Aug 10th, 12:00 AM

Aligning Business Analytics Programs with Industry Required Knowledge, Skills and Abilities

Kathleen S. Hartzel
Duquesne University, hartzel@duq.edu

Pinar Ozturk
Duquesne University, ozturkp@duq.edu

K. Bryan Menk
Duquesne University, menkk@duq.edu

Aligning Business Analytics Programs with Industry Required Knowledge, Skills and Abilities

Completed Research

Kathleen S. Hartzel
Duquesne University
hartzel@duq.edu

Pinar Ozturk
Duquesne University
ozturkp@duq.edu

K. Bryan Menk
Duquesne University
menkk@duq.edu

Abstract

This paper describes results from a topic modeling analysis of online data analytics-related job advertisements. Five distinct clusters emerged, each with a focus on different knowledge, skills and abilities (KSAs) profiles. We labelled these clusters big data, systems analysis, business, healthcare, and technical research. Identification of these clusters provides a framework that can be used by information systems and business analytics faculty to offer customized and specialized information systems and business analytics programs that prepare graduates to fill specific roles in the data ecosystem of the workplace.

Keywords

Business analytics, data analytics, curriculum design, skill requirements, KSA clusters.

Introduction

The job market for individuals with business analytics and data analytics skills is booming (Centre for the New Economy and Society, 2018). Individuals in roles such as business analyst, data analyst, financial analyst, marketing analyst, and business intelligence (BI) analyst use business, analytical and technical knowledge, skills, and abilities (KSAs) to analyze organizational performance and to inform decision makers and business strategy. Many information systems and technology (IS&T) faculty are responding to this trend by modifying existing IS curriculum to include business analytics techniques by adding new business analytics majors or minors and creating business analytics concentrations.

As early as 2014, discussions, lists and frameworks describing the KSAs required for BI[1] and data analytics-related positions started to surface in IS academic conferences and literature (Aasheim et al., 2014). These KSAs included a large number of specific tools and technologies. Despite these extant laundry lists of numerous KSAs being sorted into major categories, such as programming languages and communication skills, little insight exists into how the required KSAs are correlated with one another. Clusters of tightly associated KSAs could be explicitly bundled to describe the desired KSA profiles sought in potential new recruits. Analyzing industry and government job advertisements provides the opportunity to identify high value BI analytics-related KSA clusters. These KSA profiles may lead to a better understanding of when and where specific skill sets are needed.

[1] Business Intelligence is an umbrella term that includes data analysis, business analysis, data warehousing, data modeling and data mining. (White, 2019)

The exploratory research described here seeks to detect the inherent bundles (profiles) of KSAs described within job advertisement found on the commercial online job site Indeed (www.indeed.com). Using this data, opportunities may surface for IS&T faculty to better prepare the workforce to meet the evolving needs of industry.

The research questions addressed by this work are: what are the identifiable combinations of highly correlated BI and data-related analytics KSAs (KSA profiles) found in job advertisements and how could these be used to inform hiring practices, employee development, and curriculum design?

Literature Review

The literature contains a number of KSA frameworks for business analytics and data analytics, data science, and BI jobs. As described in Ozturk and Hartzel (2020), these articles originated from three different types of sources: 1) faculty opinion (e.g., Chiang et al., 2012; Johnson et al., 2020; Wymbs, 2016), 2) expert opinion (e.g., Johnson et al., 2020; Aasheim et al., 2015; Cegielski & Jones-Farmer, 2016), and 3) content analysis of on-line job description sites such as Monster.com, LinkedIn, Indeed, Career-Builder, Glassdoor and Dice.com (e.g., Luo, 2016a; Gardiner et al., 2018; Stanton & Stanton, 2020). Ozturk and Hartzel (2020) found that the KSAs were often similar across articles and were most commonly grouped into three major categories: 1) business, 2) analytics, and 3) technical (Aasheim et al., 2015; Chang et al., 2019; Chiang et al., 2012; Luo, 2016a; Radovilsky et al., 2018; Stanton & Stanton, 2020).

Although the themes were consistent across frameworks, many of the scholars differed in how they organized KSA categories. For example, Wilder and Ozgur (2015) expanded these three KSAs groups into five and suggested that new undergraduate data analytics programs should include: 1) data management, 2) analytical techniques, 3) results deployment, 4) project life cycles, and 5) a functional area. Business context should be a central component in the curriculum of analytics programs housed in business schools (Mortenson et al., 2020; Urbaczewski & Keeling, 2019). Therefore, including practitioners in program design and requiring real-world projects in the curriculum could be critical success factors (Chiang et al., 2012; Wymbs, 2016). Hefley, Parker, and Chatterjee (2019) took a step toward bundling KSAs when they analyzed internship experiences of graduate students in their business analytics program. Data analysis, data cleaning and visualization were the top activities reported by 85%, 75% and 75% of the students respectively. These results led the authors to conclude that most business analytics programs should focus on those areas.

A number of studies used online job advertisements to identify highly sought-after knowledge and skill sets. At aggregate levels, online job advertisements could be valuable indicators and reveal shifting labor demands. For example, Debortolli, Müller and vom Brocke's (2014) content analysis of 1,807 Monster.com job advertisements led to taxonomies that revealed that business skills were as important as technical skills in both BI and big data jobs. Luo's (Luo, 2016b) analysis of 924 data analyst job listings from LinkedIn, Indeed and Monster.com was used to identify required job responsibilities and skills. This study found that when seeking technical skills associated with data management and analysis, the demand for soft skills decreased. Conversely, when seeking candidates for a position of functional responsibility, the need for technical skills declined. Similarly, Cegielski and Jones-Farmer (2016) pulled 186 job ads from LinkedIn, Career-Builder and Monster.com. Their KSA analysis resulted in three major categories 1) business, 2) analytical and 3) technical skills. From Indeed, Gardiner et al. (2018) examined 1,216 job advertisements containing 'big data' in the job title. They identified and clustered 218 skill-related terms into 24 general KSAs. Business and soft skills, such as communication, domain knowledge, leadership, personal skill attributes, and team, prominently resided among the technical skill categories.

Business acumen as a necessary data analyst skill was also highlighted in the analysis of 2,786 Dice.com job listings (De Mauro et al., 2018). Johnson, Albizri and Jain (2020) scraped 5,257 business analyst entry-level Indeed listings. Their results confirmed 1) the need for graduate-level programs, 2) SQL, Python and R were key tools, 3) graduates should have some knowledge of big data platforms, and 4) both analytical and soft skills were required. Finally, Persaud (2020) used text mining on 3,009 job postings for analytics positions from Indeed, LinkedIn, Monster.com, and Procom. He categorized the resulting KSAs into 1) data analytics 2) computing, 3) business and 4) soft skills.

The existing literature does not describe which KSAs are more closely associated with others in and across business, analytic, and technical categories. The current research suggests specific KSA profiles based on the clustering of KSAs most commonly associated with each other.

Methodology

In this study, using a Python-based web crawler, we retrieved online job listings from Indeed.com during January 2021. The parameters for inclusion of a job listing in our data set are 1) it contains business analytics/analyst keywords in the job description/title and 2) the job location is within 25 miles of Pittsburgh Pennsylvania.

Based on our observations, the job listings on Indeed.com were not all new postings. Multiple listings were identical in content and were re-posted as newly added listings. Therefore, after removing the duplicates and irrelevant job listings, our dataset contains 1,500 job listings that include the following information for each position: job title, job description, location, company name and industry.

In order to prepare our dataset for cluster analysis, we removed all punctuation from the job descriptions and job titles. Then, we tokenized the description and title text into words and stemmed the words. We also removed the commonly used stop words (such as 'a', 'the', 'is') which are unlikely to have any benefit in natural language processing.

Clustering a set of documents is a standard problem addressed in machine learning and natural language processing. Clustering allows us to automatically organize many documents into a small number of groupings (clusters) and possibly find the underlying structure in an unlabeled document collection. K-means is one of the most popular unsupervised machine learning algorithms used for clustering problems. In order to cluster a set of documents with K-means, each document is first quantified as a numeric vector where each component represents a corresponding feature in the document. Then, a distance is used to measure the difference between two documents.

In this work, we first quantified each document using Latent Dirichlet Allocation (LDA). In LDA, each document is considered as a mixture of different topics and each topic is characterized by a probability distribution over a finite vocabulary of words. We then used the probability distribution of topics as the input for the K-means clustering algorithm (Bui et al., 2015, 2017). For each cluster identified in K-means, we gathered the most frequently mentioned topics of that cluster and created a word cloud showing the most frequently identified topics. The font size of each topic is proportional to the observation of that topic in that cluster.

Results

The five KSA clusters resulting from the topic modeling reveal some interesting patterns.

Big Data

Cluster 1, shown in Figure 1, has a strong technical focus. The topics in this cluster seem to have a greater focus on data analytics – especially big data analytics, given the tools present in the job descriptions, such as aws, Teradata and Hadoop. Thus, we refer to this KSA cluster as *big data*.

The big data cluster includes 16.5% of the job listings and the most common job titles include "Data Scientist", "Data Engineer", and "Big Data Analyst". The technical nature of this cluster suggests that, in the business school, the IS&T faculty may be the best prepared to deliver this type of courses. This may require that traditional IS&T majors be revamped or new majors focusing on data and analytics be developed.

Systems Analysis

Cluster 2, which also has a technical focus, seems for be the most closely aligned with the traditional systems analyst position. The call for big data technologies is not as salient in this cluster as it is in cluster 1. Furthermore, the topics shown in Figure 2 highlight terms such as management, requirements, process, project and product. This shifts the emphasis more toward of a business focus. We named this KSA cluster, *systems analysis*.

The systems analysis cluster includes 50.5% of the job listings and the most common job titles include "Business Analyst", "Financial Analyst", and "Business Intelligence Analyst". Again, the IS&T faculty are most likely well-qualified to prepare students for employment in these types of positions. Traditional IS&T KSAs must still be part of the curriculum and understanding the business domain is also highly desirable. Data analytics content could be added to the program by embedding it in the existing major or alternatively a new major (minor) could be developed.

Figure 1: Big Data

Figure 2: Systems Analysis

Business

Cluster 3, shown in Figure 3, we call *business*. The only tool that emerged in this KSA cluster is Excel. Given the popularity of Microsoft Excel and its applications, Excel is primarily taught in core business courses and used by all business majors. The topics in this cluster are general business concepts and include words, such as credit, mortgage, account, insurance, personnel, and portfolios.

The business cluster includes 22% of the job listings and the most common job titles are "Business Manager", "Sales Executive" and "Portfolio Manager". The absence of any tools, other than Excel in the topic modeling, indicates that these positions do not demand deep technical expertise. Rather the business domain KSAs are the drivers of job candidates acquiring successful employment.

Healthcare

Cluster 4, shown in Figure 4, is *healthcare*. None of the common data analysis tools emerged in the most frequent topics. Manually spot checking the associated job titles shows that the employers are seeking medical personnel, not business people. Four percent of the job listings in our data are under this cluster. This cluster suggests that the curriculum in health sciences should include data analytics skills. Specialized tools found in medical fields could be emphasized.

Technical Research

We labeled the final cluster *technical* research, since majority (67%) of the job listings in this cluster are from institutes of higher education. Both academic research and market research skills tend to dominate this cluster. Technical skills, such as coding, are also prevalent. This cluster also has a strong emphasize

on statistical tools which are sparsely referenced in the other KSA clusters. Overall, 7% of the job listings in our data were under this cluster.

Marketing research could be addressed in undergraduate business programs by marketing faculty, IS&T faculty or some combination both faculties. However, much of the statistical, coding and data skills might be better served by a graduate-level degree in the area of analytics and information management than an undergraduate offering.

Figure 3: Business

Figure 4: Healthcare

Figure 5: Technical Research

Conclusion

In the information era, enormous amounts of data are available to decision makers. This data, coupled with new methods and technologies, requires the development of new interdisciplinary competencies spanning from business knowledge and communication skills to technical and analytical skills. Preparing the future generation of data analytics professionals to meet the evolving needs of industry poses a significant challenge for IS&T educators.

This data-driven analysis informs IS&T curriculum design by presenting five distinct data analytics KSA clusters. These specific clusters were found in Indeed listings for business analytics-related positions within 25 miles of Pittsburgh, PA.

One of the KSA clusters is specific to healthcare professionals and outside the purview of the business school.

However, the other four clusters each merit their own business school curricular options. The *business* cluster emphasizes the need for all business people to understand, value and analyze data. This encourages faculty in all functional areas to embed domain specific analytic content in their curriculum. The *big data* and *systems analysis* clusters suggests that IS&T curriculum should not ignore the emerging trend of analytics as a service and must introduce more analytics focus into their programs. The *technical research* cluster highlights the importance of developing deeper, higher-level KSAs to advance learning and business strategy. Graduate-level programs may best fill this niche.

This paper makes a number of contributions to the practice of data analytics. First, our results could support educators and curriculum design in multiple ways. Educators could develop new curricula or improve existing curricula by focusing on KSA profiles that closely match current job demands. In addition, understanding different groups of skills associated with data analytics jobs could help employers to develop job requirements by leveraging the sets of skills identified in this study.

This research could be extended in several ways in the future. The current study is exploratory in nature as it is based on 1,500 job listings collected during January 2021. Further studies may require longitudinal data to verify the job trends. The KSAs found in the job listings could also be analyzed in conjunction with other job characteristics such as required education level, years of experience and position level. Additional company related information, when available, could also be analyzed to see how different job markets might vary.

REFERENCES

Aasheim, C. L., Williams, S., Rutner, P., & Gardiner, A. (2015). Data analytics vs. data science: A study of similarities and differences in undergraduate programs based on course descriptions. *Journal of Information Systems Education*, *26*(2), 103.

Aasheim, C., Williams, S., Rutner, P., & Gardiner, A. (2014). Big Data Analytics and Data Science Undergraduate Degree Programs. *Proceedings of the Decision Sciences Institute Annual Meeting*, 338–359.

Bui, Q. V., Sayadi, K., Amor, S. Ben, & Bui, M. (2017). Combining Latent Dirichlet Allocation and K-means for documents clustering: effect of probabilistic based distance measures. *Asian Conference on Intelligent Information and Database Systems*, 248–257.

Bui, Q. V., Sayadi, K., & Bui, M. (2015). A multi-criteria document clustering method based on topic modeling and pseudoclosure function. *Proceedings of the Sixth International Symposium on Information and Communication Technology*, 38–45.

Cegielski, C. G., & Jones-Farmer, L. A. (2016). Knowledge, skills, and abilities for entry-level business analytics positions: A multi-method study. *Decision Sciences Journal of Innovative Education*, *14*(1), 91–118.

Centre for the New Economy and Society. (2018). *The Future of Jobs Report 2018*. World Economic Forum. https://www3.weforum.org/docs/WEF_Future_of_Jobs_2018.pdf

Chang, H. C., Wang, C. Y., & Hawamdeh, S. (2019). Emerging trends in data analytics and knowledge management job market: extending KSA framework. *Journal of Knowledge Management*, *23*(4),

664–686. https://doi.org/10.1108/JKM-02-2018-0088
Chiang, R. H. L., Goes, P., & Stohr, E. A. (2012). Business intelligence and analytics education, and program development: A unique opportunity for the information systems discipline. *ACM Transactions on Management Information Systems (TMIS)*, *3*(3), 1–13.
De Mauro, A., Greco, M., Grimaldi, M., & Ritala, P. (2018). Human resources for Big Data professions: A systematic classification of job roles and required skill sets. *Information Processing & Management*, *54*(5), 807–817.
Debortoli, S., Müller, O., & vom Brocke, J. (2014). Comparing business intelligence and big data skills. *Business & Information Systems Engineering*, *6*(5), 289–300.
Gardiner, A., Aasheim, C., Rutner, P., & Williams, S. (2018). Skill requirements in big data: A content analysis of job advertisements. *Journal of Computer Information Systems*, *58*(4), 374–384.
Hefley, B., Parker, J., & Chatterjee, S. (2019). Entry-Level Data Science Work Practices and Environments. *Proceedings of the International Conference on Frontiers in Education: Computer Science and Computer Engineering (FECS)*, 134–139.
Johnson, M. E., Albizri, A., & Jain, R. (2020). Exploratory analysis to identify concepts, skills, knowledge, and tools to educate business analytics practitioners. *Decision Sciences Journal of Innovative Education*, *18*(1), 90–118.
Luo, W. (2016a). Responsibility and skills requirements for entry level analytics professionals. *Journal of Organizational and End User Computing*, *28*(4), 1–14. https://doi.org/10.4018/JOEUC.2016100101
Luo, W. (2016b). Responsibility and skills requirements for entry level analytics professionals. *Journal of Organizational and End User Computing (JOEUC)*, *28*(4), 1–14.
Mortenson, M. J., Doherty, N. F., & Robinson, S. (2020). Creating a typology of analytics Master's degrees in UK universities: Implications for employers and educators. *Journal of the Operational Research Society*, *71*(9), 1327–1346.
Ozturk, P., & Hartzel, K. S. (2020). Business Analytics: Addressing the Real Skill Requirements of Employers. *Proceedings of the EDSIG Conference*, *2473*, 4901.
Persaud, A. (2020). Key competencies for big data analytics professions: a multimethod study. *Information Technology & People*.
Radovilsky, Z., Hegde, V., Acharya, A., & Uma, U. (2018). Skills requirements of business data analytics and data science jobs: A comparative analysis. *Journal of Supply Chain and Operations Management*, *16*(1), 82–101.
Stanton, W. W., & Stanton, A. D. (2020). Helping Business Students Acquire the Skills Needed for a Career in Analytics: A Comprehensive Industry Assessment of Entry-Level Requirements. *Decision Sciences Journal of Innovative Education*, *18*(1), 138–165.
Urbaczewski, A., & Keeling, K. B. (2019). The Transition from MIS Departments to Analytics Departments. *Journal of Information Systems Education*, *30*(4), 303–310.
White, S. K. (2019). *What is a business intelligence analyst? A role for driving business value with data*. CIO. https://www.cio.com/article/3387619/what-is-a-business-intelligence-analyst-a-role-for-driving-business-value-with-data.html
Wilder, C. R., & Ozgur, C. O. (2015). Business analytics curriculum for undergraduate majors. *INFORMS Transactions on Education*, *15*(2), 180–187.
Wymbs, C. (2016). Managing the innovation process: Infusing data analytics into the undergraduate business curriculum (lessons learned and next steps). *Journal of Information Systems Education*, *27*(1), 61.