

: *What are the skills of a data scientist?*

by Ciara Heavin

Managing Editor, Journal of Decision Systems

and

Daniel J. Power

Editor, DSSResources.COM

Like many of the IT professional roles, much jargon and many buzzwords are used to describe the role of a data scientist. The term was first discussed by Cleveland (2001). Davenport and Patil (2012) extolled the data scientist as the “Sexiest job of the 21st Century”. They noted a data scientist is “a high-ranking professional with the training and curiosity to make discoveries in the world of big data.” Today, organisations small and large across industry sectors continue to be challenged to define the role and skills of a data scientist. Employers need to ask some difficult questions about data scientists, including what and how much domain expertise does a data scientist need to have? What can a data scientist do for my organisation? and How can I attract and keep a good data scientist?

Almost 400 job titles have been identified for the role of data scientist (Granville, 2014). Many of these job titles include the word “analytics”. The volume and variety of job titles could be an indicator that there is a lack of consensus around the skills and role of a data scientist. It could also be an indicator that the job of a data scientist differs from one organisation to the next depending upon the business need.

Some confusion also exists in terms of the qualifications and skills of a data scientist and about how one can find and hire a data scientist. There seems to be a broad consensus that the role of the data scientist connects a range of expert domains (or disciplines) including computer science, mathematics and business. Data scientists generally develop skills at the intersection and more generally across these three broad areas, see Figure 1. While quantitative analysis and technical skills are integral to the role, managers need to appreciate that data scientists who can speak the language of business are highly valuable. These experts will help senior management to rethink business strategies to extract real value from ‘big data’ and data analytics.

: What are the skills of a data scientist?

Heightened global awareness of data analytics and big data has meant that the "role" data scientist is increasingly important across a wide range of business sectors including technology, financial services, retail, hospitality, healthcare and military/defense forces. For each of these domain areas the requirements of a data scientist and even the name of the role can differ dramatically.

Data scientists are highly paid, skilled employees, therefore managers need to know exactly what a data scientist will do and what value they can bring to the business (Power, 2013). It is useful to understand the key skills of a good data scientist, this will help to write an appropriate job description and evaluate candidates for the role. For the most part, data scientists work with very large unstructured data sets. According to Jonathan Hassell, the following are the 4 qualities to look for in a data scientist:

1. Understands Statistics and the Law of Large Numbers
2. Inquisitive
3. Familiar With Database Design and Implementation
4. Has a Baseline Proficiency in a Scripting Language

Nine categories of data scientists are proposed by Vincent Granville. As a data scientist with over 20 years' experience, Granville is an expert in mathematics, statistics, machine learning and business and that influences his categories. His categories include:

● Data scientists strong in statistics. Experts in statistical modeling, experimental design, predictive modeling.

● Data scientists strong in mathematics. Experts in forecasting, pricing optimization, and quality control.

● Data scientists strong in data engineering, Hadoop, database/memory/file systems optimization and architecture, API's, Analytics as a Service, optimization of data flows, data plumbing.

● Data scientists strong in machine learning/computer science (algorithms, computational complexity)

● Data scientists strong in business, ROI optimization, decision sciences, involved in some of

: What are the skills of a data scientist?

the tasks traditionally performed by business analysts in bigger companies (dashboards design, metric mix selection and metric definitions, ROI optimization, high-level database design)

• Data scientists strong in production code development, software engineering and programming languages

• Data scientists strong in visualization

• Data scientists strong in GIS, spatial data, data modeled by graphs, graph databases

• Data scientists strong in multiple skill categories

The categories Granville identified are not mutually exclusive. An 'outstanding' data scientist encompasses some, if not all, of these skill categories. Nevertheless, Patil (2011) has a somewhat opposing view. While he acknowledges that technical/scientific competencies are important, curiosity, storytelling and cleverness are personal characteristics he has identified in his top performing data science hires to date. Excellent communication and documentation skills with audiences of varying degrees of technical expertise is also often cited as a needed skill. Similarly skills in working on a team are important.

Overall, the skills needed by a data scientist are an interesting and unique mix that draw from a diverse range of disciplines. From an employer's perspective, amid the hype of the 'big data' phenomenon it may seem appealing to hire a new employee with some or all of the skills identified. Pause a moment. Given the relatively new organisational role of the data scientist and the lack of consensus around the skills and boundaries of such a role, it is also useful to consider what level of expertise is really needed along with the challenges of hiring and keeping 'good' data scientists. Hire the data analyst or data scientist that is needed now and build a team to meet organizational needs.

References

Cleveland, W. S. (2001). "Data Science: An Action Plan for Expanding the Technical Areas of the Field of Statistics," *International Statistical Review*, Volume 69, No. 1, April

Davenport, T. H., & Patil, D. J. (2012). "Data scientist: The Sexiest Job of the 21st Century". *Harvard Business Review*, 90, October, 70-76.

Granville, V. (2015) "400 Categorized Job Titles for Data Scientists", 8 Feb 2015 at

: *What are the skills of a data scientist?*

<http://www.datasciencecentral.com/profiles/blogs/400-categorized-job-titles-for-data-scientists>

Granville, V. (2014) "16 analytic disciplines compared to data science", 24 July 2014 at <http://www.datasciencecentral.com/profiles/blogs/17-analytic-disciplines-compared>

Hassell, J. (2014) "4 Qualities to look for in a data scientist", 15 April 2014 at <http://www.cio.com/article/2377108/big-data/4-qualities-to-look-for-in-a-data-scientist.html>

Patil, D. J. (2012). Data Jujitsu: the art of turning data into product. "O'Reilly Media, Inc.

Power, D. J. (2013). "What is a data scientist?" Decision Support News, Vol. 14, No. 13 at URL <http://dssresources.com/newsletters/347.php>

Power, D. J. (2014). "Using 'Big Data' for analytics and decision support," Journal of Decision Systems, 23(2), 222-228 at URL <http://www.tandfonline.com/doi/full/10.1080/12460125.2014.888848>.

Power, D. J. (2014). "What are examples of data science jobs?" Decision Support News, Vol. 15, No. 4, February 16 at URL <http://dssresources.com/newsletters/364.php>

Power, D. J. (2016). "Data science: supporting decision-making," Journal of Decision Systems, 25 (4) 345-356 at URL <http://www.tandfonline.com/doi/full/10.1080/12460125.2016.1171610>.

Author: Daniel Power

Last update: 2016-11-25 06:05