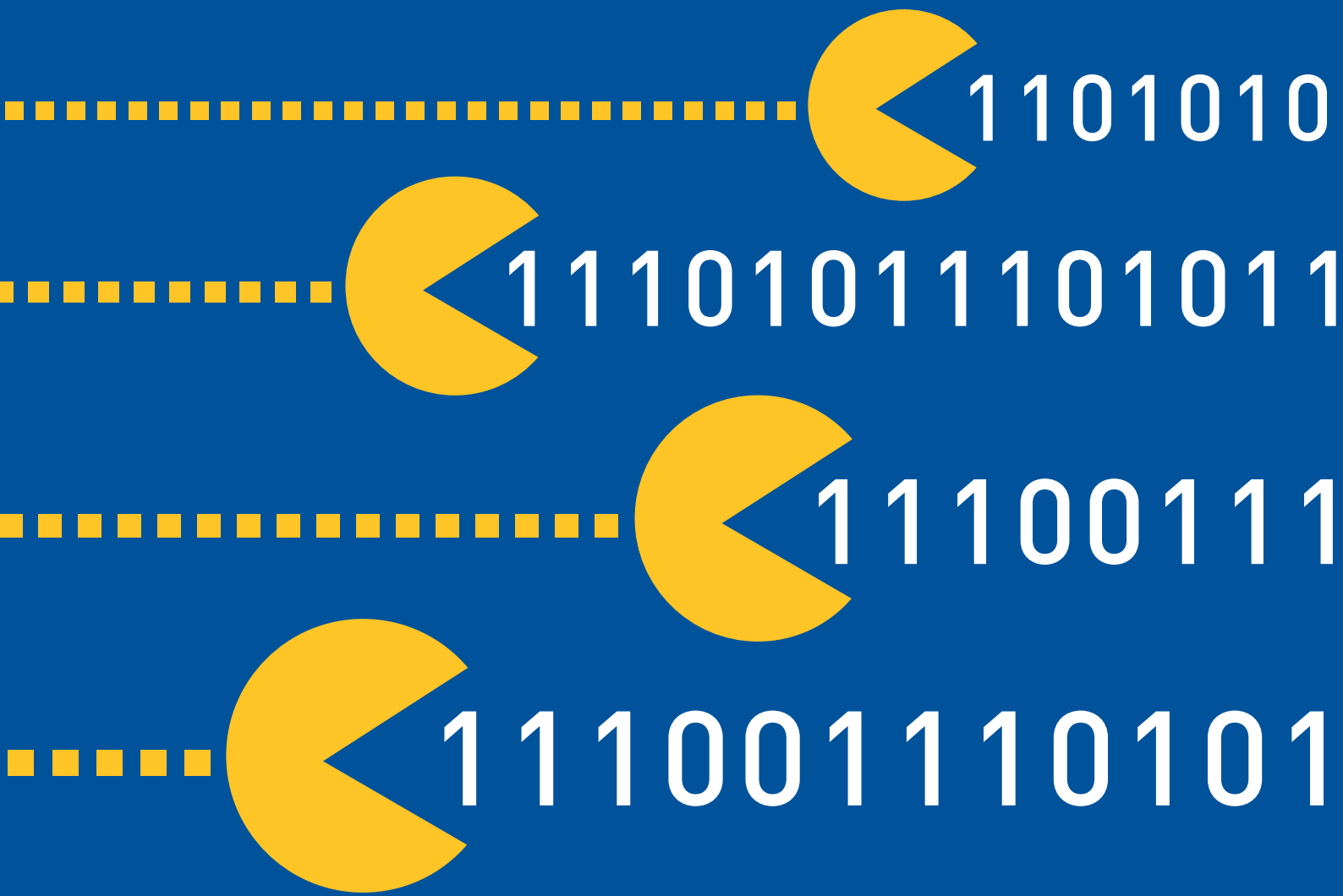
Nesta...

RISE OF THE DATAVORES

HOW UK BUSINESSES ANALYSE AND USE ONLINE DATA

Hasan Bakhshi and Juan Mateos-Garcia
November 2012



About Nesta

Nesta is the UK's innovation foundation. We help people and organisations bring great ideas to life. We do this by providing investments and grants and mobilising research, networks and skills.

We are an independent charity and our work is enabled by an endowment from the National Lottery.

Nesta Operating Company is a registered charity in England and Wales with company number 7706036 and charity number 1144091. Registered as a charity in Scotland number SC042833. Registered office: 1 Plough Place, London, EC4A 1DE

www.nesta.org.uk

EXECUTIVE SUMMARY

The Internet economy is growing, and so is the amount of data that it generates. Are UK businesses making the most of this data to innovate and grow? Or are they missing out on the opportunities of online data?

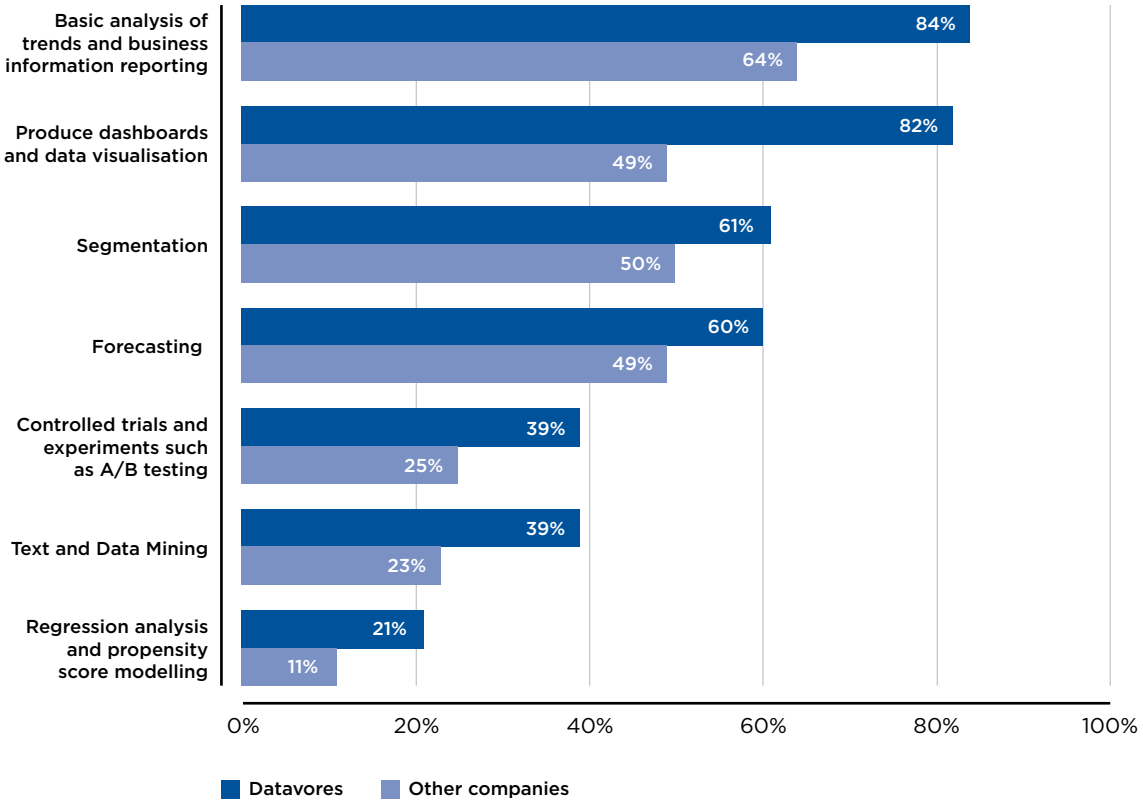
A business survey by Nesta shows how a small but significant group of UK companies – the datavores – are making sophisticated use of data to drive business decisions.

Datavores make up 18 per cent of the firms with active online operations that we surveyed. They have a number of defining features. They gather online customer data intensively, subject this data to sophisticated analyses (such as controlled trials and data and text mining), and use what they learn to improve their business. They also report that they are more innovative than their competitors, in products as well as processes.

They are very different from the majority of businesses in the Internet economy, which monitor their customer information only casually, and rely on intuition and experience when looking for new ways to grow their sales.

Adoption of Online Analytics techniques by type of company

Datavores experience a visible data dividend – they are almost twice as likely as the

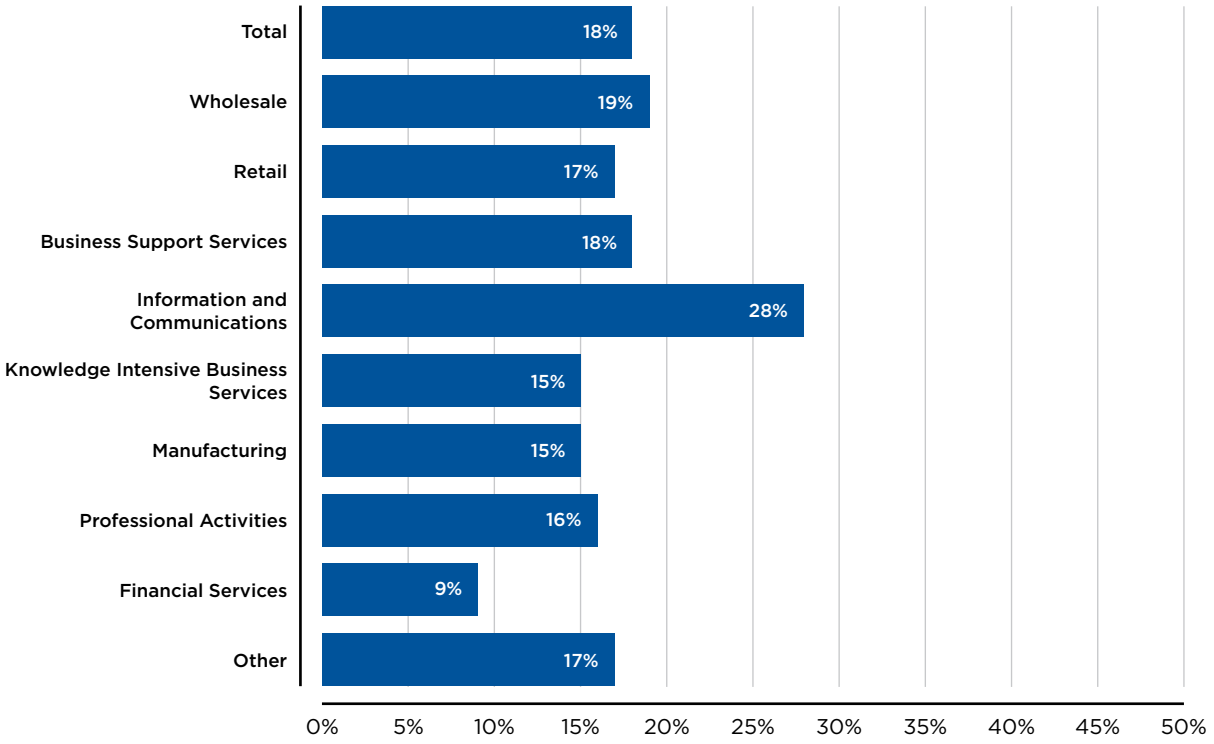


average to report that their investments in analytics have already made a significant or very significant contribution to their business performance. There is considerable potential for other businesses to adopt analytic techniques: even among the online-active businesses we surveyed, over 40 per cent were only casual users of data.

The positive connection between Online Analytics adoption and business impacts is statistically significant, and holds after we control for variables such as online revenues, innovation, company size and sector. Our statistical analysis also supports the idea that collecting and analysing online customer data is not enough – those businesses that want to benefit from their data need to put it to work, by using it to make practical decisions. This is what datavores do.

Datavores are present across all company sizes, and in every sector that we have surveyed – although perhaps unsurprisingly, they are more common in Information and Communication Industries (where they represent over a quarter of respondents). This widespread appetite for data suggests that the potential of Online Analytics and data is not confined to any industry in particular.

Proportion of datavores across different industries



There appears to be a widening gap between datavores and the rest of our sample. Datavores are investing more in Online Analytics and in integrating analytics into their businesses than other companies. For example, almost three-quarters had trained their staff compared with 55 per cent in the rest of the sample, and two-thirds had re-engineered their business process compared with 45 per cent of other firms.

In the last two years, our sample has seen a boom in the use of A/B testing and controlled trials, where companies test product launches or changes in their website as if they were scientific experiments to determine what works and what doesn't. Half the companies in our sample currently running these tests only started doing so in the last two years. Google Chief Economist Hal Varian's quip that the Internet is a *"lab experiment that got loose"* appears to be true in more than one way.

Our research has implications for businesses, government and investors in the UK and beyond.

Businesses are faced with a major opportunity to harness the power of Online Analytics to improve their product development, pricing, marketing and, ultimately, their profitability. This may require changes in business culture, and empowering employees to use online data to make the right decisions. Failing to do this could mean being left behind by the datavores, who are deepening their investments and making them pay off. A number of UK businesses who think of themselves as 'data–driven' may need a reality check. Our analysis reveals that some respondents who say they are all about data aren't walking the walk in the way that 'true' datavores do.

Government should recognise the power of Online Analytics in driving economic growth and innovation. Good analytics helps companies determine quickly what their customers want, and use this information to innovate more effectively. To take advantage of this emerging trend, businesses will need to be able to recruit skilled talent with a good mix of data and business skills – and policymakers should ensure that the education system produces them. Clear, proportionate regulatory frameworks on how customer online data can be stored and analysed will also be essential. Policymakers should also reflect on the potential of the Internet as a 'sandpit' where UK companies dip their toes into data, developing analytics capabilities which they then apply to other 'offline' parts of the business: these innovation spillovers could further justify efforts to get UK businesses online.

Investors are already seeing the power of Online Analytics. Some of the hottest venture capital–backed businesses in the UK are intense users of analytics, including Zoopla and Mind Candy. Venture capital firm Accel Partners has established a $100 million global fund to finance big data companies.

We will be following up this research report with analysis of the effects of Online Analytics on business profitability, which we expect to publish early in 2013, once this year's company accounts are available for all the firms in our sample.

To research this report, we surveyed by telephone 500 UK businesses of different sizes in a range of different industries with active online operations. By this, we mean businesses that either use their website to generate revenue (e.g. through e–commerce, as well as revenues from advertisers), and/or advertise online (so as to avoid excluding from our sample companies that don't actually transact online but use their websites to generate 'leads' which translate into offline sales).

CONTENTS

1. INTRODUCTION

The new oil?

TV producer Richard Curtis recalls that, in the early 1980s, he would wander around Shepherd's Bush, *"looking in people's windows, particularly people in basement flats, to see whether or not anyone was watching Blackadder [series] one … because I didn't know whether or not it had been successful otherwise."*[1]

Things have changed a lot since then. The digitisation of commerce and communications, together with cheap access to processing power, data storage and statistics tools, means that organisations can now draw on vast troves of data to learn from the past, make sense of the present, and shape the future.

For example, and in contrast to Richard Curtis three decades ago, today's TV producers have real time information about their audiences, not only in terms of 'quantity' – how many people are watching their programmes – but also 'quality': from where and on which platform, what they say about them to their social networks, and even what they do afterwards. If information is power, then today's managers are mighty indeed. As David Abraham, Channel 4's CEO puts it, data is 'the new oil'.[2]

Or the new sugar?

Analytics expert Ron Shevlin recently pointed out, however, that data can also be compared to sugar: it tastes good but it is unhealthy when unrefined.[3] The rhetoric about the impact of data outstrips hard evidence about its magnitude, the mechanisms through which it is generated, and the best practices to harness it.[4]

A growing data dictionary

Increasing interest in data as a driver of competitive advantage has gone hand–in–hand with a proliferation of concepts and terms referring to its different dimensions:

- **Its size, 'Big Data':** Big Data refers to data sets of such 'volume, velocity and variety' that they cannot be easily managed using conventional tools such as relational databases or desktop statistics packages.[5]

- **Its processing, 'Analytics':** Analytics is the measurement, collection, analysis and reporting of data to inform decision making. This is a broader category than 'Web Analytics', which draws on Internet data to understand and optimise web usage.[6] In this report, we use the term 'Online Analytics' to refer to the collection, analysis and use of online customer data not only to improve website usability and performance, but also in other areas of decision making in a business (e.g. pricing, product development and business strategy).

- **Its users: 'Data–Driven Companies' (DDCs):** These are businesses that predominantly ground their decisions in the rigorous analysis of data rather than in the instinct and experience of their managers. It is sometimes said that they follow a 'Data–Driven Decision Making' (DDDM) approach.[7] In this report, we also call them datavores.

The Internet is the main 'Big Data Generator'

There are now more than three billion global users online, creating, sharing, searching and transacting.[8] When we add to them the devices connected to 'the Industrial Internet' (or the 'Internet of Things'), and consider the dwindling costs of storing and combining data, it becomes clear that the recent data deluge (IBM's famous claim about how 90 per cent of the world's data was created in the past two years) is set only to intensify.

And UK consumers and advertisers are migrating there

Sir Tim Berners–Lee's World Wide Web Foundation estimates that the UK is third in the world in its use of the web (after Sweden and the US).[9] Other research from The Boston Consulting Group claims that the UK's Internet economy is the largest in the G20 as a percentage of GDP (8.3 per cent compared with an average of 4.1 per cent for developed economies as a whole).[10]

UK consumers lead Europe in their use of e–commerce – in 2011, 64 per cent reported having purchased goods or services online in the previous three months – almost twice the average for the countries in the Eurozone.[11] In the UK 8.5 per cent of all retail spending already takes place online,[12] and The Boston Consulting Group expects this percentage to reach 23 per cent by 2016 (the highest in the G20).[13]

Internet advertising is also growing very rapidly in the UK (by as much as 14.4 per cent between 2010 and 2011). The UK was the first country in the world where online advertising spend overtook TV advertising, in 2009.[14]

Data–driven growth and innovation appear to be easier on the Internet

Every single click a user makes when browsing a website is a data point containing potentially valuable information that a business can collect and analyse in order to make sense of its market, and act accordingly. The Internet also provides a low–cost platform for the adoption of powerful analytical techniques such as randomised control trials.[15]

Google's Chief Economist Hal Varian has described the Internet as a *"lab experiment that got loose"*. Although he was referring to the origins of the commercial Internet in scientific networks such as ARPANET, he was also touching on its ability to produce rigorous, generalisable knowledge, and on the economic returns from acting on this knowledge – in other words, from adopting a data–driven decision making approach.[16]

Another prominent US economist, Richard Nelson, has speculated that one of the reasons for rapid rates of innovation in fields such as computation and communications compared with others like education and health, lies in the ability of the former to support this kind of experimentation and learning in 'on–line' environments, that is, in the use of analytics.[17] Erik Brynjolfsson at MIT has argued that the ability to measure online behaviour and sentiment at ever increasing levels of resolution will have an economic and societal impact comparable to that of the microscope four centuries ago.[18]

Are UK businesses harnessing the potential of online customer data?

UK businesses across all industries have no option but to follow their consumers, clients and advertisers into highly competitive online marketplaces. This will require more than investments in technology: management and organisation matter too. For example, as much as one–half of the productivity growth gap between the US and Europe in recent years has been attributed to the adoption of management practices that help US companies exploit IT better than their European counterparts.[19]

We know from past research that UK businesses in sectors such as Manufacturing and Retail lag behind their international competitors in the adoption of good 'offline' management practices, including performance management – an important element of which is the setting and tracking of targets, where Analytics could be expected to play an important role.[20] Could this be happening online as well? Recent claims that the UK's businesses may already be wasting hundreds of millions of pounds in lost sales because of poor website design and optimisation suggest that this is a serious possibility.[21]

We need more and better data about data

As Table A1 in the Appendix shows, some of the highest-profile, publicly-available executive surveys of recent years haven't examined the collection and deployment of online customer data in any great detail – and neither have they focused their attention on UK businesses. Yet, as we have just noted, the Internet is a big data generator, and also a platform where UK companies are increasingly competing and innovating. We need more and better data to answer questions such as:

- To what extent are UK Internet businesses adopting Online Analytics practices?
- How much are they benefiting from doing so?
- What other changes do they need to make in their businesses in order to put their online customer data to work?
- What are the barriers standing in the way of a better use of online customer data?

Our research seeks to address these important questions

By providing a rigorous understanding of the return on investing in Online Analytics, and a clearer idea on how to do it effectively, we hope to help remove some of the uncertainties about benefits and costs that could discourage businesses from adopting Online Analytics technologies and practices.

Assessing the contribution of Online Analytics to business innovation and growth is also important for regulatory policymakers who must trade off any benefits against privacy-related costs of allowing businesses to collect and make use of their online customer data.[22]

This briefing is the first milestone of our programme of research on Online Analytics

In it, we discuss the emerging findings of a telephone survey of Online Analytics adoption and impact in 500 UK companies. It was designed by us in collaboration with researchers at Google, and carried out by Ipsos MORI.[23]

Section 2 provides an overview of who we interviewed and what data we collected.

Section 3 describes some of our preliminary findings – including the levels of adoption of Online Analytics practices by UK businesses, the impacts they perceive, the complementary investments they make (such as in skills and management practices) and barriers to the more effective use of online data. We pay particular attention to the differences between those companies that we identify as 'Data-Driven' and those that rely on experience and intuition in their decision making.

Section 4 explores the relationship between Online Analytics and self-reported indicators of business performance, controlling for other important company characteristics.

Section 5 concludes with some implications of our findings, and sets out the next steps for our research project – in particular, how we propose to get a firmer handle on the causal relationship between Online Analytics and business performance.

2. A SURVEY OF ONLINE ANALYTICS PRACTICES

In Autumn 2011, we commissioned Ipsos MORI to carry out a telephone survey of Online Analytics practices by UK companies. Our questionnaire design and sample frame were informed by qualitative interviews with Online Analytics users[24] and input from leading business survey experts, as well as reviews of the relevant Economics and 'grey' literatures (see Box 1, and Table A1 in the Appendix for summaries of the better known studies).

Our goal was to look at those business–to–consumer (B2C) and business–to–business (B2B) businesses that are more active in online markets – that is, where Online Analytics could be expected to play a stronger role as a driver of competitive advantage. This important consideration shaped the design of our sample frame and the approach we followed when screening respondents. It also means that our results should not be interpreted as representative of the UK economy. Rather, they give us a snapshot of the UK's 'Internet Economy'.

Sample frame

We used the results of the Office for National Statistics' (ONS) E–Commerce Survey, 2010 to identify those broad industry groupings (technically, 'SIC 2007 Sections'[25]) with the highest proportion of firms reporting transactions over a website[26] and, informed by our qualitative interviews, added Financial and Insurance Activities to this list (this Section wasn't included in the sample frame of the E–Commerce Survey). We present our findings at a finer level of resolution than SIC 2007 Sections (Table 1).

We restricted our sample frame to companies with 50 employees or more. The reason for this is that we had decided to collect data about our respondents' general financial performance from FAME (Financial Analysis Made Easy, the financial database we also used to source our sample) rather than including specific questions about this in our survey. The coverage of those data for small companies in FAME is, however, not comprehensive, hence our decision to focus on medium and large ones.[27]

Screening respondents

Given that our interest is in the Internet economy, we included in our telephone survey a screener question to determine whether respondents used their website to generate any kind of revenue (e.g. through e–commerce, as well as revenues from advertisers), and/or advertised online (so as to avoid excluding from our sample B2B companies that don't actually transact online but use their websites to generate 'leads' which translate into offline sales).[28]

Given the specialised nature of some of our questions, we made sure to ask respondents whether they were definitely the right individuals for the survey. If they were not, we sought a referral to the right person within the company. The feedback from our qualitative interviews led us to first target senior individuals in Marketing functions (the Chief Marketing Officers in sufficiently large companies).

Box 1: The Economics and Management of ICT adoption, Big Data and Analytics

The academic literature

One key finding of research looking at the connection between ICT investments and productivity is that companies need to amass complementary resources and redesign their processes and organisational structures in order to reap the benefits from ICT. ICT investments tend to increase demand for highly-skilled workers (Autor, Levy and Murnane, 2003, Bresnahan et al., 1999) and organisational change (Brynjolfsson and Saunders, 2009) – in particular, with the adoption of decentralised ways of working (Bresnahan et al., 1999, Bloom et al., 2009).

To the best of our knowledge, only one published academic study has looked specifically at the impact of Analytics and Data-Driven Decision Making (DDDM) on company performance. This 2011 paper by Erik Brynjolfsson and his co-authors uses survey and financial data for a sample of 179 publicly listed companies. Their key finding is that higher levels of DDDM are associated with increased productivity – between 5 per cent and 7 per cent – even after holding fixed other investments (including IT spend) (Brynjolfsson et al., 2011). They also find that DDDM leads to better financial performance, as measured by return on equity and stock market value.

The 'grey literature'

In the past two years alone, there has been a sharp increase in the number of large-scale studies by consultants, ICT providers and business magazines looking at data as a driver of competitive advantage. With few exceptions, they are based on international surveys of corporate executives.

These studies suggest that companies can reap substantial benefits from adopting data-driven approaches to decision making. They indicate that the barriers to better use of data in businesses are organisational and cultural as well as technical (in line with the academic findings focusing on wider ICT investments). There is also some evidence of shortages of Analytics skills in the labour market.

References: Autor, F., Levy, F. and Murnane, R. (2003) The Skill Content of Recent Technological Change: An Empirical Exploration. 'The Quarterly Journal of Economics.' 118(4): 1279–1333. Bresnahan, T., Brynjolfsson, E. and Hitt, L. (2002) Information Technology, Workplace Organization, And The Demand For Skilled Labor: Firm-Level Evidence. 'The Quarterly Journal of Economics.' 117(1): 339–376. Brynjolfsson, E. and Saunders, A. (2009) 'Wired for Innovation.' Boston, MA: MIT Press. Bloom, N., Garicano, L., Sadun, R. and Van Reenen, J. (2009 'The distinct effects of Information Technology and Communication Technology on firm organization.' NBER Working Paper No. 14975. Brynjolfsson, E., Hitt, L. and Kim, H. (2011) 'Strength in Numbers: How Does Data-Driven Decision making Affect Firm Performance?' Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1819486

Testing

We pre-tested our questionnaire to ensure that its language was meaningful for companies in different industries, and that the set of options it offered respondents was comprehensive. We did this through cognitive testing (one-on-one qualitative pilots) of the questionnaire with 15 individuals and a quantitative pilot of our first 50 responses. The survey was in the field between 16 April 2012 and 25 June 2012.

The final version of the questionnaire contained 24 questions (excluding the three

screeners on online activity and respondent suitability), and took on average just under 20 minutes to complete.

Table 1: Sector definitions

What sectors did we survey?

Sector	Including	N
Wholesale	Wholesale Activities.	109
Retail	Retail Sale Activities.	35
Business Support Services	Renting and Leasing, Other Business Activities Not Elsewhere Classified, Travel Agencies.	88
Information and Communications	Publishing, Information Services, Computer Programming and Consultancy, Software Publishing, News Agencies, Film and Video.	72
Knowledge Intensive Business Sevices	Employment Placement, Business and Management Consultancy, Technical Testing and Analysis, Market Research and Opinion Polling, Advertising.	59
Manufacturing	Manufacturing and Industrial Activities.	40
Professional Activities	Research and Development Activities, Other Professional and Technical Activities, Lawyers.	32
Financial Services	Security Dealing, Activities of Holding Companies, Insurance, Monetary Intermediation.	23
Other	Real Estate and Construction Activities, Health, Education, Creative and Entertainment Activities.	42
		500

Note: the sector breakdown of our sample is broadly similar to the make up of the FAME database of firms from which the sample is drawn, though Financial Services is under-represented and sectors like Information and Communications are slightly over-represented.

Our questionnaire captures different parts of an 'Online Customer Data Value Chain' (depicted in Figure 1), where a business uses Online Analytics to transform data generated by the online activities of its actual (and potential) customers ('inputs') into business insights that inform decision making, potentially improving company performance ('impacts').

Specifically, the questionnaire asked respondents about:

- **Online Data Inputs:** The purposes of their website and levels of online activity;
- **Online Data Collection:** How comprehensive they are in collecting different types of online customer data;
- **Online Data Analysis:** What methods and software packages they use to analyse the data, and for how long have they been using these methods;

- **Online Data Use:** How important online customer data is for decision making in different areas of the business, and

- **Online Data Impacts:** What the perceived contribution of Online Analytics is to their business performance, and how this is measured.

Regarding attitudes towards data more widely, our questionnaire included three questions on Data–Driven Decision Making adapted from Brynjolfsson et al., 2011.[29] These questions asked respondents to what extent they:

- Depend on data for decision making.

- Have all the data they need to make decisions.

- Base their decisions on 'data and analysis' over 'experience and intuition' when attempting to grow sales.

We also looked at other company factors that might augment the impact of Online Analytics on our respondents' performance – including complementary investments in areas such as skills, training and new business processes – as well as at the barriers to more effective use of online customer data, both internal (e.g. cultural and technical) and external (such as those arising from data security and privacy considerations).

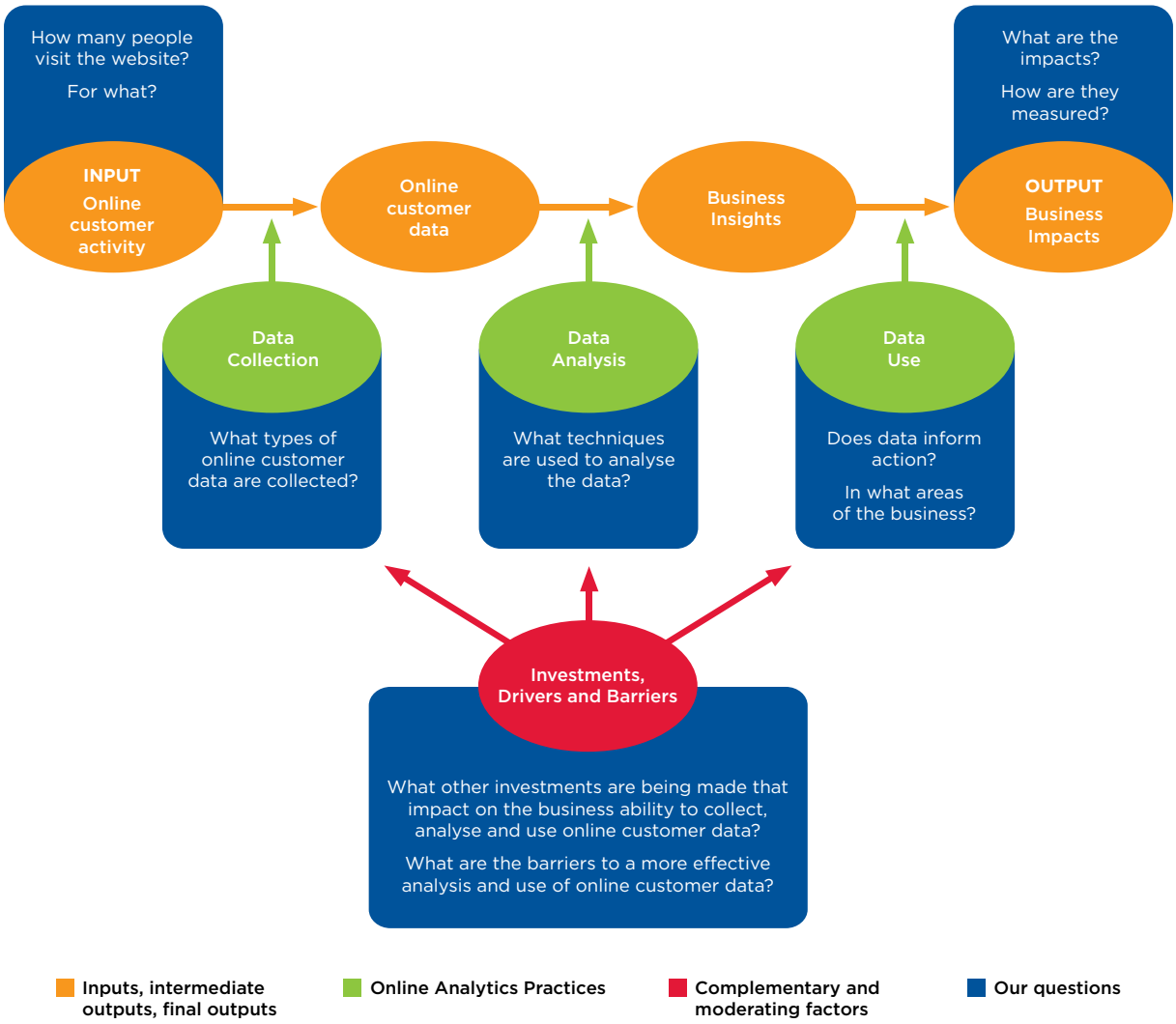
Importantly, our questionnaire captured other company characteristics and behaviours beyond Online Analytics practice that might influence or interact with our variables of interest. These included self–reported measures of innovation, overall information technology (IT) spend, marketing spend, the organisation of their workplace, reliance on other information and communications technologies (ICTs) and digital platforms, and the importance of the web as a source of revenues (which, other things being equal, may provide an incentive to invest in Online Analytics).

What's still missing: IT spend and timely data on financials

In the event, large numbers of our survey respondents, many of whom are working in marketing functions, were unable to provide detailed quantitative information on their firms' IT spend. As IT investment is likely to be an important driver of company performance, and one that might intuitively be correlated with firms' analytics practices, we are now conducting a brief follow–up telephone survey of IT teams in our sample firms to plug this gap in our data set.

As discussed earlier, our sample design was informed by the need to extract data on financial performance for all our respondents from the FAME database. However, given the variation in companies' financial accounting year–ends and the time lag that exists between the submission of these data to Companies House and their incorporation in the FAME database, we are still collating this data for some of the firms in our sample. Analysis of this data will appear in our next research paper.

Figure 1: The Online Customer Data Value Chain



3. SURVEY FINDINGS

In this section, we describe the emerging findings from our survey and present two boxes with illustrative case studies that we have conducted independently from the survey exercise.[30]

I. A DESCRIPTION OF OUR SAMPLE

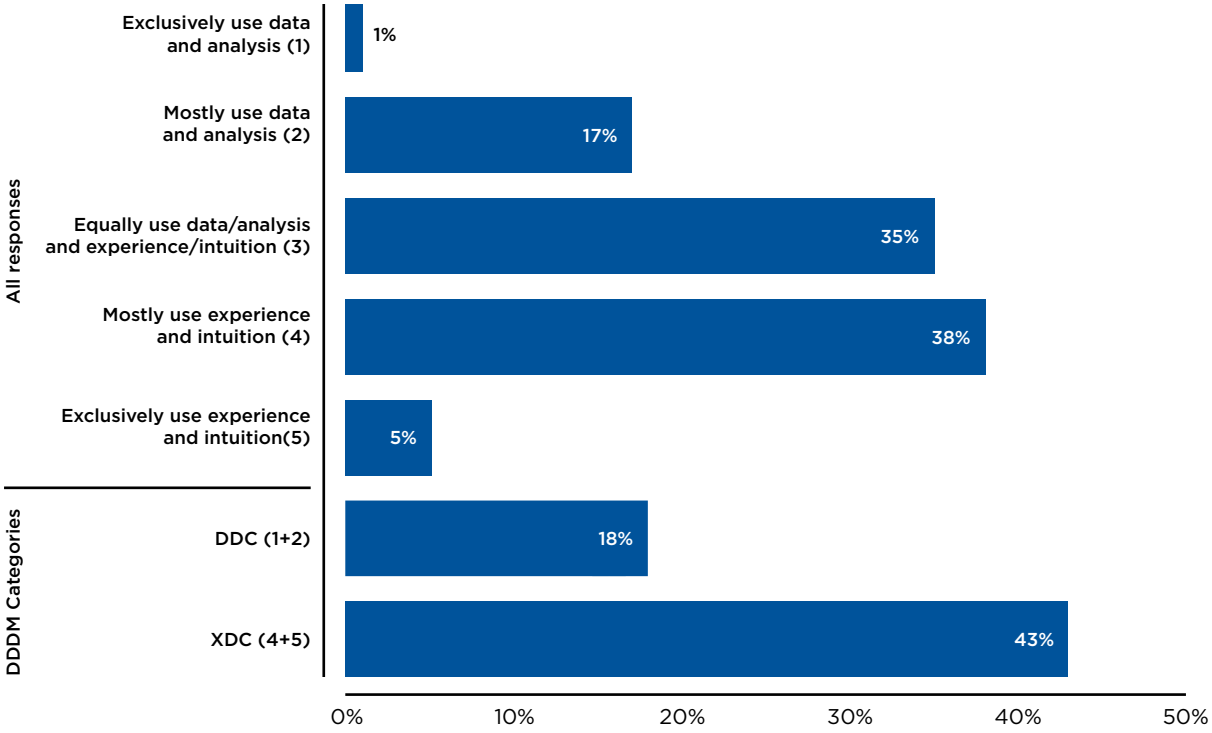
We use three key company characteristics to segment our responses – size, sector and approach to decision making.

a. **Size:** We consider three size bands – medium-sized companies of between 50 and 249 employees,[31] large companies between 250 and 999 employees, and very large companies with 1,000 employees or more.[32] Approximately two-thirds of businesses in our sample are medium-sized, one-quarter are large and just under one-tenth are very large; the median company has 155 employees.

b. **Sector:** Although we decided not to stratify our sample (that is, we did not allocate response quotas by sector), we did seek minimum thresholds of responses for all sectors so that we could use them for statistical testing.[33] In the event, the sector distribution of firms in our sample by SIC 2007 Section corresponds broadly to that for all firms in the FAME database which have more than 50 employees.

c. **Use of data for decision making:** We asked respondents whether, when attempting to grow their sales, they base their decisions on data and analysis, on experience and intuition, or a mixture of both (note that this question refers not only to the collection, analysis and use of online customer data, but to data and analysis more generally). Figure 2 presents the distribution of responses across the five options we offered respondents.

Figure 2: Use of data to make decisions when attempting to grow sales (percentage of firms in the sample)



Note: The percentages do not add up to 100 per cent in the case of 'All responses' because we have excluded those businesses which said they 'Don't Know'. The percentages in DDDM categories do not add up to 100 per cent because we exclude 35 per cent of responses 'in the middle' (Category 3) and 4 per cent who said they 'Don't know'.

We used a subset of the responses to this question to create two categories of companies – 'Data–Driven Companies' (DDCs), which report that when attempting to grow their sales they make decisions based on data and analysis more than on experience and intuition (we also call these companies 'datavores') and 'Experience–Driven Companies' ('XDCs'), to refer to those companies who say that they base their decisions on experience and intuition above data and analysis. Although this classification does not include companies 'in the middle' (around a third of our sample), the findings we report are robust to their inclusion.[34]

Data–Driven Companies are in a minority

As Figure 2 shows, our respondents are much more likely to ground their decisions primarily on intuition and experience than on data and analysis when attempting to grow their sales – 18 per cent fall within our DDC category, compared with 43 per cent of XDCs. It seems that despite of all the 'buzz' around Big Data, Analytics and Data–Driven Decision Making, its adopters are still in a minority, even in the Internet economy where our respondents trade.

We find few significant differences in the use of data for decision making across different company sizes

The exception is that medium–sized companies in our sample tend to rely on intuition and experience more often than larger and very large ones when attempting to grow their sales (46 per cent of them are XDCs compared with 29 per cent – still a significant proportion – of very large companies) (Figure 3).

Figure 3: Use of data for decision making in companies of different sizes

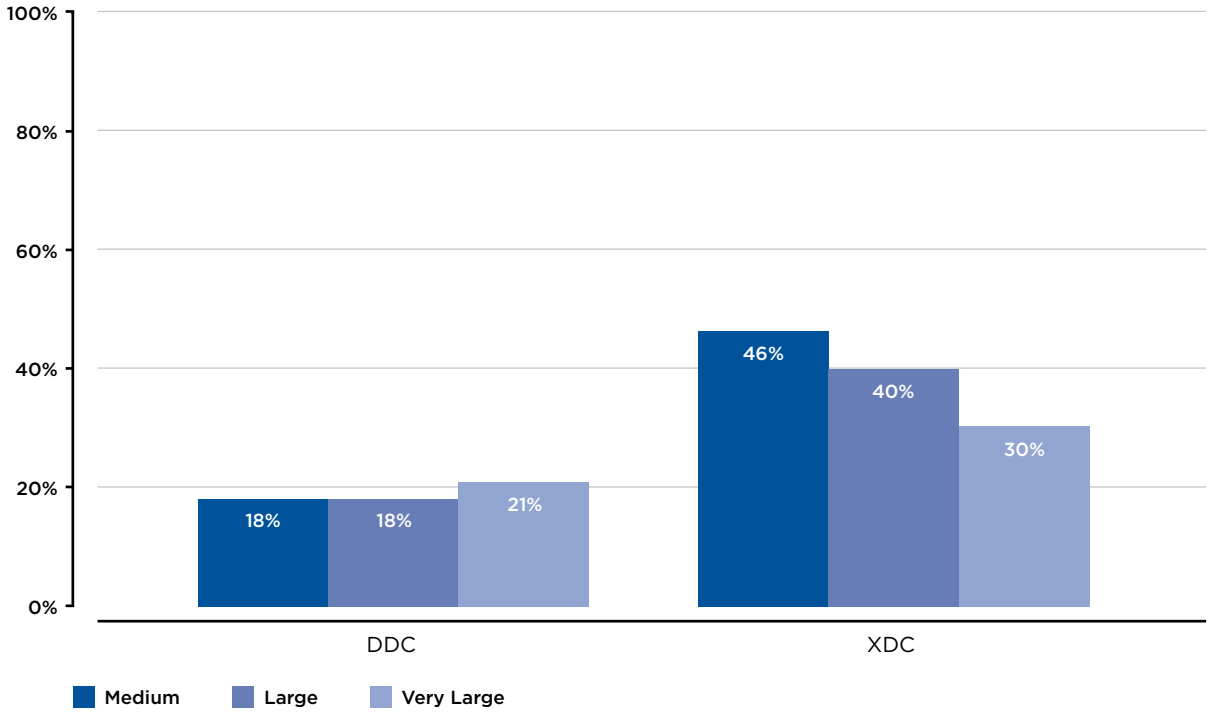
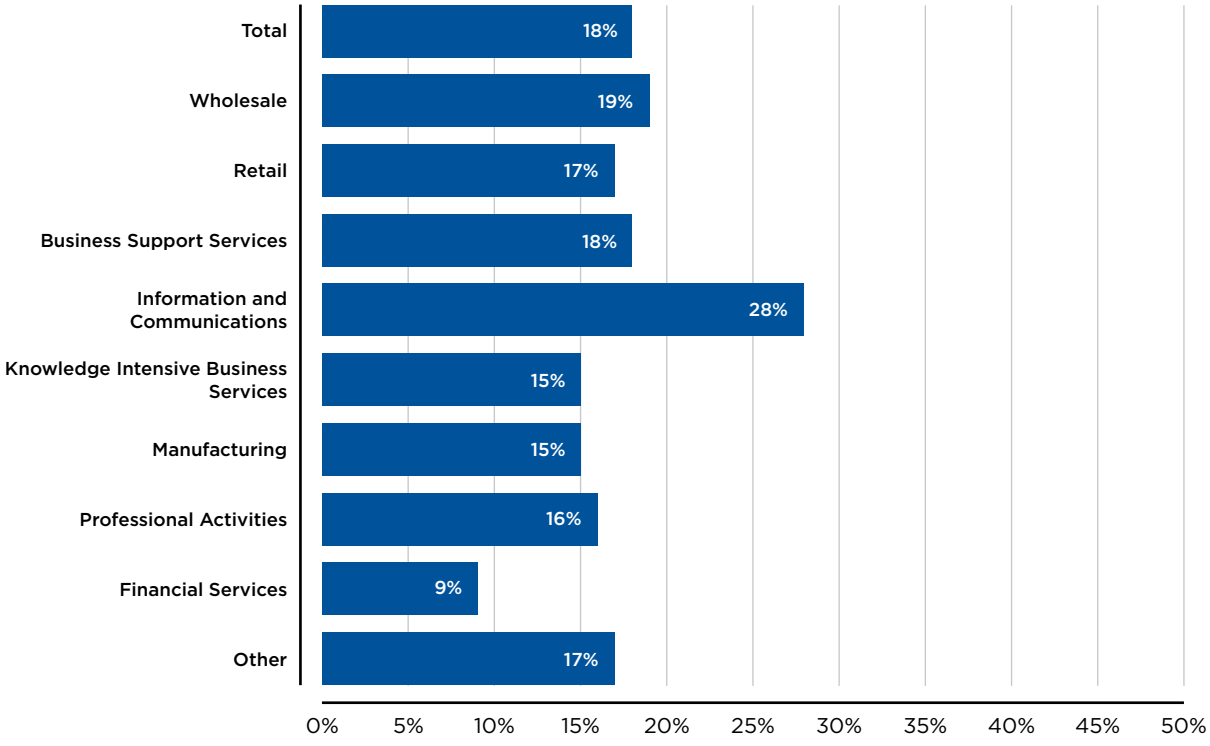


Figure 4: Proportion of DDCs across different industries



Box 2: Online Analytics aren't just about online

When talking about Online Analytics, Big Data and Data-Driven Decision Making, most people immediately think of Internet platforms like Google, Amazon, eBay or Facebook. It is easy to forget that the types of online data these businesses use can also be collected by firms in other parts of the economy. Many of the tools with which they analyse data are relevant across industries, and so are the urgent questions that they seek to answer – how to increase their visibility, return on marketing investment or customer satisfaction? And so forth.

Arguably, the Internet is making all sectors of the economy digital – or, as David Kirkpatrick recently pointed out in Forbes Magazine, *"Now every company is a software company."*

This means that Online Analytics is becoming a source of competitive advantage regardless of industry. Elisa Interactive, an agency specialising in digital marketing, performance tracking and website optimisation, recently worked with UK Flooring Direct, a floor specialist that has now been featured twice in the *Sunday Times* Fast Track 100, to redesign its sample offers in order to increase sales. They also used performance data from Photobox, Europe's number one personalised online photo service, to determine why their mobile visitors were significantly less likely to buy than other groups, and redesigned their website to resolve this problem.

The potential impacts of online data go beyond website variables and metrics – they can improve the efficiency of 'offline' parts of the business, something that is particularly important for companies that don't engage in e-commerce but do use their websites to generate business leads and *"bricks and mortar"* sales. As Robert Jackson from Elisa Interactive put it, *"merging online with other types of data is where the market is going"*.

T W White & Sons, a Surrey and Kent car dealer group that has now been in operation for almost half a century is a good example of this. In 2008, they moved away from a system where they logged test drives and sales on slips of paper and spent most of their advertising budget on print media. With the help of Calltracks, a call analytics company, they started measuring which advertising campaigns – including keyword search – lead to individual calls, as well as sales. This helped them work out, of the publications where they advertised, which generated most revenues, and which salespeople were performing better. They have used such information to identify members of staff that may require training and professional development.

According to Ali White, the group's Marketing Communications Manager, using data means that the firm *"can cut on what doesn't work, and invest what we save on the things that work".* As a consequence of this, between 2008 and 2012 TW White & Sons have decreased their marketing costs per sale by 55 per cent (the percentage for used cars is 82.7 per cent).

Sources: Robert Jackson, Elisa Interactive, and Ali White, T W White & Sons http:www.twwhiteandsons.co.uk. For more information, see: http://blogs.telegraph.co.uk/technology/robjackson/100007363/the-digital-gold-rush-why-business-is-hungry-for-big-data/ Kirkpatrick, D. 2011.,'Now every company is a software company.' available from http://www.forbes.com/sites/techonomy/2011/11/30/now-every-company-is-a-software-company/

Information and Communication companies are forging ahead in their use of Data-Driven Decision Making.

Information and Communications companies in our sample are proportionately the most reliant on data for their decision making (over one-quarter of them are 'DDCs' in contrast to 18 per cent for the sample overall), and Financial Services businesses the least (Figure 4). [35]

Bearing this in mind, it is important to note that we find DDCs across all sectors of the Internet economy (see Box 2 for a case study of a data-driven car dealer group).

II. MAPPING ONLINE ANALYTICS ADOPTION IN OUR SAMPLE

In this sub-section we navigate the different parts of the Online Customer Data Value Chain we presented in Figure 1, examining how our respondents collect, analyse and use the online data generated by their customers, and highlighting differences across size-bands and sectors.

Our respondents have significant online operations, and use their websites to learn about their customers

Forty per cent of businesses in our sample report using their websites to transact with end-consumers and other businesses (this compares, for example, with 15.3 per cent of respondents in ONS's 2010 E-Commerce survey).[36] Almost three-quarters of our respondents use their website to generate business leads.

On average, our respondents earn 15 per cent of their business revenues through their website, a substantial amount compared with the wider business population. Although this proportion is skewed by the presence of a handful of 'pure-play' companies generating the majority or all of their revenue online, our median company still generates around 8 per cent of its business via its website, compared with an economy-wide average of 4.2 per cent.[37]

In addition to using their websites to generate sales and new business opportunities, firms clearly see them as a source of valuable intelligence about their customers – 72 per cent of them report that one function of their websites is to gather data about customers (including through pages or products viewed, as well as from registration forms), and 63 per cent say that their website also helps them to collect customer feedback.

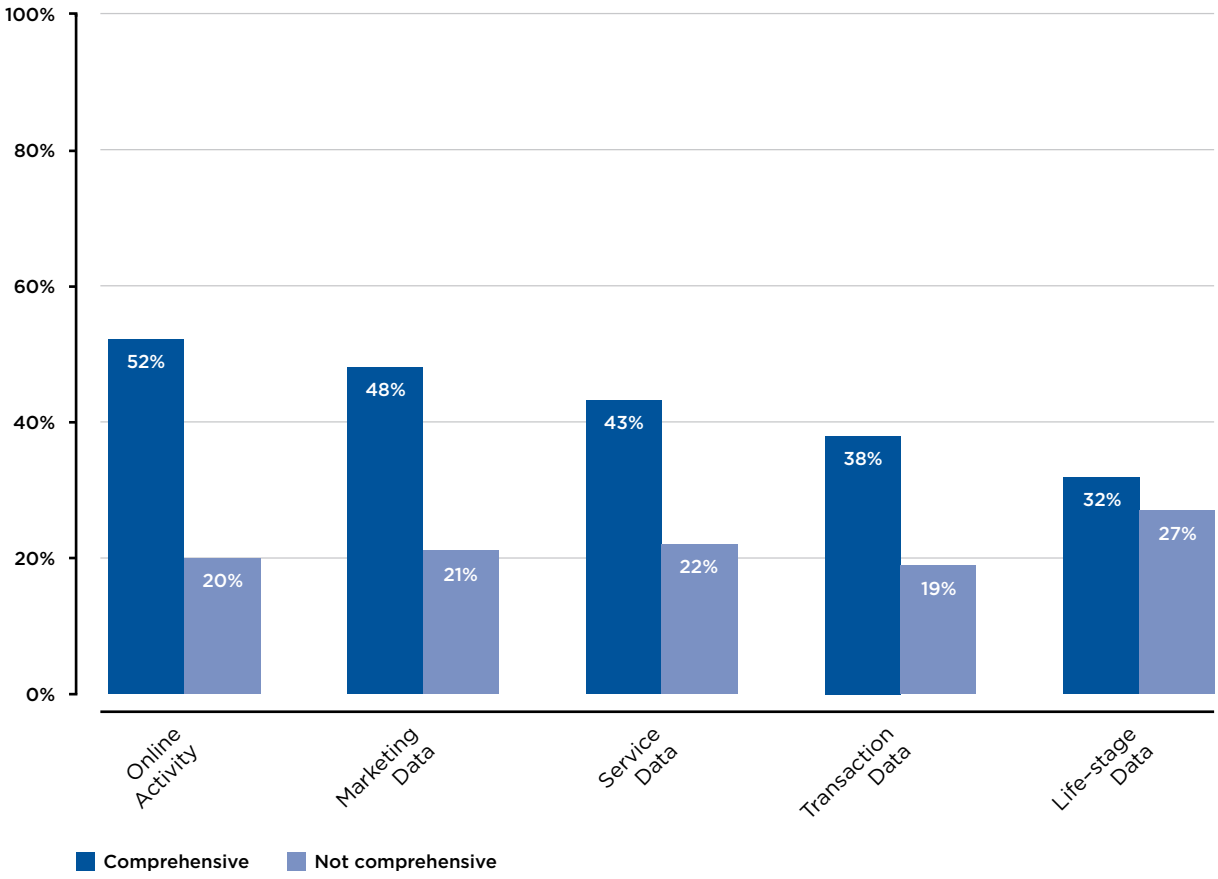
They measure website activity, and use these metrics to improve website performance

Figures 5, 6 and 7 respectively illustrate the adoption of Online Analytics practices in terms of online data collection, analysis and use.

They show that Web Analytics is dominant among our respondents' online data activities – more than half report that they are comprehensive in collecting online metrics such as web pages visited, incoming links etc. (Figure 5), and over 70 per cent claim that online customer data are important for decisions about how to improve their websites (Figure 7). It is interesting that the types of analysis most frequently reported in our sample (Figure 6) – basic trends and reports, and dashboards and visualisations – are standard features of popular Web Analytics packages such as Google Analytics (which just under 80 per cent of our respondents report using).[38]

These findings are consistent with the 'Web Analytics Maturity Model'.[39] According to this model, companies generally begin their Analytics activities by collecting and deploying simple website data, before they move into collecting other types of data, which they analyse with more advanced techniques. In this respect, the Internet acts as a low-cost, low-risk, high-visibility environment where companies can experiment with the collection and use of customer data, hone their Analytics capabilities and gauge their benefits, possibly helping to build the case in the organisation for their wider adoption.[40]

Figure 5: Comprehensiveness in the collection of different types of online customer data (percentage of firms in the sample)



Note: The figures do not add up to 100 per cent because the base excludes those firms that answered neither 'comprehensive' nor 'not comprehensive' or 'Not relevant' or 'Don't know'.

Figure 6: Current types of online customer data analysis (percentage of firms in the sample that report using the technique)

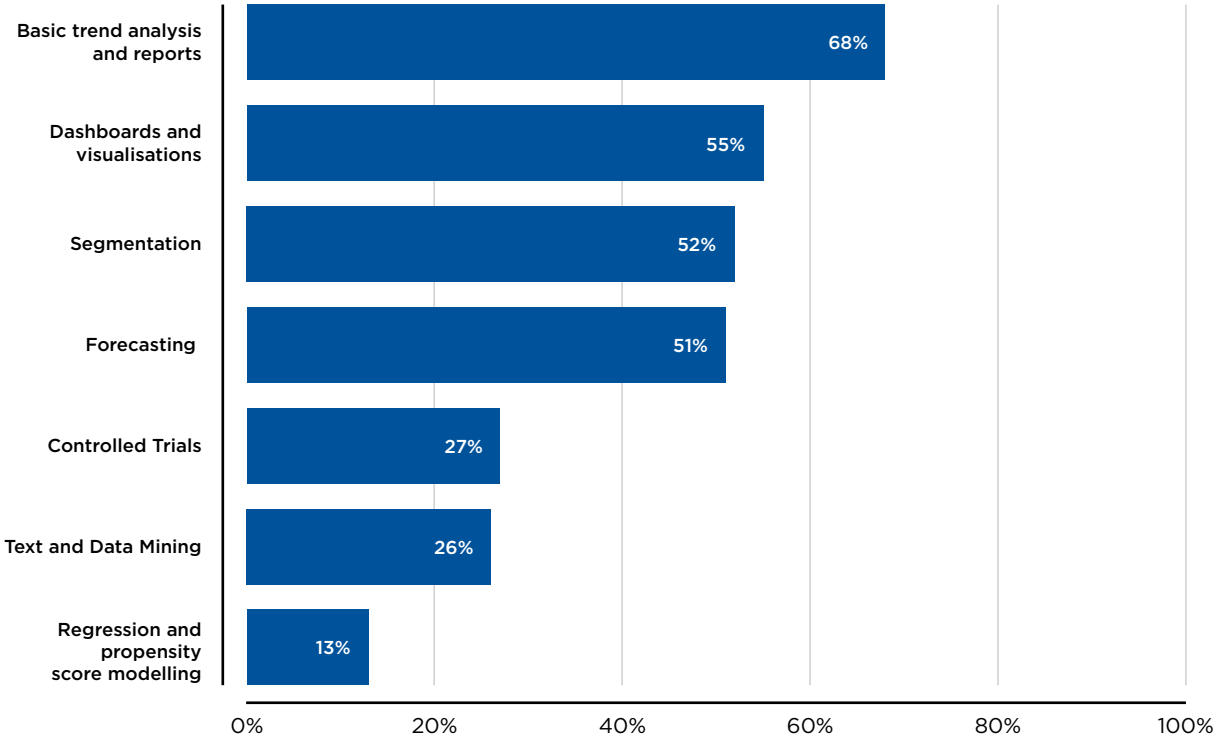
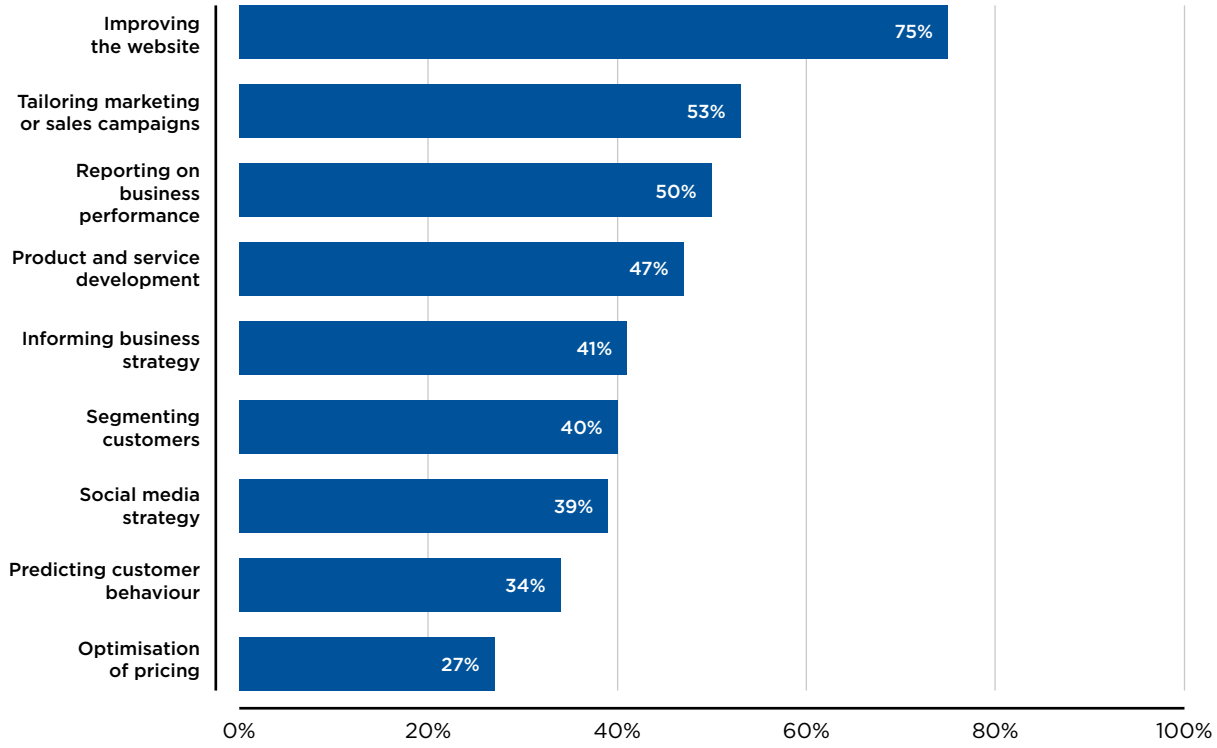


Figure 7: Importance of online customer data for decision making in different areas (percentage of firms in the sample saying online customer data is important for that area of the business)



Outside website metrics, there are large gaps in the collection of other types of online customer data

Less than 40 per cent of firms in our sample collect transaction data comprehensively (including products and services purchased by customers), and even fewer have comprehensive life-stage (that is, historical) data on their customers (Figure 5). Note that these numbers exclude respondents who reported that a given type of data is not relevant for their business. In other words, in both of these cases, the majority of companies fail to collect comprehensive data even when it is deemed relevant to their business.

Correlation isn't causation

Figure 6 shows that the majority of our respondents use basic and descriptive tools to analyse their online customer data – at best, these should help them to identify patterns, or correlations, in their data. Use drops off significantly when it comes to more sophisticated techniques (including controlled trials, e.g. A/B tests, data and text mining and regression and propensity score modelling[41]) that could help them to determine what the causal relationship is between behaviours and impacts, that is, extract meaning from their data.

Somewhat surprisingly, a significant proportion of online-active businesses (31 per cent) fail to perform even elementary analyses of their online customer data, such as basic trend analyses and reports.

Other than for website optimisation, the use of online data to inform decision making is patchy

As Figure 7 shows, with the exception of the website improvements we have already mentioned, the majority of businesses in our sample do not feel that online customer data is important for decision making in key business areas. This includes price-setting, where less than 30 per cent of firms say that online customer data is important. That percentage is still only 39 per cent even for companies which generate more than 50 per cent of their revenues online.

More generally, it seems, data is being used operationally and retrospectively, rather than to inform company strategy or to try and predict customer behaviour.

We see some variation in Online Analytics adoption across sectors

Unsurprisingly, retailers are much more likely than the rest of the sample to sell to consumers on their websites (80 per cent compared with 43 per cent overall), and therefore collect transaction data more comprehensively (63 per cent report that this is the case). Information and Communications companies (including digital publishers) that are particularly active in gathering customer data through their websites also collect more marketing data (such as campaign activity), and say that it is important for developing tailored marketing campaigns, and for designing and evaluating their social media strategy (Figure A1 in the Appendix).

Does size matter for Online Analytics adoption?

There is a theoretical ambiguity about the relationship between company size and the adoption of Online Analytics that is reflected in our findings.[42] We find few substantive differences across company size. The only noteworthy exception is that very large companies are much more likely to use the more sophisticated analysis tools – for example, 26 per cent of them say they make use of regression analysis or propensity score modelling compared with 13 per cent of the sample overall; over 40 per cent say that they run controlled trials compared with 27 per cent for all businesses (Figure A2 in the Appendix).

III. ONLINE DATA DIFFERENCES AND DIVIDENDS

Having looked at the adoption of Online Analytics practices in our sample of firms overall, we move on to examine the differences between Data-Driven Companies – our 'datavores' – and experience-driven companies.

Data-Driven Companies are walking the walk...

We see a clear contrast between DDCs and XDCs across the whole Online Analytics piece: from their collection of online consumer data (Figure 8) to their use of analysis tools (Figure 9), and deployment of data for decision making (Figure 10).

Overall, DDCs are much more comprehensive in their collection of online customer data, much more likely to subject it to rigorous analysis, and rely on it more heavily to inform decision making across the business. The differences are particularly dramatic in their use of sophisticated analysis tools such as controlled trials, data and text mining, regression analysis and propensity score modelling (which DDCs are two to three times more likely to use than XDCs, Figure 9), and in the use of online customer data to inform business strategy (which 61 per cent of DDCs report doing, compared with just over a quarter of XDCs, Figure 10).

Even here, however, we find large numbers of Data-Driven Companies that do not make any use of controlled trials (as many as 60 per cent), while 80 per cent do not employ regression analysis or propensity score modelling (Figure 9). And less than 40 per cent of DDCs see online customer data analysis as important to setting prices for their products (Figure 10).

Figure 8: Comprehensiveness of online customer data collection (by type of company)

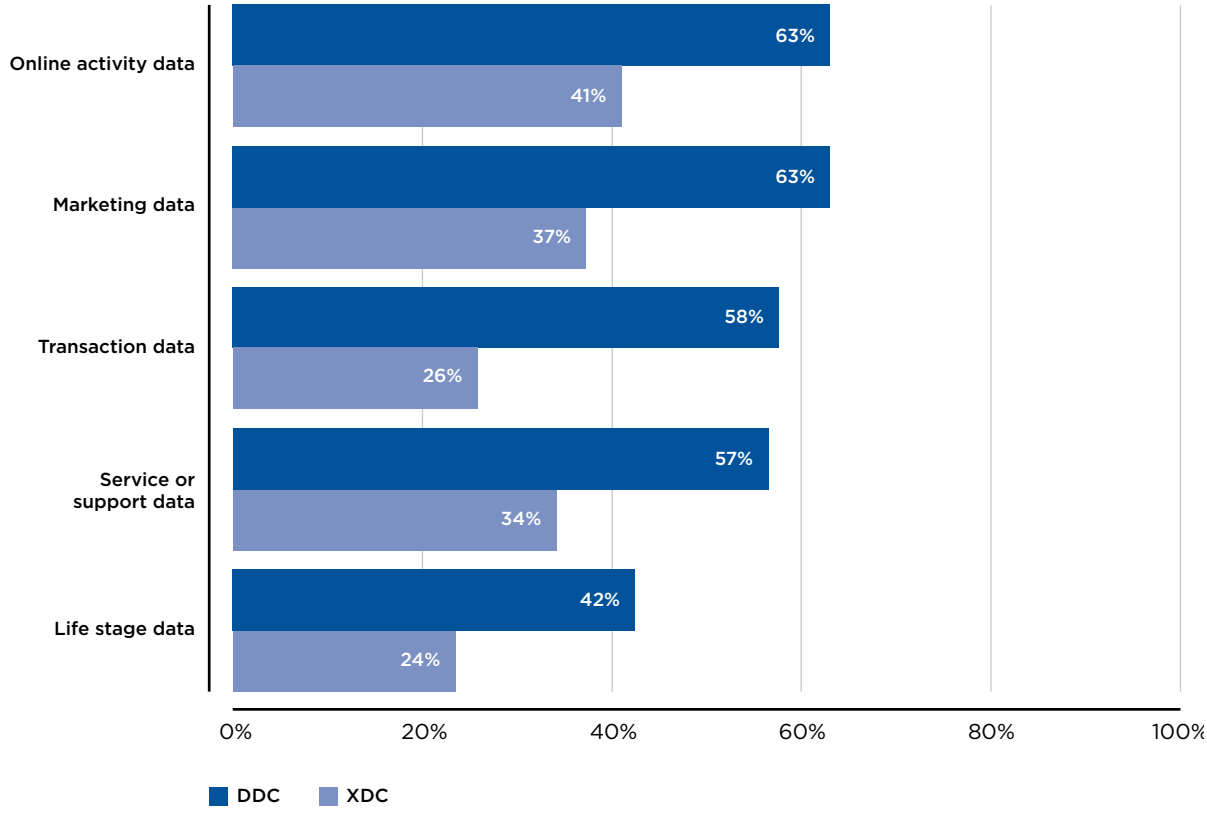


Figure 9: Current types of online customer data analysis being done (by type of company)

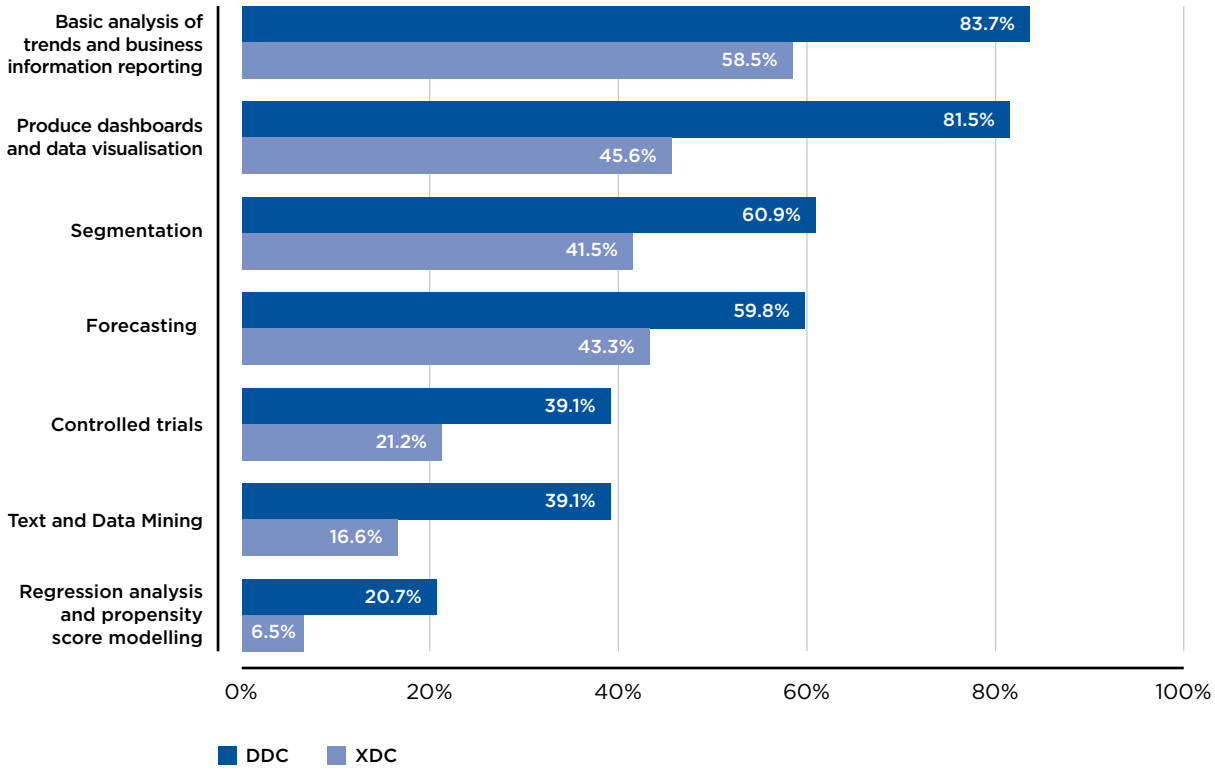
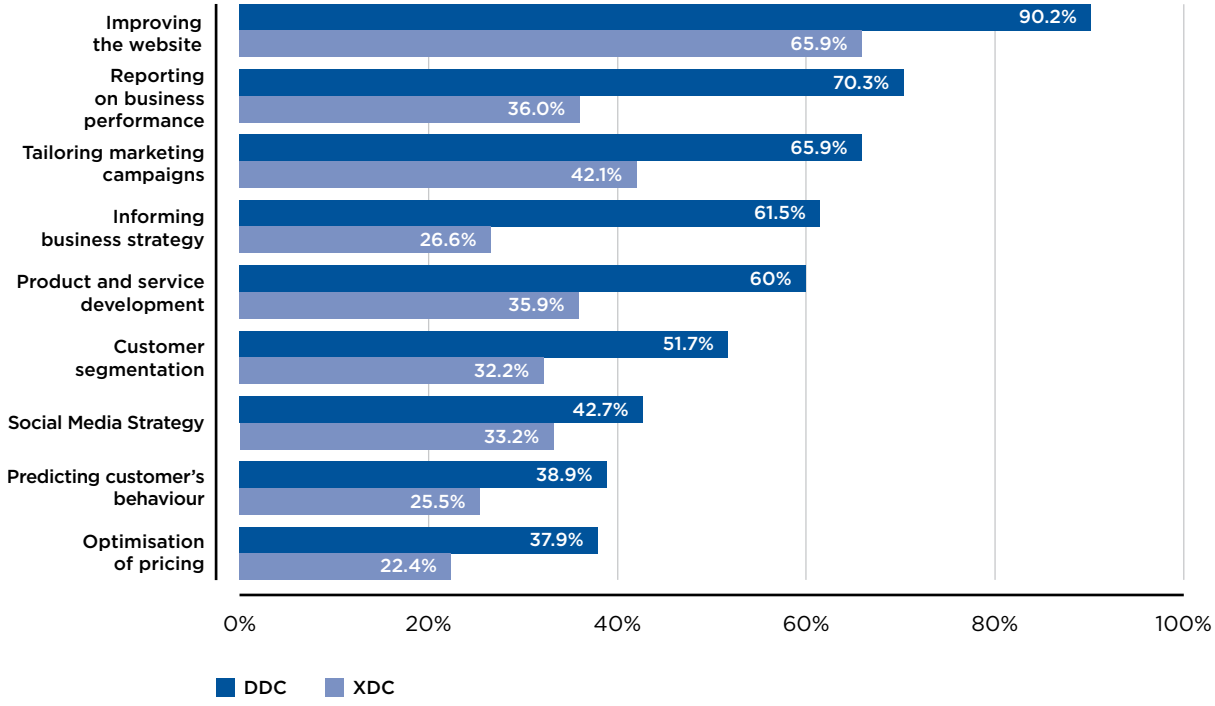


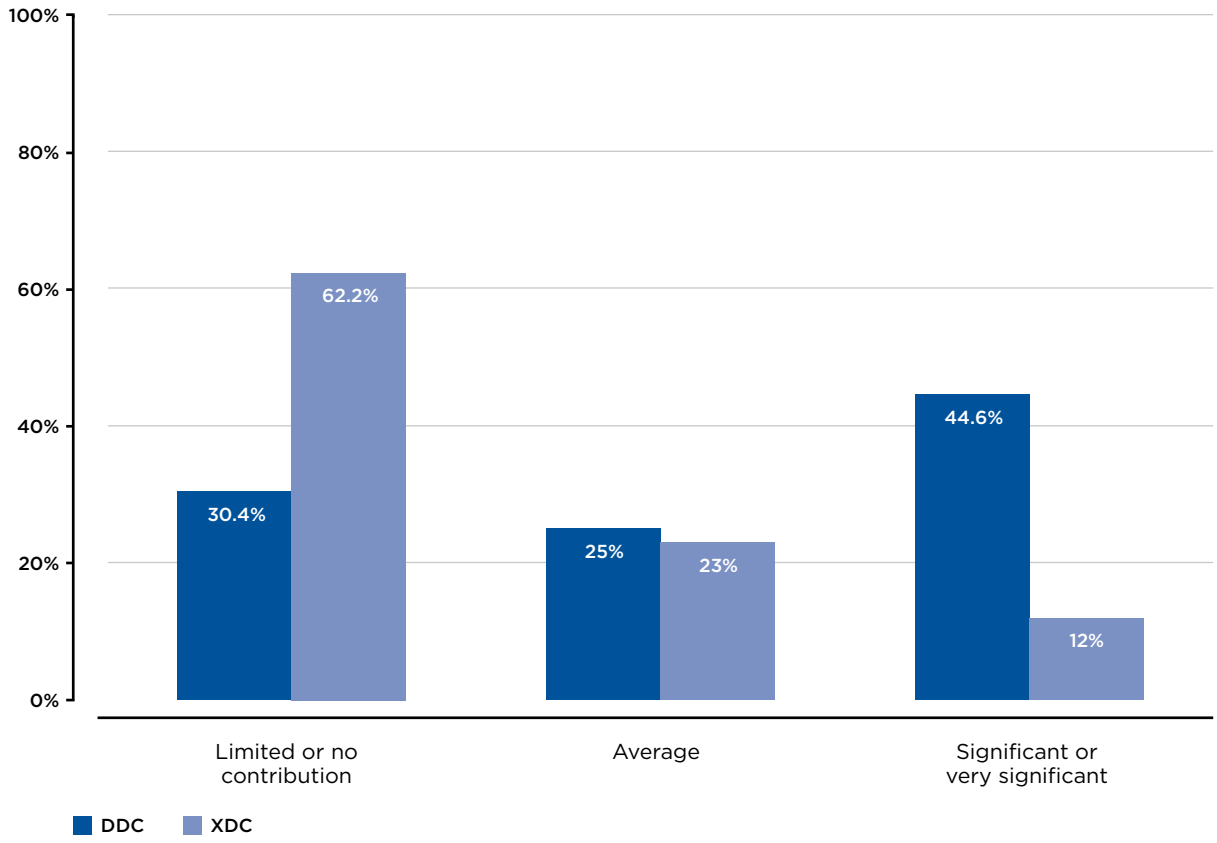
Figure 10: Importance of online customer data for decision making in different areas (by type of company)



...And apparently reaping – and measuring – the benefits

DDCs report that they are reaping high returns on their Online Analytics investment – when asked about the contribution of their ability to collect and analyse online customer data to their business performance, they are almost four times as likely as XDCs to say that this has been significant or very significant (see Figure 11).

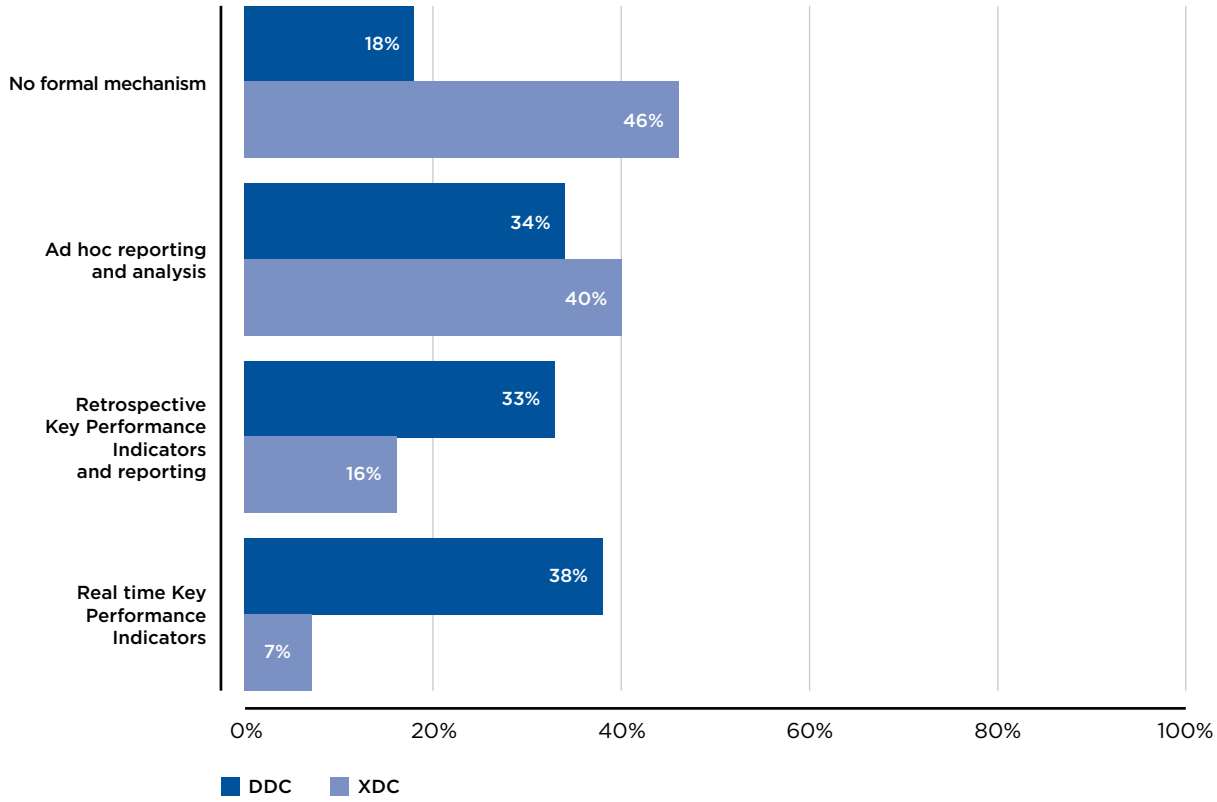
Figure 11: Contribution of online customer data to business performance (by type of company)



These results need to be interpreted with caution however; it may simply be that DDCs operate in markets – or use business models – where an efficient use of online customer data is particularly important.[43] We attempt to tackle this issue in the multivariate analysis we present in Section 4.

The results may also reflect the possibility that XDCs do benefit from Online Analytics but that they are less effective or thorough at measuring those impacts. Figure 12 shows how businesses evaluate (or do not) their Online Analytics: it is clear that the majority of XDCs in our sample fail to measure the effectiveness of their Online Analytics in any formal or systematic way, in contrast to many DDCs (though, again, a significant minority of these do not formally evaluate their use of Online Analytics). Note in particular the stark difference in the percentage of XDCs using real time Key Performance Indicators versus DDCs.

Figure 12: Evaluation of Online Analytics (by type of company)



Note that the categories in this figure are not mutually exclusive, and as such add up to more than 100 per cent.

What is happening with those DDCs where Online Analytics fails to contribute to business performance?

We were intrigued to find that a significant minority of DDCs (30 per cent) report that Online Analytics has made limited or no contribution at all to their performance (the equivalent percentage for XDCs is 62 per cent).

We investigate the possible reasons for this by probing further the differences in Online Analytics adoption between those DDCs saying that Online Analytics contributes weakly to business performance – the 'lower impact' group' – with those reporting a stronger performance – 'the higher impact' – group.[44]

We find that in a large number of instances, the 'lower impact' group of DDCs have a much lower propensity to use sophisticated analytics techniques such as controlled trials, regression analysis and propensity score modelling, and text and data mining than the 'higher impact' group of DDCs.

The differences between our two groups are particularly stark in their reliance on online customer data for decision making across the businesses, where a significantly lower percentage of 'lower impact' DDCs rely on online customer data than 'higher impact' businesses in seven out of the nine key decision making areas considered. Put simply, the DDCs that are not experiencing benefits from Online Analytics do not appear to analyse or use their online customer data for decision making.

Again, this result could be explained if Online Analytics plays a less important role as a driver of competitive advantage in the markets that 'lower impact' DDCs operate in. Still another possibility is that there are reporting biases in the responses to the question we have used to construct our DDC variable if, for example, there are a substantial number of companies who think of themselves as 'data-driven' when in fact they are not (which would mean that the actual proportion of DDCs in our sample is lower than the 18 per cent we report). This echoes McAfee and Brynjolfsson's (2012) criticism of some companies for pretending to be more data-driven than they in fact are, and *"[spicing] up their reports with lots of data that supported decisions they had already made using the traditional HiPPO (Highest Paid Person in the Office) approach."*[45]

IV. MAKING BETTER USE OF ONLINE CUSTOMER DATA: INVESTMENTS AND BARRIERS

If companies can generate such benefits from using their online customer data, why don't we see more of them doing it?

In this sub-section, we look at the investments in resources and capabilities that should help companies in our sample use online customer data more effectively, and at their adoption of new techniques for data analysis. We also explore what barriers they face when trying to use their online customer data more effectively.

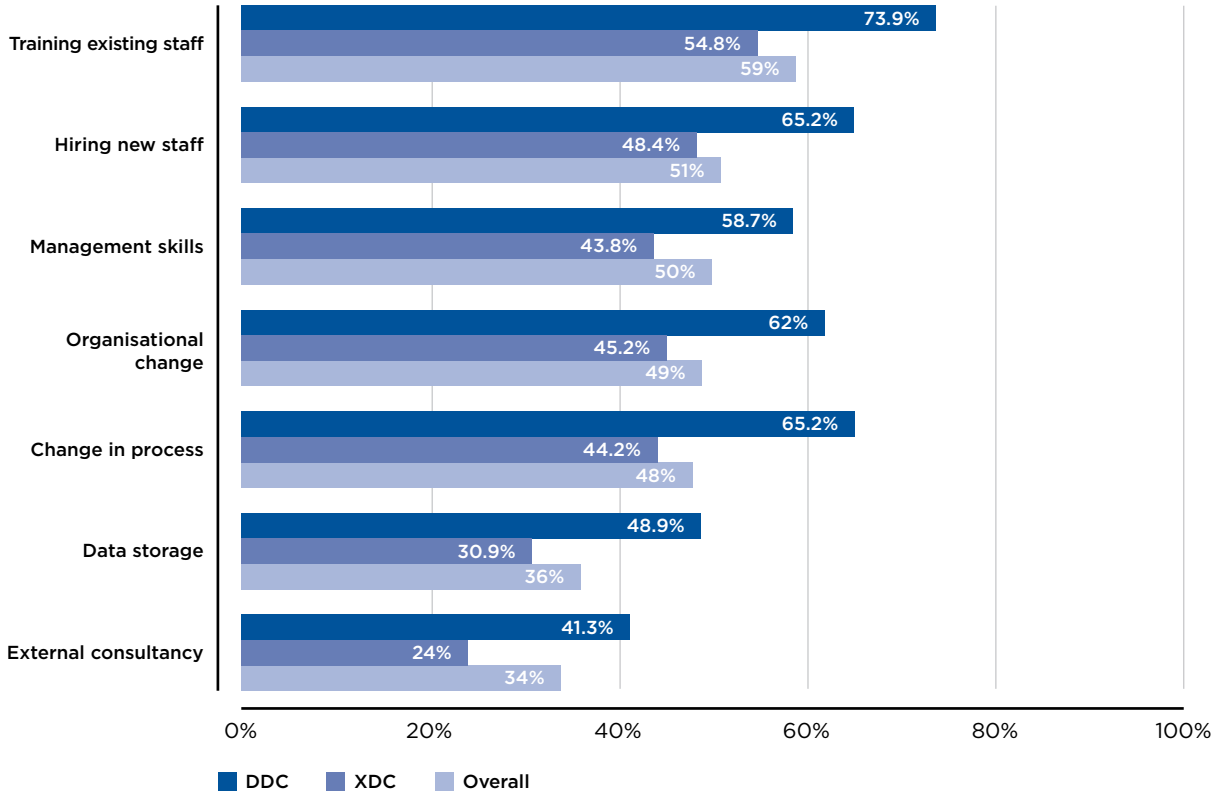
Our respondents are making substantial investments in resources and capabilities that complement Online Analytics

We asked our respondents whether, over the last two years, they had undertaken significant investments that affected their ability to gather and use insights from online customer data. These included investments in data storage capacity, workforce skills, external expertise (the use of consultants), and in business processes and organisational change (Figure 13).

Overall, businesses in our sample appear to be taking the need to build up their Online Analytics capabilities seriously – the average company reported having undertaken significant investments in at least three of the areas we asked them about. But this masks a great deal of variation – around 18 per cent of all respondents to our survey had, in fact, undertaken no investments at all complementing Online Analytics over the last two years.

Most of the activity has concentrated on skills development – including training and recruiting new staff, and developing management skills: more than half of our sample reports investments in these areas. This is followed by changes in organisation and processes, and, less often, acquisitions of data storage and in the use of external consultants.

Figure 13: Complementary investments supporting the collection and use of online customer data (percentage of firms in the sample reporting having invested in the different areas)



The divide between DDCs and other companies appears to be widening

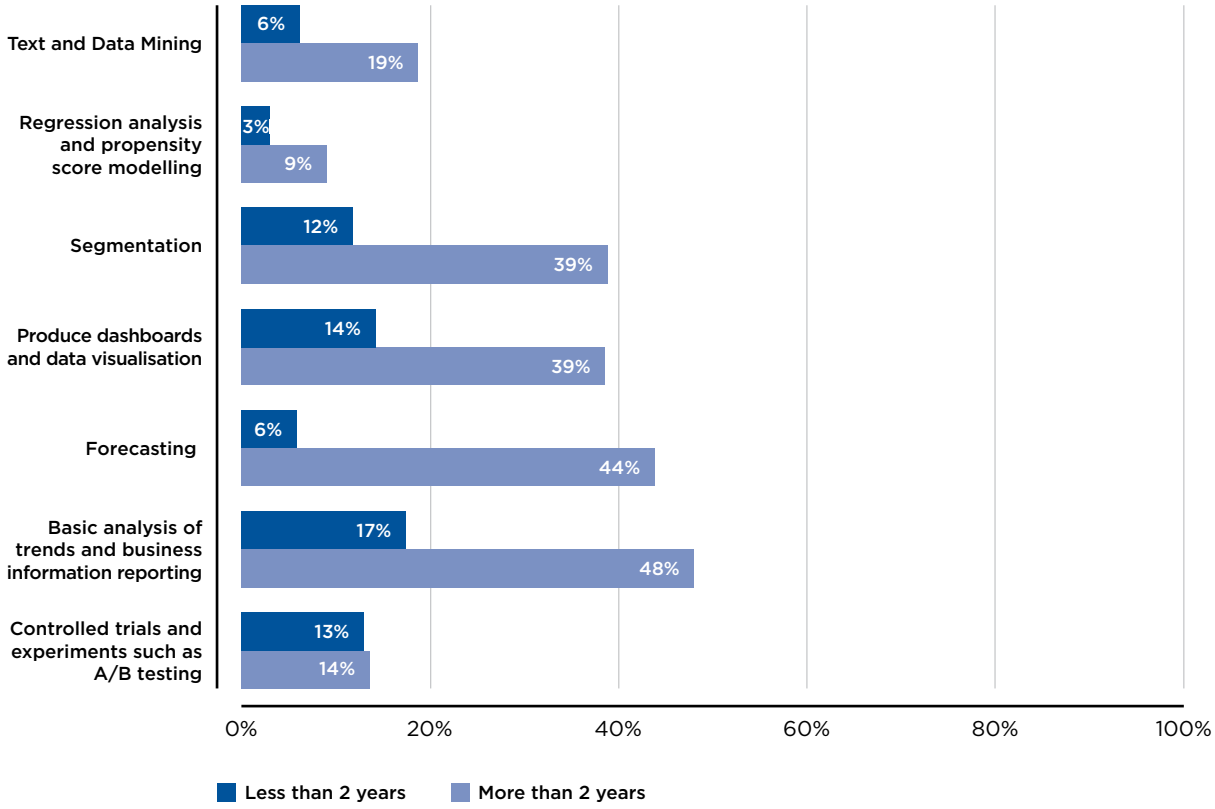
The average DDC has made over four substantial investments in areas relevant for Online Analytics over the last two years – as many as three-quarters are training their staff to improve their ability to collect and analyse online customer data, for example, and almost two-thirds have taken on new staff with expertise on analytics. This compares with investment in just under three areas by the average XDC.

An implication is that DDCs are investing more heavily in those capabilities and resources that should help them generate further benefits from their data – that is, there appears to be a 'widening divide' between DDCs at the forefront of the online data revolution and others, which is consistent with the results from previous research.[46]

Heading towards an experimental web?

We also looked at the pace of adoption of Online Analytics over the last two years, finding that most of those companies using Online Analytics techniques have in fact been doing so for some time (Figure 14).

Figure 14: Adoption of Online Analytics techniques in the last two years (percentage of firms in the sample)



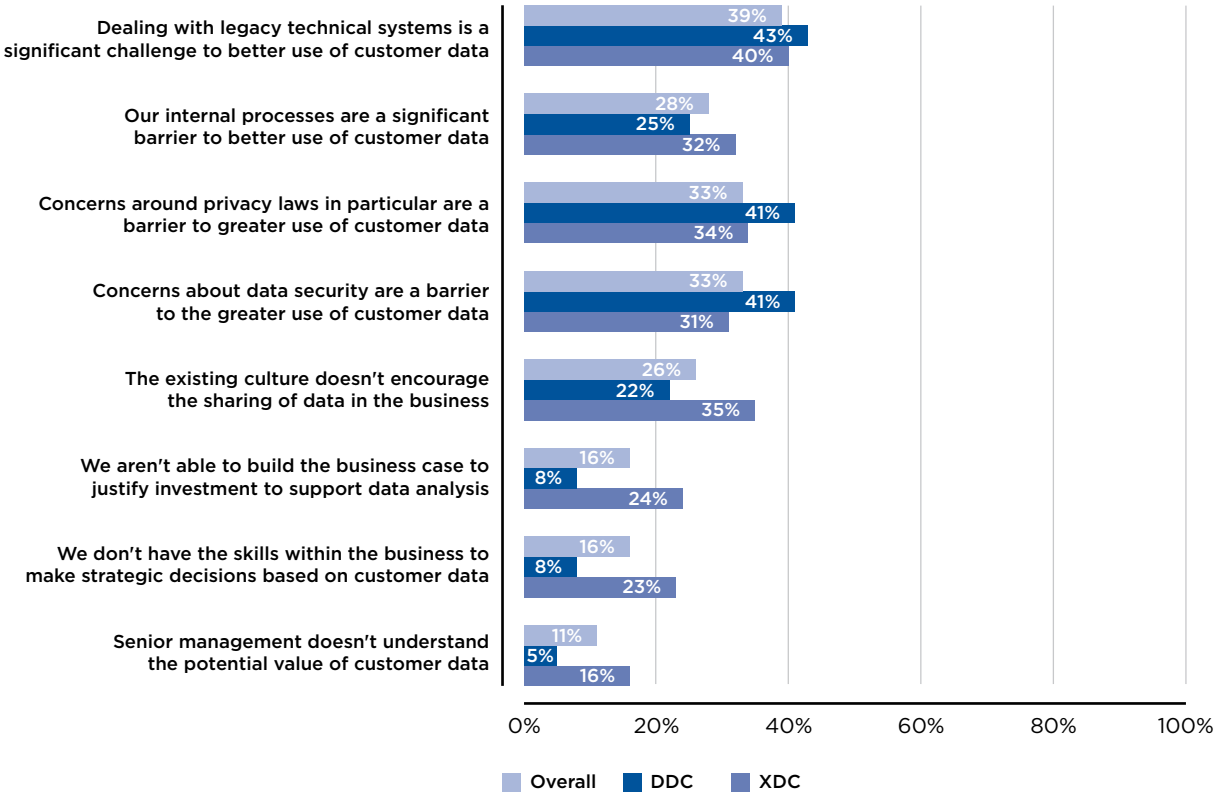
A noteworthy exception is in the use of controlled trials and experiments which has grown very rapidly in the past two years: almost half of those companies in our sample currently running controlled experiments online only started doing them in the last two years.

This is an important finding – a great deal has been made of the Internet's potential to be a 'lab' for business experimentation, and in particular through the use of randomised controlled trials that can help businesses measure the impact of their activities more accurately, quickly and more cheaply than before.[47] Most of the evidence for this 'randomisation movement' has been impressionistic, however, or based on case studies of leading Internet companies like Google, Amazon and Facebook. Our findings suggest that experimentation practices are rapidly diffusing across the population of UK businesses (though, admittedly, starting from a low base), with potentially significant returns for their business efficiency and innovation.

What are the barriers to a better deployment of online customer data by UK companies?

Figure 15 illustrates the barriers that our respondents say they face when seeking to make more effective use of their online customer data. Lock-in to legacy technical systems is the most common barrier reported. When we compare DDCs and XDCs, we find that proportionately greater numbers of DDCs are obstructed by what one could describe as 'external' barriers – in particular, concerns with data security and privacy, while XDCs mostly face 'internal' barriers related to business processes, culture and, to a lesser degree, lack of skills and management support.

Figure 15: Drivers and barriers to Online Analytics adoption (percentage who report that the statement applies to their business)



Overall, and somewhat surprisingly, the majority of businesses in our sample do not report a lack of skills at the strategic level and lack of management support as barriers to the adoption of Online Analytics (though which may perhaps be easier to understand given the strong investments in these areas over the last two years).

V. DATA, INNOVATION AND DECENTRALISATION

We conclude this Section by exploring the complex relationships between the adoption of DDDM and Online Analytics, other business innovation strategies and the organisation of the workplace in our sample.

Data-Driven Companies claim to be more innovative

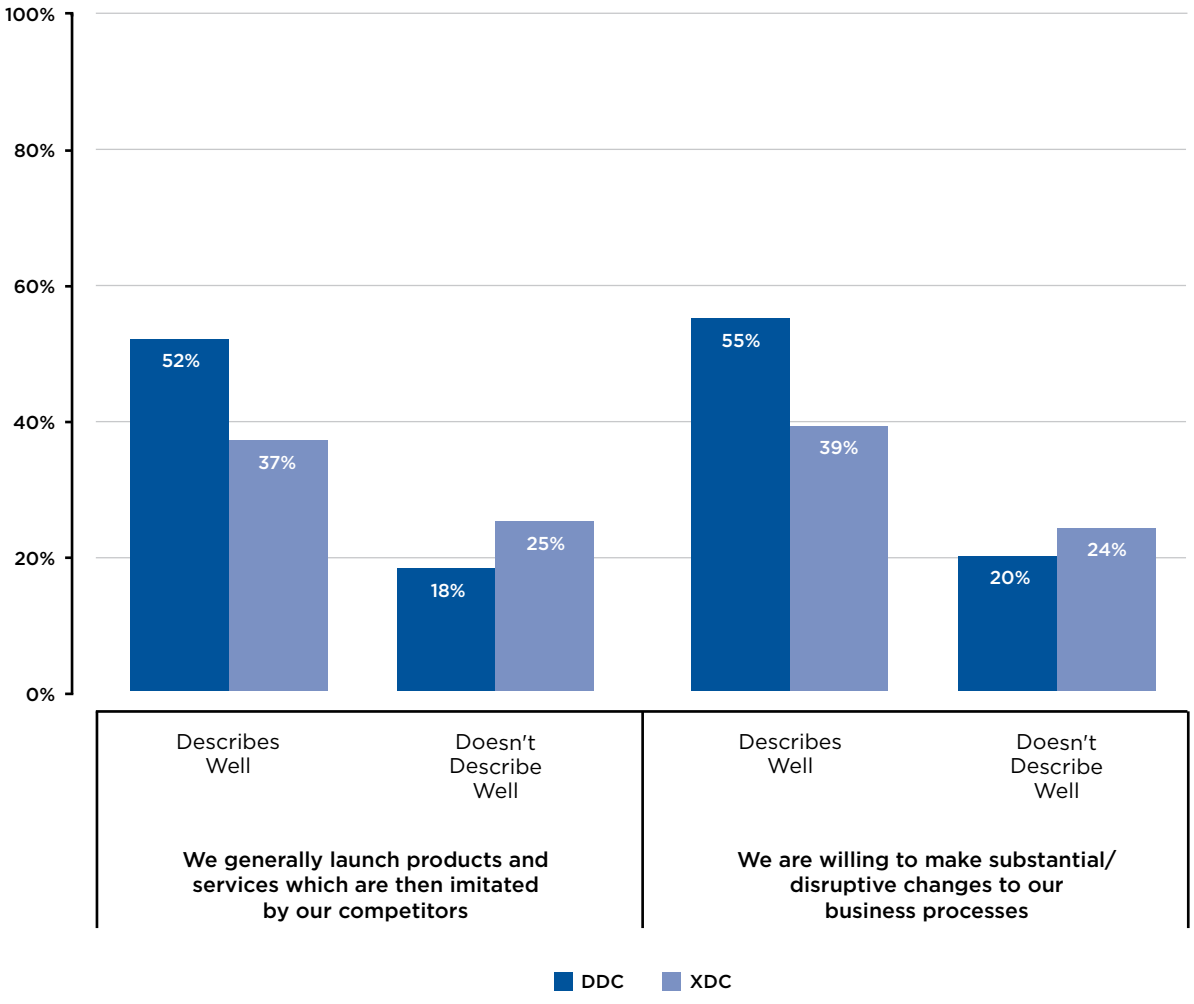
We asked our respondents about their product/service and process innovations. Most obviously, online customer data collection and analysis might be expected to support the design and execution of business innovations (see the case study in Box 3). At the same time, if a business reports high levels of innovative activity, this may indicate that it operates in a highly competitive market, which gives it no choice but to adopt management practices and technologies that increase its productivity.[48] The causality could also in fact go the other way too, if innovative companies have a stronger propensity to experiment with data, and change their business practices.

Although each of these three different (and not necessarily mutually exclusive) hypotheses mean different things in terms of the drivers and impacts of DDDM and Online Analytics

adoption, they all have one implication – where we see innovation, we should see DDDM and investment in Online Analytics.

Our findings confirm this: DDCs are significantly more likely to report that they launch products and services that are then imitated by their competitors, and to make substantial and disruptive changes to their business processes than their XDC counterparts (Figure 16). Disentangling this relationship – and in particular, the direction of causality – between innovation and DDDM is a key question for our research going forward.

Figure 16: Innovation and Data-Driven Decision Making



Box 3: Online as a complement to creativity in Albion London

Albion London is a digital advertising agency based in Shoreditch, London. It works with clients such as Telefónica, Skype and Wonga, and has a strong reputation for its creativity, innovation, and digital capabilities. Albion follows an agile approach to advertising development based on rapid iteration where the company 'lights a hundred fires', and 'pours petrol' on those that work, using social media data and analytics in real time to determine which ideas fall into different categories.

In a recent campaign for user–powered mobile network giffgaff, Albion used data from Facebook, as well as forums used by 'superusers' of the giffgaff community to test which ideas for the campaign had traction, and develop further those that did. Albion set hard metrics at every decision point of this development process in order to choose whether to progress an idea or not.

Glyn Britton, Albion's Managing Partner, points out that this is a very different approach from the traditional brand development process where *"an idea can gestate for months, shaped by subjective feedback from agency and client, and research which, no matter how rigorous, is always guesswork. After all of this expensive swirl, the Big Idea is launched with a Big Bang. When it works it can be spectacular, but too often it doesn't, and the campaign is met with indifference and shifting blame."*

The use of hard data neither deters creativity during the ad development process, nor leads to *'bland and formulaic work'.* Albion's goal for the giffgaff campaign was to help people who may be interested in joining the mobile network but lacked information about how to unlock their phone (a prerequisite for doing so). The agile, data–informed method followed by Albion produced innovative results that could be shown to work. They included the 'giffgaff Illegal Theatre', the 'Unlockapedia', the 'Unlock a phone, Unlock a chicken' campaign, and even a rap video starring 80s ventriloquist Keith Harris and his duck Orville. If anything, Albion's ability to use online data to determine which ideas worked and which didn't encouraged more experimentation during the campaign.

This case study illustrates how online data is transforming the innovation process in creative industries such as advertising, without taking the creative spark out of it. As Erik Brynjolfsson points out in a recent article about 'Big Data' in *Harvard Business Review*, data helps determine what the answer to a question is, but not what question to ask in the first place – that is the domain of business and innovation strategy, intuition, imagination and creativity.

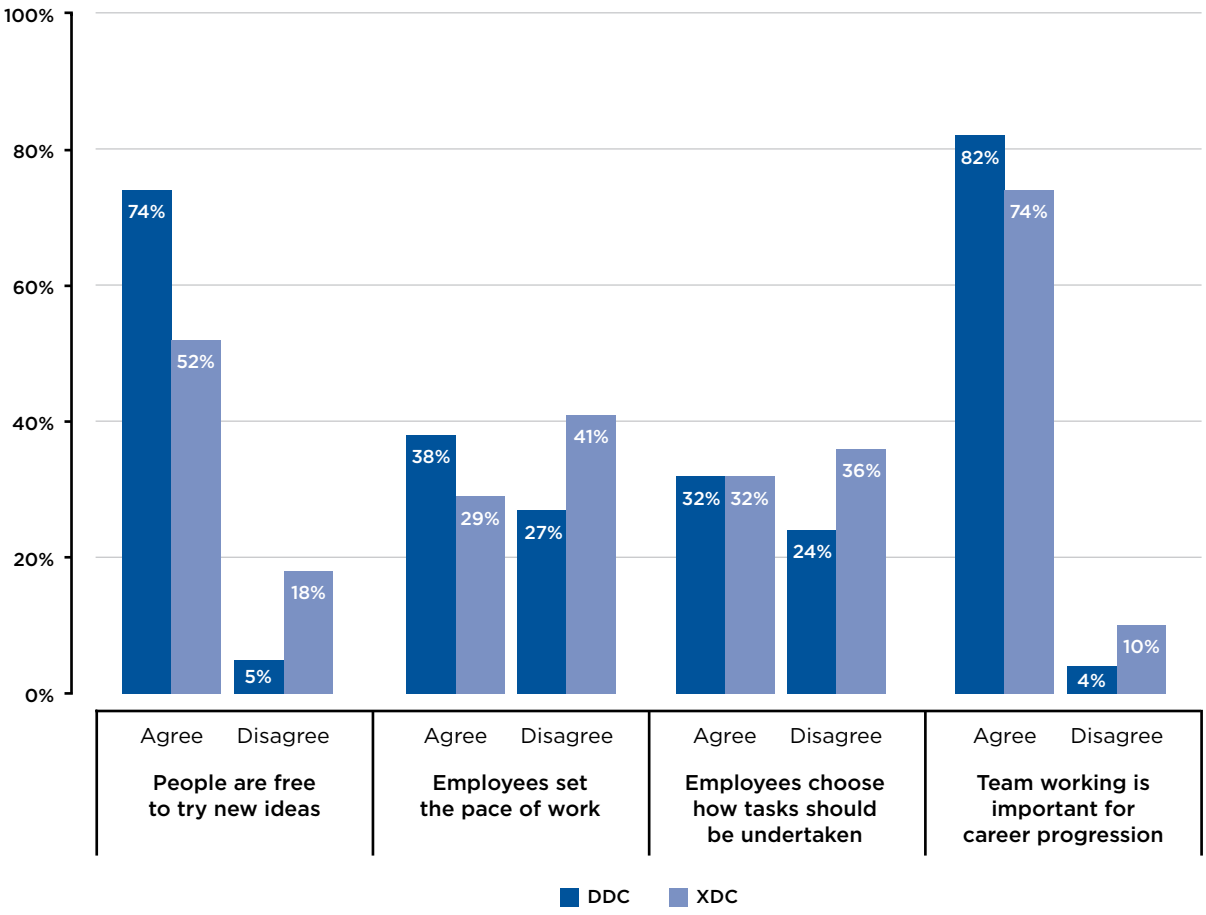
Sources: Glyn Britton and Sam Wander, Albion London. For more information see: http://www.albionlondon.com/blog/orville–and–agile–advertising/

They also have more decentralised organisational structures, and empowered workers

We also asked companies about the organisation of their workplace. Previous research has identified this as a critical factor for understanding the impacts of ICT investments on more general business performance.[49] It also relates to long–standing debates about whether ICT adoption acts as a force for employee empowerment, or on the contrary, creates highly monitored and hierarchically–organised workplaces.

Different digital technologies have been found to have different impacts on the organisation of the workplace (quantified using measures of organisational decentralisation and delegation). While Information Technologies such as Enterprise Resource Planning (ERP) and Computer Aided Design (CAD) and Manufacturing (CAM) increase decentralisation by helping workers access the information they need to resolve complex problems without resorting to their managers, 'communication' technologies (for example, digital networks connecting different sites) reduce the costs of transferring information 'up the command chain', thus moving decision making further up.[50] We adapted questions used in previous research in order to explore the nature of the connections between DDDM (and therefore Online Analytics) and organisational structure (Figure 17).

Figure 17: Decentralisation and Data-Driven Decision Making (percentage who agree or disagree with the statement)



On the whole, we see a positive connection between the adoption of DDDM (which is, as described throughout this report, correlates with the use of Online Analytics), and decentralisation in the workplace: respondents in DDCs are significantly more likely to agree than those in XDCs that employees in their companies are free to try new ideas and set the pace of their work. While the proportions of DDCs and XDCs saying that their companies allow employees to choose how tasks should be undertaken are similar, a significantly higher proportion of XDCs actively disagree with this statement. Likewise,

a significantly higher proportion of XDCs report that team working is not important for career progression in their organisations.

All this suggests that businesses perhaps need to take a hard look at their organisational and management structures if they are to fully reap the benefits of data. They need to ensure that data is available wherever it can be deployed more effectively, and that those employees who have the right data are empowered to act on it.[51]

4. MULTIVARIATE ANALYSIS

A limitation of previous studies exploring the connections between Analytics and business performance is that they have usually looked at the relationship between pairs of variables without controlling for other determinants of company performance. If we do not consider these factors explicitly in our analysis, we are at risk of identifying spurious relationships between our variables, which are in actual fact caused by other 'omitted' factors (for example, a company's business model, its sector, or how innovative it is).

In this section, we take a first step to explore this issue by looking at the significance and magnitude of the relationship between the main measure of performance impact in our survey – Question 16 – on how significant was the contribution of our respondent's ability to collect and analyse online customer data to their business performance – and measures of Data-Driven Decision Making (DDDM) and Online Analytics adoption. This also helps assess the robustness of some of the findings we presented before, by making use of all our responses (including information from businesses that identified themselves as neither data-driven nor experience-driven).

DDDM has a significant and strong association with the contribution that online customer data makes to business performance, even after controlling for other variables

Table 2 presents the outputs of our multivariate analysis. This analysis looks at the relation between DDDM and the perceived contribution of online customer data to business performance, after controlling for other variables.

DDDM is the variable we used to produce our DDC and XDC categories in Section 3. It ranges between 1 (decision making is exclusively based on experience and intuition) and 5 (decision making is exclusively based on data and analysis). Q16 (our dependent variable) ranges from 1 to 5, with 1 indicating insignificant business benefits from using online customer data and 5 indicating very significant benefits.[52]

- **Model 1** in Table 2 shows that the coefficient of DDDM on the impact of online data on business performance is positive and significant – this confirms the result that we already reported in Section III: there is a positive and strong relationship between a company's use of data for decision making, and the contribution of online customer data to business performance.

The other models introduce control variables in the analysis to see whether the relationship holds once we account for other factors.

- **Model 2** includes the proportion of revenues a respondent generates online (as we may expect online data to be a stronger contributor to performance for companies that are more commercially active online).

- **Model 3** adds further controls for self-reported levels of product and process innovation (which may capture market competition as well as some hard-to-observe attributes of our respondents, such as their willingness to take risks and try new ideas).

- **Model 4** includes eight sector dummies as control variables,[53] as well as the logarithm of employee size.

Also note that the coefficients we report are standardised (that is, comparable between variables irrespective of their original scales).

The results show that once we introduce our controls, the size of the Data-Driven Decision Making (DDDM) coefficient becomes smaller, but remains positive and statistically significant. This means that the connection between a company's reliance on data and analysis for decision making, and the contribution that online customer data makes to its performance cannot be wholly explained by the market where it operates, and other important company-level characteristics: a company's approach to data and analysis appear to matter in its own right.

Table 2: Relationships between DDDM and the contribution of Online Analytics to business performance (standardised coefficients)

	Model 1	Model 2	Model 3	Model 4
	Q16: How significant a contribution has your ability to collect and analyse online customer data made to your business performance?			
DDDM	0.329*** (7.39)	0.237*** (4.91)	0.208*** (4.23)	0.198*** (3.87)
Online sales as a percentage of revenue		0.369*** (7.81)	0.331*** (6.68)	0.326*** (6.5)
Product innovation			0.125** (2.62)	0.129** (2.71)
Process innovation			0.125* (2.52)	0.125* (2.46)
Number of employees (log)				0.041 –0.93
N	472	402	376	376
adj. R-sq	0.106	0.233	0.252	0.256
Sector controls	N	N	N	Y

Standardised beta coefficients; t statistics in parentheses
* p<0.05, p<0.01, *** p<0.001

What is the connection between specific Online Analytics practices and the perceived benefits from Online Analytics?

Throughout this research paper, we have focused on DDDM – that is, whether a company is data-driven or not – as an important indicator of its reliance on online customer data for decision making, and of the impact of these data on business performance.[54] Here, we look in further detail at the connection between the adoption of particular Online Analytics practices (in online customer data collection, analysis and use) and the perceived impacts of online customer data.

These results can be seen as providing evidence on the possible 'mechanisms' through which the impact of DDDM on performance is realised. Because they are more 'tangible' measures of behaviour, they also help us to test the robustness of our findings, bearing in mind the concerns about reporting biases in the Data–Driven Decision Making questions we raised in Section 3 (recall that many companies who said they were data–driven appeared to have adopted few Online Analytics practices).

Data reduction

Our questionnaire included three questions respectively covering the collection, analysis and use of online customer data by our respondents – the three key areas of Online Analytics adoption of our online customer data value chain. In each of these questions, we provided respondents with a set of options about what types of data they collected (five options), what techniques they analysed these data with (seven options), and in what areas of the business did they use the insights from these data to make decisions (nine options).

Including all of these options in our analysis would have been unwieldy, so we have used Principal Components Analysis, a data reduction technique, to create factors that summarise our respondent's Online Analytics responses in the areas of online customer data collection, analysis and use.[55]

It turns out that the comprehensiveness of our respondents' online customer data collection is best summarised by two factors (respectively capturing respondents' collection of e–commerce and online marketing data). A single factor summarises our respondents' use of online data analysis techniques, and the same is true for their use of online customer data for decision making.

This is the same as saying that, with the exception of online data collection – where we see a divergence between e–commerce companies that collect customer transaction data and other companies that focus on marketing and web activity data – we do not tend to see our respondents specialising in a subset of data analysis techniques, or using their online customer data to make decisions in some areas of the business at the expense of others. Instead, they tend to adopt a range of Online Analytics techniques, and use them to inform decision making in many parts of the business.

Use it or lose it

Table 3 presents the results of our multivariate analysis.[56] As before, our dependent variable is the perceived contribution of online customer data to business performance. Higher values of our four factors would represent 'more online marketing data collection', 'more transaction data collection', 'more data analysis', and 'more data use'.

We introduce the factors we have estimated with our Principal Components Analysis in sequence.

- **Model 1** includes the two factors we have extracted from responses to our question on online data collection (the 'collection of marketing data' factor, and the 'collection of transaction data' factor).
- **Model 2** includes our 'online customer data analysis' factor.
- **Model 3** includes the one on 'use of online customer data'.
- **Model 4** introduces all of the control variables we used previously.

Our findings clearly support the idea that collecting online customer data by itself is not enough to generate business benefits – and neither is analysing it. To put it simply, those businesses that score highly in their use of data for decision making (F4) also report higher impacts from online data. When we incorporate this factor into our analysis (**Model 3**), it goes as far as explaining almost 15 per cent of the variation of our performance measure (see the change in overall explanatory power – represented by the adjusted R^2 – between Models 2 and 3). The obvious interpretation is that, in order to fully benefit from their online customer data, businesses have to actually put it to work.

Table 3: Online Analytics practices and online data contribution to business performance (standardised coefficients)

	Model 1	Model 2	Model 3	Model 4
	Q16: How significant a contribution has your ability to collect and analyse online customer data made to your business performance?			
F1: Online customer marketing data collection factor	0.445*** (12.49)	0.304*** (7.89)	0.115** (2.88)	0.057 (1.24)
F2: Online customer e-commerce data collection factor	0.268*** (7.33)	0.182*** (5.14)	0.083** (2.61)	0.049* (1.41)
F3: Online customer data analysis factor		0.363*** (9.5)	0.197*** (5.15)	0.170*** (4.09)
F4: Online customer data use for decision making factor			0.486*** (11.3)	0.483*** (10.13)
N	483	483	483	382
adj. R-squared (R^2)	0.265	0.369	0.5	0.531
Sector, Innovation, Business Model and Size controls	N	N	N	Y

Standardised beta coefficients; t statistics in parentheses
* p<0.05, p<0.01, *** p<0.001

5. CONCLUSIONS AND NEXT STEPS

The data wave

This paper has illustrated both the opportunities and challenges that online data presents for UK businesses.

Clearly, the majority of UK businesses, including those that are already active in the Internet economy, are at an early stage of exploiting their data. This is evidenced by the predominance of companies driven by experience and intuition rather than data and analysis in our sample, and the relatively low levels of online data analysis and use even by companies that describe themselves as data-driven in their approach to business decisions. This is echoed in the low levels of adoption of Online Analytics practice outside of Web Analytics.

Where data is collected and used intelligently, it does however appear to generate substantial rewards, at least according to the businesses we have surveyed. It is not a coincidence therefore that many of these companies are investing in the resources and capabilities they need to exploit online customer data even more effectively.

Our analysis suggests that UK businesses may also have to undergo substantial (and what in some cases could be painful) transformations in their business processes and organisational structures in order to fully realise the benefits from online customer data. For example, managers may have to empower their employees to use online customer data – as suggested by the relationships we have uncovered between Data-Driven Decision Making and organisational decentralisation.

Our findings also suggest that a deeper engagement with Online Analytics can help UK companies develop good practices and capabilities – for example, in the use of rigorous experimental techniques – that can afterwards be applied in other parts of the business.

Policymakers for their part need to think hard about how to create a regulatory environment that strikes the right balance for consumers between potentially costly intrusions into their data privacy and the potential benefits that accrue, from the more rapid development of innovative products to more efficient pricing of products. In this regard it is notable that concerns about data privacy as well as security ranked highly as a barrier to greater use of customer data for the Data-Driven Companies in our sample. We hope that our research in this area will provide independent and rigorous evidence to inform debates in this controversial area.

Next Steps

UK companies may be deterred from adopting Online Analytics practices through a lack of knowledge about their impacts, about the complementary investments they need to make for them to work, and an uncertain policy environment. The preliminary findings from our survey lend some weight to these concerns. Our intention now is to subject them to more rigorous testing in the next stages of our research.

In the coming months, we will integrate our survey data with other indicators of financial performance collected from the FAME database, as well as data on our respondents' IT investments (which we are separately collecting with the help of Ipsos MORI), in order to carry out a more thorough econometric analysis of the causal relationship between Data-Driven Decision Making and company performance.

But, econometric techniques can only go so far in identifying causal relationships between variables such as online data analysis and use and business performance, given that both are likely to be influenced by common factors. For this reason, we are also exploring the possibility of running a controlled experiment to measure the impact of an Online Analytics 'intervention' on a randomly selected sample of UK companies. That way, we hope we will be able to establish robustly the direction of causality between Online Analytics adoption and business growth, the mechanisms through which such impacts are realised, and their magnitude.

APPENDIX

Table A1: Summary of reports on Big Data and Data Analytics.

Reference	Method	Headline findings	Online aspects
IBM/MIT Sloan (2010), 'Analytics: The New Path to Value.' http://goo.gl/SP702	Survey of almost 3,000 executives, and analysts working across more than 30 industries and 100 countries	• Top performing respondents use analytics five times more than lower performers. They are twice as likely to use analytics to inform future strategy, and day-to-day operation. • Respondents report that the most significant barriers to adopting analytics relate to management and culture, rather than technology.	Internet mentioned once, as an increasingly important source of data.
Economist Intelligence Unit (2011), 'Levelling the Playing Field: How Companies Use Data for Competitive Advantage.' http://goo.gl/janMw	Survey of 602 business executives around the world in a wide range of industries	• 70 per cent of respondents report that data is extremely valuable to competitive advantage. • The majority (two-thirds) of respondents produce information for the benefit of senior management. Poor organisation of data and poor processes create the biggest challenges to using data.	Online search and social media expected to become more important as sources of data in the coming years.
McKinsey Global Institute (2011), 'Big Data: The Next Frontier for innovation, competition, and productivity.' http://goo.gl/6mHsr	Use of secondary data to produce indices of Big Data, potential impacts by industry, labour demand etc.	• Big Data can generate big gains across the economy – for example, 60 per cent increase in margins for retailers, $300 Billions in the US Health Sector. • The researchers estimate that demand for talent with 'deep analytical skills' in the USA will outstrip supply by 140–190,000 positions in 2018.	Internet and Internet of Things presented as major 'big data' generators.
IBM/MIT Sloan (2011), 'Analytics: The widening divide. How companies are achieving competitive advantage through analytics.' http://goo.gl/h8CPJ	Use of insights gathered from more than 4,500 managers and executives	• Growth in the number of companies using data to create competitive advantage, with a widening gap between 'transformed', 'experienced' and 'aspirational' organisations. • Respondents are twice as likely (44 per cent versus 24 per cent) to rate organisational challenges to adopting analytics as 'extremely difficult to resolve' by comparison to technological ones.	Internet mentioned once in the document.
Economist Intelligence Unit/Cap Gemini (2012), 'The Deciding Factor: Big Data & Decision Making.' http://goo.gl/UCfwW	Survey of 607 executives working in 20 industries across the globe	• On average, respondents believe that the performance improvement to be experienced from Big Data in the next three years is 41 per cent. • Two-thirds say that their organisation is data-driven, but 55 per cent report data is not viewed strategically at senior levels of the organisation. • More than half of respondents report that organisational silos and analytics skills shortages hinder the use of data for decision making.	Includes a case study on General Electric (GE) and 'The Industrial Internet'.

Figure A1 : Online customer data collection across sector (percentage of firms in the sample reporting comprehensive data collection)

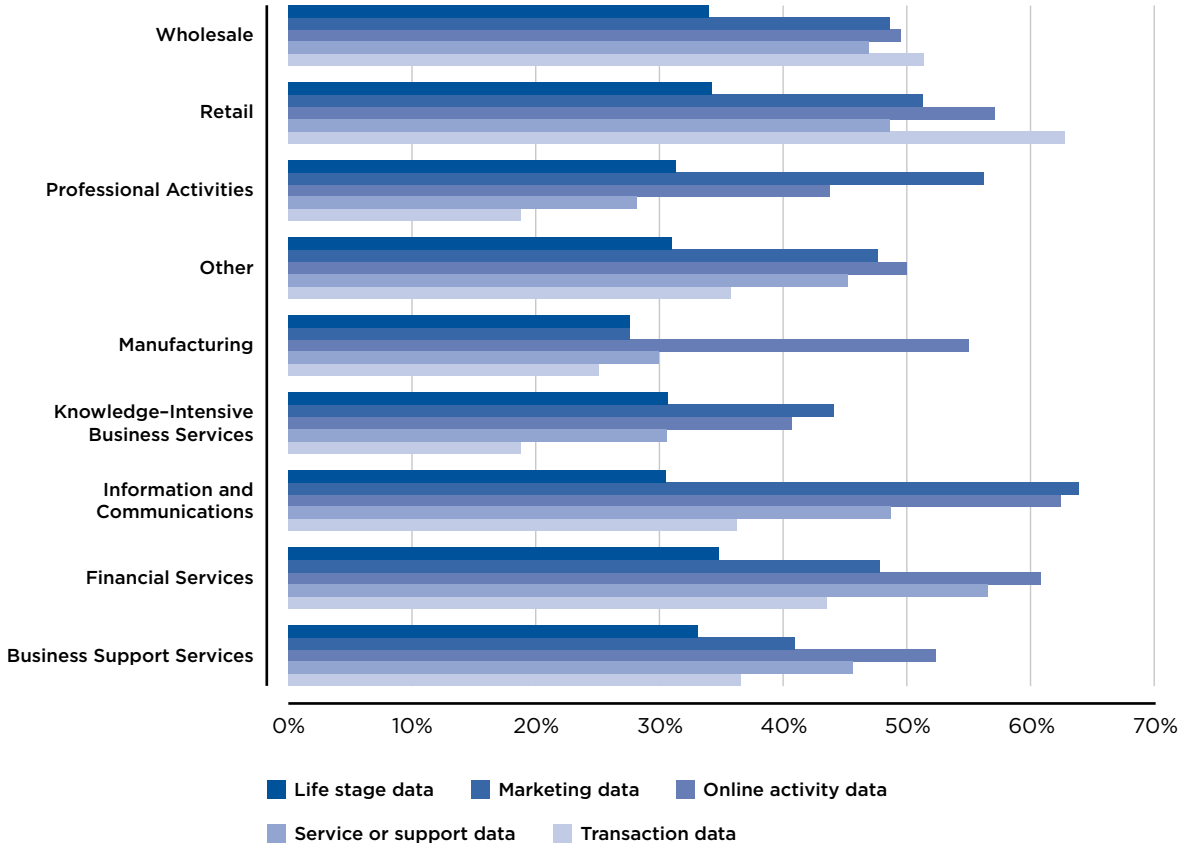
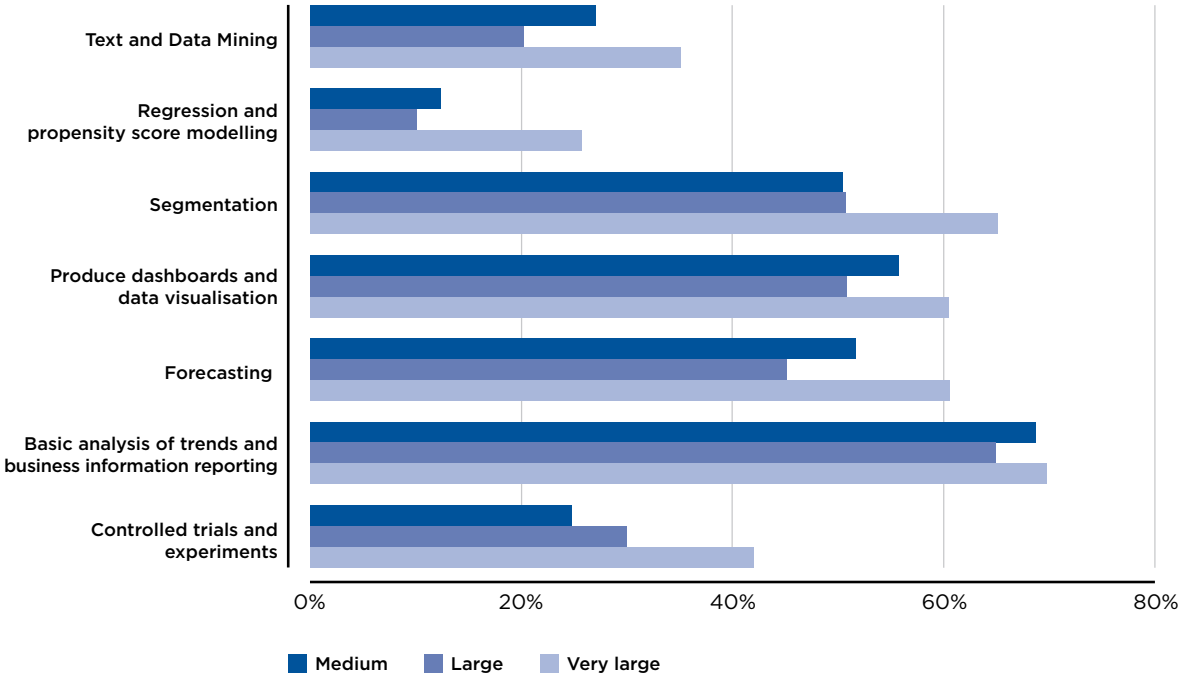


Figure A2 : Adoption of Analytics technique by size-band (percentage of firms in the sample who use the different techniques)



ENDNOTES

1. Locke, M. (2011) 'Next on TV: Data Driven Programming.' See: http://www.wired.co.uk/magazine/archive/2011/12/ideas-bank/next-on-tv
2. http://www.guardian.co.uk/media/2011/may/24/david-abraham-royal-television-society
3. http://blogs.telegraph.co.uk/technology/robjackson/100007363/the-digital-gold-rush-why-business-is-hungry-for-big-data/
4. Case studies of 'lead adopters' abound but it isn't clear how generalisable their findings are, given their focus on large and successful – often multinational – corporations. Previous quantitative surveys are based on self-selected samples, and only use self-reported indicators of general business performance. There is also always a niggling concern that employees best placed to answer questions on data and analysis are not the ones with the most accurate knowledge of their firm's general performance.
5. http://www-01.ibm.com/software/data/bigdata/
6. http://www.digitalanalyticsassociation.org/?page=aboutus,
7. Brynjolfsson, E., Hitt, L., and Kim, H., (2011) 'Strength in Numbers: How Does Data-Driven Decision making Affect Firm Performance?' See: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1819486
8. Web Index 2012, World Wide Web Foundation. See: http://thewebindex.org/2012/09/2012-Web-Index-Key-Findings.pdf
9. Ibid.
10. https://www.bcgperspectives.com/Images/Internet_Economy_G20_tcm80-100409.pdf
11. Eurostat. See: http://www.epp.eurostat.ec.europa.eu
12. http://www.ons.gov.uk/ons/rel/rsi/retail-sales/july-2012/stb-rsi-july-2012.html
13. Op. cit.
14. http://www.guardian.co.uk/media/2009/sep/30/internet-biggest-uk-advertising-sector
15. Manzi, J. (2012) 'Uncontrolled: The Surprising Payoff of Trial-and-Error for Business, Politics, and Society.' New York: Basic Books. Also Kohavi, R., Henne, R. and Sommerfield, D. (2007) 'Practical Guide to Controlled Experiments on the Web: Listen to Your Customers not to the HiPPO.' See: http://www.exp-platform.com/Documents/GuideControlledExperiments.pdf
16. Varian, H. (2010) ' Computer Mediated Transactions.' http://people.ischool.berkeley.edu/~hal/Papers/2010/cmt.pdf
17. By 'on-line' Nelson meant a broader class of live environments which enable testing, of which the Internet is an obvious example. Nelson, R. (2003) On the uneven evolution of human know-how. 'Research Policy.' 32(6), 909-922.
18. http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html?pagewanted=all
19. Bloom, N., Sadun, R. and Van Reenen, J. (2012) Americans Do IT Better: US Multinationals and the Productivity Miracle. 'American Economic Review.' 102(1), pages 167-201.
20. Bloom, N., Genakos, C., Sadun, R. and Van Reenen, J. (2011) 'Management Practices Across Firms and Countries.' CEP Discussion Paper No 1109. Boston MA: Harvard Business School.
21. http://www.qubitproducts.com/services/research/
22. Goldfarb, A. and Tucker, C. (2011) 'Privacy and Innovation.' NBER Working Paper 17124. Cambridge MA: NBER. http://www.nber.org/papers/w17124.pdf
23. In a future paper, when this year's financial accounts data are available for all of the companies in our sample, we will explore the impact of Online Analytics using company accounts-based indicators of business performance.
24. These qualitative interviews were led by Will Silverwood and Sally Taplin at Concise and Matthew Chatterton at Ipsos MORI.
25. The Standard Industrial Classification 2007 (SIC 2007) for establishing business establishments is divided into 21 Sections. See: http://www.companieshouse.gov.uk/infoAndGuide/sic/sic2007.shtml.
26. http://www.ons.gov.uk/ons/rel/rdit2/ict-activity-of-uk-businesses/2010/stb-e-commerce-and-ict-activity-2010.html
27. We have also used FAME to source other basic demographic information on firms, such as their size, sector, legal status and location.
28. In the event, over 41 per cent of firms with websites that had agreed to complete the survey were screened out at this stage, consistent with the finding from the ONS's E-Commerce Survey that the number of 'web-active' companies in the UK is still a minority, even in sectors with higher rates of e-commerce adoption.
29. The language of Data-Driven Decision Making has been adopted by the US Census Management and Organizational Practices Survey. See: http://bhs.econ.census.gov/bhs/mops/about.html.
30. All the differences between groups that we report in this section are statistically significant at least at the 5 per cent level.
31. Corresponding to the OECD definition of a medium-sized enterprise. See: http://stats.oecd.org/glossary/detail.asp?ID=3123
32. This is a higher resolution classification than that used by the OECD (which considers all companies above 250 employees as 'large'). The 250-999 and >1000 size-bands are also used in the 2010 ONS E-Commerce Survey.
33. We met the rule of thumb of 30 or more observations for all but financial services, so the results that we report for that sector should be taken with particular caution.
34. We have established this by calculating pairwise correlations between the full set of responses to this question and Online Analytics practices and impacts. We have also made use of the full set of responses in the multivariate analysis presented in Section 4.
35. This is particularly true of Financial Services companies in our sample. However, and as we said, caution is advised in interpreting these results given the small number of Financial Services companies for which we have data.
36. http://www.ons.gov.uk/ons/rel/rdit2/ict-activity-of-uk-businesses/2010/stb-e-commerce-and-ict-activity-2010.html
37. Ibid.
38. This market share for Google Analytics is in line with the adoption rates for Google Analytics reported in eConsultancy's Online Measurement and Strategy Reports. See: http://www.advanced-web-metrics.com/blog/2012/02/24/google-analytics-market-share/
39. Hamel, S. (2009) 'The Web Analytics Maturity Model: A strategic approach based on business maturity and critical success factors'. http://immeria.net/oamm/WAMM_ShortPaper_091017.pdf
40. This is in line with the views of the Internet as a 'site of learning' that we highlighted in Section (I), with the added benefit that some of the lessons it generates can be applied in 'offline' parts of the business.

41. Propensity Score Modelling uses measurable characteristics in observations (for example, individuals) in order to match them with others. This creates a control or counterfactual that helps to test the impact of an action (for example, a marketing campaign) on one of the matched observations. http://ec.europa.eu/regional_policy/sources/docgener/evaluation/evalsed/sourcebooks/method_techniques/counterfactual_impact_evaluation/propensity/propensity_details_en.htm
42. This reflects the theoretical ambiguity about this issue in the literature. On the one hand, insofar as Online Analytics is seen as a good management practice, we might expect it to be more widely embraced by larger companies which other things being equal, tend to be better managed (see Bloom, N., Genakos, C., Sadun, R. and Van Reenen, J. (2012) 'Management Practices Across Firms and Countries.' Academy of Management Perspectives.) Large companies may also enjoy economies of scale in their use of online data – for example, if it is possible to draw more robust inference from larger data sets. On the other hand, Online Analytics adoption may require the re-engineering of business processes, organisational structures and ICT infrastructures – and the organisational barriers to doing this could be greater in larger companies. This has led some researchers to speculate about the potential of digital data to be a driver of new entry and competition in online markets Goldfarb and Tucker, 2011. Op cit.
43. In this case, the weaker contribution of Online Analytics to XDCs' business performance could simply be a consequence of its lack of importance as a driver of competitive advantage in these companies' markets.
44. We have done this using the χ^2 test of goodness of fit to compare the observed distribution of responses in both groups across different levels of adoption of Online Analytics practice in terms of online data collection, analysis and use.
45. McAfee, A. and Brynjolfsson, E. (2012) Big Data: The Management Revolution. ' Harvard Business Review.' October 2012. See: http://hbr.org/2012/10/big-data-the-management-revolution/ar/5
46. MIT Sloan/IBM (2011) 'Analytics: The Widening Divide.' See: http://sloanreview.mit.edu/feature/achieving-competitive-advantage-through-analytics/
47. Greenstein, S. (2009) 'Glimmers and Signs of Innovative Health in the Commercial Internet.' Mimeo and Manzi, J. (2012) Op cit.
48. That is, a third factor may account for an observed relationship between innovation and DDDM/Online Analytics.
49. Bresnahan, T., Brynjolfsson, E. and Hitt, L. (1999) 'Information Technology, Workplace Organization and the Demand for Skilled Labor: Firm-Level Evidence.' NBER Working Paper 7136. Cambridge MA: NBER. See: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1819486
50. Bloom, N., Garicano, L., Sadun, R. and Van Reenen, J. (2011) 'The Distinct Effects of Information Technology and Communication Technology on Firm Organization.' CEP Paper No. 927. London: CEP.
51. This is also one of the key messages in EIU, 2011. 'Levelling the Playing Field: How Companies Use Data for Competitive Advantage.' London: EIU. http://goo.gl/janMw
52. We have estimated this model using ordinary least squares. We present the result of our estimations with robust standard errors.
53. Using Financial Services as the reference class.
54. This is justified by the close association between our respondents' use of DDDM and their embrace of Online Analytics that we have reported throughout the paper.
55. These factors can be intuitively understood as 'weighted averages'. Technically, we report the findings of our analysis using factors produced with a varimax rotation, and eigenvalue above 1. Our findings are robust to changes in the rotation method. In the results below, we report the findings using factors that replace missing values with the variable mean. They are in line with the results using factors produced excluding missing values.
56. Once again estimated using Ordinary Least Squares, and robust standard errors.

ACKNOWLEDGEMENTS

We are grateful for advice and suggestions on this research from Mario Callegaro, Adam Ghobarah, Hal Varian, Sarah Hunter, Timo Josten and Wiesner Vos at Google, John Van Reenen at the London School of Economics, Ed Steinmueller at the Science Policy Research Unit at University of Sussex, and E.J. Reedy (formerly at the Kauffman Foundation) and Albert Bravo–Biosca and Stian Westlake at Nesta. Their comments materially improved the survey instrument and this paper. Thanks also to Matthew Chatterton and Paola Ricoldo at Ipsos MORI who were instrumental in the design of the survey, and the fieldwork. Sally Taplin and Will Silverwood at Concise also contributed significantly to the project. Sam Wander and Glyn Britton at Albion London, Rob Jackson at Elisa Interactive and Ali White at TW White and Sons kindly helped with the case studies. Wherever possible we adopted survey questions from existing business surveys such as those reported in Brynjolfsson, Hitt and Kim (op. cit. Box 1) and Bresnahan, T., Brynjolfsson, E., and Hitt, L. (1999) 'Information Technology, Workplace Organization and the Demand for Skilled Labor: Firm–Level Evidence.' NBER Working Paper 7136 (http://www.nber.org/papers/w7136).

Nesta...

Nesta

1 Plough Place
London EC4A 1DE

research@nesta.org.uk
www.twitter.com/nesta_uk
www.facebook.com/nesta.uk

www.nesta.org.uk

November 2012