DECISION SCIENCES INSTITUTE

Should Business Analytics Be Taught to Every Undergraduate? – We are Already Doing It

Zsolt Ugray
Utah State University
Email: Zsolt.Ugray@usu.edu

ABSTRACT

One prediction of a 2011 McKinsey Global Institute study on Big Data is a shortage of "… 1.5 million managers and analysts with the know-how to use the analysis of big data to make effective decisions." While many specialized graduate and undergraduate programs training data analysts with deep analytical skills have been started recently, there is an absence of initiatives to train the projected 1.5 million managers who can apply results of analytics towards effective decision making. I report on our Business School's initiative to require an analytics acumen course for all undergraduate business students to close this gap.

KEYWORDS: Business Analytics, Business Curriculum Development, Course Design, Analytics Instruction

INTRODUCTION

Business Analytics

During the past several years, numerous programs have been created to prepare future employees and entrepreneurs with the skills necessary to utilize business analytics and manage big data. While the number and specifics of the topics and courses included in such programs is not uniform and there is no clear consensus about the content of related curricula, there are clearly observable patterns about what constitutes the knowledge necessary for practitioners of business analytics.

The need for handling and analyzing volumes of data is well supported. The amount of data created, managed and consumed is expected to reach 4 ZB (zettabytes) by 2020 (Gantz & Reinsel, 2012). Handling this vast amount of data with available technology is stretching organizations' abilities (McAfee & Brynjolfsson, 2012). On the one hand, there is continuous pressure to innovate and sustain competitive advantage (Barton & Court, 2012; Bell, 2013). On the other hand, for some organizations, there is the need to stay competitive with early adopters of related technologies (Davenport & Harris, 2010). The need to find, keep, and train business analytics experts, or data scientists, has become an unceasing challenge (Davenport, 2013; Davenport & Patil, 2012; Santaferraro, 2013). It is predicted that by 2018 in the US there will be a shortage of between 140,000 and 190,000 people with deep data science experience and 1.5 million managers with data analytical skills able to convert analytical insights into practice (Manyika et al, 2011).

Business Analytics Education

Academia and industry are focused on addressing the forecasted shortage of 140,000 to 190,000 people with deep data science experience Even with the relatively fast acknowledgement of this gap, the emergence of programs and students graduating do not keep pace with the growing demand for skilled data scientists: in 2010 only 15 full time programs (only 3 them at the undergraduate level) were identified in a review (Wixom et al, 2011). A few

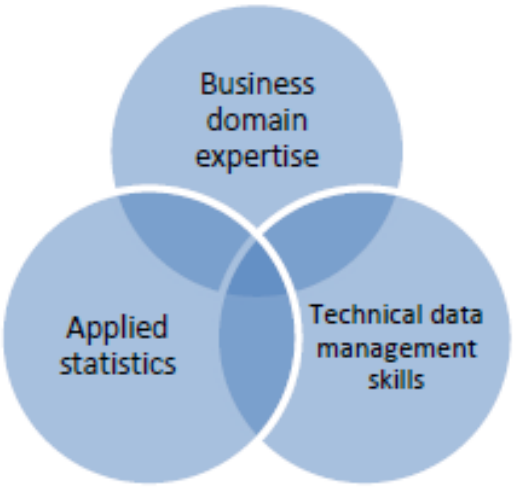
years later the number of programs has dramatically increased to 131 full time programs, including 47 at the undergraduate level (Wixom et al, 2014). The increase in interest and demand for business analytics, however, still outpaces the supply of students. Additionally, this recent survey found that prospective employers found university graduates' practical experience and foundational skills lacking.

Very little attention has been paid to the other forecasted shortage: that of 1.5 million managers who will need skills to interpret the results of analytics and convert insights into actions (Manyika et al, 2011). Existing curricular recommendations barely address the general education of business students towards the basic understanding and appreciation of big data and business analytics concepts. Based on the prevalence of large volumes of data in so many fields and areas of business, as well as the recognition of the importance of analytics to effectively compete in those areas (Davenport & Harris, 2010), organizations need people who are conversant in the language of analytics even if they won't become experienced practitioners of data analytics themselves. They ought to understand the basic concepts and capabilities of analytics. They should also be able to identify business problems that can be addressed with analytical tools and – at the other end of the process – they should be able to understand and interpret the results and make actionable recommendations based on them.
My business school was swayed by the argument that it essential to prepare all business students with skills to become data literate and decided to implement a big data analytics acumen course that is part of the required general preparation for all business undergraduate students. Steps of the development of the course, its initial roll out, and first experiences are described below.

COMPETENCY-BASED COURSE DEVELOPMENT

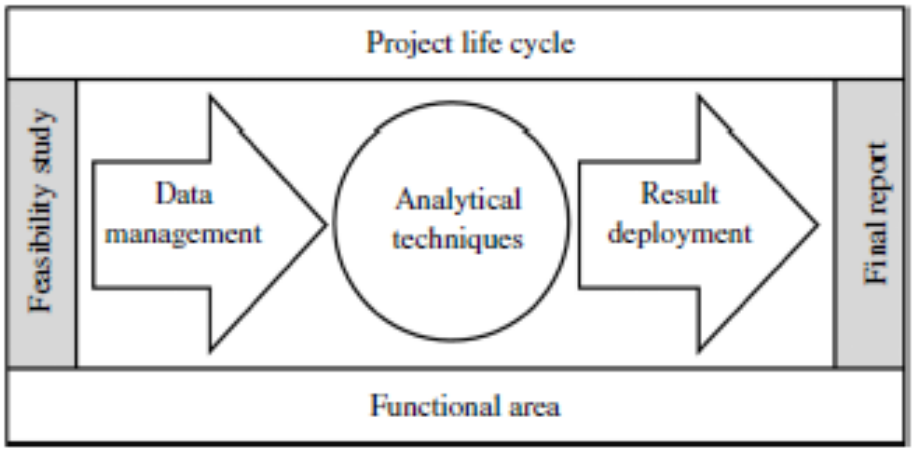
To determine the content for a course that is required by all business undergraduates, a course steering committee was set up. Its goal was to create a course at the sophomore/junior level that is comprehensive, addresses anticipated business needs towards a high level understanding of big data analytics, and suitable for a broad audience. Following (Bowden, 2004), the major steps in competency based development is 1) identifying a common set of skills which are in demand in practice and 2) developing an academic curriculum that empowers students with the corresponding skillset.
The committee relied on several sources of information regarding what is considered essential skills by employers, practitioners, and academic educators (Wilder & Ozgur, 2015; Wixom et al, 2011; Wixom et al, 2014). NCSU (North Carolina State, 2014) maintains a detailed list of graduate programs in analytics. For another good, but less comprehensive list of university program reviews, see the references in (Mamonov et al, 2015).
An often cited understanding of the components of business analytics skillset, based upon (Conway, 2013), is illustrated in Figure 1. Since the target audience is business undergraduates, the need for business domain expertise is only addressed in specific examples. Coverage of applied statistics are dispersed throughout the course and are discussed as necessary for the comprehension of the actual concept, relying on concepts previously covered in the prerequisite Business Statistics course. Technical data management skills are demonstrated and taught throughout the course. Some of the skills are referred to as they are taught in another prerequisite course, Principles of MIS. The limited content that is possible for a single course forced the steering committee to make compromises. As a result, several skills embedded in the three skillset categories are covered only briefly and there is a reliance on synergies with other required business acumen courses.

Figure 1: Business analytics skillset



A different perspective of the topics belonging to the business analytics knowledge base is summarized in Wilder & Ozgur (2015). They are discussed as they relate to proposed courses in an undergraduate business analytics curriculum. The proposal is comprised of 5 knowledge domains that include: 1) Project life cycle; 2) Data Management; 3) Analytical techniques; 4) Result deployment and 5) Functional areas (Figure 2.) Components of many of these domains might appear in multiple courses that are more specialized to analytics topics. The list of recommended specialized courses are the following: Data Management; Descriptive Analytics; Data Visualization; Predictive Analytics; Prescriptive Analytics and Data Mining, augmented with an Analytics Practicum and the options for Electives.

Figure 2: Knowledge domains



An inescapable decision in business analytics courses is the selection of software tools. While industry and practitioner surveys (Information Week, 2014; KDNuggets, 2015) give a snapshot of the popularity of the available tools they are only one of several factors that play a role in the final decision. There are multiple tools available for all of the tasks and elements of analytics. Data management and retrieval tools, such as SQL and Microsoft Excel / Access are widely

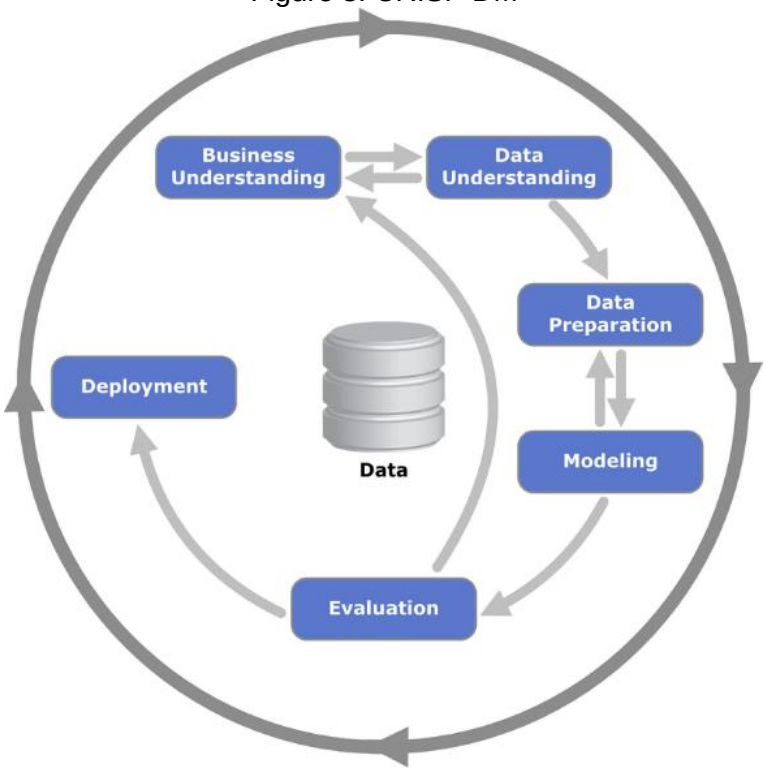
available common tools. For data visualization two obvious choices are Tableau (Tableau, 2016) and Power BI by Microsoft (PowerBI,2016). Several of the most widely used Data Mining software platforms, R and Python, require extensive coding (Python, 2016; R, 2016). Others, like KNIME, RapidMiner, and SAS Enterprise Miner allow the user to follow a process oriented view of performing analytics (KNIME, 2016; RapidMiner, 2016; SAS Enterprise Miner, 2016) which has the advantage of a less steep learning curve, making the process of doing analytical work more available to an audience without coding experience.

COURSE CONTENT

Topics

The guiding principles for selecting topics to include into the course content were based on two ideas: 1) to demonstrate the whole process of data analytics 2) to keep the topic coverage relatively simple but rigorous. Steps of the Cross Industry Standard Process for Data Mining (CRSP-DM, 2016; Figure 3.) are often referred in discussion and examples to illustrate not only the individual tasks but also the interrelatedness of tasks and their reliance on each other.

Figure 3. CRISP DM



Accordingly, the following topics are covered in the semester-long course: 1) Data acquisition; 2) Data Cleansing; 3) Data Summarization, 4) Data Visualization; 5) Data Mining. The last topic, Data Mining is covered in the most detail. Besides a general Overview, Prediction, Classification, Validation, Unsupervised, and Supervised Learning are covered with several data mining techniques, such as decision trees, logistic regressions, neural networks, cluster analysis, and association analysis.

Software Tools

The variety of topics require that students gain exposure to several software tools. SQL, Microsoft Excel and Power BI are used in the Data Acquisition module. They were chosen partly because of the prevalence of these tools in practice and partly because students already had exposure to them in an earlier prerequisite course with the exception of Power BI (PowerBI, 2016), which is new to them. Power BI is also used to demonstrate Data Cleansing, Data Summarization and Data Visualization. All of the data mining techniques are introduced in concept at first and then RapidMiner (RapidMiner, 2016) is used to put the concepts to work and analyze relevant data. RapidMiner was chosen for its process oriented view of analytics: it is relatively easy to learn and if one does not want to or need to get into the detailed workings of the algorithms it is a very simple tool to use.
Across the introduction of all of the techniques, every details of the steps of the CRISP-DM framework is emphasized and demonstrated, emphasizing the importance of not only the technical details, but the steps of Business understanding evaluation and Deployment as well.

COURSE DEPLOYMENT

The course was first piloted during the fall semester of 2015 with two instructors and a total of 70 students in 3 sections taught in a lab environment. It was followed by 5 sections taught by 3 instructors in the spring of 2016. Detailed review of student responses and course evaluation are in progress but initial responses from students ranged from great excitement to quiet reservation. It is a course that is definitely stretching the skills and experiences of the typical business undergraduate student. Despite the demands of the course, evaluations were similar to other required business acumen courses and students responded with the typical spread of opinions about how they perceive the usefulness and relevance of topics.
Between the two semesters several minor changes were made to the course that did not affect its original design. Some topics got more time and attention while others were either deemphasized or omitted. The software tools used were upgraded to their latest version with the greatest change of replacing Microsoft's PowerView, PowerQuery and PowerPivot with Power BI which integrates the capabilities of component previously spread across multiple software pieces.
More detailed review of the course, based on instructor evaluations, student course evaluations, and surveys conducted are performed between the semesters of two academic years. At this point no major changes are foreseen.

CONCLUSIONS

With the development and deployment of a new Big Data Analytics course required to be taken by all business undergraduate students our Business School is committed to equipping students with skills that are forecasted to be in great demand for future managers. In contrast to most data analytics program that concentrate on a sequence of several course to prepare students with deep analytical and data science skills, the course presented in this paper is aimed to provide the basic knowledge necessary for understanding the essence of analytics, the way it is used, and the type of results and solutions it provides. Only with these skills will future managers be able to interpret results and put them into practice for improved decision making in organizations.

REFERENCES

Barton, D. & Court, D. (2012). Making Analytics Work for You. *Harvard Business Review*, October, 79-83.

Bowden, J. (2004). Capabilities-driven curriculum design. *Effective learning and teaching in engineering*, 36–47.

Bell, P. (2013). Creating Competitive Advantage Using Big Data. *Ivey Business Journal*, May, 1-5.

Conway, D. (2013). THE DATA SCIENCE VENN DIAGRAM. Retrieved from http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram

CRISP-DM, (2016). Cross Industry Standard Process for Data Mining (CRISP-DM) https://en.wikipedia.org/wiki/Cross_Industry_Standard_Process_for_Data_Mining

Davenport, T. H. (2013). Keep Up with Your Quants. *Harvard Business Review*, July-August, 120-123.

Davenport, T. H. & Harris, J. G. (2010). *Analytics at Work: Smarter Decisions, Better Results*. Harvard Business School Press.

Davenport, T. H. & Patil, D. J. (2012). Data Scientist: The Sexiest Job of the 21st Century. *Harvard Business Review*, October, 70-76.

Gantz, J. & Reinsel, D. (2012). The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East. *IDC Analyze the Future*, 3, 1-16.

Information Week (2014). 16 Top Big Data Analytics Platforms, Retrieved from: http://www.informationweek.com/big-data/big-data-analytics/16-top-big-data-analytics-platforms/d/d-id/1113609

KDNuggets, (2015). R leads RapidMiner, Python catches up, Big Data tools grow, Spark ignites; Retrieved from: http://www.kdnuggets.com/2015/05/poll-r-rapidminer-python-big-data-spark.html

KNIME (2016). Retrieved from: www.knime.org

Mamonov, S., Misra, R. & Jain, R. (2015). Business Analytics in Practice and in Education: A Competency-based Perspective, *Information Systems Education Journal, 13 (1)* 4-13.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung-Byers, A. (2011). *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute, 1-3.

McAfee, A. & Brynjolfsson, E. (2012). Big Data: The Management Revolution. *Harvard Business Review*, October, 61-68.

North Carolina State University (2014). Survey of Graduate Degree Programs in Analytics. Retrieved from: http://analytics.ncsu.edu/?page_idD4184.

PowerBI (2016). Retrieved from: powerbi.microsoft.com

Python (2016). Retrieved from: https://www.python.org/

R (2016): R: The R Project for Statistical Computing. Retrieved from: https://www.r-project.org/

RapidMiner (2016). Retrieved from: www.rapidminer.com

SAS Enterprise Miner (2016). Retrieved from: www.sas.com/en_id/software/analytics/enterprise-miner.html

Santaferraro, J. (2013). Filling the Demand for Data Scientists: A Five-Point Plan, *Business Intelligence Journal, 1), (4*), 13-18.

Tableau (2016). Retrieved from: www.tebleau.com

Wilder, C. R. & Ozgur, C. O. (2015). Business Analytics Curriculum for Undergraduate Majors. *INFORMS Transactions on Education, 15, (2)*, 180-187.

Wixom, B., Ariyachandra, T., Goul, M., Gray, P., Kulkarni, U. & Phillips-Wren, G. (2011). "The Current State of Business Intelligence in Academia: Results from the BI Congress," *Communications of the Association for information Systems, 29,* 300-312.

Wixom, B., Ariyachandra, T., Douglas, D., Goul, M., Gupta, B., Lyer, L., Kulkarni, U., Mooney, J. G., Phillips-Wren, G., & Turetken, O. (2014). "The Current State of Business Intelligence in Academia: The Arrival of Big Data," *Communications of the Association for information Systems, 34,* 1-13.