

## **ROYAL STATISTICAL SOCIETY RESPONSE TO CONSULTATION ON THE NATIONAL DATA STRATEGY**

9 December 2020

This is the Royal Statistical Society's (RSS) individual response to the consultation on the National Data Strategy.

As part of our work on the National Data Strategy, we worked in partnership with the Ada Lovelace Institute, the Centre for Public Data, the Institute for Government and the Open Data Institute to organise a series of events covering the four pillars of the National Data Strategy. At these meetings, experts and practitioners discussed how the aims of the National Data Strategy could be realised practically and sustainably.

We have jointly published a document [Getting data right: perspectives on the UK National Data Strategy 2020](#), which details the key insights gathered through the series. We support the key findings outlined in this document.

Here, in our individual response to the consultation, we cover two key areas where we most obviously have a remit: first regarding data skills and statistical literacy and second regarding data ethics (where our response in part addresses Q11 in the consultation document).

We make five recommendations:

- Recommendation 1. Reintroduce a one-year post-GCSE maths qualification with a substantial statistical element to increase the number of people studying the subject beyond GCSE.
- Recommendation 2. Invest in training teachers in non-scientific subjects to confidently use statistics as part of their lesson plans and promote courses in social sciences that incorporate statistical aspects.
- Recommendation 3. Improve teaching of statistics and data skills in post-18 education – especially in social science subjects and journalism.
- Recommendation 4. Encourage joint appointments between universities and academia and ensure that research councils recognise diverse outputs so that this career path is valued.
- Recommendation 5. Recognise the profession of data analysts and support people with data and statistical skills to move to leadership positions.

### **Data Skills and statistical literacy**

In [Getting Data Right](#) we highlight four key points (pp.5-6):

1. Data skills are valued by a wide range of employers across all sectors. It is important not just to focus on the growing demand for workers with specialist data skills and to ensure that people from a wide range of backgrounds are given the opportunity to develop these skills. This might mean: reintroducing something like AS-Level mathematics to increase the number of people taking maths past GCSE level and training teachers in all subjects so that they are comfortable in teaching data skills that are relevant to their particular subject.
2. There is a need for greater collaboration between universities and industry to produce graduates with appropriate skills. This could be done by encouraging joint appointments for people to work at the interface of industry and academia. Which would require research councils to recognise diverse outputs so that this career path is properly valued.
3. Career pathways for people with data skills could be improved in, eg, operational research. To take the NHS as an example, this employs a large number of people with data skills to

analyse things like waiting lists and where there is need for new services – but these roles are classified as admin/clerical and they lack career pathways and opportunities for advancement. These roles should be properly classified as scientific/technical and leadership training should be provided.

4. It is important to develop a professional framework for data scientists with shared codes of practice. Part of this means professional accreditation for people working in this area to help establish professional standards for data science. This is the objective behind the Royal Statistical Society's new Data Analyst accreditation.

There are two related but distinct issues here: first, around increasing the number of people developing professional data analytical skills and pursuing careers in this field across industry, academia and the public sector; second, as data becomes increasingly important to the world around us, improving statistical and data literacy among the wider population.

The National Data Strategy clearly recognises the importance of data skills as a professional pathway and we warmly welcome this emphasis – we believe that the Royal Statistical Society has a role to play in helping to advance the government's agenda in the area.

### **Foundational skills and literacy**

It is also welcome that the Strategy recognises that foundational data literacy will be required by all. From our perspective, the importance of statistical and data literacy has been clearly demonstrated in the pandemic. demonstrated the importance of being able to engage with and understand data. Public understanding has been important both for navigating a sometimes-confusing mass of information about the pandemic, including misinformation on various topics, and for understanding how to manage personal risk. The importance of this is underlined by evidence that the greatest predictor of a person's resilience to fake news about Covid is their level of numeracy.<sup>1</sup>

A recent report into economic statistics by NIESR, [Public Understanding of Economics and Economic Statistics](#), has also highlighted the problem in relation to economic statistics. This research, using focus groups, suggests that part of the problem stems from economic statistics being communicated in an inaccessible way – using unnecessary economic jargon, which is detrimental to people's engagement. This highlights the importance of communication of economic statistics specifically as well as statistics and data more generally. This requires academics, the government, the media and organisations like the Royal Statistical Society with a mission to promote public understanding to think about how statistics can be presented in the most accessible way possible.

Improving both literacy and professional skills begins at school: statistics and data skills are currently not taught widely enough at this level. Ending AS Level maths as a stand-alone qualification reduced the number of people taking statistics modules and the core maths programme is only being offered by a minority of schools. Students who do not take mathematics at school do not get enough exposure to statistics and data skills. The problem is exacerbated by the fact that many schoolteachers, especially in non-scientific subjects, do not have the background to be able to teach the aspects of statistics that are relevant to their subject.

A similar issue is apparent at university level: data skills are increasingly important in non-STEM subjects, but they are not widely taught. As with schools, this is partly because teaching staff are not confident in doing so.

In terms of non-STEM subjects, the RSS's view is that too many social science programmes, at both undergraduate and postgraduate level – including those focusing on topics in social statistics,

---

<sup>1</sup> Rozenbeek et al, [Susceptibility to misinformation about COVID-19 around the world](#)



advanced quantitative methods, social analytics and social data science – do not currently give students sufficient training in key statistical foundations. Data analytical skills need to be more deeply embedded in programmes and should promote hands-on experience. At doctoral level in particular, it is worth looking at US social science programmes which routinely include more advanced features such as Python programming and webscraping. UK programmes should aspire to this level as well.

It is perhaps also worth saying a word about journalism, which has a particularly important role to play in improving statistical literacy (although, of course, not all journalists go through university). The importance of the media in communicating statistics has been clearly shown over the past year in reporting around the pandemic. Specialist journalists – health or science and technology correspondents – have been generally very good at reporting data accurately, including graphically, and this has helped people understand the pandemic. However, general reporting on statistical aspects of the pandemic has been less positive and, especially in the early stages, contributed to public misunderstanding of, for example, the death figures that were being presented at the daily briefings.

To improve foundational skills and literacy, the RSS makes the following recommendations.

*Recommendation 1. Reintroduce a 1-year post-GCSE maths qualification with a substantial statistical element to increase the number of people studying the subject beyond GCSE.*

*Recommendation 2. Invest in training teachers in non-scientific subjects to confidently use statistics as part of their lesson plans and promote courses in social sciences that incorporate statistical aspects.*

*Recommendation 3. Improve teaching of statistics and data skills in post-18 education – especially in social science subjects and journalism.*

### **Professional skills and pathways**

There are three key elements for improving the development of professional skills and providing career pathways for data analysts:

1. Ensuring high quality statistics and data education at university level.
2. Developing relationships between academia and industry to ensure that universities are producing graduates with the skills needed by industry
3. Valuing the profession of data analysts and providing clear pathways for them to advance in their careers – including to leadership positions outside of the specific discipline.

In terms of ensuring high quality education at university level, one of the key challenges is retaining staff. Academics with strong data analytical skills are highly valued by industry – where they are typically paid more and can spend more time focusing on R&D and innovation. It is hard for universities to hold onto these staff. If too many academics leave universities for industry, this clearly impacts on the quality of education that undergraduates and postgraduates will receive.

We believe that part of the answer to this lies in promoting joint appointments between universities and academia – allowing staff to be based for part of their time in universities and for part of it working in industry. This could have at least three positive impacts: reducing the number of academics leaving university entirely to move to industry; promoting collaborations between industry and academia that will direct university research towards innovation; and, enabling more academics to gain an insight into industry need so that they can adapt their teaching to help students develop more relevant skills.

It is important that data analysts are properly valued and, currently, is not always clear that this is the case. There are two ways that this shows itself: first, when data analysts work in operational research they are sometimes treated as admin or clerical staff rather than scientific or technical staff. Ben



Goldacre has detailed the impact of this in the NHS – both in RSS meetings and a research paper<sup>2</sup> – but there are general lessons here as well. Second, the value of statisticians and data scientists as leaders – and the potential that background brings to help organisations take advantage of new opportunities available through the innovative use of data is underappreciated. For example, in the civil service people will advance within the Government Statistical Service, but it is comparatively rare for them to move to senior leadership roles outside that sphere.

It is also important to develop a professional framework for data scientists with shared codes of practice. As noted in the National Data Strategy, the Royal Statistical Society has a role here to work with employers and universities to identify the skills needed for data scientists and accredit courses so that students and professionals can be confident in their quality. We have begun this work by developing an accreditation for Data Analysts and we will be able to start recognising people for their skills in this area in 2021. Our aim is to ensure that this accreditation is valued by industry and academia and we are working closely with stakeholders in both sectors to ensure that the framework we are developing works well from both perspectives. On this matter in particular we would welcome further discussion as the National Data Strategy develops.

*Recommendation 4. Encourage joint appointments between universities and academia and ensure that research councils recognise diverse outputs so that this career path is valued.*

*Recommendation 5. Recognise the profession of data analysts and support people with data and statistical skills to move to leadership positions.*

### **Data ethics and the role of the Centre for Data Ethics and Innovation (CDEI)**

The RSS would like to see the National Data Strategy say more about what we take to be a very important issue: trust in the responsible and appropriate use of personal data by government, academia and the private sector for a wide variety of purposes (ranging from the desire to increase profitability, improve service efficiency and/or to address major research issues where there is a public interest). We welcome the creation of the CDEI and the government's data ethics framework: however, we believe that there are ways in which the data ethics agenda could be strengthened.

As the Cabinet Office has now assumed responsibility for government use of data, we would like to see its ethics functions more strongly linked to the UK Statistics Authority (UKSA). In particular, we would like to see the adoption of some of UKSA's good practice – as set out in the Code of Practice for Statistics – of being open and external facing.

We view the self-assessment tool that UKSA have developed and that is mentioned in the strategy as exemplary and would strongly encourage an uptake in its usage. This agenda could be developed further by extending it to include:

- The provision of online ethics training for the research and statistical community across Government, academia, and the commercial sector.
- An ethics user support service for the research and statistics community to provide ethics advice at the research design phase.
- The publication of guidance on cross cutting ethical issues in research and statistics.

We believe that there is an opportunity for the UK to be distinctive and offer international leadership on data ethics. The ambition of ensuring “that UK values of openness, transparency and innovation, as well as the protection of security and ethical values, are adopted and observed globally” is praiseworthy: signalling our commitment to independent ethics and use of data in the public interest

---

<sup>2</sup> Goldacre et al, [‘Bringing NHS data analysis into the 21st century’](#)

may have soft power benefits in the international scene. Indeed, this has been recognised in the independence of our regulation of statistics by OSR, which is not something other countries have, and the role of professional bodies in establishing the data science accreditation standards.