



The Career Development Organisation

Evaluation of National Data Skills pilots

Interim briefing report

For the Office for Students

By the Careers Research & Advisory Centre (CRAC)

March 2022

© Office for Students 2022

Contact: Robin Mellors-Bourne
Director, Research and Intelligence, CRAC

Careers Research & Advisory Centre (CRAC) Ltd
22 Signet Court
Cambridge CB5 8LA

robin.mellors-bourne@crac.org.uk

www.crac.org.uk

Introduction

Through the Office for Students, the Department for Digital, Media, Culture & Sport (DCMS) has provided funding for an exploration of practice by universities in enabling higher education (HE) students to develop foundational data skills.¹ The focus is on provision for undergraduates who are studying courses without a significant data science element.

The funding is supporting seven UK universities to showcase current provision of this type, together with support from the Careers Research & Advisory Centre (CRAC) to consider its effectiveness and identify good practice. CRAC is also undertaking a wider evidence review to identify how other higher education providers are teaching these data skills to non-cognate students (i.e. those on courses without a significant data science element) or supporting such learning, including outside the UK.

This briefing summarises progress in the project to date and signals that more substantial findings will be shared in summer 2022.

Context

The National Data Strategy (NDS), published in September 2020, set out a vision to harness the power of responsible data use to boost productivity, create new businesses and jobs, improve public services, support a fairer society, and drive scientific discovery, positioning the UK as the forerunner of the next wave of innovation.² It is a framework for the action required to maximise the power of data across the UK. One of its four key pillars is data skills – with a stated desire that the right skills are developed through the education system and that people can continue to develop the data skills they need throughout their lives.

The need for data skills continues to grow across the economy. The Royal Society has reported that demand for specialist data skills has more than tripled since 2013,³ while analysis of over 9 million online job adverts predicted that data analysis skills will be the fastest growing digital skills cluster over the next five years. Recent research has found that UK companies were seeking to fill up to 234,000 roles requiring hard or technical data skills and almost half of businesses have been recruiting for data-related roles.⁴ Research from

¹ By 'foundational' we do not mean learning specifically at foundation degree level (Level 5), but rather a range of data-related skills. These could include: management, modelling, cleansing and enrichment of data; data visualisation; statistical methods and data analysis; data quality assurance, validation and linkage; and some other skills within the broad concept of 'data literacy'.

² <https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy>

³ <https://royalsociety.org/-/media/policy/projects/dynamics-of-data-science/dynamics-of-data-science-skills-report.pdf>

⁴ <https://www.gov.uk/government/publications/quantifying-the-uk-data-skills-gap/quantifying-the-uk-data-skills-gap-full-report>

Forrester in 2021 has found that recruiters rank data literacy as the skill highest in demand for entry-level candidates.⁵

Those and other reports suggest exponential growth in the demand for advanced applications of data science and machine learning across all sectors of the economy. Growth in artificial intelligence (AI) and cyber specialisms also drives demand for a broader supply of data skills, specifically at foundational level, as these help to feed the pipeline of advanced skills and provide businesses with the ability to work with data.

In 2015, Universities UK urged the HE sector to ‘do more’ to embed data skills across all degrees.⁶ However, Forrester’s 2021 research believes only half the UK’s academic institutions have data literacy skills initiatives in place. With UK universities operating in a competitive international market, how data skills learning is embedded within or supplements curricula could increasingly inform the choices of overseas students and sponsors.

This all suggests that UK HE institutions need to find ways to embed foundational data skills and data literacy across their programmes – not just in technical courses – as almost all professions will utilise data in one form or another over coming years. This project aims to showcase and demonstrate approaches being made to teaching these foundational data skills, and where possible which approaches work best, so that students at more UK HE institutions learn them as part of a wide range of degree courses.

Project progress and emerging results

Seven pilot projects received funding from October 2021 to March 2022 and will conclude delivery of activities during May 2022, after a short extension was provided to allow more time to evaluate their activities. Project teams have been working with CRAC to establish measures of success in relation to students’ learning and skill development.

The wider evidence review conducted by CRAC has encompassed a rapid literature search together with a web search exercise to identify provision of data skills modules and other learning offered by universities and/or third parties that is relevant to this project. Most of the practice found has been through the latter, rather than evidenced in more formal literature, although a few published studies were found that contributed to the understanding of approaches being taken in certain countries.

Analysis of documented examples of provision found in our desk research suggests it may be helpful to establish a framework with which to illustrate and consider different approaches. In Figure 1 we suggest such a framework, which considers how provision is

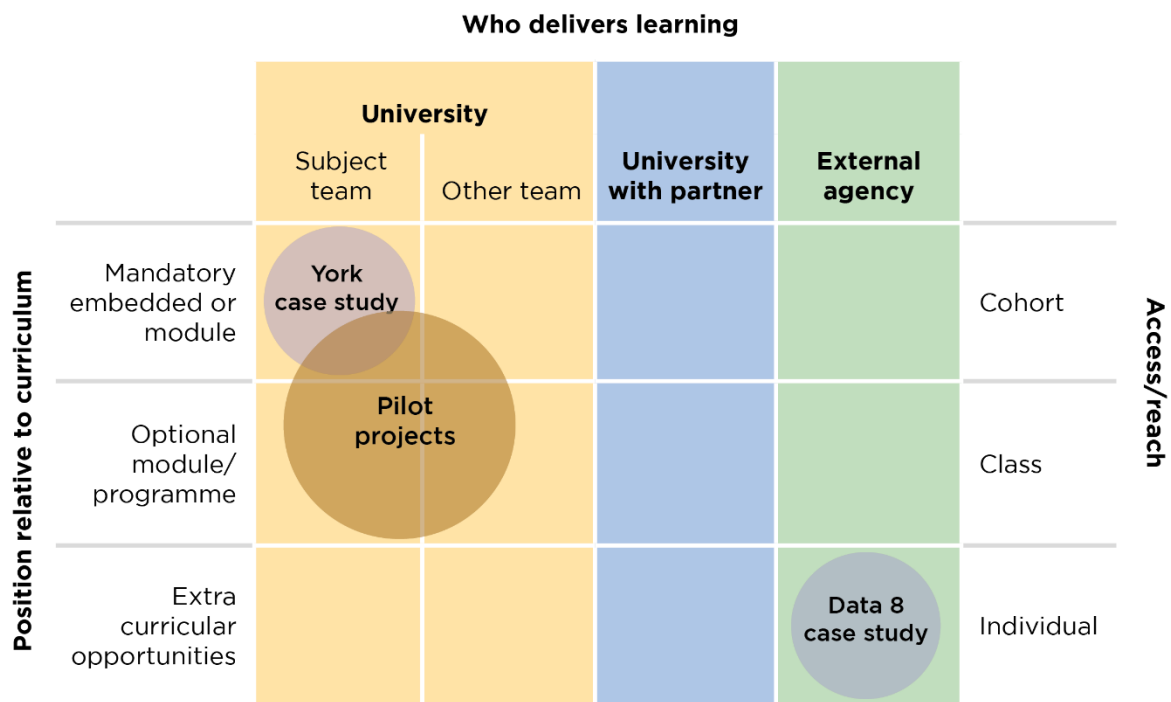
⁵ https://www.tableau.com/sites/default/files/2021-06/Tableau_Data_Literacy_Report.pdf

⁶ https://nanopdf.com/download/making-the-most-of-data-data-skills-training-in-english-universities_pdf

positioned in relation to the student’s subject curriculum, who delivers the additional provision or learning, and whether it is accessed on a cohort or individual basis.

Approaches we have identified from this evidence review range from compulsory data skills modules embedded within the core curriculum of a subject course (delivered by those teaching the subject), to optional modules provided by other parties in the university (such as on a ‘service teaching’ basis), through to optional learning opportunities from external providers that are made available to all students (and staff).

Figure 1: Framework for approaches to development of foundational data skills



We are also able to plot the approaches taken by the pilot projects within the framework. The pilot projects’ aims included establishing evaluative insights into the effectiveness of their approaches, while some are specifically comparing different approaches. The pilot projects are:

- **Birmingham City University** and **Teesside University** are both undertaking comparative studies of the effectiveness of different pedagogical approaches to teaching mandatory modules on data skills in a wide range of disciplines.
- **Lancaster University** is assessing the effectiveness of its newly developed teaching of R skills to undergraduates within its core psychology curriculum.

R is a programming language for statistical computing and graphics, now part of a wider open-source software environment as users have created packages to augment the language itself. It is one of the most widely used languages in data mining.

- **The University of Hull** is also comparing approaches across different non-cognate courses, where data skills are taught as standalone, embedded or scaffolded modules.
- **Newcastle University** offers a largely standard data skills module to a wide range of courses, rather than bespoke embedded approaches, and is assessing this across three different faculties.
- **The University of Wolverhampton** is looking at the effectiveness of offering an existing module (for students in engineering and science) to non-cognate students on a non-credit-bearing basis.
- **Solent University** has developed an optional data skills development programme to precede its curriculum modules, and is testing its efficacy as a means to enhance student learning and attainment.

We are aiming for the framework to help us showcase different possible approaches in order to assist other universities in considering how best to enable their students to develop these foundational data skills.

From the evidence review, at this interim stage, and purely for illustration, we include here two abbreviated case studies of different approaches taken to development of data skills, to signal that a range of case studies (of a fuller nature, of which there will be several for each type of approach) will be provided within the final output from this project. These will encompass both pilot projects and initiatives discovered in the desk research.

Case study examples

University of York: Big Data Biology

This core, 10-credit Year 2 module aims to extend student capability with statistics and the R programming language in particular, teaching biology-relevant data concepts, analysis and handling skills. Essentially it enables students to use large tables of data to understand biology, while developing programming language skills that can be used in many kinds of jobs.

The online module includes introductory video lectures; video lectures on datasets, data analysis and R skills; exercises to practise R skills; elective data exercises; further exercises supported by online workshops; and an assessed analysis and report.

Data 8: University of California Berkeley

This is a course open to all undergraduates at the university, offered by the Division of Data Sciences, now taken by several thousand students. It starts with concepts like inference and random sampling, based on real-world problems, rather than computational theories. Students use real data to understand relationships and patterns while being taught critical concepts and skills.

It is a 15-week course consisting of face-to-face or online lectures and labs, supported by self-study assignments and final examination, all designed for students without a statistics or computer science background. Examples used are located in disciplines across sciences, sports and literature. Half of those who enrol are women, which is higher than traditionally seen on data-focused courses, suggesting that the approach reaches out widely.

Next steps

The evidence review will shortly be concluded, incorporating certain examples of practice from outside the UK. Work with the pilot projects is aiming to gain insight into their reach and effectiveness, which will deepen understanding of different approaches and their potential. In summer 2022 we will publish a final showcase report illustrating the range of approaches with a variety of case study examples for each (across different HE contexts and disciplines).