

Association for Information Systems

AIS Electronic Library (AISeL)

AMCIS 2022 Proceedings

SIG ED - IS in Education, IS Curriculum,
Education and Teaching Cases

Aug 10th, 12:00 AM

What Really Matters? Evaluating the Importance of Skills for Data Analysts

Cassandra Artman Collier
Southern Illinois University Edwardsville, cassaco@siue.edu

Anne L. Powell
Southern Illinois University Edwardsville, apowell@siue.edu

Follow this and additional works at: <https://aisel.aisnet.org/amcis2022>

Recommended Citation

Collier, Cassandra Artman and Powell, Anne L., "What Really Matters? Evaluating the Importance of Skills for Data Analysts" (2022). *AMCIS 2022 Proceedings*. 20.
https://aisel.aisnet.org/amcis2022/sig_ed/sig_ed/20

This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in AMCIS 2022 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

What Really Matters? Evaluating the Importance of Skills for Data Analysts

Emergent Research Forum (ERF)

Cassandra A. Collier

Southern Illinois University Edwardsville
cassaco@siue.edu

Anne L. Powell

Southern Illinois University Edwardsville
apowell@siue.edu

Abstract

Data analysis is becoming an increasingly critical activity in today's organizations as companies collect more data and rely on data analysis to drive decision-making. To meet this need, many universities now offer data analytics programs to train entry-level data analysts. However, the field is changing rapidly, representing a challenge to academics determining which skills and software platforms to include in a data analytics curriculum. Information systems researchers have investigated this problem, but gaps remain in our knowledge. Specifically, it is unclear how important different technical skills and tools are to managers hiring entry-level data analysts and what is expected of graduates of university data analytics programs. In this emergent research forum paper, we review the existing research on data analyst skillsets and outline a plan to bring the field up-to-date by building on the extant literature with a new focus not just on frequency of skills listed in job postings but also on their importance to today's organizations. This work will contribute valuable findings to IS scholars developing data analytics curriculum, entry-level data analysts competing on the job market, and hiring managers staying up to date on the field's demands.

Keywords

IS education, data analytics, data analyst, data visualization, skillsets, career readiness, mixed methods

Introduction

As the amount of data generated in our world increases, so does the importance of data analysis in our organizations. Data analytics is posited to have a significant impact on a wide variety of industries and applications, rendering the skills of those working with data highly important (Dwivedi et al., 2021). Data analysis has become crucial for today's corporations to achieve a competitive advantage – but only when the firm has individuals in-house with data analytics skillsets (Dahiya et al., 2021). To help provide these skillsets, universities offer undergraduate and graduate programs intended to train entry-level information systems (IS) analysts to take on data-related challenges (*The Best Business Analytics Programs, Ranked*, n.d.). Quickly, “the number of university-based analytics programs... exploded” with at least 400 analytics degree programs being offered across 220 international business schools (LeClair, 2013).

However, the field is changing rapidly (Saha, 2020), leaving academics challenged to offer relevant curricula to adequately train students for the constantly-evolving demands of their data careers while students also seek training via certifications (*Which Certification Is Best for Data Analytics?*, n.d.), bootcamps (Waguespack et al., 2018), and various non-academic online learning tools (e.g. a search for “data analytics” on the popular online learning platform Coursera renders over 2,000 results; *Top Data Analytics Courses - Learn Data Analytics Online*, n.d.). The data field is inundated with dozens of software platforms and techniques for collecting, storing, cleaning, manipulating, analyzing, and visualizing data. It is unclear which would be best utilized in a university program to help students build a relevant and competitive skillset. Scholars studying this problem have reported that employers “lament that IT students do not graduate with the ‘last mile’ training and skills they need to be employable” (Setor & Joseph, 2021).

But what are those “last mile” training and skills that entry-level data analysts need to be employable – or, better yet, competitive against their peers on the job market and impressive to their potential employers? What skills are best transferred in the university environment to help entry-level data analysts hit the

ground running in their new organizations? In this research-in-progress, we describe a plan to investigate which skills are most in-demand for today's entry-level data analysts and how these are ranked in importance by individuals hiring college graduates and advising universities on curriculum development.

Literature Review

Several scholars have focused on uncovering the skillsets that will make students graduating from IS programs more employable. Many of these focus on IS students as a homogeneous group. For example, Setor and Joseph (2021) report that cooperative education, internship, and mentorship experiences increase the likelihood of initial IT employment. In their biennial report of knowledge and skills required in the IS field, Cummings and Janicki (2021) report an increased need for students to have experience with the MacOS platform, cloud and virtualization technologies, network security, a wider variety of database platforms, and a working knowledge of different programming languages. Plice and Reinig (2007) highlight the importance of a balance between business and technical content for IS students.

Developers of IS curriculum may benefit from a narrower view of IS student employability. The IS field is broad, and graduates of university IS undergraduate and graduate programs may ultimately work in any number of specialized fields from networking to cybersecurity to project management. It stands to reason that each of these specializations requires unique knowledge and skills, which we see reflected in specialized programs and certificates developing in the academy now. One such example is a data analytics specialization, intended to prepare IS students for a specialized role as a data analyst, data scientist, data visualization analyst, analytics analyst, or similar role (Clayton & Clopton, 2019).

In 2020, Dong and Triche published a study investigating longitudinal trends of skills for entry-level data analysts between 2014 and 2018. By scraping and analyzing data from Indeed.com, the authors report which tools and skillsets gained popularity in job postings over the four-year span (SQL Server, NoSQL, Tableau, Power BI, Python, Pig, Hadoop, Salesforce, Azure, Hive, Google Analytics, R, SAS, and SPSS), which grew less popular (Microsoft Access, Cognos, and SAP), and which stayed the same (Oracle, Microsoft Office) (Dong & Triche, 2020).

Dong and Triche's work gives IS curriculum designers important insight into which technical skills are most frequently cited in job postings for data analysts. However, a few gaps remain in our understanding. First, the job postings utilized in Dong and Triche's study were not comprehensive but rather only those captured by the Common Crawl tool to facilitate the longitudinal nature of their study. Second, the wildcard searches used by Dong and Triche, which included titles like data analyst and business intelligence analyst, excluded some other valuable titles such as analytics analyst and data visualization analyst. Finally, their results tell us the frequency with which certain skills were listed in job postings but cannot shed light on which of the listed technical skills were most important to any given job. Our work aims to build on, extend, and update the findings from Dong and Triche by addressing the research question: *What skillsets are 1) most frequently listed in entry-level data analytics job listings and 2) of those skillsets most frequently listed, which among them are most important to industry professionals when hiring entry-level data analysts?*

Theory

The Resource-Based View of the Firm (RBV, Barney, 1991) posits that sustained competitive advantage is derived from the resources and capabilities a firm controls that have several characteristics, including being valuable, rare, imperfectly imitable, and not substitutable (Barney et al., 2001). A chief capability regarding competitive advantage is human resources (HR), which is argued to be "the crux of competitive strategy" because "the ability of a firm to innovate is majorly influenced by the knowledge and skills of its employees" and the knowledge and skills in question are posited to be evaluable by the firm itself (Mishra et al., 2019, p. 1737). In the specific context of data analytics, human skills are particularly important for the adoption of big data and predictive analytics, and the hiring and more importantly, retainment of appropriately skilled individuals is key to achieving a competitive advantage (Dubey et al., 2019). Drawing on the RBV and its application so far in data analytics, we predict that firms are capable of identifying the skills needed in entry-level data analytics professionals and that some skills are valued over others by the firms.

Proposed Methodology

Our work aims to build on the previous work done by Dong and Triche (2020) and fill in gaps remaining in the literature by employing a mixed method approach. First, we will collect data in a similar way to Dong and Triche by web scraping and analyzing data from a popular job posting website. We will focus on collecting a snapshot of today's job postings to get a more comprehensive view, rather than the limited but longitudinal view collected by Dong and Triche. We will then supplement that data with qualitative data collected from individuals hiring entry-level data analysts and advising IS departments on curriculum. Our proposed methodology is summarized in Figure 1 and described in detail next.

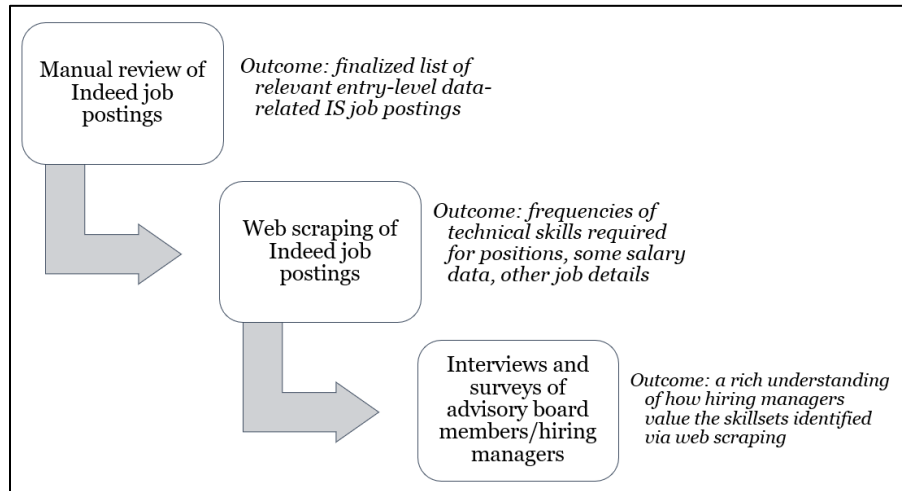


Figure 1. Summary of proposed methodology.

Web Scraping Data Collection

We will utilize web scraping techniques to extract job posting data from Indeed.com, which remains the number one career website by traffic and keywords (*Top 10 Career Websites by Traffic (And How They Did It!)*, 2020). We will follow Dong and Triche (and previous literature) by utilizing a wildcard search for *business analy**, *data analy**, and *business int**, but will also include *visualization analy** and *data visualization analy** to hone in on these visualization-specific career paths. Before beginning data collection, we will also employ a manual examination of Indeed.com to evaluate any other commonly utilized job titles that fit within the purview of our study.

Indeed.com includes some information on salaries associated with job postings as well. We plan to collect and explore this data and, if viable, examine potential correlations among specific skillsets listed and job salary levels to investigate whether some skills are valued more highly than others in terms of posted salaries or salary ranges.

Interviews with Employers and Advisory Boards Data Collection

To supplement our snapshot of how frequently skills are listed in job postings, we will also employ interviews with the advisory board to an information systems department at a large midwestern university. These individuals are representatives of IT and HR departments at major companies who additionally volunteer time to advise the department on curriculum development. They regularly hire graduates of the IS department and are therefore likely to be well-versed in the needs of the business community and in the curriculum of the IS department.

Previous research has drawn on the wisdom of the advisory board to inform curriculum decisions. Hammond & Moser detail the use of advisory board surveys to adjust the requirements of a general business degree (2009). Advisory boards can “serve as an effective mechanism for transferring best practices and ‘lessons learned’ from industry to academics” (Viswanathan, 2012, p. 1) and are effective in “providing

specific direction for revising and redesigning existing programs to meet the changing needs of industries and technologies” (Viswanathan, 2012, p. 10).

In this study, we will utilize a combination of semi-structured interviews and surveys with members of the advisory board to ascertain which skillsets drawn from the earlier web scraping exercise are perceived to be most important or valuable for a new data analyst to obtain prior to college graduation. These questions will investigate which skills are expected as a baseline for applicants, which would set an applicant apart from his or her competitors on the job market, and which are not expected or perhaps even seen as a negative attribute of a given candidate. We will also investigate which skills are explicitly demanded in job postings and which are implicitly expected. Additionally, we will evaluate the value that these industry professionals place on non-academic trainings such as bootcamps, self-directed online learning, and professional certifications. Based on feedback and data collected from the Advisory Board, we will refine the survey questions as needed and broaden the interview and survey collection to a larger population of industry employees in a minimum of two large metropolitan areas of the US.

Conclusion

Our work will continue to build on a stream of research investigating the skills needed for IS professionals entering the workforce today. Specifically, we will contribute a snapshot of the technical skills desired for data analysts in today’s job market and a rich understanding of how those skills vary in importance to different organizations and hiring managers. This fills in gaps left in extant literature, which has primarily focused on a longitudinal view, limited job titles, and simple frequencies of desired skills rather than the importance related to individual skillsets. In addition, our work will provide an updated view of the research problem, since the most recent research to investigate it is based on data from over four years ago.

The contributions of our work will be meaningful to both academia and practice. Individuals developing curriculum in higher education can draw from our results to inform changes to curriculum to ensure programs remain relevant, a well-known challenge for curriculum development in IS (Cummings & Janicki, 2021). Practitioners can utilize our results to inform how they should invest training funds to prepare IS graduates and/or compare their own technical skillset requirements to those required by other organizations. Students or recent grads can also benefit from these findings by seeking to adjust their own skillsets to match what is desired by organizations today. Ultimately, persistent attention to this stream of research as we suggest will produce programs that are better able to prepare our students for organizations and result in more effectiveness for organizations and more career satisfaction for data analysts.

Our work has several limitations that can be addressed with future research. First, we limit our focus to entry-level data analysts in the United States; future research may investigate a broader set of job titles and levels and incorporate a comparison to other global locations. Second, our work is limited to job postings by organizations, which leads to two potential limitations. This constricts our findings from applying to more broad contexts such as economies and societies and limits us to understanding jobs only from the firm’s point of view. Future research may investigate the skillsets necessary to influence economies and societies at a broader scale, and may incorporate other data collection (e.g. interviews of entry-level analysts themselves) to understand jobs from multiple points of view rather than just the firm’s. Finally, the IS department in the academy prepares not only IS majors, but also offers curriculum to the entire business school and sometimes even beyond. A future project may investigate what IS- or data-related skills are expected and valued for graduates of other university programs, to inform the development of general education IS courses.

REFERENCES

- Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17(1), 99–120.
- Barney, J., Wright, M., & Ketchen, D. J. (2001). The resource-based view of the firm: Ten years after 1991. *Journal of Management*, 27, 625–641.
- Clayton, P. R., & Clopton, J. (2019). Business curriculum redesign: Integrating data analytics. *Journal of Education for Business*, 94(1), 57–63. <https://doi.org/10.1080/08832323.2018.1502142>
- Cummings, J., & Janicki, T. (2021). Survey of Technology and Skills in Demand: 2020 Update. *Journal of Information Systems Education*, 32(2), 150–159.

- Dahiya, R., Le, S., Ring, J. K., & Watson, K. (2021). Big data analytics and competitive advantage: The strategic role of firm-specific knowledge. *Journal of Strategy and Management, ahead-of-print*(ahead-of-print). <https://doi.org/10.1108/JSMA-08-2020-0203>
- Dong, T., & Triche, J. (2020). A Longitudinal Analysis of Job Skills for Entry-Level Data Analysts. *Journal of Information Systems Education, 31*(4), 312–326.
- Dubey, R., Gunasekaran, A., Childe, S. J., Blome, C., & Papadopoulos, T. (2019). Big Data and Predictive Analytics and Manufacturing Performance: Integrating Institutional Theory, Resource-Based View and Big Data Culture. *British Journal of Management, 30*(2), 341–361. <https://doi.org/10.1111/1467-8551.12355>
- Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., ... Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management, 57*, 101994. <https://doi.org/10.1016/j.ijinfomgt.2019.08.002>
- Hammond, K. L., & Moser, E. (2009). Curriculum Advice from Your Advisory Board: A Survey Instrument. *Academy of Educational Leadership Journal, 13*(3), 15.
- LeClair, D. (2013, February 7). *Analytics Demand Reshaping Business Education | AACSB [Blog]*. AACSB Insights. <https://www.aacsb.edu/insights/articles/2018/02/an-insatiable-demand-for-analytics-talent-is-reshaping-business-education>
- Mishra, D., Luo, Z., Hazen, B., Hassini, E., & Foropon, C. (2019). Organizational capabilities that enable big data and predictive analytics diffusion and organizational performance: A resource-based perspective. *Management Decision, 57*(8), 1734–1755. <https://doi.org/10.1108/MD-03-2018-0324>
- Plice, R. K., & Reinig, B. A. (2007). Aligning the Information Systems Curriculum with the Needs of Industry and Graduates. *Journal of Computer Information Systems, 48*(1), 22–30.
- Saha, D. (2020, May 20). *How The World Became Data-Driven, And What's Next*. Forbes. <https://www.forbes.com/sites/googlecloud/2020/05/20/how-the-world-became-data-driven-and-whats-next/>
- Setor, T. K., & Joseph, D. (2021). College-Based Career Interventions: Raising IT Employability and Persistence in Early Careers of IT Professionals. *Journal of Information Systems Education, 32*(4), 262–273.
- The Best Business Analytics Programs, Ranked*. (n.d.). Retrieved February 22, 2022, from <https://www.usnews.com/best-graduate-schools/top-business-schools/business-analytics-rankings>
- Top 10 Career Websites by Traffic (And How They Did It!)*. (2020, October 7). The Scripted Blog. <https://www.scripted.com/content-marketing/top-10-career-websites-by-traffic-and-how-they-did-it>
- Top Data Analytics Courses—Learn Data Analytics Online*. (n.d.). Coursera. Retrieved April 20, 2022, from <https://www.coursera.org/search?query=data%20analytics&>
- Viswanathan, S. (2012). Importance of Advisory Boards in Program Development and Management. *2012 ASEE Annual Conference & Exposition Proceedings, 25.739.1-25.739.11*. <https://doi.org/10.18260/1-2--21496>
- Waguespack, L., Babb, J., & Yates, D. (2018). Triangulating Coding Bootcamps in IS Education: Bootleg Education or Disruptive Innovation? *Journal of Information Systems Education, 16*(6), 48–58.
- Which Certification Is Best for Data Analytics?* (n.d.). Default. Retrieved April 20, 2022, from <https://www.comptia.org/blog/best-data-analytics-certifications>