



# CONTENTS

Foreword	1
Summary: Findings	2
Summary: Recommendations	4
Introduction	7
1. Data skills: demand and supply	9
2. Teaching within subject areas	11
3. Sector initiatives to expand provision	16
4. Challenges	17
5. Opportunities	20
6. Implications for policy	22
Acknowledgments	23

# FOREWORD

Data presents a significant opportunity for the UK. Businesses are increasingly collecting and analysing data to enhance their productivity. Policy makers are considering how best to use data to transform public service delivery. Researchers are using data to advance scientific research and engineering. The potential value of data is far reaching. Across a range of sectors and industries, developing capacity in data analytics can advance knowledge, improve services and drive innovation.

Universities have played – and will continue to play – an important part in the data agenda. They allow for the opportunities presented by data to be realised and exploited, propelling technological advancements and spreading analytical skills. Universities also play an important part in overcoming some of the challenges associated with data, such as capability issues, skills shortages, and the ethical and legal issues posed by increased data usage.

As the use of data increases, a sufficient supply of data skills becomes critical to the UK's labour force. Therefore it is timely that we examine data skills provision at universities, considering what skills our graduates need to succeed in a range of sectors. This report presents research that I have led to determine the extent of data skills training across a range of subjects at undergraduate level in England.

Training in data analytics is found in most STEM and social science courses, yet the extent of provision varies by institution and degree subject. To meet the current and future needs of the UK economy we must do more to embed data skills as an essential component of many degrees. We must reflect the opportunities that data presents in our STEM degree courses and we must ensure that all social science graduates are able to conduct quantitative research in their field. To do so, we need to look across the sector at data skills provision, opportunities and challenges. I want this report to offer a point of guidance as we evaluate our course offerings and continue to train skilled and highly-employable graduates.



**Professor Sir Ian Diamond**

Principal and Vice-Chancellor, University of Aberdeen  
Chair, Universities UK Data Skills Steering Group



## SUMMARY: FINDINGS

The collection and analysis of quantitative data is becoming increasingly important across a range of sectors. As business and research interest in data expands, so too does the demand for workers able to analyse and interpret datasets. The potential to utilise data hinges on the supply of skilled individuals. However, research suggests that employers are struggling to find suitable candidates for data roles.<sup>1</sup>

In recognition of this challenge, the government asked Universities UK to 'review how data analytics skills are taught across different disciplines and assess whether more work is required to further embed these skills across disciplines.'

This report aims to engage with both the immediate shortage of data analysts, and the need for greater data literacy. As organisations become more data driven there is a need for all workers to be able to interpret data and to undertake basic analysis.

Taken together, this report and Nesta's report, *Skills of the datavores: talent and the data revolution*, set out a coherent picture of both the supply and the demand for data analysts and data-literate graduates. In a joint briefing statement in July 2015, Nesta and Universities UK will present findings, implications for policy makers, and recommendations.

### Findings

- 1 Although the skills shortage is widely reported the skills that entry-level data analysts should have is not clearly set out.** This has restricted positive action. In order to move forward these skills should be clearly set out, both by employers and by educators in their description of course content.
- 2 The data skills shortage is not simply characterised by a lack of recruits with the right technical skills, but rather by a lack of recruits with the right combination of skills.** The shortage of technical skills widely reported in the media is an over simplification of what is, in reality, a more complex issue. Employers report that there is a shortage of graduates with the right combination of skills. The combination of skills required includes a range of technical skills and domain knowledge, but also the ability to transform data outputs into something valuable to employers.
- 3 Usually, a combination of technical skills is achieved through multi-disciplinary teams, with every team member possessing deep skills in several areas and basic knowledge in others.** This shows that data skills needs cannot be boiled down to a simple list of skills that all undergraduates should acquire. Rather, there may be a number of core skills that should be shared by all members of a data team, and individual, specialist skills that may be developed in particular disciplines. The development of data teams emphasises the need for data analysts to possess strong teamwork and communication skills.
- 4 There is no consistent method for identifying the extent of data analysis teaching within undergraduate programmes.** A scheme to identify courses with significant data analysis components would provide valuable information to both prospective students and employers.
- 5 Many undergraduate degree programmes teach the basic technical skills needed to understand and analyse data.** Data can be gathered and analysed to enhance knowledge and understanding. This is largely reflected in undergraduate degree courses, where data analysis skills are taught across many programmes. This is also recognised by employers, who recruit data analysts from a range of subject areas, most commonly from those science, technology, engineering and mathematics (STEM) and social science courses where data analysis training is most prevalent and advanced.

<sup>1</sup> McKinsey Global Institute (2011) *Big data: The next frontier for innovation, completion and productivity* available at: [http://www.mckinsey.com/insights/business\\_technology/big\\_data\\_the\\_next\\_frontier\\_for\\_innovation](http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation); Deloitte (2013) *Market assessment of public sector information* available at: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/198905/bis-13-743-market-assessment-of-public-sector-information.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/198905/bis-13-743-market-assessment-of-public-sector-information.pdf), commissioned by BIS; and TDWI (2011) *Best practices report: Big data analytics* available at: <http://tdwi.org/research/2011/09/best-practices-report-q4-big-data-analytics.aspx?tc=page0>

## 6 The extent of training in data analytics provided across STEM and social science undergraduate courses varies by institution and subject.

More advanced data analysis is taught in some subjects than in others, there is also significant variation in the extent of data analysis teaching between similar courses at different institutions. In some cases, data analysis is isolated from the main curriculum. This risks treating data analysis as a specialist skill, rather than a valuable tool and research method with which to explore a range of research questions.

## 7 Data management is not always sufficiently embedded in courses teaching data analysis.

Data analysis requires a number of mathematical and computational skills. Stakeholders reported that students often do not possess basic skills in both. Many stakeholders were particularly concerned by a lack of knowledge and training in data management, which they felt was not sufficiently embedded in many courses using quantitative data.

## 8 There are a number of challenges facing academic departments wishing to expand their data analytic provision.

- Students entering higher education have varying knowledge in statistics and computer science. The challenge that this poses is twofold. First, course convenors are uncertain about the extent of knowledge of basic statistical and computer science skills that they can assume incoming students to have. Second, students often lack confidence with numbers and software.
- There is a shortage of academic staff confident in teaching data analytics.
- There are limitations as to what can be taught within a three-year curriculum.

## 9 There are a number of opportunities that put universities in a good position for expanding data analytics training in the future.

- Many universities are collaborating actively with employers. The nature of this collaboration varies. In some cases universities and employers offer work placements to undergraduates or provide data and research problems for postgraduates. In others, employers are actively involved in the development of curricula.

- The government has developed a coding curriculum for schools to support the early development of programming skills. This initiative could allow for more advanced teaching in statistics and data analysis to be offered in undergraduate degrees than is currently possible.
- The announcement of the new Core Maths course is welcomed by stakeholders. This course could provide post-16 mathematics education to students entering courses that traditionally do not require an A-level in mathematics.
- Statistics, including the use of software to explore real large datasets, will become a compulsory part of A-level mathematics in accordance with a new specification for the subject that comes into effect in 2017.

## 10 In recent years the higher education sector has led a number of initiatives to expand data analysis teaching.

- The research councils are developing initiatives to support data analysis teaching. The Nuffield Foundation, British Academy, the ESRC<sup>2</sup>, and the Higher Education Funding Council for England (HEFCE) have collectively set out a vision that all social science graduates should be quantitatively literate.<sup>3</sup>
- The EPSRC<sup>4</sup> has developed a number of doctoral centres focused on data analysis. Doctoral studies at these centres aims to unite graduates from mathematics, computer science and physical science degrees.
- Universities are also reacting to the demand for graduates with a broad set of skills through, for example, creating interdisciplinary courses, such as undergraduate data science degrees.
- The majority of universities have established mathematics and statistics support centres to provide targeted support for students enrolled on a wide range of programmes (not just STEM degrees). The centres are necessary because they help bridge the skills gap that exists between 16–18 education and higher education.

2 ESRC: Economic and Social Research Council

3 The British Academy's position statement is set out in *Society Counts* (2012), available at: [http://www.britac.ac.uk/policy/Society\\_Counts.cfm](http://www.britac.ac.uk/policy/Society_Counts.cfm) Further information on the ESRC's activities and position is set out on its website at: <http://www.esrc.ac.uk/research/skills-training-development/qmi/index.aspx>

4 EPSRC: Engineering and Physical Sciences Research Council

## SUMMARY: RECOMMENDATIONS

The recommendations arising from this work are set out in a joint policy briefing that draws on this research report and *Skills of the datavores* published by Nesta<sup>1</sup> These 'twin' reports consider both the supply and demand for analytical skills in the UK. The recommendations are set out in detail within this section.

### About our recommendations

The recommendations span the whole analytical talent pipeline, including schools, colleges, universities and the labour market and industry. The recommendations aim to remedy skills shortages in the short term, while ensuring a sustainable supply of excellent analytical talent in the longer term. Additionally, the recommendations encourage cross sector collaboration so that knowledge about how to create value from data and awareness of analytical skills shortages are not trapped in siloes, but are widely shared.

The data revolution has implications not only for experts with advanced analytical skills (such as data scientists), but for the entire workforce. We all need to become more data literate to operate successfully in increasingly 'data rich' environments. This is a key lesson from *Count Us In*, the British Academy's review of the UK landscape for quantitative skills.<sup>2</sup> Our recommendations reflect the diversity of analytical skills levels which are needed, and also suggest creating early 'touch points' between young people and data, acknowledging that in some cases these will mark the beginning of a lifelong analytical career, while in others it will involve raising awareness and confidence in using data, whatever the occupation.

Regarding the overall scope of our recommendations, our assessment of the current situation in the analytical talent pipeline is broadly optimistic. The data revolution has created analytical skills shortages, which are being particularly felt by those companies that are most innovative with their data, but there is a rapidly increasing awareness of those issues, and a willingness by educators and the government to act upon them. Much good work is already taking place in schools, universities and skills development agencies. The challenge, therefore, is to step up significantly the response so that it is commensurate with the scale of the opportunities. Our recommendations set out to adapt, repurpose and prioritise existing initiatives and programmes, rather than overhaul them.

We are calling for the following:

### Schools and colleges

#### 1 We need stronger teaching of mathematics and statistics in schools and colleges

The pipeline of data talent starts in schools, where we need to ensure that the teaching of analytical skills is embedded across curricula, and that, consistent with the Government's ambition, more young people study mathematics and statistics after the age of 16. Regarding the mathematics curriculum, **we endorse the policy agenda set out by the Royal Statistical Society (RSS) in its Data Manifesto<sup>3</sup> where it is suggested that we need to prepare for the data economy by 'skilling up' the nation in maths and statistics.** As the RSS recommends, basic data handling and quantitative skills should be an integral part of the taught curriculum across most A-level subjects. The new Computing curriculum in English schools is a significant development, but it needs to be supported by sustained investment in teacher training.<sup>4</sup>

1 NESTA (2015) *Skills of the datavores: talent and the data revolution*

2 The British Academy. (2015) *Count us in: Quantitative skills for a new generation*. Available at: <http://www.britac.ac.uk/news/news.cfm/newsid/1285>

3 Royal Statistical Society (2014) *The Data Manifesto*. Available at: [http://www.rss.org.uk/RSS/Influencing\\_Change/Data\\_manifesto/RSS/Influencing\\_Change/Data\\_democracy\\_sub/Data\\_manifesto.aspx?hkey=5dd70207-82e7-4166-93fd-bcf9a2a1e496](http://www.rss.org.uk/RSS/Influencing_Change/Data_manifesto/RSS/Influencing_Change/Data_democracy_sub/Data_manifesto.aspx?hkey=5dd70207-82e7-4166-93fd-bcf9a2a1e496)

4 NESTA (2011) *Next Gen: Transforming the UK into the worlds' leading talent hub for the video games and visual effects industries*. A review by Ian Livingstone and Alex Hope.

Initiatives like the Department for Education's (DfE) extension of the Further Maths Support Programme to boost mathematics participation and provision by providing advice, guidance and teacher training, and the extension of funding for the Network of Excellence programme in Computing teaching until March 2016 are also important. The same is true for the development of Core Maths to ensure a greater uptake of mathematics post-16 beyond the 20 per cent that currently study it after their GCSE.

## 2 More and better information about analytical career prospects and role models in schools and colleges

We need to improve access to information in schools and colleges about the career opportunities for data analysts, and identify and promote appropriate role models. Otherwise, young people may be disinclined to pursue a career in data analysis, despite the rich career prospects.

The Tech Partnership provides a wealth of information about digital careers. In particular, its TechFuture initiatives (including TechFuture Careers, TechFuture Girls and TechFuture Classroom) bring together practical lesson guides, group activities and industry experience to educate students about the types of tech jobs available and the paths they need to follow to get them. Data analytics occupations could have a stronger presence in these initiatives, in particular focusing on the range of sectors and organisations utilising data analytics.

**We call on the Tech Partnership to further highlight data analytics as a great career in the TechFutures Programme, and build up resources to support its teaching.** We also recommend convening a workshop including the recently established careers and enterprise company for schools, employers and local authorities (including teaching and management staff) to identify the best ways to embed data skills in schools across all subject areas, and make the link to careers in data analytics in a way that captures the imagination of young people and their parents or carers.

We also need to create more data analyst role models in schools, using networks of practitioners such as STEMNET (especially through the Tech Partnership's TechFuture Ambassadors network). There should be a drive to attract analysts who are deploying STEM skills to analyse data in commercial contexts, and in sectors such as retail and creative media, as well as traditional science and tech industries. Finding a range of role models will also contribute to diversity, and help reduce the gender gap in the profession.

## 3 Embed data analysis in other subjects

There is much scope to embed data analysis in school subjects other than mathematics and statistics. For example, the Urban Data School, run by the Open University in collaboration with Milton Keynes schools, develops data skills in young people by enabling them to use open and big datasets, such as energy, transport and satellite data, to analyse local issues. The benefits to students include an increased understanding of the use of this data and its impact, and data analytic and visualisation skills.<sup>5</sup>

The Urban Data School is currently developing pedagogical resources, an online platform for sharing data and good teaching practice, and a network of teachers engaged in data analytics and visualisation.

**The Government should keep a watching brief on innovative initiatives like the Urban Data School, and consider potential interventions to support their expansion where they are shown to have a beneficial impact. Local government should make these initiatives more attractive by opening up local datasets for students to work with.** Giving students the opportunity to use data to understand – and even make a difference – in the environment around them is after all an excellent way of engaging them with data. It can also have significant benefits to local authorities as they rethink services aimed at young people, and, through 'reverse mentoring' can even help build capacity within the local government workforce.

## 4 Support the development of extracurricular data activities

Ther Extracurricular and afterschool activities can help bring data and analysis to life for young people, and highlight the extent to which data now plays a central role in a wide range of industries, some of which – retail and creative media, for example – have not traditionally been seen as recruiters of analytical talent.

To make more of this happen, **the Skills Funding Agency should fund an experimental Data Summer Schools pilot in collaboration with industry, schools, colleges and universities to give more students the opportunity to engage with data, and develop their hands-on data skills. This would be aimed at students who require additional preparation for university study, or employment (including through Apprenticeships).**

<sup>5</sup> See: <http://urbandataschool.org/> Wolff, A. and Kortuem, G. (2015) 'Urban Data in the primary classroom: bringing data literacy to the UK curriculum.' Data Literacy Workshop at ACM Web Science Conference 2015, Oxford. Wolff, A., Kortuem, G., Cavero, J., Gooch, D. and Mir, U. (2015) 'Removing Barriers for Citizen Participation to Urban innovation.' Digital Cities 9 Workshop, 27 June 2015, University of Limerick.

For students who do not take a specialist data analytics course at university, or for those pursuing an alternative vocational route, the Data Summer School will provide guidance on embedding these skills in other disciplines and across industries, developing domain knowledge, and team working.

This scheme should be closely aligned with the Trailblazer Apprenticeship standard for Data analysts and, as such, provide the skills and guidance that students need to enter the labour market, pursue further study, or embed data analytics skills in other disciplines.<sup>6</sup>

### Universities and vocational education

#### 1 Increase the visibility of strong data analytics courses

The technical skills required to analyse data can be developed in a number of different degree programmes. Research undertaken by Universities UK suggests that this is not always made clear, both to prospective applicants and employers. More needs to be done to verify the extent and professional relevance of data analytics skills in university courses in the UK. **We recommend a ‘kite marking’ scheme, developed by the Tech Partnership and the Royal Statistical Society to identify relevant courses to prospective applicants and employers, based on input from the breadth of industries being transformed by data.** Importantly, this should consider vocational courses as well as undergraduate ones, and draw on relevant information, such as the latest IT professional standards.

#### 2 Embed quantitative analysis across disciplines

Many of the companies we surveyed in our research said they were concerned about insufficient domain (industry) knowledge in the analytical talent pool. An implication of this is that, in addition to setting up specialist data science courses focused on technical skills, universities should build on existing programmes and find new ways to embed analytical, computing and visualisation skills in them. We find that the higher education sector has led a number of initiatives to meet this challenge.

The Q-step initiative in particular is addressing this issue for the social sciences. Funded by the Nuffield Foundation, Economic and Social Research Council (ESRC) and the Higher Education Funding Council for England (HEFCE), Q-step aims to address the shortage of quantitatively-skilled social science graduates through 15 Q-step centres. The Q-step

centres support the development and delivery of specialist undergraduate programmes, including new courses, work placements and pathways to postgraduate study. We suggest that Q-step should be considered as a model for other funders to support, and further, **that Research Councils UK (RCUK) explore additional funding to support methodological and skills development in analytics so that the Q-step model can be extended to other disciplines.**

Universities themselves need strong data analytics practitioners, both to support research across disciplines and to educate tomorrow’s graduates. We find that the current research funding model has led to the concentration of disciplines in particular universities. **We recommend that RCUK maps data analytics capabilities across universities and assess whether more needs to be done to address potential skills shortages within universities.**

#### 3 Boost the business and soft skills of graduates from data analytics courses

Data analysts apply theoretical knowledge and hands-on technical skills in commercial environments, often as part of multidisciplinary teams. To speed up and smooth the transition of graduates from data courses into the commercial and public sectors, **we propose that a graduate data skills accelerator programme should be convened by university/industry partnerships across the UK, tapping into the experience and expertise of university–industry brokers like Pivigo and The ASI.**

There are opportunities to develop additional pathways to higher level skills in data analysis, such as degree apprenticeships, which combine academic education with work in industry. There are currently 13 such degree apprenticeships, including one in ‘Digital & Technology Solutions’. Designed by employers of tech talent in all sectors and supported by the Tech Partnership, this degree apprenticeship includes a common core of technology skills with options relevant to six different occupational areas, including data analyst.

**Our recommendation is that relevant stakeholders actively promote the data analyst degree apprenticeship option, encouraging employers to offer it and young people to apply for it.**

<sup>6</sup> An industry defined standard for entry requirements, technical competencies, and technical knowledge needed for a Data analytics apprenticeship role, for more information see: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/384258/DIGITAL\\_INDUSTRIES\\_-\\_Data\\_Analyst\\_standard\\_-\\_Final\\_081214.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/384258/DIGITAL_INDUSTRIES_-_Data_Analyst_standard_-_Final_081214.pdf)

Our research has shown the importance of team diversity and multidisciplinary collaboration in analytical work; however, this is not so easy to simulate within traditional undergraduate programmes, delivered by a single academic department. The challenge is similar to that faced in courses feeding talent to other innovative sectors, like videogames.<sup>7</sup> In order to encourage more multidisciplinary collaboration in that industry, Abertay University put in place its Dare to Be Digital Competition, where teams of university students from different disciplines (coding, arts, and design) compete for the opportunity to win a BAFTA prize.<sup>8</sup> We think that the same model could be used in data analytics, encouraging students with different analytical skillsets (e.g. programming, statistics, domain knowledge and data visualisation) to build an innovative product, or address a challenge using data. **Universities with a strong presence in the analytics space should consider organising this prize as a pilot, possibly in collaboration with the Alan Turing Institute and a high visibility big data conference like O'Reilly Strata.**

#### 4 Increase the supply of high-end analytical talent

Demand for highly-educated data analysts is acute, both in research and development, and in business more generally. This demand is two-fold: first, for practitioners capable of engaging with the current state of the art; second, for practitioners capable of developing new techniques and methods. In addition, universities need highly educated staff to lecture future generations of analytical talent.

Centres for Doctoral Training (CDTs) bring together diverse areas of academic expertise to train engineers and scientists with the skills, knowledge and confidence to tackle today's evolving issues. The Engineering and Physical Sciences Research Council (EPSRC) has recognised the challenge that data analytics brings through the funding of dedicated CDTs that train analysts with cutting-edge skills. The further commissioning of a CDT by the ESRC in New Forms of Data is also a positive and major step forward. **We support the approach of the research councils to this challenge, and suggest that the government devise a funding model for the expansion of such initiatives where they prove to be successful in meeting current and future needs.**

#### 5 Foster interdisciplinary research and skills development programmes

Many breakthroughs in the development of analytical methods and tools have happened at the intersection between different disciplines. An implication is that we need to support interdisciplinary, innovative research projects involving advanced data analytics, statistics and quantitative skills, and that calls for cross-research council collaboration and funding.

**Our recommendation is for a top slice of the RCUK budget to establish a strategic fund through which interdisciplinary research is funded. RCUK could itself take a strategic and convening role in this space.**

Skills development programmes are another area where a more 'joined up' approach would serve researchers better than initiatives conducted in silos. RCUK, working with individual research councils, has already taken steps to realise this. We note the work that RCUK has undertaken to facilitate a policy network on data and its applications, as well as networking opportunities for researchers from different disciplines to share best practice, and encourage the expansion of this activity.

#### Labour market and industry

##### 1 Create a cross-cutting taskforce around data analytics

There are currently many agencies in the UK exercising leadership to address the skills shortages arising in industry from the data revolution. However, no single body has all the answers to what are system-wide challenges. Collaboration is needed to address the national challenges identified in our research.

**We call on relevant stakeholders, including the Tech Partnership, the Royal Statistical Society, the UK Commission for Employment and Skills, the E-Infrastructure Leadership Council, the Digital Economy Council, techUK, the ODI, HEFCE and the research and sector skills councils to set up a cross-cutting taskforce around data analytics to identify good practices for education and skills provision, and spur collaboration across industry.**

7 NESTA (2011) 'Next Gen: Transforming the UK into the world's leading talent hub for the video games and visual effects industries. A Review by Ian Livingstone and Alex Hope.' London: NESTA.

8 See: <http://www.daretobedigital.com/>

### 2 Actively convene industry analytics networks

Our research has demonstrated the potential value of industry crossover in data applications and talent flows, some of which is being realised at the grassroots level: at data analytics events, meetups and hackathons.

**We believe that UK universities can play a stronger role in convening such events, potentially making space freely available to their organisers.** Not only would this support skills upgrades in the local analytical community, but it would also strengthen local analytical networks, including those between universities and analysts in industry and government, and those between academic researchers doing analytical work in different university departments. By working with local public services and communities, such data partnerships could help support public service transformation and build analytical capacity within civil society.

These grassroots activities could be usefully complemented by a higher visibility network following the example of an organisation like the US National Consortium for Data Science, or Scotland's Data Lab, a £11.3 million initiative supported by the Scottish Funding Council, Highlands and Islands Enterprise and Scottish Enterprise, which 'enables new collaborations between industry, public sector and universities driven by common interests in the exploitation of data science, provides resources and funding to kick start projects, deliver skills and training, and helps to develop the local ecosystem by building a cohesive data science community.'<sup>9</sup>

**We believe that a data science network along these lines should be developed with involvement from existing communities of analytical practice, the Alan Turing Institute, the Data Lab, the Catapults, and major Data Science institutes at universities like Imperial College, UCL, Manchester and Warwick.**

### 3 Support innovative interventions enabling local authorities to boost local analytical skills

Local authorities and Local Enterprise Partnerships (LEPs) can play a role in upgrading the analytical capabilities of the economies and digital clusters around them, not least through the release of open data. Two examples of this include the Silicon Abbey initiative in St. Albans (a social enterprise partnering with the University of Hertfordshire, Oaklands College and the local authority, aimed at building digital capacity in St. Albans through opening up data sets, training, internships and placements, mentoring and hack days), and the Connecting Data Programme (where Bath and North East Somerset Council, Bath University and Bath and North East Somerset CCG (NHS) are, through collaborations, hacks and multidisciplinary projects, working with communities, academics and public service organisations to bring analytical capabilities to a wider audience and increase local analytical capacity.)

**We call on the Department for Communities and Local Government (DCLG), through its Public Service Transformation Network, to convene a workshop to bring together local government analysts and others with an interest in this work to discuss and promulgate good practices. DCLG should consider financially supporting the wider roll-out of promising initiatives** (for example, the Connecting Data project was initially supported by DCLG's Transformation Challenge Award).

These interventions would build on the Cabinet Office's 2015 Open Data Champions initiative, which highlighted the role of 16 Local and Regional Councils as data pioneers that are leading in opening up their data to create opportunities for innovation, economic and social growth and better public services.

<sup>9</sup> See: <http://data2discovery.org/>

#### 4. Raise awareness of the value of data for business

A third of the companies we surveyed in Nesta's research – so-called 'dataphobes' – displayed low levels of data activity, in spite of its benefits for innovation and productivity. While we feel that other skills interventions might be premature for these organisations, we need to look for ways to encourage them to explore the opportunities of data, and build up their analytical capabilities.

**We call for a targeted campaign in the Government's Great Business website to raise awareness about the value of data for business.<sup>10</sup> The website should also provide information about ways in which 'dataphobe' businesses can start building their analytical capabilities. Local Growth Hubs that act as a meeting point for business and support and advice providers can also help to raise business awareness of data as a driver of growth, and provide leadership and management training to encourage its use.**

In doing this, it's worth paying attention to the experience of an organisation like techUK which has been leading a series of business-focused seminars exploring the value of big data analytics in key sectors, and DataKind, which has already developed effective models to boost impactful data use in charities.

#### 5. Deliver innovative solutions for data analytics training

Nesta's research suggests that while many companies are up-skilling their analytics workforce intensively, there remain serious gaps to be addressed. One particular challenge is to combine technical, analytical, teamwork and communication skills, and industry knowledge and business nous in one individual. **The UK Commission for Employment and Skills could usefully organise a competition as part of its Futures Programme with the goal of developing innovative models to address the challenge of upskilling analytical workers in the UK.**

As our research shows, data analysts are harnessing innovative learning models based on peer-to-peer learning, Massive Open Online Courses (MOOCs) and online communities to keep their skills fresh in the face of a rapidly changing technology landscape. The recently founded European Data Science Academy (an European Commission funded project with involvement from the Open University, University of Southampton and Open Data Institute) is also seeking to do this. We think that many of these tools and platforms could play a role in the innovative solution delivered by the competition we are calling for. The Futures Programme competition could also challenge businesses to consider the ways that they organise their work and designs job roles to make best use of multi-skilled teams. It is important to develop the right skills but also to ensure that the workplace is organised in such a way as to effectively utilise and capitalise on the skills developed.

The Tech Partnership's Training Fund, which co-funds employee tech training is also helping to upgrade the data analytics skills of the UK workforce, particularly among SMEs. This fund has, in the year to March 2015, supported training for 1,400 people in 200 businesses. Through the fund, employers can focus on those skills areas that are a priority for them (including big data), using those training options that are most relevant. **More should be done to improve the visibility of the fund as a mechanism to upgrade analytics skills (particularly among non IT/tech firms), and the government should consider options to extend it after the funding runs out in 2016.**

<sup>10</sup> See: <http://www.greatbusiness.gov.uk/>

## Conclusion

The all-pervasive reach of the data revolution explains why a variety of disciplines and skills need to come together if the UK is to fully benefit from it. As a system challenge, it can only be addressed with a systemic programme of actions like the one we set out in this policy briefing. We believe that if our recommendations are acted upon as a group, they will make the UK a stronger analytic nation, best placed to assume a leading role in the data economy.

### Summary of top line recommendations

Area	Top line recommendations
<b>Schools and Colleges</b>	<ol style="list-style-type: none"> <li>1. Stronger teaching of mathematics and statistics in schools and colleges</li> <li>2. More and better information about analytical career prospects and role models in schools and colleges</li> <li>3. Embed data analytics in other subjects</li> <li>4. Support development of extracurricular data activities.</li> </ol>
<b>Universities and vocational education</b>	<ol style="list-style-type: none"> <li>1. Increase the visibility of strong data analytics courses</li> <li>2. Embed quantitative analysis across disciplines</li> <li>3. Boost the business and soft skills of graduates from data analytics courses</li> <li>4. Increase the supply of high-end analytical talent</li> <li>5. Foster interdisciplinary research and skills development programmes</li> </ol>
<b>Labour market and industry</b>	<ol style="list-style-type: none"> <li>1. Create a cross-cutting taskforce around data analytics</li> <li>2. Actively convene industry and community analytics networks</li> <li>3. Support innovative interventions enabling local authorities to boost local analytical skills</li> <li>4. Raise awareness of the value of data for business and local public services</li> <li>5. Develop innovative solutions for data analytics training</li> </ol>

# INTRODUCTION

The Department for Business, Innovation and Skills launched the Data Capability Strategy in October 2013 with the publication of the report *Seizing the Data Opportunity*.<sup>11</sup> The strategy sets out the government's commitment to the sustained development of the UK's data agenda through strategic investment in capital and skills.

Several reports have identified the skills shortage as a significant barrier to businesses wishing to expand their use of data. One of the key findings of Nesta's *Model Workers*<sup>12</sup> is that four in five of the companies interviewed were struggling to find the talent they needed, and skills shortages appear to be more acute outside of London. Indeed, a wide range of reports cite the skills shortage as one of the key barriers to exploit the opportunities presented by data analytics, including those published by McKinsey, Deloitte and the research consultancy TDWI.<sup>13</sup> Against this, demand is only expected to increase.

In most cases, the shortage of skills is seen as relating specifically to a shortage of employees able to extract value from datasets.<sup>14</sup> The term 'data analysts' is used throughout this report to describe those who develop or manage data analysis software, analyse data and undertake research using quantitative data.

A shortage of data analysts should, however, not be the only concern. The demand for data-literate graduates extends beyond the specific demand for graduates with advanced data analytics skills. As organisations become more data driven there is a need for all workers to be able to interpret data and to undertake basic analysis. This will be essential in order to understand decisions that are made on the basis of data.

## Research and scope

This report examines data analysis skills teaching at English universities at undergraduate level. A literature review followed by a consultation exercise provided insight into the nature of the skills shortage and employer needs.

The findings outlined in this report are based on a survey of universities' heads of academic departments and professors of statistics, as well as a series of stakeholder interviews. Seventy-seven responses to the survey were received and a total of 35 stakeholders were interviewed, including academic staff, research councils, academies, and employers. Findings from the survey and interviews were discussed and reviewed with experts and recommendations developed with appropriate bodies and policymakers.

The focus of this report is on the technical skills required by graduates. These skills might include, for example, research design, data collection, data management, data analysis, modelling, interpretation, and the presentation of research findings. The precise activities that a data worker will engage in will vary within and between organisations and job roles: not all data workers will be required to demonstrate all of these skills.

---

11 HM Government (2013) *Seizing the data opportunity: A strategy for UK data capability* available at: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/254136/bis-13-1250-strategy-for-uk-data-capability-v4.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/254136/bis-13-1250-strategy-for-uk-data-capability-v4.pdf)

12 Nesta (2014) *Model Workers: How leading companies are recruiting and managing data talent* available at: [http://www.nesta.org.uk/sites/default/files/model\\_workers\\_web\\_2.pdf](http://www.nesta.org.uk/sites/default/files/model_workers_web_2.pdf)

13 McKinsey Global Institute (2011) *Big data: The next frontier for innovation, completion and productivity* available at: [http://www.mckinsey.com/insights/business\\_technology/big\\_data\\_the\\_next\\_frontier\\_for\\_innovation](http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation); Deloitte (2013) *Market assessment of public sector information* available at: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/198905/bis-13-743-market-assessment-of-public-sector-information.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/198905/bis-13-743-market-assessment-of-public-sector-information.pdf), commissioned by BIS; and TDWI (2011) *Best practices report: Big data analytics* available at: <http://tdwi.org/research/2011/09/best-practices-report-q4-big-data-analytics.aspx?tc=page0>

14 Nesta (2014) *Model Workers: How leading companies are recruiting and managing data talent* available at: [http://www.nesta.org.uk/sites/default/files/model\\_workers\\_web\\_2.pdf](http://www.nesta.org.uk/sites/default/files/model_workers_web_2.pdf)

Graduates need a range of competencies beyond technical skills, particularly communication and teamwork skills. However, there are a number of measurement issues that make it hard to determine the extent to which these skills are developed across subjects and institutions. Our examination of data analytics teaching by subject areas focuses primarily on technical skills and domain knowledge. However, we discussed the development of soft skills with sector stakeholders and this is recognised by academics as an important aspect of curriculum design across subject areas.

For some data roles, employers seek graduates who have experience in analysing large and complex datasets, many of whom have postgraduate qualifications. Provision at postgraduate level is not reviewed in detail in this report. However, the transition of graduates into postgraduate research positions was examined to assess whether or not undergraduates were sufficiently prepared.

---

# 1. DATA SKILLS: DEMAND AND SUPPLY

The technical complexity of working with what is often termed 'big data' is well recognised.<sup>15</sup> New technologies and methodologies allow for the collection of increasingly large datasets. This has led to new opportunities across all disciplines from astrophysics to zoology, genome sequencing to business analytics. These large datasets are often unstructured, dynamic or both. Data may be incomplete or corrupted. Each of these characteristics represents a challenge to analysts.

Some data opportunities require one to determine which questions to ask using a large dataset that has already been collected. In these cases, the scientific method is reversed. Rather than establishing a research question and collecting appropriate data to answer it, one has to establish what the data captures and how it might be useful.

The skills shortage could be attributed to the advanced level of technical skill needed to extract value from complex datasets. However, research undertaken by Nesta suggests that the skills shortage is more complicated. In fact, Nesta finds that employers struggle to find recruits with the right combination of skills. A recruit with the right combination of skills would possess a range of technical skills needed to process data and the domain knowledge required to design data projects and interpret outputs. Strong communication skills and business know-how are also important so that data outputs can be turned into outcomes for the employer.

## Data analysts

The breadth of skills that a data worker must possess means that the solution to the shortage is not simply to increase the number of people with appropriate technical skills, but also to ensure that they have the domain knowledge and soft skills to convert data outputs into business outcomes.

For most data roles, employers recruit from a range of academic disciplines, assessing graduates for their technical abilities on an individual basis. Most commonly they recruit from the computer sciences, mathematical sciences, physical sciences and social sciences. It is often the case in industry that multidisciplinary teams are essential to unlocking maximum value from data, because they have the breadth of knowledge to generate most insight. In a team, a data analyst will have particularly advanced skills in a certain area of analysis. However, there is a need to ensure that all members of a data team have a common, core skillset so that teamwork can be effective.

In Nesta's *Model Workers* report, the so-called data scientist is likened to a unicorn. This is because recruits who possess all possible data skills at an advanced level simply cannot be found. Nesta finds that companies struggle to recruit data talent because recruiting experienced staff is costly, and inexperienced staff require significant training before they can contribute effectively to the company.<sup>16</sup>

This suggests that the core skills a junior data worker might reasonably be expected to possess must be distinguished from the ideal attributes of an established data worker or the skills required of a data team. There is, however, a degree of uncertainty among stakeholders as to the skills that should be acquired for entry-level data analysis jobs. This can partially be explained by the complexity of data skills needs and the various positions that data analysts may adopt within a data team.

<sup>15</sup> See for example: Fan J, Han F, Liu H (2014) 'Challenges of Big Data Analysis' *National Science Review*. 00: 1–22. There have also been a number of conferences held that focus on the challenges of big data. For example, the Institute of Mathematics and its Applications held a conference on the Mathematical Challenges of Big Data.

<sup>16</sup> Nesta (2014) *Model Workers: How leading companies are recruiting and managing data talent* available at: [http://www.nesta.org.uk/sites/default/files/model\\_workers\\_web\\_2.pdf](http://www.nesta.org.uk/sites/default/files/model_workers_web_2.pdf)

## Data literacy

Training data workers who can analyse complex datasets is not the only concern. There is also a need to improve data literacy. A range of professions now require some knowledge of data uses, an ability to assess the quality of data and its application, and the skills to conduct basic analysis.

The Royal Statistical Society launched its data manifesto in 2014; it sets out the potential of data to improve policy and business practice, as well as enhance citizen engagement.<sup>17</sup> Understanding basic statistical concepts is essential to make sense of the data that increasingly informs decision-making.

Graduates, almost irrespective of subject area, are educated with a wide range of professions in mind. As such, consideration must be given to how to develop the pedagogy of using data and digital technologies to enhance teaching in all subject areas.

---

<sup>17</sup> Royal Statistical Society (2014) *Data Manifesto* available at: <http://www.rss.org.uk/Images/PDF/influencing-change/rss-data-manifesto-2014.pdf>

## 2. TEACHING WITHIN SUBJECT AREAS

A wide range of, but not all, undergraduate degrees typically offer some of the skills needed to use data. In some cases, students learn a number of basic analysis skills; in others they develop more advanced ability in a number of specific skills. What is considered to constitute basic analysis skills varies by subject area, with certain subjects emphasising statistical skills but not providing much training in data management or analysis software or vice versa. Certain techniques are specific to particular disciplines, yet prepare graduates for work across a range of sectors.

This section describes the extent of quantitative data teaching at undergraduate level across a range of subject areas. There is significant variation in the extent of data teaching not only within subject areas but also between academic institutions. The aim of this section is not to explore this variation, but to offer some broad insights into technical teaching by subject area, and the types of skills acquired by graduates.

The skills shortage is complex and illustrates the need for a range of skills. Rather than listing desired skills and the extent of provision, this section identifies the extent to which students from a range of subject areas are subjected to more or less complex datasets and the nature of training that they are likely to receive.

To categorise subject areas, the Joint Academic Coding System (JACS) adopted by the Higher Education Statistics Agency (HESA) was used throughout the research process. JACS categorises subjects into 20 groups.

In some cases, significant variation in data analytics teaching was found within JACS categories. As a result, geography is discussed separately from other areas, as are economics and business studies. Mathematical sciences and computer sciences are also discussed independently. Where appropriate, reference is made to particular disciplines within JACS categories.

### Mathematical sciences

Not all mathematical courses will encompass data. A clear example is pure mathematics. However, statistics and applied mathematics graduates are well placed to develop the core technical skills required as data analysis in all forms necessarily relies on mathematical concepts. Furthermore, mathematical concepts are being used in innovative ways to overcome the complexity of working with large and complex datasets.

Some students who study mathematical sciences will choose to take advanced statistics or applied mathematics degrees, pathways or modules. These courses will train them to understand many of the underlying principles used to prepare and analyse data, often using analytical packages. In many cases, modules in statistics are assessed through coursework in which students present their results graphically and numerically. Mathematics graduates who have been trained in statistics and data analysis are particularly well placed to recognise the nuances and challenges of particular datasets and respond appropriately to them.

Although many students of statistics or applied mathematics do use analytical packages, several heads of mathematics departments felt that knowledge of data management and presentation was not sufficiently acquired by their students. Data management is sometimes taught more as theory than as practice in the absence of extensive datasets.

A number of stakeholders suggest that the UK currently produces great computer scientists and great mathematicians, but not a combination of both. This is problematic because advanced computer science and mathematical skills are both necessary to derive value from complex data. Although it may be that within data teams individuals have a particular area of expertise, working knowledge of statistics and basic computational and programme skills is necessary to critically engage in data analysis and make sense of data challenges, opportunities and outputs.

### Computer sciences

Computer science courses are valuable in offering students the skills needed to develop appropriate software for collecting, managing and analysing datasets. Some computer science courses may also offer training on the interpretation and visualisation of complex datasets.

Computer science graduates are not trained to use a single technology or piece of software. They are taught how to use a variety of programming languages, enabling them to understand the underlying logic that underpins digital processes. As a result, computer science graduates are able to manipulate software and adapt to new techniques.

Computer science is a mathematical subject. However, a number of employers expressed concern that computer science courses did not sufficiently contextualise data within a particular domain. This makes it more difficult for computer scientists to discern the meaning behind data and identify missing or faulty data that cannot be identified through technological or mathematical means.

Despite the skills shortage in digital technologies generally<sup>18</sup> and data analytics specifically, the Destination of Leavers from Higher Education (DLHE) survey, carried out annually by HESA, shows that computer science has for several years had the highest level of unemployment of the 20 JACS subject areas.<sup>19</sup>

The Council of Professors and Heads of Computing (CRAC) commissioned Robin Mellors-Bourne to investigate this trend. This investigation found that the relatively high unemployment of computer science graduates cannot be explained through a single factor, but that there are multiple contributing factors including gender, ethnicity and degree classification.<sup>20</sup>

Several employers suggested that they have often employed physical or social science graduates to do data work because they were more aware of real world applications and possessed a broader combination of skills. Increasingly, multi-disciplinary courses in computer science are being developed. These are likely to be highly valuable in providing students with the breadth of knowledge and applied skills needed to derive value from data.

### Physical sciences and engineering

New technology and data analysis techniques are being used to explore the physical world. Areas of each of these fields use data-driven modelling techniques and computational analysis to identify patterns and discern characteristics. Many physical science and engineering students will develop some of the technical skills needed to collect and analyse data, as well as experience in using data analysis methods and computational modelling in research. Training is also often provided in the estimation and analysis of experimental errors.

In some courses, data analysis is introduced in practical laboratory modules. In other courses, data analysis is taught in core computational components that constitute a significant proportion of a degree programme. In others still, they are embedded within substantive courses.

Undergraduate curricula in many engineering and physical science disciplines are designed to provide students with a breadth of knowledge and methodologies. Our research shows that data analysis and the use of evolving computational technologies is considered an important tool across physical science and engineering subject areas. However, the extent of data analysis training at undergraduate level is limited by what is possible in a three-year course. Specialisation is generally only possible at postgraduate level.

18 UKCES (2013) *Technology and Skills in the Digital Industries* available at: <https://www.gov.uk/government/publications/technology-and-skills-in-the-digital-industries>

19 Further information on methodology and data can be found on HESA's website at [https://www.hesa.ac.uk/index.php?option=com\\_content&view=article&id=1899](https://www.hesa.ac.uk/index.php?option=com_content&view=article&id=1899)

20 The Council of Professors and Heads of Computing (2012) *CS Graduate Unemployment Report 2012* available at: [https://cphcuk.files.wordpress.com/2013/12/cs\\_graduate\\_unemployment\\_report.pdf](https://cphcuk.files.wordpress.com/2013/12/cs_graduate_unemployment_report.pdf)

## Geography

Some areas of the discipline of geography have a strong grounding in quantitative methods. Geography graduates may have been taught spatial statistical approaches, process and numerical modelling and the use of Geographic Information System (GIS) technologies. These skills are highly desirable to employers across a range of sectors.

The extent of quantitative teaching within the discipline of geography does vary significantly both by subject area and institution. It is common for quantitative methods to be embedded within physical geography modules, but is often taught to a far lesser extent in human geography modules.

A report by the Higher Education Academy (HEA) on mathematics training in geography suggests that there is no evidence that any geography degrees are purely qualitative.<sup>21</sup> At a most basic level, geography students are taught data analysis in fieldwork skills modules or similar using Microsoft Excel. The HEA suggests that this is rare and only for students in a Bachelor of Arts programme. At the most advanced level, quantitative methods are embedded throughout the degree programme. Students are then taught a range of quantitative methods, including inferential statistics, GIS content, spatial modelling and visualisation.

The Royal Geographical Society has worked with the Quality Assurance Agency to reframe the subject benchmark statement for geography.<sup>22</sup> The aim was to ensure that the benchmark statements reflect developments in the discipline. One key outcome was greater specificity around quantitative methods in a geography degree.

## Biological sciences

Students of biological sciences are taught the foundations of statistical hypothesis testing and basic statistics, often up to and including general linear regression models. Most manage data using a software package. Often these skills are primarily taught in research modules in the first and second year.

The extent of quantitative data training in biological sciences varies. Although some areas, such as

biostatistics, rely heavily on quantitative research methods, other biological science courses do not offer such advanced training.

Across subject areas, particular techniques, such as population modelling, population genetics and bioinformatics, are taught as components of substantive modules. Specialisms, such as biostatistics and biochemistry, will offer students an opportunity to learn more advanced data analytical skills. These specialisms are taught at both undergraduate and postgraduate level but are highly dependent on student choices.

## Medicine and dentistry, and subjects allied to medicine

Digital advancements have led to the development and use of new analytical techniques in medical and genetics research. Several institutions run a one- or two-year statistics course for all medical students which can be taken as an intercalated year of study. Some students will choose or be expected to pursue quantitative research in their third year.

Several institutions have extended statistics and quantitative teaching in pre-clinical medicine. In these cases, students develop a good understanding of principles across a range of statistical areas, including simple linear and non-linear regression, multiple linear regression and the basics of resampling statistical approaches such as bootstrapping. However, in many medicine courses students receive little statistics teaching and therefore have a limited understanding.

Specialisation in medical data analysis primarily takes place at postgraduate level due to the advanced level of skills needed to analyse, apply and interpret medical data. At doctoral and post-doctoral level there has been targeted support for research using data analytics in the fields of, for example, biostatistics, bioinformatics and health economics.

Even after doctoral level, further work is needed to ensure that early-career research opportunities are supported. The Medical Research Council (MRC) has partnered with the EPSRC to create doctoral programmes and early research career opportunities which draw on advanced mathematics and biomedicine.

<sup>21</sup> Higher Education Academy (2014) *Skills in mathematics and statistics in geography and tackling transition* available at: [https://www.heacademy.ac.uk/sites/default/files/resources/TT\\_Maths\\_Geography.pdf](https://www.heacademy.ac.uk/sites/default/files/resources/TT_Maths_Geography.pdf)

<sup>22</sup> Recommendations for changes in the benchmark statements were made in the Royal Geographical Society's report *Quantitative methods in geography: making the connections between schools, universities and employers*, available at: [http://www.researchgate.net/publication/256191493\\_Quantitative\\_Methods\\_in\\_Geography\\_Making\\_the\\_Connections\\_between\\_Schools\\_Universities\\_and\\_Employers](http://www.researchgate.net/publication/256191493_Quantitative_Methods_in_Geography_Making_the_Connections_between_Schools_Universities_and_Employers). More information on the revisions to the benchmark statement can be found in a statement from the Royal Geographical Society, which can be accessed at: <http://www.rgs.org/NR/rdonlyres/AC15BCE0-C58C-4345-9753-2FCA103C7268/21498/GeographyQAAbenchmarkstatementdraftRGSIBGconsultat.pdf>

### Economics and business and administrative studies

The business and administrative studies category is broad and encompasses a large number of subjects with very varying levels of quantitative training. Students in most courses will take quantitative methods courses in which they develop basic knowledge of descriptive statistics, and in some cases basic analysis using SPSS and Excel. In some cases far more advanced teaching in data analysis and statistics is provided.

Data analysis training is generally most advanced in economics and finance courses. Many employers consider graduates with a strong background in areas such as numerical modelling and econometrics as highly desirable candidates for data roles. In most economics and finance courses algebra, statistics and data analysis are integral to undergraduate courses and underpin much of the training and application of subject area research. Data analysis skills are taught in specialised modules as well as in substantive courses throughout the degree programme.

A key challenge in each of these subject areas is managing student expectations and facilitating the transition from school to university. There is considerable variation in the level of mathematical knowledge of students enrolling to degree programmes in economics, and business and administrative studies. Students who have not taken A-level mathematics may feel ill-prepared for the mathematical components of their course.

### Sociology, social policy, mass communication and documentation, politics, and international relations

Social sciences offer substantial domain knowledge in areas of great importance to many employers. Knowledge of social context is particularly valuable in determining business strategy on the basis of data, as well as unpacking policy issues within the public and private sector. As such, many employers see great potential in training social science students in data analytics.

Several institutions offer quantitative modules in which students learn multivariate analysis techniques to analyse trends and patterns in the social world. These students are often trained to use statistical analysis software, primarily SPSS and Excel. When quantitative methods are taught, social science students typically learn to understand data within a social or economic context.

However, the extent of quantitative methods training varies considerably between subjects and institutions. The variability in quantitative methods content in UK undergraduate social science courses is often not reflected in degree titles and course descriptions. This makes it harder for students and employers to discern intended learning outcomes and the extent of quantitative training.<sup>23</sup>

Stakeholders reported that there is a tendency within social sciences to distinguish between qualitative, quantitative and theoretical reasoning. This has led to social scientists being labelled (sometimes by themselves) as either 'quantitative' or 'qualitative' people. There is an increasing awareness among some academics that this distinction is unproductive and that possessing the ability to adapt to the most appropriate analytical method to answer a research question is important.

In many courses, quantitative methods training is offered in discrete modules usually taught in the first or the second year. Many stakeholders felt that this isolates quantitative methods and reinforces the preconception that knowledge of quantitative techniques is only important to quantitative experts. This can result in students viewing quantitative methods as a course to 'get through' rather than as an integral component of their field.

The Nuffield Foundation, ESRC and the Higher Education Funding Council for England have funded the development of Q-step centres in fifteen universities to support the development of more in-depth and embedded teaching in quantitative skills in social science undergraduate degrees. Q-step centres will be delivering new courses and modules. Students taking a sufficient number of quantitative courses will receive a 'with quantitative methods' degree title to signal their proficiency.<sup>24</sup>

<sup>23</sup> Similar findings were reported in a review of mathematics in sociology by the Higher Education Academy (2014) *Skills in mathematics and statistics in sociology and tackling transition* available at: [https://www.heacademy.ac.uk/sites/default/files/resources/TT\\_Maths\\_Sociology.pdf](https://www.heacademy.ac.uk/sites/default/files/resources/TT_Maths_Sociology.pdf)

<sup>24</sup> Recommendations were developed by the ESRC's strategic advisor for quantitative methods, Professor John MacInnes. Findings and recommendations were set out in: John MacInnes (2009) *Proposals to support and improve the teaching of quantitative research methods at undergraduate level in the UK* available at: [http://www.esrc.ac.uk/\\_images/Undergraduate\\_quantitative\\_research\\_methods\\_tcm8-2722.pdf](http://www.esrc.ac.uk/_images/Undergraduate_quantitative_research_methods_tcm8-2722.pdf)

### Philosophical studies, historical studies, creative arts, languages, linguistics and law

Humanities, languages and law subjects generally have limited provision in quantitative methods. Several areas of linguistics do rely heavily on data analysis and this is reflected across undergraduate linguistics degree courses. The extent of data analysis teaching across these subject areas is set out in more detail in Table 1. Several employers stated that they did employ humanities graduates, providing that the graduate can demonstrate some experience and skill in analysing data. Employers generally determine on an individual basis whether a humanities graduate is appropriate for

a data role based on individual assessment, which could include consideration of the graduate's thesis topic and chosen modules.

In a number of interdisciplinary arts and humanities courses, such as liberal arts degrees, students will have to take quantitative research courses. Often, these courses are designed to expose students to issues such as sampling, probability, measurement, causation and uncertainty. Yet in some courses, students also develop more advanced skills in data mining, preparation and analysis. These courses help students to develop basic numeracy and data skills, and prepare them for future research methods courses.

**Table 1: Quantitative methods teaching in the humanities, languages, linguistics, arts and law**

Subject area	Description
Law	Law degrees do not usually include training in data and statistical skills. Students who take criminology as an optional subject will develop some understanding of statistics, though not usually the manipulation of data. There has been some debate regarding the importance of numeracy in a law degree. Numeracy skills are important to ensure that legal practitioners can understand and interpret numerical information relevant to a case. Generally, statistics or mathematics is not taught as part of a law degree. When they are, they are often taught in short, discrete modules.
Modern languages	Quantitative skills are generally not taught in undergraduate courses in modern languages.
Music	Generally, data and statistics are not taught in music courses. Several dance courses focus on the physiology of dance and use data in their research. An example of this is the dance science undergraduate course at the Trinity Laban Conservatoire of Music and Dance.
Philosophy	Quantitative skills are generally not taught in undergraduate courses in philosophy, though might be incorporated in some elements at postgraduate level.
History	Statistics is taught in several optional courses in history undergraduate degree programmes, often in modules focusing on economic or population history.
Linguistics	Statistical methods are strongly embedded in a number of areas in linguistics, including psycholinguistics, neurolinguistics, sociolinguistics and corpus linguistics. In these courses, advanced statistics and data analysis skills are taught.

## 3. SECTOR INITIATIVES TO EXPAND PROVISION

A number of initiatives have been developed by sector bodies and higher education institutions to expand data analysis provision at undergraduate and postgraduate level. Initiatives within STEM subjects have primarily focused on uniting advanced skills at doctoral level, whereas a number of initiatives have been developed for social sciences that look to embed more quantitative training in undergraduate courses.

### New courses

Several institutions are offering or developing cross-disciplinary undergraduate programmes which are jointly taught by mathematical sciences and computer sciences departments. For example, in 2014, eight universities announced data science degree programmes to start in either 2015–16 or 2016–17.<sup>25</sup> All of the data science courses announced emphasise the importance of developing a full range of skills, teamwork and effectively communicating findings to specialists and non-specialists alike.

### The Q-step initiative

Within the social sciences the Q-step initiative has been established by the Nuffield Foundation, ESRC and the Higher Education Funding Council for England. This initiative is seen by its advocates as a first step towards expanding the provision of quantitative methods training in the social sciences. Fifteen centres have been established to provide advanced training in quantitative methods. Students who have completed a Q-step undergraduate programme will receive a 'with quantitative methods' degree to signal their proficiency. In February 2015 a further three institutions were awarded Q-step Affiliate status in recognition of their high level of commitment to the aims of Q-Step.<sup>26</sup> Each institution will receive a grant to fund short courses and student bursaries until 2018.

### EPSRC doctoral training centres

The EPSRC has established cross-discipline doctoral centres to unite computer science and mathematical science disciplines and skill at postgraduate level. Computer science, mathematical science and physical science graduates are able to apply to the doctoral centres. This approach enables students to develop a discipline focus at undergraduate level, which can be applied and expanded through interdisciplinary postgraduate work. The centres focus on various areas of data analysis, including financial computing and analytics; data science; cloud computing for big data; and statistical applied mathematics.

<sup>25</sup> These universities are the University of Bedfordshire, Bournemouth University, the University of Essex, the University of Nottingham, Sheffield Hallam University, Teesside University, the University of Warwick and the University of the West of England, Bristol.

<sup>26</sup> These universities are the University of Essex, University of Nottingham and University of Southampton. See: <http://www.nuffieldfoundation.org/news/universities-essex-nottingham-and-southampton-awarded-q-step-affiliate-status>

## 4. CHALLENGES

Sector stakeholders are positive about current initiatives, but they suggest that a number of challenges remain to be addressed. Failure to properly define the skills shortage or failure to recognise the barriers to an expansion of data training will hinder initiatives aimed to tackle the shortage of qualified data analysts.

### Defining the skills shortage

The research undertaken for this report revealed that the data skills shortage is not always well understood and that the data needs of organisations can vary both between and within industries. Due to the value of quantitative data in academic research across subject areas, universities are able to identify needs based on developments in the use of certain methods within disciplines and through the identification of postgraduate skill shortages. This, however, does not necessarily align directly with the needs of businesses.

There is a need to identify the core skills that graduates require for entry-level data roles. Once a core set of entry-level skills is identified, it is possible to consider which roles or sectors may require more advanced ability in a particular technique or domain. In addition, establishing a set of core skills allows for educators and employers to develop pathways of progression. These pathways can encompass training from schools through to further and higher education, and into employment. Universities could have an important role to play in providing professional development courses.

### Facilitating transitions to higher education

Although the focus of this report is on the transition from university to employment, it is important to consider the transition from school to higher education. Course providers consistently identified three areas where this transition was problematic and at times constrains the level of training that universities can provide to undergraduates.

#### 1 Variable mathematical knowledge and training

Almost all social science degree courses require an applicant to possess GCSE mathematics at grade C and above, but few set stricter requirements. As a result, the aptitude of students in mathematics and statistics varies considerably within these courses. This makes it difficult for academics to design appropriate quantitative research courses that are stimulating for all.

Within the current A-level curriculum, statistics is taught as an optional module rather than as an integral part of the course. Change is on the way, with the introduction of a new specification in 2017. Following a recent review of A-levels undertaken by the A-level Content Advisory Board, the specifications have been revised to strengthen the statistical content of the mathematics A-level.<sup>27</sup>

#### 2 Variable computer science and coding skills

Many students do not take computer science courses at GCSE or A-level and efforts to use digital technologies across subject areas vary between schools and colleges. The latest statistical release from the Department for Education shows that in 2013–14, 3,451 students entered for A-level computing (covers computing and software systems development). This compares with 48,401 entries for A-level mathematics.<sup>28</sup>

<sup>27</sup> Department for Education (2015) *Reformed GCSE and A-level subject content consultation: Government response* available at: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/397672/Reformed\\_GCSE\\_and\\_A\\_level\\_subject\\_content\\_Government\\_Response.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/397672/Reformed_GCSE_and_A_level_subject_content_Government_Response.pdf)

<sup>28</sup> See: <https://www.gov.uk/government/statistics/a-level-and-other-level-3-results-2013-to-2014-revised>

### 3 Data rarely embedded in substantive courses

A number of initiatives have been designed to teach coding in schools, provide more individuals with post-16 mathematics education and expand teaching in statistics. The next step is to embed these skills in substantive courses. Currently, quantitative methods training is largely excluded from most social science subjects at A-level. This is a missed opportunity to show the importance of quantitative methods in these subject areas and to prepare pupils for quantitative methods training at university.

The challenge of embedding quantitative and mathematical methods also applies within universities, where methods courses can be taught in isolation from the rest of the curriculum. Failure to change this will result in the perception that these methods are not a core component of a range of disciplines.

### Gender divides in key subject areas

In 2013–14, women constituted just under 40% of entrants for A-level mathematics. The gender divide is even more acute for entrants for further mathematics, where less than 30% of entrants are female.<sup>29</sup> This trend is apparent despite the fact that there are no significant differences in the performance of girls and boys at GCSE level.

Gender divides are also evident in undergraduate mathematical and computer sciences, where often

some of the most advanced teaching in data analysis is provided. In 2013–14, 39% of full-time entrants to mathematical sciences were women. The imbalance is more acute in computer sciences, where just 15% of 2013–14 entrants were female. The effect of this gender imbalance should be considered and addressed.

### Making the case for data analytics

In order to develop a curriculum which successfully incorporates the complex set of hard and soft skills required to extract value from data, students and academic staff must be aware of its potential. Yet the case for the value of quantitative data is not universally accepted within a number of key academic fields.

### Recruiting lecturers to teach data analytics

Stakeholders said that a significant barrier to providing data training across all subject areas was a lack of lecturers with the necessary skills. This skill shortage not only impacts businesses but also creates a substantial barrier to universities wishing to expand provision.

Stakeholders have tested three possible solutions to this problem: upskilling current staff, drawing on teaching from the statistics (or equivalent) department and developing quantitative methods lectures centrally. The benefits and drawbacks of each approach are described in Table 2.

**Table 2: Possible solutions to alleviate staff shortages**

Possible solution	Pros	Cons
Upskilling current staff	Methods teaching embedded as part of the wider curriculum Relatively low amount of resource and qualified staff required	In some instances, staff members do not want to learn data analysis skills Requires pre-existing institutional expertise Potentially long lead-in time
Service-level teaching from the statistics or equivalent department	Relies on existing institutional resource Relatively short lead-in time	More difficult to embed methods teaching within wider curriculum In some institutions, existing staff may have limited capacity to do this Staff members may not want their students to be taught methods they themselves do not feel confident using High cost might mean that home department takes teaching in-house to save money
Developing quantitative methods lectures centrally	Relatively low number of qualified staff required	More difficult to embed methods teaching within wider curriculum Potentially long lead-in time as it requires the careful design of both plenary lectures and smaller seminars with access to computers

<sup>29</sup> See: <https://www.gov.uk/government/statistics/a-level-and-other-level-3-results-2013-to-2014-revised>

To maximise skills and future benefits, students should be taught how to use new data, methodologies and technologies. For students to learn to adapt to the state of the art, academic staff must feel comfortable and familiar with new technologies themselves. This means fostering an environment within universities in which continuous learning is both encouraged and enabled.

### Maximising the potential of a three-year curriculum

To obtain the skills set needed to derive value from data, students across subject areas must be taught a range of technical skills as well as the domain knowledge required to develop appropriate research designs and critically evaluate data. A major challenge facing academics is how to balance training in various skills and topics within a three-year curriculum.

Another problem involving curriculum design is that the technical skills needed to analyse data are cumulative. This requires students to gradually develop more advanced skills as they progress in their degree programmes. However, this also means that certain quantitative techniques cannot simply be pulled out to fit an area of the curriculum if they have not been preceded by more basic techniques.

To achieve an appropriate balance, standards would be helpful in order to define core skills. Standards could offer insights into future career progression, and into how new skills can be gained or old ones developed through both training and practical experience. This will enable a distinction to be made between the core skills that should be taught at undergraduate level, and the targeted training and support that could be offered afterwards. In addition, standards could be used to develop a scheme that identifies courses in which core skills are acquired by students. This would provide an important signal to employers.

### Declining numbers of mature and part-time students

There may be opportunities for employees to develop and expand data analysis skills by studying for university degrees alongside their career. However, figures show that the number of mature and part-time students has declined considerably in the last five years. The number of mature undergraduate entrants in England, a significant proportion of whom would study part-time, fell by 36.5% between 2007–08 and 2012–13.<sup>30</sup>

A call for evidence by Universities UK identified a number of factors that may have contributed to this trend. Notable among these were the impact of reforms to undergraduate student funding, including increased fees and restricted access to student loans, and the impact of economic decline and with it increased unemployment and reduced employer funding.

---

<sup>30</sup> Universities UK (2014) *Trends in undergraduate recruitment* available at: <http://www.universitiesuk.ac.uk/highereducation/Documents/2014/TrendsInUndergraduateRecruitment.pdf>

## 5. OPPORTUNITIES

There are a number of opportunities that put universities in a good position to overcome some of the challenges outlined in the previous section. In addition to initiatives established by foundations, universities and research councils to tackle shortages, policy makers and businesses have also supported and developed a range of actions.

This includes a number of changes to pre-university provision, including the introduction of the new coding curriculum in schools and the development of core A-level mathematics. The government has also repeatedly signalled the importance of data and digital opportunities to the UK economy, investing in the development of the Alan Turing Institute, which is headed by five UK universities, to promote the development and use of big data.

### University-employer collaborations

Universities continue to play an important part in preparing students for their careers by developing their skills and helping them to interact with potential employers.<sup>31</sup> Effective collaboration between universities and employers plays, and will continue to play, an important role in promoting data science as a career. It encourages more students to consider pathways into data careers and research, subsequently stimulating interest in relevant degree courses.

Universities and employers have collaborated in a number of ways, for example through the development of work placements for students, consultations around curriculum design, and engagement through careers services. Many academics commented that work placements in many ways 'made students better students', stimulating interest in their field and helping them to develop their skills.

In almost all cases, employers and academics suggested that managing expectations was instrumental in successful collaboration. For a work placement to be most effective, students should receive appropriate on-the-job training, benefit from sufficient supervision and learn via an effective feedback system. This requires the employer to invest a significant amount of time. To make this investment worthwhile, employers need to know that a student meets a minimum level of competence in core skills. Some stakeholders suggested that this could be ensured through transparent application procedures and job interviews.

### Mathematics and statistics support centres

Many institutions have mathematics and statistics support centres that provide assistance to students across disciplines. Support centres play an important role in supporting students as they transition to higher education. However, these centres should often not simply be seen as remedial, but as centres of expertise for the whole academic community.

More students are now seeking help with data analysis and statistics at support centres. However, supporting students with questions around data analysis can be challenging as it requires an appropriate dataset to illustrate techniques and sometimes does not align well with the expertise of staff in the centre.

A network for support centre staff, the Sigma network, plays an important role in creating a forum for communication between trainers. The network coordinates discussions and events to identify emerging needs and challenges. This is playing an important part in expanding capacity to support students with a range of enquiries from across disciplines. Regional hubs and the national Sigma network also offer courses and resources for trainers and present awards to support centre staff in recognition of innovative practice.

Mathematics and statistics support centres are well established and are increasingly utilised by students and staff from a range of academic departments. There may

31 UKCES, Universities UK (2014) *Forging Futures: Building higher level skills through university and employer collaboration* available at: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/356749/FF\\_FinalReport\\_Digital\\_190914.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/356749/FF_FinalReport_Digital_190914.pdf)

be opportunities for institutions to consider adapting this model to develop computer science, programming, data management support centres or tools to support students in the development of these core skills.

### **Data teaching networks**

A quantitative teaching network for social scientists was established by John MacInnes, the strategic advisor in quantitative methods teaching to the ESRC. Members of this group share teaching practices and resources, working together to promote quantitative training in the social sciences and enhance the quality of provision. There may be opportunities for the establishment of more cross-institutional academic networks on teaching data analytics within and across subject areas.

### **School and A-level reforms**

Stakeholders have welcomed the introduction of the new Core Maths curriculum. Many felt that the programme would help to prepare students for the quantitative components of their course by applying key techniques and helping them to develop greater confidence with numbers. The effectiveness of these courses should be determined in consultation with teachers and academics from a range of subject areas.

Similarly, the coding curriculum is a positive step forwards, enabling young people to develop more advanced computer skills and become aware of the possibilities of technology. The next step will be to embed both computer science and mathematical skills across subjects taught at primary and secondary level.

---

## 6. IMPLICATIONS FOR POLICY

UK plc is already extracting value from data, and will continue to do so without intervention. However, there is an existing skills shortage, and it is to be expected that without action the gap between the demand for skills and the supply of them will continue to widen.

According to a recent report by Deloitte, 35% of existing jobs in the UK are at high risk from automation.<sup>32</sup> Advances in technology will likely see lower-skilled jobs, such as those requiring repetitive processing, disappear. As a result, clerical and support services may be replaced by roles requiring digital, data management and creative skills.

This challenge is not limited to those jobs at high risk from automation. Some 84% of London businesses say that the skills of their employees will need to change over the next ten years. 'Digital know-how' is just one of the skills that businesses will increasingly need.

Having reviewed current provision of data analysis training at undergraduate level, we conclude that some of the technical skills needed to analyse data are indeed found across a range of subject areas – but that the extent of training is variable and not clearly articulated.

Robust action is needed to get most value from data. The higher education sector has taken several steps to expand provision, but more work is needed. To meet current and future demand for data skills, collaboration between schools, universities, employers and policy makers will be crucial in defining what core skills are needed and in embedding them in a framework of skills development and progression. The recommendations developed by Universities UK and Nesta will support this.

The consequences of failure are clear: underutilisation of current and future investment in information technology, and a workforce that is less competitive and productive than those of our competitor countries.

---

<sup>32</sup> Deloitte (2014) *One-third of jobs in the UK at risk from automation* available at: <http://www2.deloitte.com/uk/en/pages/press-releases/articles/deloitte-one-third-of-jobs-in-the-uk-at-risk-from-automation.html>.

# ACKNOWLEDGMENTS

This research was overseen by a steering group led by Professor Sir Ian Diamond and involving Professor Richard Davies, Dr. Phil Richards, Dr. Hetan Shah, Professor Sir Adrian Smith.

**UUK would like to thank the universities who participated in the survey and/or interviews:**

University of Bath, Birkbeck University of London, University of Bristol, Brunel University, University of Cambridge, City University London, University College London, University of Essex, University of Exeter, University of Greenwich, University of Hull, Keele University, King's College London, Kingston University, University of Leeds, Leeds Trinity University, University of Leicester, University of Manchester, Manchester Metropolitan University, University of Newcastle, University of Nottingham, Nottingham Trent University, The Open University, University of Oxford, University of Portsmouth, Plymouth University, Queen Mary University of London, University of Reading, University of Roehampton, London School of Economics, University of Sheffield, University of Southampton, University of Warwick, University of the West of England

---







This publication has been produced by Universities UK, the representative organisation for the UK's universities. Founded in 1918, its mission is to be the definitive voice for all universities in the UK, providing high quality leadership and support to its members to promote a successful and diverse higher education sector. With 133 members and offices in London, Cardiff (Universities Wales) and Edinburgh (Universities Scotland), it promotes the strength and success of UK universities nationally and internationally.

Woburn House 20 Tavistock Square London WC1H 9HQ

**Tel:** +44 (0)20 7419 4111

**Email:** [info@universitiesuk.ac.uk](mailto:info@universitiesuk.ac.uk)

**Website:** [www.universitiesuk.ac.uk](http://www.universitiesuk.ac.uk)

**Twitter:** @UniversitiesUK

July 2015

**ISBN: 978-1-84036-339-5**



This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

To download this publication, or for the full list of Universities UK publications, visit [www.universitiesuk.ac.uk](http://www.universitiesuk.ac.uk)