



## › Survey Summary Report

# What Makes a Great Data Scientist?

What are the personality traits of today's data scientists?

Do they have the skills needed to support businesses' changing needs?

How can senior management build cohesive teams to drive the most value from big data?

Is this fast evolving discipline causing stress to its practitioners?

# Contents

- Foreword ..... 1
  - Summary of Findings ..... 2
- About the Survey ..... 2
  - The DISC Personality Profiling Methodology ..... 3
- The Data Scientist Personality Mix ..... 3
  - Distribution of Personality Types in our Survey Sample ..... 4
  - Data Science: An Equal Opportunity ..... 6
- Are We Making or Breaking Data Scientists? ..... 6
  - A Stressed Generation? ..... 7
- Conclusion: The Multi-Faceted Data Scientist ..... 8

## Foreword

Data science is a relatively new but rapidly evolving discipline, as more organisations focus on unlocking the insight held in ‘big data’ - the vast amounts of information collected through systems and monitors; automated business processes; the increased use of smart devices such as mobile phones and tablets; social media and GPS tracking. The volumes of big data and the possibilities for insight will only get bigger as ‘the internet of things’ provides more sources of information.

Data science is the discipline that can derive vital insights from the analysis of that big data. The impact of data science is already felt across many domains: from medical research to marketing, financial services to fire prevention. And, with modern technologies making it cheaper and faster to process and analyse huge volumes and varieties of data, that influence is only set to increase.

This has huge implications for the talent pool. The Centre for Economic Business Research (Cebr) forecasts that the uptake of big data analytics will create almost 70,000 jobs in the UK and 61,000 new jobs in Ireland, by 2017<sup>1</sup>. This is supported by e-skills UK, the IT skills development organisation, which puts the figure at 69,000 new jobs for the UK<sup>2</sup>.

McKinsey estimates that for every data scientist, organisations will need ten data savvy managers with the skills and understanding to make decisions based on data analysis<sup>3</sup>. Yet with 57 per cent of large UK organisations already struggling to hire individuals with the required data analytics skills<sup>4</sup>, how can managers find and develop the right people to drive the most value from data?

SAS has been in the business of big data analytics for almost 40 years, providing the software which makes the analysis of vast volumes of data possible to provide meaningful insights. SAS is dedicated to helping businesses, governments, schools and universities across the UK and Ireland to develop the skills required to thrive in the information economy. Over the past 15 years, we’ve invested around £100 million to provide access to our industry-leading solutions to around 80 universities.

In 2014 SAS launched a survey to explore what makes a great data scientist, specifically: which skills and personality types are needed to fulfil pivotal roles in the future?

This report is based on analysis of our initial 596 respondents who identified themselves as part of the data science profession. The survey highlights 10 key personality types as well as some of the challenges data scientists are facing today. We hope it will encourage more data scientists to take part in the survey. This will provide an even bigger platform for analysis, debate and exploration of some of the important issues that will affect the rate at which data science grows and performs to meet the increasing demands of organisations, regulators, citizens and consumers.

---

1 Data Equity: Unlocking the Value of Big Data (UK and Ireland), SAS and Cebr, June 2013

2 Adoption and Employment Trends 2012-2017, SAS and e-skills UK, November 2013

3 Big Data: The Next Frontier for Innovation, Competition and Productivity, McKinsey Global Institute, May 2011

4 Adoption and Employment Trends 2012-2017, SAS and e-skills UK, November 2013

## Summary of Findings

- Data scientists with “traditional” traits (analytical, logical, technical) make up the largest group (41 per cent), yet a significant proportion of data scientists display other traits, such as strong communication and creativity skills.
- We expect a lot from data scientists: to be technically proficient, mathematically agile, business savvy and good at communicating. Yet 55 per cent of data scientists have fewer than three years of experience in the discipline, and more than a quarter are adapting their behaviours to fulfil roles that are not well matched to their skills or work personality profiles.
- In most personality types there is a 70:30 split between men and women. The Geeks, the profile with technical bias, strong logic and analytical skills, shows a higher percentage of women at 37 per cent compared to 63 per cent for men.
- Data scientists are exhibiting high levels of work-related stress, with a total of 55 per cent of the respondents showing a level of stress, with 1 in 4 male and just under a third of female data scientists being heavily stressed.
- Organisations must better identify and define what they need from data scientists. Only then can they build and develop multi-faceted teams with the complementary skills needed to realise the full value of big data.

## About the Survey

We invited data scientists from UK and Ireland to complete a confidential, online psychometric survey that comprised 24 questions. Each participant received a tailor-made profile describing their personality type and highlighting the key attributes they are likely to bring to the data science professions.

The survey was based on the well-established DISC profiling methodology, which has been used for more than half a century to categorise respondents into a range of personality types with recognised characteristics. To conduct the analysis, we worked with an independent psychologist who is an expert in the DISC profiling methodology.

This report looks at the aggregated data of the initial 596 responses, received from 405 men (68 per cent) and 191 women (32 per cent). Respondents’ experience ranged from less than a year to a handful of data science pioneers who have committed more than 21 years to the discipline. By using DISC, we were able to create the typical profiles of today’s data scientists, and explore how these compare to skills and personalities required by organisations.

## The DISC Personality Profiling Methodology

The DISC methodology has been widely used to assess personality types for decades. This information is often utilised to analyse and build effective teams, improve communication and increase productivity. DISC is based upon a two-axis matrix - one ranges from Proactive to Reactive (vertical axis) and the other Introvert to Extrovert (horizontal axis). This generates four possible combinations - named D, I, S and C. Individuals possess measures of all four of these combinations but in varying amounts. The values for each combination is calculated and plotted on a graph to illustrate the relative strengths of each. These values can be matched to known behavioural characteristics to describe an individual's personality profile.

**Dominance** (Proactive Introverts)

*Relating to control, power, aggression and the drive to achieve*

**Influence** (Proactive Extroverts)

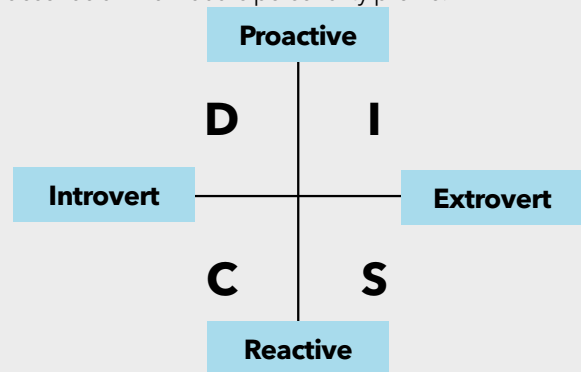
*Relating to social and communication skills*

**Steadiness** (Reactive Extroverts)

*Relating to patience, persistence, constancy and thoughtfulness*

**Compliance** (Reactive Introverts)

*Relating to precision, order and caution*



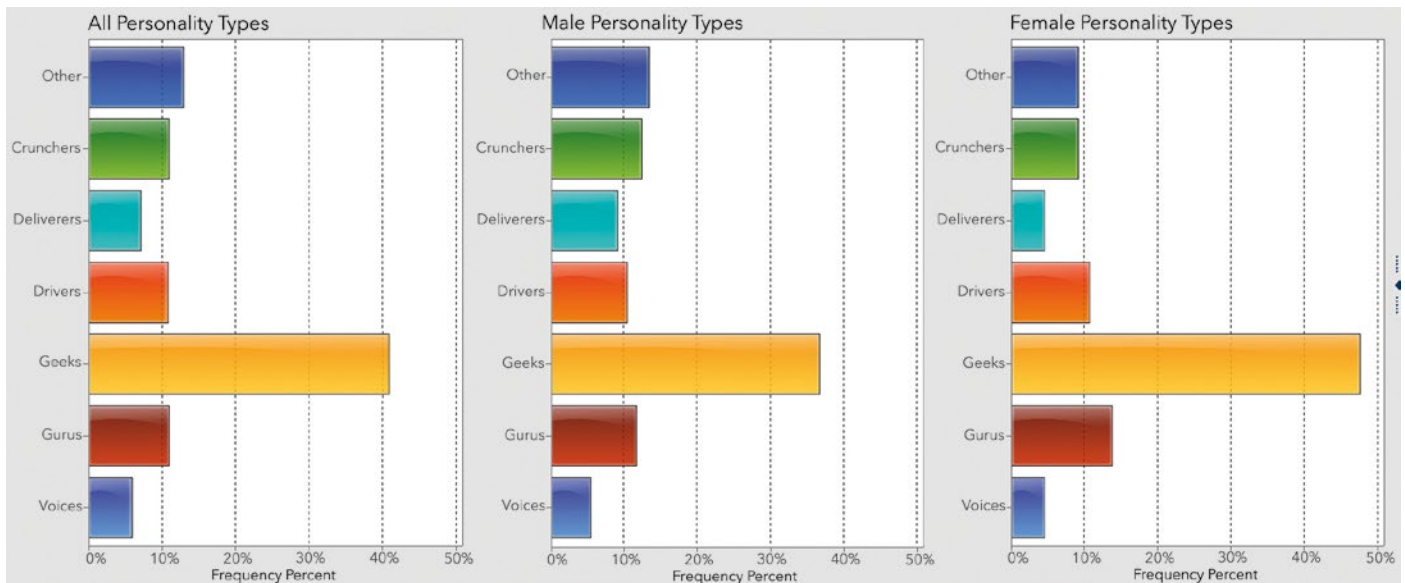
## The Data Scientist Personality Mix

Our analysis identified ten psychometric profiles evident within the data scientist community, based on our responses so far. These profiles are characterised by definable patterns that are well known to scientists and were consistently apparent in our survey sample.

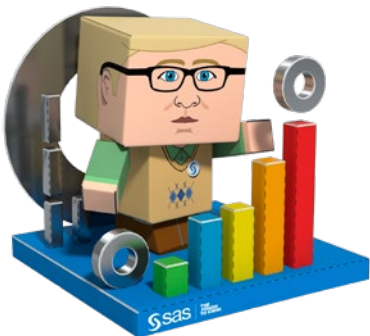
The 'traditional' traits associated with data scientists - such as technical, analytical and logical skills - still dominate. However, other less technical traits - such as project management, creativity and good communication skills - are also present. Organisations need data savvy individuals who are technically proficient, mathematically minded, business oriented and strong communicators.

It's unlikely that any individual will have all of the skills required to maximise the value of big data. So it's important for managers to identify the particular skills needed and build a cohesive team of individuals with complementary skills and traits. As we will see later in the report, failure to do so can result in individuals trying to fulfil roles to which they are not suited - which may lead to stress and burn-out.

## Distribution of Personality Types in our Survey Sample



\*Data analysed using SAS® Visual Analytics. Better analytics. Faster insights. Built for everyone.



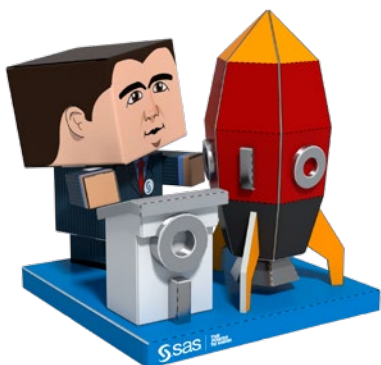
### 1. The Geeks 41%

The Geeks are the largest group in our sample and have the largest female membership of all the groups at 37 per cent. They have a naturally technical bias, strong logic and analytical skills. Essentially “black and white” thinkers, they like to speak plainly and stick to the point - don’t expect them to be moved by emotionally charged arguments. With their attention to detail and fondness for the rules, the Geeks are well suited to roles such as defining systems requirements, designing processes and programming.



### 2. The Gurus 11%

The next largest group, the Gurus, has a measure of reactive introversion, like the Geeks, which pre-disposes them to scientific and technical subjects. Yet they also display a diametrically opposite characteristic: the strong presence of proactive extroversion, including solid and often highly persuasive communications and social skills. The Gurus can play a very important role by using their enthusiasm, tact and diplomacy to promote the benefits of the data sciences to those holding the purse strings, or who have the authority to give projects the green light.



### 3. The Drivers 11%

The Drivers are proactive introverts: highly pragmatic individuals who use their determination and focus to realise their goals. Self-confident and results-oriented, they are ideal project managers and team leaders, who excel at prioritising, monitoring and driving projects to a successful conclusion.





#### 4. The Crunchers 11%

This category is probably one of the least self-promoting groups. Strongly reactive - rather than proactive - personalities, the Crunchers like routine and constancy. They display high technical competence and consistency, making them superb in a range of technically-oriented support roles including data preparation and entry, statistical analysis, monitoring of incoming data and quality control.



#### 5. The Deliverers 7%

Like the Drivers, these individuals are proactive and well suited to project and man management. This is also the group with the largest proportion of men at 80 per cent. However, the Deliverers also have a strong pre-disposition towards acquiring and/or applying technical skills. So, while they are capable of bringing focus and momentum to ensure project success, they are also likely to understand the finer technical details and devise solutions in much greater technical depth.



#### 6. The Voices 6%

The Voices are strong communicators with less apparent detailed technical knowledge than the Gurus. The presence of this group suggests a strong demand for natural promoters who have the ability to generate enthusiasm for the potential of big data and the data sciences at a conceptual level - rather than the practical or technical level. The Voices are strongly valued for their positive outlook, and may be engaged in presenting the results of big data projects as well as supporting their implementation.



#### 7. Other Personalities 13%

A smaller number of respondents displayed a range of other traits. Phase two of our study will give us a bigger sample to analyse, when we'll be able to learn more about how these traits are emerging in the data science community.

- **The Ground Breakers:** offer new approaches, new methods and new possibilities, drawn from a mix of inspiration and dogged logical thinking. Roles include: system design and algorithm development.
- **The Seekers:** combine superb technical knowledge and understanding with inquisitiveness and a drive to find solutions. Roles include: research.
- **The Teachers:** skilled at imparting knowledge and inspiring others to want to learn. Roles include: training and mentoring.
- **The Lynchpins:** important team players who may not have a depth of technical knowledge but provide essential support services. Roles include: co-ordination and administration.

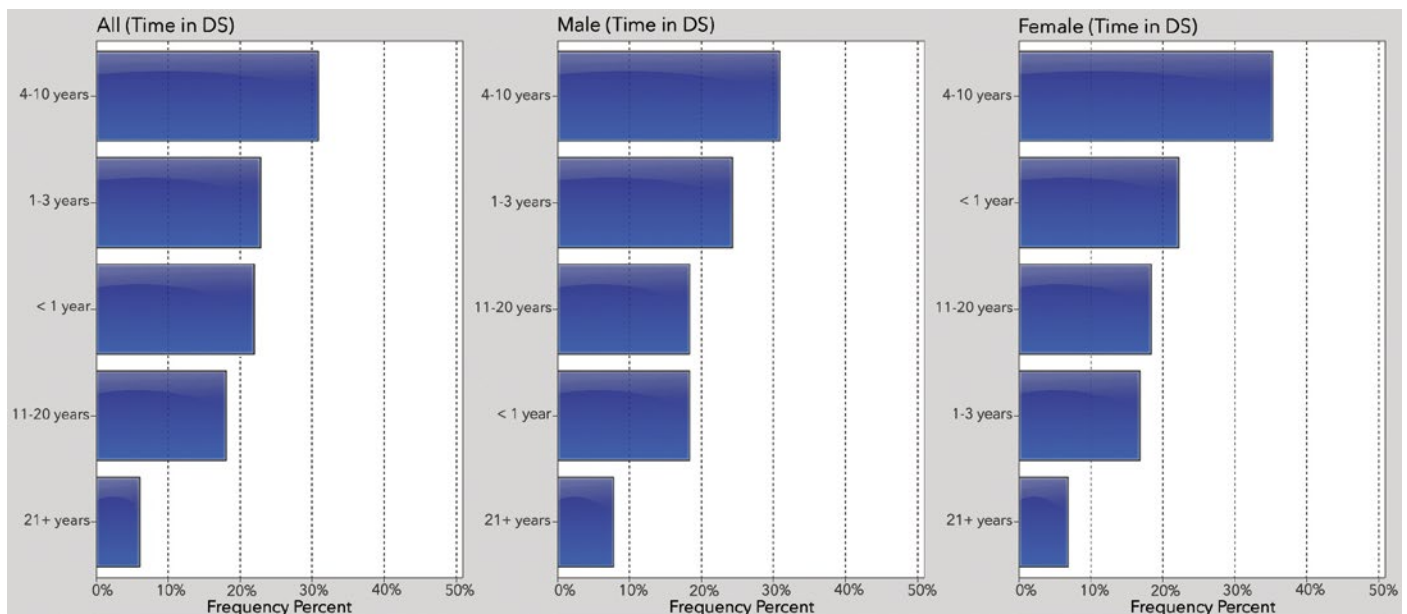
## Data Science: An Equal Opportunity

It is interesting to note that the proportional representation of each of these personality profiles is closely matched when we compare the male and female samples. So, whilst the men do outnumber the women in the sample by an approximate ratio of 2:1, men would not appear to have established a disproportionately high representation in any particular data science related activity or role (except for the Deliverers group which is split 80:20). It appears that women are actively suited to, and capable of, performing every specialist role we have identified – making data science a genuinely equal opportunity career.

## Are We Making or Breaking Data Scientists?

As data science is a relatively new discipline, that is not as well defined as traditional professions like accountancy and sales, it's not surprising that organisations struggle to hire and develop people with the appropriate skills and experience that match their rapidly evolving business needs. Our survey found that 76 per cent of respondents have fewer than ten years of experience in data science – while 45 per cent have a maximum of three years' experience.

### Percentage of Total Sample Size by Years of Experience in Data Science (DS)



\*Data analysed using [SAS® Visual Analytics](#). Better analytics. Faster insights. Built for everyone.

In a rapidly evolving discipline, individuals can find themselves in roles to which their personality type is not suited, or equipped, to fulfil. The survey revealed that a high proportion of respondents are regularly adopting behaviours that do not match their personality profiles. In other words, they are adjusting their behaviours to match what they believe is needed in their role.

The psychometric analysis shows that 26 per cent of male respondents and 27 per cent of females are adjusting their behaviours to suit their perception of their work



responsibilities. Some of these individuals may be unclear of what is expected of them, and are therefore trying to fulfil an imagined role. However, this still leaves a substantial proportion struggling to cope with the demands of working responsibilities to which their skills are poorly matched.

While this is common in fast growing specialist disciplines like data science, substantial amounts of behavioural adjustment are not sustainable and can lead to burn out or a failure to realise an individual's potential. This is evidenced in the high levels of work-related stress indicated in our sample.

## A Stressed Generation?

The DISC personality profiling methodology can also highlight stress levels. Some personality types are more susceptible to work-related stress and some are better able to cope with stress.

We compared the levels of stress demonstrated by our respondents against the maximum recommended for each personality profile identified by our independent psychologist.

### Levels of Stress Amongst of Survey Respondents

	Calculated by Personality Type		
	Men %	Women %	Male and Female Combined %
Not stressed	48	42	45
Mildly stressed	27	28	27.5
Heavily stressed	25	30	27.5

It is clear that work-related stress levels are very high among our respondents. Half of men and almost two thirds of women indicate some stress, while over a quarter are identified as being heavily stressed. This latter level is unsustainable and will usually affect performance and attendance - and have consequences on the individual, other team members and beyond.

While we cannot know for sure the causes of stress experienced by our respondents, our independent psychologist has identified the four main origins of work-related stress as:

- **Communication issues:** with the immediate manager
- **Poorly defined role:** often due to new activities being officially or unofficially added, leading to no objective means of measuring performance
- **Responsibility without authority:** leading to no direct control over factors that affect an individual's ability to perform the task at hand
- **Lack of skills:** either because the job is too difficult or because a person has not yet been adequately trained.

It's clear that more focus is needed on defining the various job responsibilities within data science so managers can recruit individuals with the appropriate personality types to fulfil the role without undue stress. The development of well-defined benchmarks should be a priority, as it will likely lead to increased job satisfaction and levels of productivity in the long-term.

## Conclusion: The Multi-Faceted Data Scientist

Phase one of our survey has shown that data science is a multi-faceted and evolving discipline that will benefit from recruiting individuals with a range of different personality attributes. There is huge pressure on data scientists to display a range of skills: technical, mathematical, creative, business aptitude and communication. Unsurprisingly, organisations struggle to find this mix of skills in just one person, particularly when the majority of data scientists have only a few years of experience.

It is therefore important for managers to build cohesive, diverse teams that can answer all the needs of the organisation - not just data scientists but also data savvy managers to interpret the insights and transform them into actions. Organisations will need to:

- **Manage** the business' expectations from data scientists
- **Find and develop** individuals with the appropriate skills and personality traits
- **Create teams** of individuals with complementary skills and experience
- **Encourage** peer to peer support, learning and development
- **Match training** to the business' requirements.

It's possible that, as technology develops, there will be less of a weighting towards the more traditional, technical personality traits - while demand for the less frequently found profiles such as the Ground Breakers, the Teachers and the Seekers will grow. As more respondents take part in our survey, we hope to be able to explore this interesting development further.

### TAKE PART IN THE SURVEY

To take part in the survey and receive your free personal profile report visit [www.sas.com/uk/ds](http://www.sas.com/uk/ds) today.

### ABOUT SAS

SAS is the leader in [business analytics](#) software and services, and the largest independent vendor in the business intelligence market. Through innovative solutions, SAS helps customers at more than 70,000 sites improve performance and deliver value by making better decisions faster. Since 1976 SAS has been giving customers around the world THE POWER TO KNOW®.

[www.sas.com/uk](http://www.sas.com/uk)

SAS UK & IRELAND WITTINGTON HOUSE HENLEY ROAD MARLOW BUCKS SL7 2EB  
+44 1628 486933 [www.sas.com/uk](http://www.sas.com/uk)

---

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies. Copyright © 2014, SAS Institute Inc. All rights reserved. 1022821UK0914

