USING MOOCS TO EDUCATE AND ENTHUSE THE DATA PROFESSIONALS OF TOMORROW

Luke Burns

University of Leeds (UNITED KINGDOM)

Abstract

This research pulls together the increasing need for specialist numerate data professionals in the contemporary workforce with an appreciation that these skills need to be embedded far earlier in the student education lifecycle.

With 90 per cent of the world's data generated in the past two years, the data needs of industry are changing. Core quantitative skills such as descriptive statistics and probability are still important but are being overtaken by the technical skills needed to handle 'Big Data' and other new forms of data.

This research undertakes a skills audit of industry before developing pedagogically-influenced Massive Open Online Courses (MOOCs) aimed at pupils and teachers at the GCSE and A-Level stage of education in the United Kingdom.

The interactive courses, which will be made available via the FutureLearn educational platform, will enable pupils to see first-hand how data can be used to solve problems in a wide range of application areas such as crime, health and business whilst at the same time introducing new data forms such as 'Big Data' and social media.

The resources will be designed in a way such that they can be used by teachers for whole-class sessions or taken by interested pupils in their own time to develop their skills and interests and enhance university applications.

All resources will be fully supported and endorsed by Q-Step, a national initiative that seeks to get more students using and embracing data.

Whilst this research is about enhancing data and quantitative skills, it is not directed by probability and descriptive statistics – topics which have been part of quantitative curricula for many years. It is about using engaging examples and interactive forms of teaching to introduce twenty-first century data skills to young people at a time when their academic interests are first developing.

This presentation will provide an update on this twelve-month project, ending December 2017.

Keywords: MOOCs, data science, quantitative, employability, education.

1 INTRODUCTION

This ongoing research explores the changing dynamic of supply and demand between universities and industry with regards to data skills. With 90 per cent of the world's data generated in the past two years (Wall, 2014 [1]), contemporary graduates are exiting university with a need for new skills. This research, culminating in December 2017, explores this shift in demand and seeks to devise innovative ways to nurture these data skills in pre-university curricula using learning methods commensurate with education stage.

The research is broadly structured around the following four research questions:

>> What 'data' skills are regarded as core and desirable in twenty-first century graduate employment?

>> How have these required/desired skills changed over the past ten years?

>> What proportion of graduates are exiting university with these skills?

>> How can these skills be embedded earlier in the education life style?

The end goal of the research is two-fold; firstly, to fully understand the current relationship between university teaching and industry demand. Secondly, to develop a series of MOOCs designed to enthuse young professionals about data science topics whilst at the same time providing an introductory education to the new forms of data and skills needed to work in this sector.

2   RATIONALE

We are surrounded by more data than ever before, whether these be more traditional datasets such as school performance statistics or store revenue estimations or newer forms of what we now term 'Big Data'. These new and exciting datasets include social media interactions like Facebook check-ins or geolocated tweets and data generated from a wide range of daily activities such as loyalty card swipes, mobile phone usage, credit card spends and internet searches. Such datasets are now used as a matter of course by leading industries and the public sector to make business decisions (Yin and Kaynak, 2015 [2]). There is no doubt that we live in a data rich society, so much so that more data about the population (behaviour, habits, interests, demographics, movement etc) has been generated, collected and stored in the past two years than all previous years in history added together (Wall, 2014 [1]). As educators, we need to be aware of this and both enthuse and nurture the next generation of the workforce where such data and analytical skills will only grow in importance. The 'data revolution' is highly apparent and with that comes a need for people with the appropriate skills to help turn this mass of data into meaningful action.

Data skills are not solely important when developing the next cohort of computer programmers, accountants or statisticians, these quantitative skills are necessary across the full breadth of industry. With such rich data at our finger tips both the public and private sectors are looking to skilled graduates to help them better understand this data and make informed decisions. The fact that "*Businesses are drowning in data but starving for insights*" (Watkins, 2016 [3]) and that we are deemed *"Data rich and knowledge poor"* (Deloitte, 2017 [4]) is a reflection on these changing times – graduates and employers need new skills and it is the role of education providers to ensure this demand is met.

In the UK, the average adult is weak when it comes to teaching and/or learning quantitative skills. Many people are open in their lack of confidence when working with numbers whereas others vastly overstate their abilities, often deliberately as a means of defence. This is somewhat highlighted by the market research company Ipsos MORI's bi-annual opinion poll of MP's. In winter 2011, the Royal Statistics Society approached Ipsos MORI and asked if they could include a question into this poll to test the basic quantitative skills of UK MP's. The question posed was a simple one: *What is the chance of getting two heads if you toss a coin twice?* MP's were given five multiple-choice responses to select from (meaning that even a random selection would lead to a 20% chance of getting the answer right) and only 40% responded correctly (Ahuja, 2015 [5]). The fact that only two-in-five MP's could conduct a simple probability is alarming in itself but even more so when close to 80% of the same group of MP's declared confidence in their numeracy skills. With half of these people being on the then government list and all very well-educated, not to mention making big decisions on risk and return on a daily basis, our approach to quantitative skills education ought to evaluated.

Universities should be held to account when it comes to producing what are often termed 'industry standard' graduates – these are graduates with strong quantitative skills but also the right quantitative and data skills needed for modern-day industry. Part of this battle is overcoming statistical anxiety and a fear of numbers as partly evidenced above and the other part is ensuring graduates have the toolkits to be able to handle some of the newer datasets that more traditional quantitative techniques can't address, even if these are still readily taught with little consideration given towards their industry value. There is a need to change tact, modernise curricula and reconsider approaches to teaching. Research suggests that students first develop their core interests (and rarely sway too far from these) during their early education years and so this process of engagement with data must start much earlier than in the university lecture theatre to achieve the best results (Schiefele, 2001 [6]).

3   RESULTS

Further details on this project can be found on the Leeds Institute for Teaching Excellence website: http://teachingexcellence.leeds.ac.uk/opinion-higher-education-and-the-data-revolution/ (see Burns, 2017 [7]). The output will be able available in a forthcoming publication and available results presented at the Edulearn Conference, Barcelona 2017.

ACKNOWLEDGEMENTS

REFERENCES

[1]   M. Wall, Big Data: Are you ready for blast-off?, 4 March, 2014, Retrieved from:
        http://www.bbc.co.uk/news/business-26383058

[2]   S. Yin and O. Kaynak, "Big Data for Modern Industry: Challenges and Trends [Point of View]," in
        Proceedings of the IEEE, vol. 103, no. 2, pp. 143-146, 2015.
        doi: 10.1109/JPROC.2015.2388958

[3]   C. Watkins, Hottest Jobs in 2016 #3 Data Scientist, 20 January, 2016, Retrieved from:
        http://blog.udacity.com/2016/01/hottest-jobs-in-2016-3-data-scientist.html

[4]   Deloitte, Information rich, knowledge poor, 5 April 2017, Retrieved from:
        https://www2.deloitte.com/us/en/pages/financial-services/articles/information-rich-knowledge-
        poor-overcoming-insurance-data-conundruml.html

[5]   A. Ahuja, UK's collective innumeracy adds up to a big problem, 19 April, 2015, Retrieved from:
        https://www.ft.com/content/feca8d32-e519-11e4-bb4b-00144feab7de

[6]   U. Schiefele, "The role of interest in motivation and learning". In J.M. Collis & S. Messick (Eds.),
        Intelligence and Personality: Bridging the Gap in Theory and Measurement. Mahwah, NJ:
        Lawrence Erlbaum Associates, pp. 163–194, 2001.

[7]   L. Burns, Opinion: Higher education and the data revolution, 15 March 2017, Retrieved from:
        http://teachingexcellence.leeds.ac.uk/opinion-higher-education-and-the-data-revolution/