs from HealthCare.gov's bumpy start.

Is Big Data Ruining Sports?
29 comments
From sailing to the NBA -- big data could increase the gap between sporting haves and have-nots (while ruining the magic in the process).

More from Robert Plant

FLASH POLL

All Polls

DIGITAL AUDIO

Data Scientist Radio
Latest Archived Broadcast
Your Big Data Initiative Starts With a Great Team
The Big Data Conference and Nemertes Research have teamed up to provide expert advice on building a big data team and finding the right leader. Join us on Tuesday, Aug. 20.
Listen

Full schedule   Archived shows   Download to calendar

BDR IN YOUR INBOX

Enter email   SUBMIT

FOLLOW US ON TWITTER

closely with the CMO and the rest of the executive team to inform and influence our brand and marketing messages as well as the features and experience within the Netflix product itself.

It also defines the desired skills in terms of the Davenport-Patil definition: *"well balanced skills," "Analytics expert," "Confident and well-reasoned"* and *"charismatic."*

Other members of the team fit the jigsaw of data analytic needs. The Senior Data Engineer -- Experimentation & Testing is a hybrid computer science, statistics, and experimental philosopher. Netflix does not even ask for a degree, although it's unlikely the skills will have been gained without formal training.

Technically, Netflix is looking for skills including:

Cloud technologies such as MapReduce/Hadoop/HIVE/Pig; Big data & production support; Strong SQL skills; Python/Perl; Analytics, statistics, and machine learning; App development, web sites, and software development; and A/B or multivariate testing is a plus.

But it's interesting that Netflix is not looking for an introverted computer scientist or management scientist, rather, for a true team member with motivation and drive. It defines the ideal candidate as:

You're tenacious, persistent, & industrious; not just talented. Micromanagement makes your skin crawl. Given context, you're capable of self-direction. You leave your desk & collaborate (not to be mistaken with extroversion, but you realize the effectiveness of face -to-face contact in communication). You move seamlessly between code and analytical concepts. You're an active listener who asks good questions and puts the company first. You appreciate beauty & simplicity. You're known as a creative, innovative, & out-of-the-box thinker.

Starting out in a data science team

From a technical computer scientist perspective, the Solutions Architect -- Big Data Platform position may well be the best starting point from which to join the team. Netflix states its technical goal for this position as follows:

The Big Data Platform team at Netflix provides a reliable platform that powers search and recommendations and surfaces product and consumer insights from a petabyte scale data-warehouse. Tens of billions of events are processed daily and the volume is climbing rapidly as we grow our global subscriber base. Our big data analytics infrastructure in the cloud is built upon open source technologies like Hadoop, Hive, Pig, Genie, and Lipstick. Particularly, Genie and Lipstick are projects open sourced by the Data Platform team at Netflix.

Thankfully, today's lattice organizational structures allow a technologist to gain adjacent skills in order to integrate deeply into the team as a whole and move onward within the organization if they should so desire.

In light of this (admittedly small sample) dataset, a data scientist as a singular company-saving entity should be considered an abstract notion. Companies would do well to realize the ability of an individual to perform all the functions required to capture, analyze, execute, and develop strategies for the company based on big data sets as extremely challenging and unlikely to work in practice.

The collective jigsaw that Netflix has adopted in its approach to talent is more likely to pay dividends, as the skill sets needed to execute a big data solution are clearly too wide and varied for an individual to own alone, and the cross fertilization of ideas between team members is an intensely important part of the process.

For an individual looking for challenging and rewarding work, the message is to be a team player, have eclectic skills, think outside the box, and enjoy the interaction with a heterogeneous set of individuals. They may come from a previous career path in techno-analytics or through a technical background (i.e. an undergrad computer science degree), with perhaps an MBA or advanced degree in management science, picking up additional skills in Hadoop and Pig along the way. Clearly, you don't have to fit a specific mold.

If you have these skills, firms will define a data scientist type role around you within their big data landscape; which can only be a good thing. And if you so desire, you can still tell people you have the sexiest job in the 21st century!

Related posts:

- What Is Machine Learning?
- That's Show Viz
- Free Big Data Education: A Data Science Perspective

Email This   Print   Comment

COMMENTS

Newest First | Oldest First | Threaded View

PAGE 1 / 3 > >>

technetronic, User Rank: Blogger
9/28/2013 | 5:00:56 PM

Login

50% 50%

Re: Several employees in one
see i don't think that universities are where day-by-day innovation will come from. unless they adapt too, universities are behemoths where we'll see the future, but not tomorrow.

to adapt a workforce -- and we need to adapt a few million people -- it'll come from private certificate programs and personal initiative, a la Khan Academy for adults

Reply | Post Message | Edit/Delete | Messages List | Start a Board

MDMConsult, User Rank: Exabyte Executive
9/2/2013 | 4:07:19 AM

Login

50% 50%

Re: Several employees in one
@AlphaEdge Yes, and as the industry grows new and better tools should make it more convenient and easier for data analysts to learn. If universities can adapt faster as areas such as big data rapidly grow it will help education.

ACCOLADES

MEDIA PIONEER AWARDS WINNER 2013   Media Pioneer Award Winner

Reply | Post Message | Edit/Delete | Messages List | Start a Board

Pradeepta Mishra, User Rank: Exabyte Executive
8/31/2013 | 2:40:37 PM

Login

Re: Several employees in one
To me data scientist as defined by Davenport is an individual with multiple qualities which can enable him to deduce and derive actionable insights for business decisions and subsequently implementing the rules to maximize the profit.

50%   50%   Reply | Post Message | Edit/Delete | Messages List | Start a Board

AlphaEdge, User Rank: Exabyte Executive
8/26/2013 | 9:23:46 AM

Login

Re: Several employees in one
That's a good point. There are always more to learn.

50%   50%   Reply | Post Message | Edit/Delete | Messages List | Start a Board

James M. Connolly, User Rank: Blogger
8/23/2013 | 10:30:39 AM

Login

Re: Several employees in one
@Dcawrey. You have a valid concern about the data scientist role, and whether someone with a heavily scientific background can take a leadership role on the business side. It looks like two things could be happening. First, the data scientist role is really still taking shape. We're setting awfully high expectations for a person, and we will see over the next year or two how the job shakes out. Second, I'll toss out the possibility that maybe the data scientist role isn't one for an individual but rather it might be an aggregate role, filled by the team.

That being said, I know people had similar doubts about the CIO being able to bridge technology and business. That role has worked out very well in some situations and not so well in others. Plus, I don't recall when the CTO role evolved. However, someone may remember whether there were similar doubts about the ability of an engineer to take the lead in meeting market needs.

Reply | Post Message | Edit/Delete | Messages List | Start a Board

technetronic, User Rank: Blogger
8/22/2013 | 7:25:50 PM

Login

Re: Several employees in one
Good point dc...I agree with that, in that the desired roles are blurring...can/should scientists need to speak directly with marketers? this is why we get bureaucracies and layers of management.

then again, "data scientist" is becoming more of a team sport anyway, so you put the right mix of people on a particular team (as the great physicists of our time often did) to get better results

50%   50%   Reply | Post Message | Edit/Delete | Messages List | Start a Board

dcawrey, User Rank: Exabyte Executive
8/22/2013 | 6:04:08 PM

Login

Re: Several employees in one
It seems almost as if the name "data scientist" is not the right term for this position. I only say that because it seems very much that these roles require someone who is outgoing and very business-focused. No offense, but some of these qualities simply aren't exuded by those under the name "scientist". I know that this is a bit of a generalization, but aren't there others that also think this is true?

50%   50%   Reply | Post Message | Edit/Delete | Messages List | Start a Board

technetronic, User Rank: Blogger
8/21/2013 | 7:59:29 PM

Login

Re: Several employees in one
One reason why I think data scientists are here to stay:

Quantum computing doesn't solve certain types of problems (http://www.quora.com/Mathematics/What-are-some-problems-that-even-theoretically-a-quantum-computer-will-find-hard-or-impossible-to-do)

The need for human data processing will be around for a while!

50%   50%   Reply | Post Message | Edit/Delete | Messages List | Start a Board

James M. Connolly, User Rank: Blogger
8/21/2013 | 9:16:13 AM

Login

Re: Several employees in one
@MDMconsult. Right, a great analyst is always looking for new problems to solve and new things to learn.

50%   50%   Reply | Post Message | Edit/Delete | Messages List | Start a Board

MDMConsult, User Rank: Exabyte Executive
8/19/2013 | 5:53:20 PM

Login

Re: Several employees in one
Agreed. Any analyst would always look for more and more to learn. Most experienced analysts have the ability to master key programs. Being able to be exposed to a plethora of knowledge is important and the "Unaccompanied" approach is definitely a successful way to further cultivate curiosity and growth.

50%   50%   Reply | Post Message | Edit/Delete | Messages List | Start a Board

BIG DATA REPUBLIC

ABOUT US | CONTACT US | HELP | REGISTER
TWITTER | FACEBOOK | LINKED IN | GOOGLE+ | RSS

in partnership with **InformationWeek**   UBM Tech   powered by **UBM DEUSM**

bigdatarepublic : /bigdatarepublic/section/2635 : /bigdatarepublic/section/2635/266472

UBM TECH

OUR MARKETS: **Business Technology** | **Electronics** | **Game & App Development**

Working With Us: Advertising Contacts | Event Calendar | Tech Marketing Solutions | Corporate Site | Contact Us / Feedback