Demystifying and Developing a Framework for “Data Analytics”

Joseph Kirby, DBA
Bellevue University
College of Business
1000 Galvin Road South
Bellevue, NE 68005
September 2021

Abstract

Big data and data analytics have fundamentally changed the way we live and work. Data literacy is a foundational element for data analytics, yet there is no agreement on what it means to be data literate. Conversations on data literacy skills range from a basic understanding of the value of data and the ability to work with data through advanced topics of Machine Learning (ML) and Artificial Intelligence (AI).

The skills necessary to make fact-based decisions using data are generally lacking across employees and recent college graduates, which some may find is interesting considering that Gen X and Millennials have grown up with technology. This gap highlights a critical distinction between being comfortable with technology in general and being data literate. A key stumbling block involves the lack of a clear definition. If you can't define it, you can not measure it, which makes it difficult to teach it.

Many conversations about data analytics focus on the technology, which this author believes is misdirected. This paper seeks to demystify some of the confusion surrounding "data analytics." Two conceptual frameworks are presented. The first framework proposes a business-centric model for "data analytics," which positions Domain Knowledge, Data Literacy, Problem Solving, and Numeracy/Statistics as critical skills and abilities necessary to support a Business Strategy, Value, or Objective. The second framework proposes a hierarchy of definitions for "data literacy," with a decomposition including basic data skills, data analysis, data analytics, and data science.

Introduction

Data has recently been described as the "new oil" (Holeni, 2020, p. 14) which speaks to the recent hype associated with data (McAfee et al., 2012). As the foundational element of the knowledge pyramid, the role of data has always been important for knowledge development and good decision-making (Frické, 2009). Yet recently, the focus has turned to not just data, but "big data" and the capabilities enabled by leveraging advanced data analytics. The term "data analytics" is overly used and largely misunderstood. Clarity will aid businesses, professional organizations, and universities to prepare individuals with the data literacy skills necessary for success in the 21st Century skill.

Data analytics has captured much attention, as it rests at the center of many technology-enabled disruptions witnessed across multiple industries in the last 10 to 20 years. What is especially worrisome for firms across many industries is the fact that many of these technology-enabled disruptions involve competitors from outside of the firm's competitive environment, catching organizations off-guard (Downes & Nunes, 2013). Further, lacking a clear understanding of "data analytics," while firms may recognize the technology-enabled disruptions occurring in their markets, they struggle to implement the necessary changes to effectively respond (McKinsey, 2019).

The lack of a clear definition of what it means to be data literate contributes to the lack of traction in enhancing data literacy skills. Conversations related to data literacy and data analytics range from the basic skills associated with the collection and analysis of data to the more advanced statistical skills involved in machine learning (Machine Learning) and artificial intelligence (AI), which are more aligned with research and science than traditional programming and data management (Alpaydin, 2016). Further, the abundance of training programs available for these high-end skills and lack of training programs for basic data skills is out of step with the current workforce demands.

At present in the U.S., there is a gap of 1,500,000 employees with solid data analysis skills to support decision-making. This compares to a gap of 150,000 high-end data scientists, which require more sophisticated technology, programming, and statistical skills. The availability of training programs seems to lean towards the more advanced technical skills, which does not address the reality of this 10x disparity in the skills gap. Assuming training programs address market demands, the request for higher-end training programs may reflect a clear understanding of the path to developing advanced data analytics capabilities.

Efforts to enhance data literacy skills and leverage data analytics are can be seen in the industry, professional associations, and across higher education. The emergence of roles such as the Chief Data Office (CDO) and Chief Analytics Officer (CAO) at firms and federal agencies is an interesting addition to the C-Suite and perhaps reflects the challenges organizations face in dealing with data within their existing organizational structures (Gartner, 2016; Shibu, 2019). Additionally, the American Institute of Certified Public Accountants (AICPA) has introduced a new format for the CPA Exam, which will have data analytics as a core focus for the redesigned exam starting in January

2021 (AICPA, 2021). Colleges and universities are moving to integrate data analytics into their curriculum through certificates, boot camps, courses, and concentrations, yet employers are still frustrated with graduates that are unprepared for 21st skills (Aasheim et al., 2015; Bichsel, 2012; Casner-Lotto et al., 2009; Dzuranin et al., 2018; Eisner, 2010; Foundation, 2021; Geerts, 2021).

Organizations seeking to enhance their data literacy skills and leverage data analytics within their operations face a confusing environment. The purpose of this working paper is to review the various approaches, demystify "data analytics," propose a conceptual framework, and a hierarchy of definitions to aid organizations in their efforts in developing the data literacy skills and abilities of their organization and workforce.

When Did Data Become Important

Data analytics has fundamentally changed the way people live and work, and recently data literacy has emerged as an essential 21st-Century skill. Why just now? Data has always been important. As the foundational element for information, knowledge, and wisdom, data has always been an essential element for success for businesses, researchers, and citizens (Rowley, 2007). Interestingly, Information Literacy emerged in the 1970s as an essential skill (Doyle, 1994), yet it took another 40 years for the importance of data literacy to take hold. To more fully understand this delay, it is helpful to look back into the 1940s, 1950s, and 1960s.

Data analytics is often used in conjunction with Artificial Intelligence (AI), for which the theoretical underpinnings trace back to the 1940s and 1950s when researchers developed theories related to learning, knowledge development, inference, and the potential of intelligent machines (Buchanan, 2005). The technology horsepower required for AI was not available at that time and it would not exist for many years. In 1965, Gordon Moore provided insight on how the technology gains might evolve through exponential increases in computing power (Moore, 1965). A tipping point arrived circa 2005, with the convergence of increased computing power, the ability to store and manage massive amounts of data, and the availability of advanced analytical tools and algorithms becoming available at a reasonable cost (Padhy, 2013). Finally, the ability to harness value from data itself through various data analytics techniques provided capabilities to achieve that which was theorized in the 1940s and 1950s.

Stepping back, it is helpful to recognize that data analytics is fundamentally different than other technological advances witnessed over the past 30-40 years, such as ERP, Data Warehousing, CRM, the internet, and eCommerce. Each of these major technologies significantly impacted organizations and created some levels of disruption, yet the technologies were effectively integrated within the organization. Disruptions related to data analytics, however, have persisted for almost 20 years, and organizations have struggled to embrace and operationalize data analytics. One sign of the unique challenges associated with data analytics is the creation of the Chief Data Officer (CDO) role, the first of which was introduced by Capital One in 2003 (Zhang et al., 2017). Since this time, 90% of large organizations were predicted to have a CDO, and a 2019 federal law mandated a CDO at federal agencies (Gartner, 2016; Shibu, 2019). Similar, persistent C-suite roles have occurred with the previously mentioned technology advancements.

The Disruptive Nature of Data Analytics

The technology-enabled disruptions from data analytics are occurring rapidly across most industries (Downes & Nunes, 2013). Beyond exemplars like Amazon, Netflix, Airbnb, and Google, mainstream organizations such as John Deere, Shell, Walmart, and Rolls-Royce leverage data analytics to enhance value to their customers, develop operational efficiencies, and identify ways to drive top-line enhancements to their businesses; these companies have invested heavily to use data analytics as a competitive differentiator, which involves far more than just technology (Marr, 2016). To leverage the full potential of data analytics involves an investment in a data-driven culture, technology infrastructure, and the development of data literacy skills across both technical and business personnel (Davenport, 2006).

Developing data analytics capabilities do not reflect one-time investments, but rather they are long-term commitments to cultural change. Building advanced data analytics capabilities takes many years and offers organizations a competitive strategic advantage that continues to grow with time and creates a widening gap between firms that possess these analytical capabilities and their less analytical peers (McKinsey, 2019). As with any significant organizational change, the role of the top management team is critical. The investments involve more than just approving technology investments; firms that have successfully embraced analytics start with a commitment from senior executives. Research has shown that firms with data and technology-savvy executive teams significantly outperform their peers (Weill, 2021).

More advanced firms recognize the need to enhance the data literacy skills of their workforce and continue to invest heavily in skills training (Brown, 2019; Landi, 2019), while the less advanced firms struggle to react. While the majority of firms recognize the impact of data analytics on their competitive environments, most firms do not know how to respond and struggle with defining the roadmap and the path to enhancing data analytics capabilities (Columbus, 2014). While many firms rely on their IT departments as they embrace new technologies, the unique characteristic of data analytics present challenges, as the skills for data analytics success are fundamentally different than that of a programmer but more akin to a scientist/statistician (Alpaydin, 2016).

Professional organizations and higher education institutions are working to integrate data literacy and data analytics. In response to a 30% decline in recent accounting graduate hiring rates and declining accounting enrollments, the AICPA recognizes the need to embrace data analytics as a profession. Effective January 2024, a new format for the CPA exam will be introduced to include data analytics (AICPA, 2021; Geerts, 2021). While using the term "data analytics" to describe key changes to the Exam, a review of the data and analytics relevant Learning Objectives in the proposed model curriculum describes basic data modeling, management, governance, and analysis skills which have been relevant for 30+ years, as well as a small amount of content related to more advanced topics such as "Artificial Intelligence" (AICPA, 2021). The use of "data analytics" to describe a broad array of data-related skills contributes to the confusion.

Business schools are rushing to integrate analytics into their accounting core to address the new CPA Exam requirements (Dzuranin et al., 2018; Geerts, 2021; PriceWaterhouseCoopers, 2017;

Qasim et al., 2020). More broadly, Universities have incorporated data literacy and Analytics in various ways, including boot camps, certificates, new courses, new majors/concentrations, and hybrid methods (Cummings & Janicki, 2021; Radovilsky et al., 2018; Rios et al., 2020; van Laar et al., 2020). These efforts are not new, and findings would suggest that new graduates are still lacking in relevant workforce skills and more effective strategies are needed (Belkin, 2015; PriceWaterhouseCoopers, 2017; Radovilsky et al., 2018; Rios et al., 2020).

Demystifying Data Analytics

The overuse of the term "data analytics" is problematic. The term is used to describe basic data skills, such as data collection, storage, modeling, governance, and ethics, as well as advanced data analysis techniques involving Machine Learning and AI. To make meaningful progress in targeted workforce training, clarity in the definition is needed. By stepping back and looking at the knowledge, skills, and abilities needed to enhance data literacy, we can take a more strategic approach to tailor training and development programs to the needs of the workforce. The disconnect between the supply of courses and demands for learning may reflect the confusion related to what organizations need to enhance the data literacy skills and analytics capabilities in the workforce.

In the United States, there is a workforce shortage of 150,000 data scientists and 1,500,000 data-savvy business analysts or managers (Manyika et al., 2011). The skills needed for each of these roles are distinctly different. Yet, the availability of training and certificate programs today is largely slanted towards the more "exciting" fields of data science and the required higher-end technical skills. One does not need to search for very long on the internet to see courses and certificate programs offered by well-known schools, with a heavy emphasis on "Data Analytics."

Many of the program offerings for data analytics focus on specific technologies or advanced analytical techniques, which seem to be missing the point on a couple of fronts. First, the advanced data analytics curriculums do not address the largest current gap in the workforce. Second, a technology-centric curriculum is missing the key point that technology is necessary, but not sufficient. Sound technology strategies are rooted in the timeless principle that technology needs to support critical business imperatives (Ackoff, 1967; Ross et al., 1996). Irrational action seems to accompany the introduction of complex and potentially disruptive technologies.

New technology trends are exciting and may lead companies to narrowly focus on the technology as the "silver bullet," as opposed to developing strategies aligning technology investments in support of critical business imperatives. It is this technology-centric thinking that contributes to the 70% - 85% failure rate for major technology initiatives; data analytics projects included (Deloitte, 2015; Tabrizi et al., 2019; Waid, 2019). Still, the key reasons technology initiative failures are not related to the technology but involve unrealistic expectations, cultural challenges, and a failure to properly plan for integrated change management (Bennett, 2016; Wade & Shan, 2020).

The lack of clarity regarding "data analytics" may lead to fear and uncertainty for organizations, which may lead to well-intentioned, yet misdirected actions. In 1965, the prediction that machines would be able to replace all work performed by humans was rejected by experts that knew the capabilities and the limitations of computers (Dreyfus et al., 2000). Today, organizations seem to exhibit a similar level of fear and uncertainty related to Machine Learning and Artificial Intelligence. While the capabilities offered through advanced analytics are exciting and impactful to organizations, they don't happen without a clear strategy and intentional actions related to creating the technical infrastructure that is architected to provide relevant and reliable data to support an organization's analytics efforts. Organizations need a more comprehensive, business-centric strategy to effectively harness the value of data analytics. Investments in technology alone are not likely to succeed.

As organizations assess their data analytics capabilities and needs, Garner provides a well-regarded progression of analytics across four levels: Descriptive □ Diagnostic □ Predictive □ Prescriptive (McNellis, 2019). While often portrayed linearly, the required investments in culture, technical infrastructure, data governance, and employees skills are more exponential as organizations progress from one stage to the next. Organizations that seek the promise of predictive and prescriptive analytics, without first possessing lower-level capabilities will likely be disappointed in their investments. In terms of achieving value, organizations that currently lack basic descriptive and diagnostic analytics capabilities can often realize solid returns by developing these foundational capabilities. Success is unlikely for firms seeking to skip the foundational levels on the analytics continuum.

The skills needed to enable an organization to advance across the above-described levels of analytics vary greatly. Great value exists in the descriptive and diagnostic analytics levels, and these levels represent the largest skills gap in the U.S. workforce. As a first step in assessing skills gaps within their workforce, there is value in assessing the organization's capabilities in terms of Gartner's levels of analytics. Organizations need to assess their current capabilities before pursuing a data analytics strategy. If, for example, an organization lacks reliable and quality data sources and the basic skills needed for successfully deploying descriptive analytics, it makes little sense to pursue more advanced applications related to prescriptive analytics. An organizational assessment provides critical input to determining the skills needed in their workforce to support existing capabilities and prepare the organization for more advanced analytical levels.

As organizations pursue efforts to increase the data literacy skills of their workforce, they are faced with the lack of a common definition of what it means to be data literate. One model describes data literacy as resting at the center of other literacies, including computational, statistical, scientific, information, media, and digital, creating a very intimidating view of developing data literacy Skills. (Bhargava et al., 2015). Many other definitions exist, often tailored to specific industries or disciplines (Corrall, 2019). Moving beyond the technical aspects of working with data (Ridsdale et al., 2015), definitions of data literacy also incorporate problem-solving within specific domains (Mandinach & Gummer, 2013), curiosity (Dykes, 2019; Markham, 2020), statistics (Schield, 2005), data privacy and security (Markham, 2020), data quality (Lawson, 2019), digital infrastructure (Gray et al., 2018), and advanced algorithms and tools (Alpaydin, 2016; Letouze, 2016). The varied definitions make it difficult for organizations to frame and focus the training needed to support data literacy.

A simplified definition is needed. Gartner provides a definition that works across multiple domains and is focused on value-enhancing activities (Gartner, 2018):

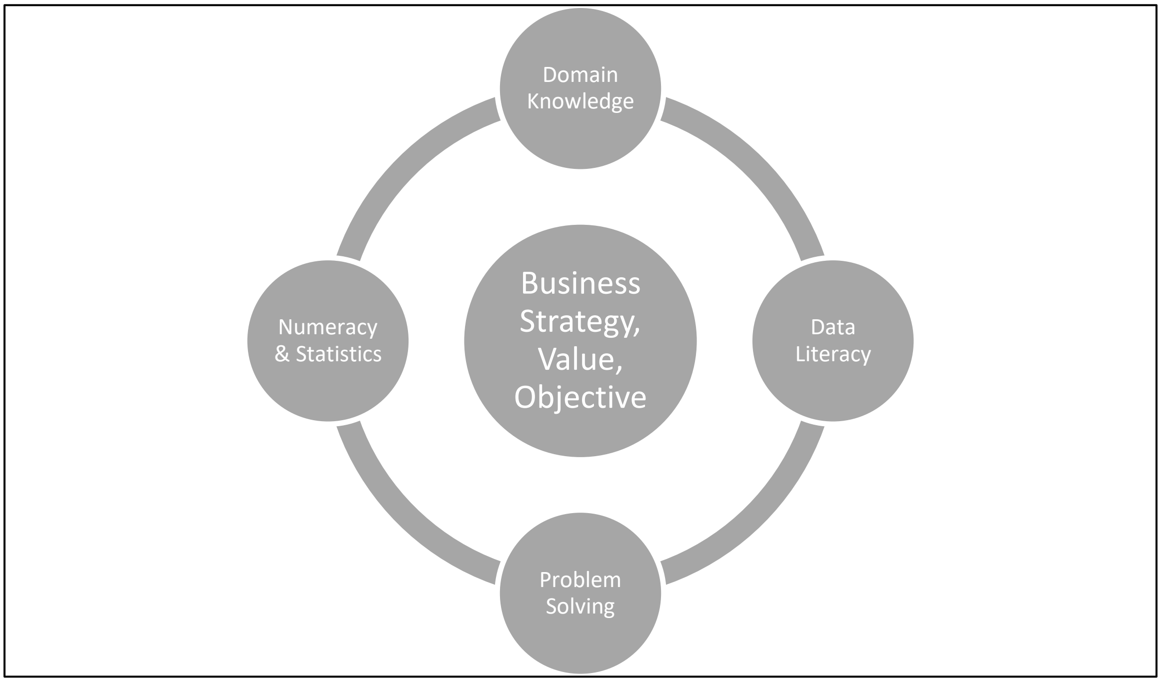
DATA LITERACY "The ability to read, write, and communicate data in context, including an understanding of data sources and constructs, analytical methods and techniques applied, and the ability to describe the use-case application and resulting value" Gartner

Gartner's definition provides a solid, value-based focus on data literacy, and touches on the need to be proficient with technology, without technology serving as the central focus. Technology should be aligned to support an organization's key objectives and efforts at enhancing an organization's workforce should be focused on driving value. It is with these concepts in mind, that the following frameworks are proposed to help an organization develop comprehensive strategies to enhance data literacy and develop data analytics capabilities.

Developing a Common Framework

This paper proposed two frameworks for organizations seeking to enhance data analytics capabilities. Figure 1 presents a contextual model for "Data Analytics." The business strategy/value/objective is at the center of the model, reinforcing the importance of data analytics efforts aligned with critical business strategies. Other critical elements of this model include Domain Knowledge, Numeracy/Statistics, and Problem Solving abilities, all of which are needed to deliver business value. Data Literacy is the fourth element for success, but not the sole driver. As organizations seek learning opportunities to enhance data literacy skills, this model supports the need to develop specific learning exercises using real-world business problems or use-cases that require individuals to use each of the four elements in the model to enhance the active learning process (Kosslyn, 2021).

Figure 1 - Data Analytics Context Model

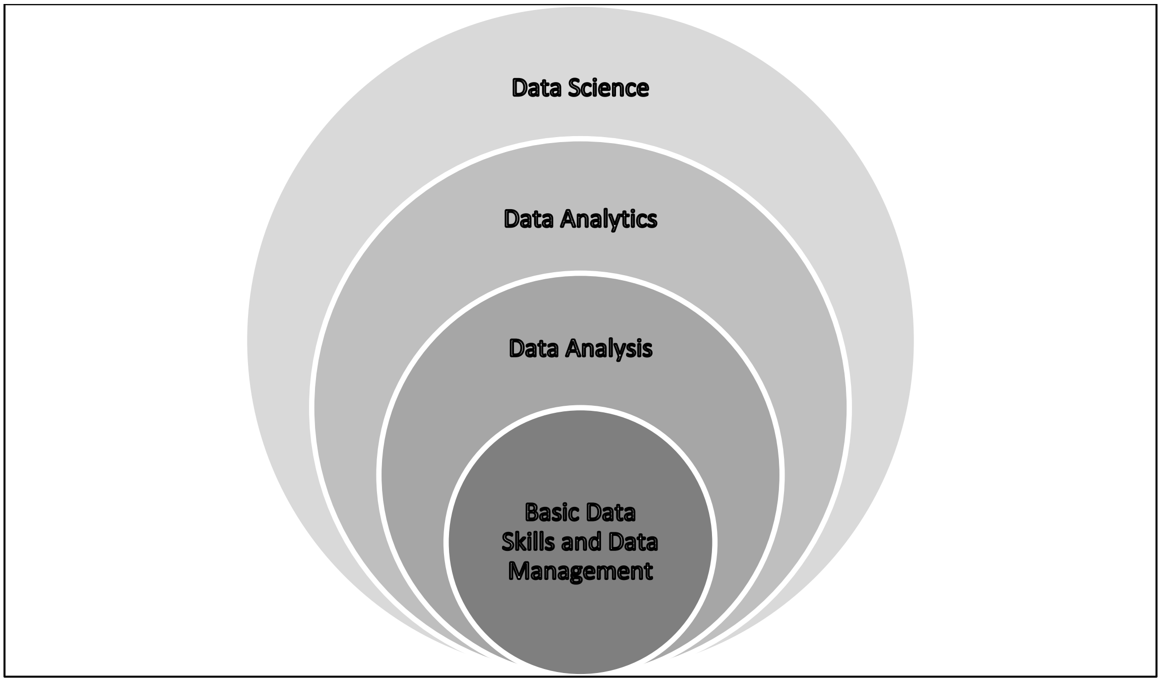


Beyond the definition and context model, further decomposition of "data literacy" is needed. When examined in greater detail, one sees that the term "data literacy" is used to describe a broad spectrum of data-related knowledge, skills, abilities, technologies, and techniques. Figure 2 proposes a decomposition of data literacy terms, to which specific and consistent meaning for each term can be used by organizations to ensure clarity in discussions and planning.

The following broad draft definitions for each level are presented for consideration:

1) Basic Data Skills and Data Management – involves the basic understanding of the value of data for decision-making, basic data structures, relationships, and the ability to conduct basic data management operations using the appropriate technologies.
2) Data Analysis – involves data collection, processing, evaluation of quality, data modeling, and data visualizations. Many of these techniques find their roots in the descriptive modeling and analysis developed during the rise of Data Warehousing in the 1990s, yet the skills are not widely taught at this point. Basic statistics and descriptive analysis of data are involved.
3) Data Analytics – leverages more advanced tools, algorithms, and statistical knowledge to identify patterns within data and is generally associated with what would be described as predictive analytics.
4) Data Science – refers to the more technical aspects of analytics, involving big data and more advanced techniques including machine learning and artificial intelligence.

Figure 2 - Data Analytics



The above definitions are informed by the literature (Alpaydin, 2016; Geerts, 2021; Olavsrud, 2021), but also reflect the practical knowledge as it relates to data-related skills that have evolved in practice over the past 30 years. The definitions offer clarity for organizations to examine their skill needs and gaps in their workforce, to ensure a more intentional approach to identifying key learning outcomes and sourcing or developing training programs for their workforce.

Appendix A provides a draft for a more granular breakdown of the concepts and techniques/models for each of the definitions, which can be used for gap analysis and planning. Appendix A offers organizations a starting point for discussions. Depending on the needs and current capabilities of the organization, items may shift from one level to another; the intent is to provide a structure that can be customized by the organization. The arrows at the bottom of Appendix A are intended to overlay the existing workforce gaps in the U.S. with the skills associated with the various levels in the data literacy hierarchy. Organizations can use this model to help define the knowledge, skills, and abilities appropriate for members of their workforce.

Conclusion

Organizations across all industries are pressured to develop data literacy skills in their organization to support data analytics efforts. The lack of a clear definition of data literacy makes it difficult to assess, measure, and enhance the skills of the workforce. Advanced algorithms that rely on complex statistics, technologies, and tools that are different than exist in a typical IT department may be intimidating for many firms. The broad use of the term "data analytics" to describe basic data management and analysis skills creates confusion. Organizations need to identify their critical skills gaps and adopt a more intentional use of the terms when in their efforts to enhance data literacy skills and data analytics capabilities.

Addressing the lack of clarity that exists today, this paper presents two frameworks to help organizations move forward in developing their data literacy Skills and data analytics capabilities. The first framework provides a business-centric contextual model of "Data Analytics," highlighting the collective role of Domain Knowledge, Data Literacy, Problem Solving, and Statistics to enhance the fact-based capabilities associated with data analytics. The second model proposes a hierarchy of definitions that decomposes "data literacy" into levels of Basic Data Skills, Data Analysis, data analytics, and Data Science, intended to support clarity in discussions and allow organizations to be more intentional in defining the data-related skills and capabilities needed to support their learning needs.

References

Ackoff, R. L. (1967). Management misinformation systems. Management Science, 14(4), B-147-B-156.

AICPA. (2021). CPA Evolution Model Curriculum.

Alpaydin, E. (2016). Machine Learning. The New AI. MIT Press.

Belkin, D. (2015). Test finds college graduates lack skills for white-collar jobs. Wall Street Journal, 25.

Bennett, J. (2016). How chief data officers can tackle formidable roadblocks including people, culture and internal resistance. https://www.gartner.com/smarterwithgartner/half-of-cdos-succeed/

Bhargava, R., Deahl, E., Letouzé, E., Noonan, A., Sangokoya, D., & Shoup, N. (2015). Beyond data literacy: reinventing community engagement and empowerment in the age of data.

Brown, S. G., D.; Herring, L.; Puri, A. (2019). The analytics academy: Bridging the gap between human and artificial intelligence., September 2019, 1-9.

Buchanan, B. G. (2005). A (very) brief history of artificial intelligence. Ai Magazine, 26(4), 53-53.

Columbus, L. (2014). 84% of enterprises see Big Data Analytics changing their industries' competitive landscapes In the next year. (Journal, Electronic), August 2019. https://www.forbes.com/sites/louiscolumbus/2014/10/19/84-of-enterprises-see-big-data-analytics-changing-their-industries-competitive-landscapes-in-the-next-year/#6a5aa2317de1

Corrall, S. (2019). Repositioning Data Literacy as a Mission-Critical Competence.

Cummings, J., & Janicki, T. (2021). Survey of Technology and Skills in Demand: 2020 Update. Journal of Information Systems Education, 32(2), 150.

Davenport, T. H. (2006). Competing on analytics. Harvard Business Review, 84(1), 98.

Deloitte. (2015). Your Guide to a Successful ERP Journey.

Downes, L., & Nunes, P. (2013). Big bang disruption. Harvard Business Review, 44-56.

Doyle, C. S. (1994). Information literacy in an information society: A concept for the information age. Diane Publishing.

Dreyfus, H., Dreyfus, S. E., & Athanasiou, T. (2000). Mind over machine. Simon and Schuster.

Dykes, B. (2019). Data Curiosity: How To Cultivate An Inquisitive Workforce. Forbes. Retrieved October 11, 2019 from https://www.forbes.com/sites/brentdykes/2019/10/10/data-curiosity-how-to-cultivate-an-inquisitive-workforce/#7d72b1735471

Dzuranin, A. C., Jones, J. R., & Olvera, R. M. (2018). Infusing data analytics into the accounting curriculum: A framework and insights from faculty. Journal of Accounting Education, 43, 24-39.

Gartner. (2016). Gartner estimates that 90 percent of large organizations will have a Chief Data Officer by 2019. Gartner. Retrieved 8/15/2019 from https://www.gartner.com/en/newsroom/press-releases/2016-01-26-gartner-estimates-that-90-percent-of-large-organizations-will-have-a-chief-data-officer-by-2019

Gartner. (2018). Gartner Glossary - Data Literacy. https://www.gartner.com/en/information-technology/glossary/data-literacy

Geerts, G. W., S. (2021). Data Analytics in Accounting: How to Help Your Students Become Better Critical Thinkers (Faculty Hour, Issue.

Gray, J., Gerlitz, C., & Bounegru, L. (2018). Data infrastructure literacy. Big Data & Society, 5(2), 2053951718786316.

Kosslyn, S. M. (2021). Active Learning Online. Alinea Knowledge, LLC.

Landi, H. (2019). Amazon plans to spend $700M to retrain a third of its workforce for data, analytics roles. Fierce Healthcare.

Lawson, R. H., T.;Desroches, D. (2019). How to Embrace Data Analytics to Be Successful. I. o. M. Accountants.

Letouze, E. (2016). SHOULD 'DATA LITERACY' BE PROMOTED? Retrieved 2/20/2020 from https://unstats.un.org/unsd/undataforum/should-data-literacy-be-promoted/index.html

Mandinach, E. B., & Gummer, E. S. (2013). Building educators' data literacy: Differing perspectives. The Journal of Educational Research & Policy Studies, 13(2), 1-5.

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.

Markham, A. N. (2020). Taking data literacy to the streets: critical pedagogy in the public sphere. Qualitative Inquiry, 26(2), 227-237.

Marr, B. (2016). Big data in pract Joseph Kirby, DBA, is an Assistant Professor at the Bellevue University College of Business. ice: how 45 successful companies used big data analytics to deliver extraordinary results. John Wiley & Sons.

McKinsey. (2019). Catch them if you can: How leaders in data and analytics have pulled ahead. McKinseyAnalytics. https://www.mckinsey.com/~{}/media/McKinsey/Business%20Functions/McKinsey%20Analytics/Our%20Insights/Catch%20them%20if%20you%20can%20How%20leaders%20in%20data%20and%20analytics%20have%20pulled%20ahead/Catch-them-if-you-can-How-leaders-in-data-and-analytics-have-pulled-ahead.pdf

McNellis, J. (2019). You're likely investing a lot in marketing analytics, but are you getting the right insights? https://blogs.gartner.com/jason-mcnellis/2019/11/05/youre-likely-investing-lot-marketing-analytics-getting-right-insights/

Moore, G. E. (1965). Cramming more components onto integrated circuits. In: McGraw-Hill New York.

Olavsrud, T. (2021). What is Data Analytics? Analyzing and Managing Data for Decisions. https://www. cio. com/article/3606151/what-is-data-analytics-analyzing-and-managing-data-for-decisions. html.

Padhy, R. P. (2013). Big data processing with Hadoop-MapReduce in cloud systems. International Journal of Cloud Computing and Services Science, 2(1), 16.

PriceWaterhouseCoopers. (2017). Investing in America's Data Science and Analytics Talent: A Case for Action.

Qasim, A., Issa, H., El Refae, G. A., & Sannella, A. J. (2020). A model to integrate data analytics in the undergraduate accounting curriculum. Journal of Emerging Technologies in Accounting, 17(2), 31-44.

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.

Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D., Kelley, D., Matwin, S., & Wuetherick, B. (2015). Strategies and best practices for data literacy education: Knowledge synthesis report.

Rios, J. A., Ling, G., Pugh, R., Becker, D., & Bacall, A. (2020). Identifying critical 21st-century skills for workplace success: A content analysis of job advertisements. Educational Researcher, 49(2), 80-89.

Ross, J. W., Beath, C. M., & Goodhue, D. L. (1996). Develop long-term competitiveness through IT assets. Sloan Management Review, 38(1), 31-42.

Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. Journal of information science, 33(2), 163-180.

Schield, M. (2005). Information literacy, statistical literacy, data literacy. IASSIST quarterly, 28(2-3), 6-6.

Shibu, S. (2019). New law mandates chief data officers at federal agencies. govloop. https://www.govloop.com/new-law-mandates-chief-data-officers-at-federal-agencies/

Tabrizi, B., Lam, E., Girard, K., & Irvin, V. (2019). Digital transformation is not about technology. Harvard Business Review, 13.

van Laar, E., van Deursen, A. J., van Dijk, J. A., & de Haan, J. (2020). Determinants of 21st-century skills and 21st-century digital skills for workers: A systematic literature review. Sage Open, 10(1), 2158244019900176.

Wade, M., & Shan, J. (2020). Covid-19 Has Accelerated Digital Transformation, but May Have Made it Harder Not Easier. MIS Quarterly Executive, 19(3), 7.

Waid, B. (2019). Solving The Last Mile Problem For Data Science Project Success. Forbes. https://www.forbes.com/sites/forbestechcouncil/2019/07/23/solving-the-last-mile-problem-for-data-science-project-success/?sh=35fa81815493

Weill, P. W., Stephanie L.; Shah, Aman M. (2021). Does your C-Suite have enough digital smarts. MIT Sloan Management Review, Spring 2021 Issue.

Zhang, H., Lee, Y., Wang, R., & Huang, W. (2017). Chief Data Officer Appointment and Origin: A Theoretical Perspective.

Appendix A

Bellevue University
College of Business - "Data Literacy" Skills

Framing the Discussuion
DRAFT

Category	Basic Data Skills and Data Management	Data Analysis	Data Analytics	Data Science
Concepts	1) Introduction to Data (Incl Data Types) 2) Data Discovery 3) Data Collection 4) Data Manipulation 5) Data Conversion (format to format) 6) Evaluating Quality 7) Data Cleaning 7) Data Organization 8) Data Modeling 9) Meta Data 10) Data Goverance (Stewardship, Security, ect) 11) Data Standards 12) Data Ethics 13) Fact-Based Decision-Making	1) Exploratory Data Analysis 2) Outlier Identification and Handling 3) Descriptive Statistics 4) Data Comparison and Trending 5) Taxonomy 6) Multidimensional Modeling 7) Visionalizations and Storytelling 8) Critical Thinking with Data 9) Evaluating Decisions with Data 10) Data Sufficiency & Relevancy 11) Querying Data 12) Probability Models	1) Research Design and Methods 2) Hythothesis Testing and Significance 3) Inferential Statistics 4) Classification 5) Regression 6) Clustering 7) Principle Factor Analysis 8) Variance Analysis 9) Forecasting	1) Pattern Recognition / Data Mining 2) Big Data - Cloud Computing, Etc. 3) Predictive Analytics 4) Prescriptive Analytics 5) Machine Learning 6) Artificial Intelligence
Techniques / Models	1) Basic Data Management / Structing 2) Data Manipulation and Formatting 3) Calculations 4) Logical Operations 5) Basic Reporting and Visualizations 6) Collection and Querying Data	1) Visualization Method Selection 2) Normalizing Data - Scaling, Log, etc… 3) SQL 4) ETL	1) Regression - Linear, Logistic, Probit 2) K-nearest neighbor 3) ANOVA, MANOVA, etc. 4) Principal Component Analysis 5) Cluster Analysis 6) Principal Factor Analysis 7) Time Series Analysis / Forecasting 8) Survivor Analysis	1) Neural Networks 2) Random Forest 3) Hueristics / Rules Engines 4) Supervised / Unsupervised Learning

Note: Beginning list of skills inventory. Not all skills will be applicable to every major.