profusion

Guide to what is data science?

Henrik Nordmark, head of data science at Profusion, discusses the philosophy of data science

Data influenced journeys

Storytelling with data

Data science is unquestionably in vogue. But what is it, and is it necessary?

When any new business resource becomes popular, it's natural and wise to question it. Similarly, it's important to look at whether existing skills or functions within an organisation to fulfil the 'new' function.

It makes sense for businesses to consider what data science actually is, whether they have people in their organisation that can do it, and what makes a good data scientist. Given that I'm a data scientist, I'm obviously going to make the case for data science. But I realise that it's important for data science to continuously prove itself to be of value to organisations and people.

With that in mind, the first question to ask is whether it even makes sense to speak of 'data science'. The phrase only rose to prominence a few years ago. Is this actually a new discipline? Or is it merely a rebranding of activities that existed already? I would argue that the answer to both is 'yes'.

From a naturalistic point of view, and in much the same way anthropologists would observe a tribe in the Amazon rainforest, let's look at what people with the job title 'data scientist' actually do. Working with a team of data scientists, I have some first-hand experience of what it's like to be a member of this particular tribe.

We use a combination of methods from a variety of different fields – including but not limited to statistics, machine learning, applied mathematics and databases. Generally, we are a very pragmatic bunch.

Data scientists realise there are many ways to solve a business problem. There are also many ways to give insight into a question that has been asked. As the statistician George E P Box once wrote: "All models are wrong but some are useful."

Data influenced journeys

"In much the same way anthropologists would observe a tribe in the Amazon rainforest, let's look at what people with the job title 'data scientist' actually do"

Data science is the pursuit of translating a question into something that could be answered using data, then applying a variety of techniques to see what happens. We are at the frontier, exploring what value we can generate from data for businesses and people.
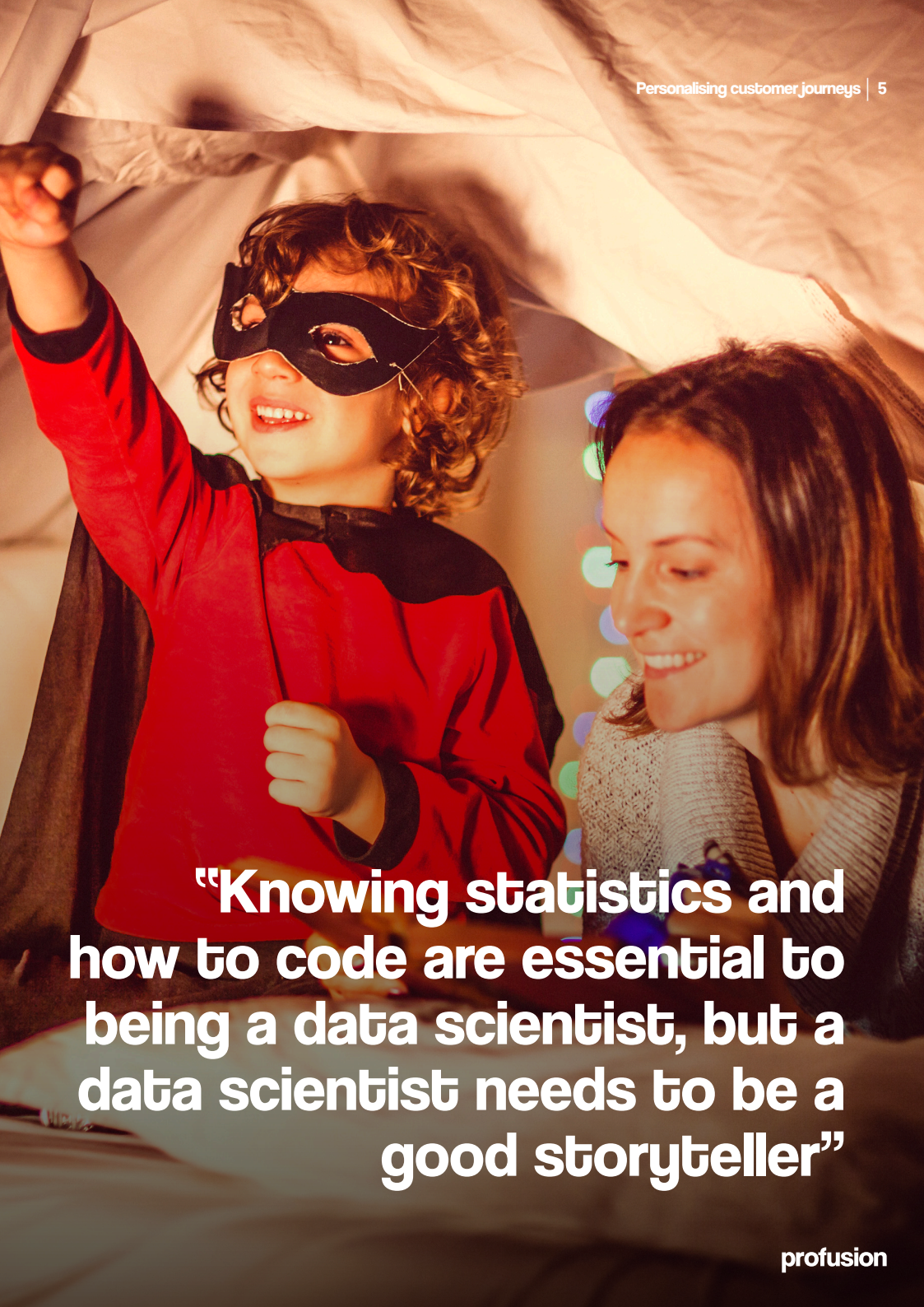
For those of you in the know, this definition of data science may not sound terribly different from what a statistician, mathematician or computer scientist does. It's tempting then to think that data science is not a field in its own right. But let's consider this influential definition by Josh Wills, a data scientist at Cloudera: "A data scientist is someone who is better at statistics than any software engineer and better at software engineering than any statistician".

Josh's definition makes it clear that your average statistician is not a data scientist and your average software engineer is not a data scientist. I've hired many data scientists and I've gone through hundreds of CVs from people aspiring to become data scientists. Knowing statistics and how to code are essential to being a data scientist, but even that is not the whole picture. A data scientist needs to be a good storyteller.

Storytelling with data often means creating beautiful visualisations that capture the main messages to be conveyed by the data. Sometimes it's not about the visualisation at all but about finding the right metaphor that will make an idea understandable to a wider audience of stakeholders. As a result, data scientists need a very niche set of expertise to succeed.

It makes sense to speak of a field named data science. It may not refer to an entirely novel set of activities but it does allude to a relatively consistent set of activities that it is useful to name. Fascinatingly, from a sociological point of view, it is because we give it a name that the field exists.

Note: this post originally appeared in the print edition of Enterprise 360.

Data influenced journeys

"**Knowing statistics and how to code are essential to being a data scientist, but a data scientist needs to be a good storyteller**"

profusion

About
the writer

Henrik Nordmark
head of data science

Henrik leads a team of data scientists, data analysts, data engineers and interns working on applying statistics and machine learning in novel ways to solve business problems. Henrik is also responsible for the research and development of data science techniques, drawing from ideas and methods in artificial and augmented intelligence, statistical learning, operational research and economics. He is a Visiting Fellow at the University of Essex and has been named in DataIQ's 100 most influential people in data-driven business.

Wired to solve problems

We're a clever team of specialists in data engineering, BI, data science and multi-channel marketing.

Blending real 1:1 communications, AI, machine learning and human intelligence, we connect people with your brand.

We were the original UK partner for Adobe Campaign. We've nearly 20 years' experience in data management, automation and customisation of enterprise data management and marketing other platforms.

Fuelled by inquisitiveness, forward thinking and excitement at the advancements of technology, we're entrepreneurial, adventurous and we make stuff happen.

We help clients get under the skin of a human being's emotional narrative – their story – and see what's important to each customer.

Our clients trust us for our responsive, personal and upfront approach to help them create highly targeted, intelligent communications.

Stay connected to discover more inspiring guides and stories within our 'knowledge think tank'.

Data influenced journeys

profusion

profusion

Winning together

Our client partners

HSBC

first direct

Rolls-Royce

TRADE POINT
Trade Building Supplies

SCREWFIX

M&S BANK

COTY

FAB

**Data
influenced
journeys**