Paper:
Ryan, L., Silver, D., Laramee, R., Ebert, D. & Rhyne, T. (2019). Teaching Data Visualization as a Skill. *IEEE Computer Graphics and Applications, 39*(2), 95-103.
http://dx.doi.org/10.1109/MCG.2018.2889526

Teaching Data Visualization as a Skill

Lindy Ryan
Montclair State University

Deborah Silver
Rutgers University

Robert S Laramee
Swansea University

David Ebert
Purdue University

Editor:
Theresa-Marie Rhyne
theresamarierhyne@gmail.com

This article explores the meaning of the term "skill" in the context of information (data) visualization and its place in the labor market. It examines the visualization skills and software competencies that are in high demand in industry today, and the ramifications for teaching Data Visualization *for professional students in higher education.*

The term "*Data Visualization*" is familiar to many in the IEEE community. As a discipline, it is a broad field consisting of many different areas, including scientific visualization, information visualization, and visual analytics, as well as statistical graphics, human-computer interaction, psychology, and several other sub-areas of research and study. In an academic sense, visualization is a subfield of computer science and involves algorithm development and new paradigm construction. Traditional educational courses on visualization required a background in computer science and engineering, and are housed within these respective departments. These courses expected and built upon advanced programming concepts, generally required building projects from scratch, and were designed for computer scientists and other research professionals.

Beyond its application in computer science and engineering, *Data Visualization* is also regarded as a critical skill for today's data-centric jobs, particularly within the analytics labor market. As such it is frequently found in job postings that span across many sectors of industry, with demand for people with this skill increasing exponentially at about 1500% in the past decade. In a recent search, *Data Visualization* as a skill was found to appear in over 30,000 different job advertisements [BurningGlass 2018][EMSI 2018]. However, unlike the types of scientific visualizations mentioned above, a significant majority of the job postings mentioning "Data Visualization" are referring to *Data Visualization* as a "skill"—in the same vein as "Java," "SQL," and "Python" are also considered skills by the employment sector. While it is clear from the prolificacy of mentions

in job advertisements that *Data Visualization* is in demand in today's labor market, what is less clear is what is expected when it is mentioned as a "skill."
While numerous textbooks on information visualization cover a wide range of topics—from traditional statistical visual representations of quantitative data to journalistic approaches to visual data storytelling—the challenge remains for educators within higher education to not only keep pace with the demand for a growing number of business analytics programs, but to also design graduate-level visualization courses that reconcile unclear and fluctuating industry expectations. Further, educators must be able to differentiate between what should be covered in professional degree programs (i.e., professional science master's, MBA) from what should be taught for research-oriented students (i.e., Ph.D., hard sciences).

In this Viewpoint, we take a novel approach of investigating labor demands in an attempt to understand what "Data Visualization as a skill" means within industry. Examining both job posting data as well as career profiles (e.g., LinkedIn), we assess what topics and technologies should be covered in new course to meet those definitions and prepare students with the skills that comprise *Data Visualization*, and how these align within the context of other critical analytical skills students need in today's labor market. Using this type of labor demand data to help inform curriculum is a new area for academia in general and for the field of visualization in particular. We show how it can be used by educators to help bridge the gap between academia and the labor market.

UNDERSTANDING LABOR DEMANDS AND INDUSTRY NEEDS

The first step in crafting new educational models that provide students with labor-ready *Data Visualization* skills is to understand current labor demands. To this end, there are a number of traditional methodologies that can be used. Professional and engineering programs have advisory boards made up of experts that can provide insight as to what industry is looking for in acquiring and retaining talent for *Data Visualization* jobs. While very useful, the information is only as good as the individuals chosen to sit on the boards. Industry surveys can also be performed to tap into the current practitioner pool [Meeks 2017]. Lastly, existing data from online job postings can be mined and analyzed, and, similarly, data from online resume postings can be mined and analyzed. In what follows, we discuss the results of using each of these methodologies to tease out what a *Data Visualization*-skill really is, and how it should be addressed in academic programs focused on teaching *Data Visualization* as a skill for industry. Interestingly, as these skills gain in popularity in the workplace, a related "visualization literacy" push has arisen to teach some of the underlying principles of *Data Visualization* [Velez 2009][Ryan and Snow 2016].

In this analysis, we use a tool called Labor Insight from Burning Glass [Burning Glass 2018] to mine data from online job postings. This tool aggregates and curates all of the current job postings across online applicant sites. To isolate information interesting to researchers for analysis, Labor Insight separates each job description posting into metadata fields that can be analyzed in a more granular fashion. A similar tool, EMSI, searches online resumes and results that are also included here [EMSI 2018].

While this form of analysis can highlight important trends in the job market, there are limitations. Although the tool attempts to remove duplicate postings, it is not always effective. Furthermore, this type of analysis can only look at advertised jobs, and does not account for current jobs or jobs that are not advertised online. The unstructured nature of job ads can make it difficult for the system to identify individual pieces of information effectively, and there may be postings that are not captured and tagged. So, there may be some irrelevant jobs caught in the net and some relevant jobs may have been left out. Mining online resumes has similar issues. For this article, we also reviewed available published white papers and data collected from surveys of executives, managers, and recruiters within big data and analytics organizations that either managed or hired IT professionals. This analysis added an industry perspective to our research that enabled us to obtain a real-time view into what skills and technical proficiencies analytics organizations are seeking in prospective job candidates.

Labor Demand Findings

From our analysis using Labor Insight, from March 2017-February 2018, out of approximately 7 million US-based job postings requiring a bachelor's degree or higher, there were 30,786 jobs that listed *Data Visualization* as a desirable skill. By contrast, in 2010, there were only 1,888 ads requesting visualization, thus resulting in a steady and very large increase in demand over time (as shown in Figure 1). There is increase of approximately 16.7% from just the March 2016-February 2017 time period. This demonstrates a strong push for this skill across many industries and a rising importance in many different employment opportunities. For perspective, out of these same 7 million jobs, about 430 thousand jobs listed SQL as a desirable skill. (Note: to help with our analysis and target the student population more effectively, we restricted our search to those jobs that require at least a BS/BA degree and required *Data Visualization* as a skill identified by Burning Glass [Burning Glass 2018].)

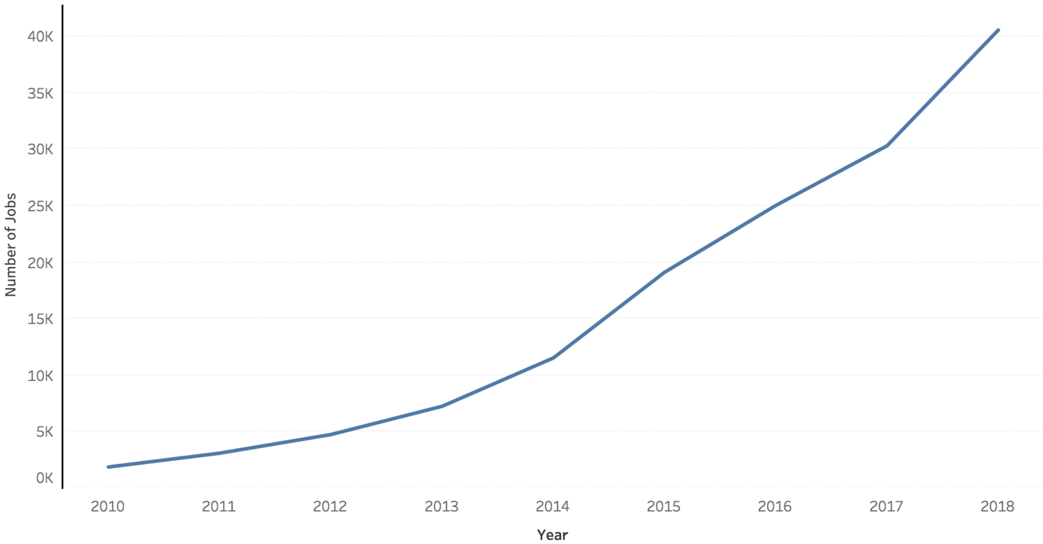

Figure 1. Visualization as a skill: The growth of jobs mentioning "data visualization" as a skill from 2010 through 2017 has steadily increased from only 1,888 jobs in 2010 to 30,327 jobs in 2018.

To get a sense of what is meant by *Data Visualization*, we wanted to see what other skills it was paired with. In the thirty thousand job postings mentioning *Data Visualization*, other high-ranking technical skills included: SQL (51%), Tableau (41%), Microsoft Excel (34%), Data Analysis (31%), and Python (30%) (Figure 2). All of these skills/technologies are consistent with a data-analysis type job. Beyond technical abilities, of note is the inclusion of specialized software packages Tableau (41%), Excel (34%), and SAS (22%), which were also identified as skills. *Data Visualization* intersects with many of these proficiencies.

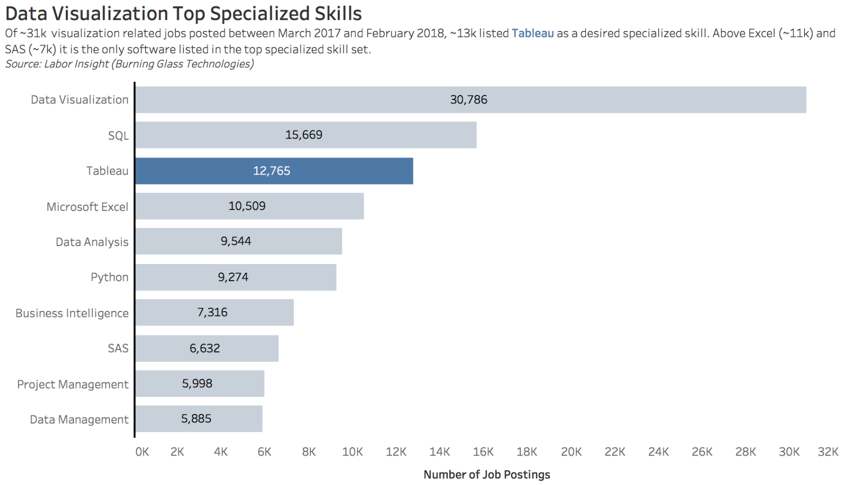

Figure 2. What is meant by visualization? Other skills also appearing when data visualization is requested. Source: Labor Insight (Burning Glass 2018)

An analysis of job titles provides another metric for exploring how industry quantifies data visualization as a skill. As is expected in industry, there were varied titles across the job postings, with some of the more popular titles being: Data Analyst (35%), Business Analyst (16%), Software Development Engineer (13%), and Business Intelligence Analyst (10%). Interestingly, only 3% (977) of postings actually had "visualization" in the title. Of these 3%, job titles included Data Visualization Specialist (14%), Senior Data Visualization Engineer (14%), Data Visualization Developer (12%), Data Visualization Engineer (11%), and Data Visualization Consultant (7%) (Figure 3). Additionally, 60% jobs listing *Data Visualization* in the title require at least 3-5 years of experience, demonstrating that the term visualization in the title was viewed by industry in favor of someone who had experience.

Looking at data from 2015-2016, the number of jobs with "visualization" in the title has been relatively consistent, but the number of jobs listing *Data Visualization* as a skill has increased exponentially. Monitoring the number of jobs with *Data Visualization* in the title itself provides an interesting opportunity for future study. Previously, visualization scientists were only found in universities and government labs. It seems that with this development the field can be seen as expanding beyond laboratories. Additionally, the low percentage of job titles that explicitly include "visualization" may be an indication of the maturation of the discipline, where visualization is now often regarded as an important aspect of a full computational/analytic skills toolbox, but not a specialized role in itself. The data does not support clear categorization of job profiles, as companies tend to approach these definitions differently and it is not up to us to set a standard. Likewise, there is not enough data to provide exact approaches to curriculum design. We believe that this is best left for curriculum designers at various institutions and in different programs to design based on their students, individual programs, and labor demands.

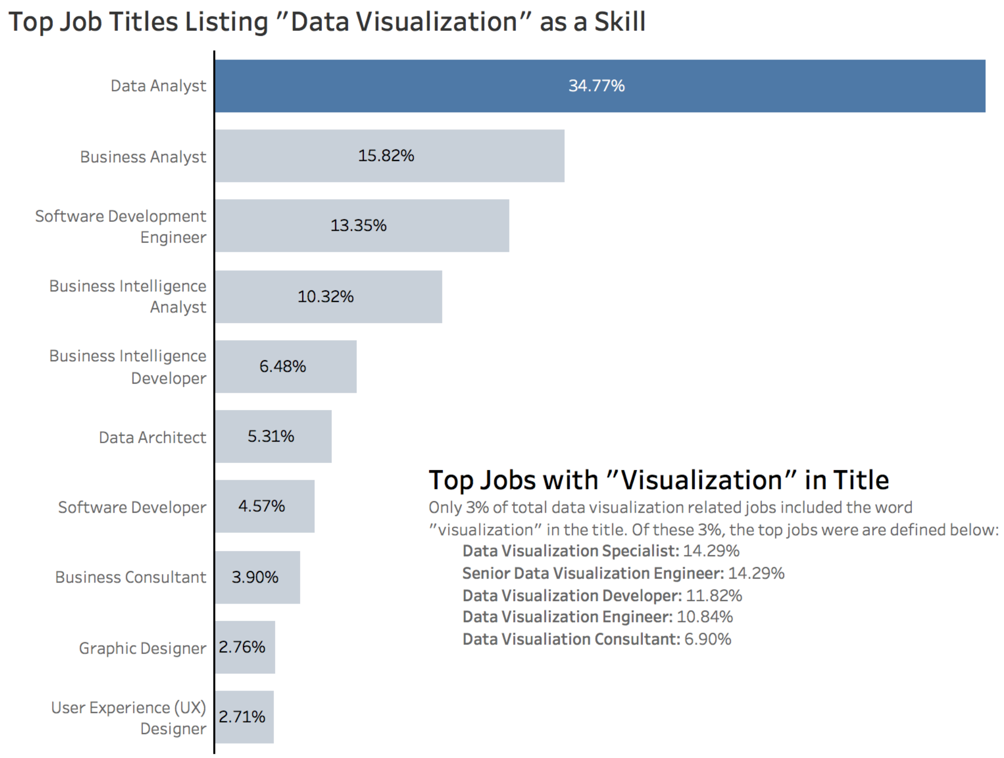

Figure 3. Job titles listed for jobs requesting "data visualization" skills and job titles that specifically list visualization in the title. Source: Labor Insight (Burning Glass 2018)

Jobs listing *Data Visualization* also include text describing desired "baseline" skills, which can be thought of as qualitative skills. These include communication skills (47%), research (37%), writing (32%), teamwork/collaboration (31%), problem solving (30%), and general project management skills, such as planning, Microsoft Excel, etc. (Figure 4). While it may seem obvious that most employers want employees to communicate effectively, these postings specifically mention this and highlight communication and qualitative abilities requisite of jobs in analytics and data communication. This reflects an increased focus on visual data storytelling as a new skill that is an aspect of *Data Visualization.*

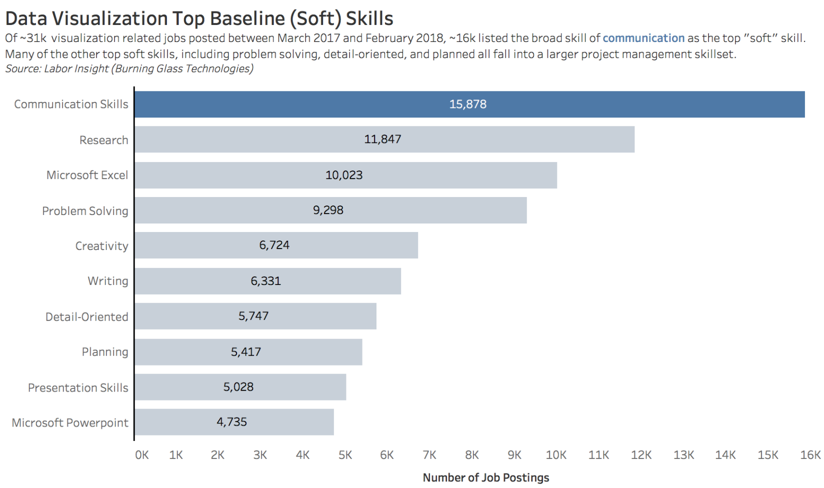

Figure 4. What other skills appear in job postings? Baseline, or "soft" skills listed for *Data Visualization* jobs. Source: Labor Insight (Burning Glass 2018)

Additionally, although they do not formally appear in the top baseline or specialized skills (these are skills defined by Burning Glass [2018]), the importance of user experience / user interface design (UX/UI) and project management in data visualization should not be understated. Specifically, we found UX/UI as an informal skill in 9,200 out of 30,852 jobs, and project management in 13,266 jobs of the same.

Although there is some overlap between what is listed as a specialized skill and what are listed as general skills regarding technologies and proficiency with preferred software packages and programming languages, about 51% of these jobs specifically listed SQL, while other common software and programming packages which were also mentioned include: Tableau (41%), Excel (34%), Python (30%), SAS (22%), PowerPoint (16%), R (16%), and Java (15%). Additionally, we also looked at the demand for skills involving JavaScript-based data-driven documents, or D3.js, and found that approximately 13% of data visualization jobs included D3.js as a keyword.

In terms of job location adjusted by population, "above average" job postings were found in the combined metropolitan areas of New York/New Jersey, followed by Washington DC, then San Francisco, Chicago, Boston, Santa Clara, and Seattle.

Looking at a search of online resumes for those already employed, including those from LinkedIn and CareerBuilder [EMSI 2018], there are about 45,000 individuals listing *Data Visualization* as a skill. The majority, 21%, have a business administration background, followed by computer science, engineering, economics, mathematics, and then statistics. Other common skills among these include: Data Analysis, Research, SQL, Management, Tableau, Business Intelligence, Data Mining, Python, and Javascript. One interesting thing to note in comparing the job postings to the profiles is that while communication skills are requested in about half of all of the job postings that also include *Data Visualization*, only about 10% of the profiles that include *Data Visualization* as a skill also mention that they have good communication skills.

Industry Demand Findings

A focus on the importance of analytics and *Data Visualization* is sweeping across industry. A 2011 survey of 3,000 executives, managers, and analysts across more than 30 industries and 100 countries highlighted the value high-performing organizations place on visual analytics to translate data into understanding [LaValle *et al.* 2011]. Likewise, academic inquiry into the need for *Data Visualization* skills has illustrated that today's organizations rely increasingly on analysts, in job titles ranging from data analyst to data scientist, to translate data into insight, and visualization tools have become a staple in the data analyst's toolkit to enable analysis and collaboration on diverse data sets [Kandel *et al.* 2012][Davenport and Patil 2012]. *Data Visualization*, has, additionally, gained value through the incorporation of visual data storytelling techniques, which has been shown to improve memory and retention [Borkin *et al* 2013], persuasion, and engagement between analysts and their audiences when communicating about business data [Dykes 2016].

Similar to the Labor Insight findings regarding the importance of communication skills, industry demands also pair communication alongside *Data Visualization* as critical for success. According to the Business Intelligence Congress's third survey (BIC3), published in 2014, over 400 recruiters from technical companies were asked what skills and competencies they looked for in new analytic hires. Not surprisingly, communication skills outrank technical skills for getting a business analysis job [Wixom *et al* 2014]. Likewise, the data industry research and advisory firm Gartner found that 99% of surveyed companies noted that half of the reasons their data analytics projects failed where due to poor organizational skills, communication in particular [Kart *et al* 2013]. Lastly, in a survey "directed at people who do professional data visualization", Meeks touched upon issues of the *Data Visualization* tools and methodologies used by 600 professional data visualizers, and found similar skills and tools as reported here [Meeks 2017]. These findings demonstrate the importance for educators not only to teach *Data Visualization* competency, but also the ability to effectively communicate insights to business audiences.

TEACHING VISUALIZATION AS A SKILL

For researchers pursuing *Data Visualization* as a career path, the current emphasis on programming within traditional computer science programs is sufficient to meet industry demand, and there are many excellent classes focusing on both scientific and information visualization coming from Computer Science and Computer Engineering departments. These cover the basics of visualization algorithms with an emphasis on applications and include a number of specialized programming APIs, such as the Visualization Toolkit (VTK) and Data-Driven Documents (d3.js). While some of these classes are interdisciplinary and include, along with scientists and engineers, individuals for whom programming is not a primary focus, the classes are created for the more technically-oriented students and thus require significant programming background. Therefore, they are not always aligned with the unique needs of today's "data professional".

A more appropriate class for professional students should focus not only on underlying principles and best practices of *Data Visualization*, but should include training on other skills, such as software specialization and communication, that are necessary for a visualization professional as described by industry job postings. While exact curricula and approaches are, of course, up to the individual schools, from this analysis we can make some suggestions for educators interested in designing graduate-level courses appropriate to professional students. First and foremost, education on data visualization best practices in terms of graphics as well as visual cognition should be instilled so that students know what different types of charts and graphs mean, how they are used to represent data, and how to appropriately show insights. Visual design principles should also be taught as a foundation for how to apply elements such as color, shape, and visual hierarchy in data visualization charts, dashboards, and stories. Further, over the term of the course, students should be tasked with learning the entire visual analytics process end-to-end—from data acquisition, to preparation, to analysis, to visualization and communication of results using tools such as reports, dashboards, and visual data stories. These concepts can and should also be incorporated into a more traditional visualization class to help the students apply advanced programming concepts within an industry setting.

Integrating Software into the Curriculum

In our class as part of the Professional Science Masters at Rutgers University, graduate students download a free Tableau Desktop license via the Tableau for Teaching program, which provides students and faculty with a renewable, free 12-month license for the software [Tableau 2018]. Classroom instruction employs a combination of: 1) Tableau's higher education curriculum to provide technical training in the software and 2) instructor-developed materials that focus on data visualization best practices, the impact of cognitive science on information visualization, and visual design principles for audience engagement, memorability, and retention. Throughout the course, students are assessed on their grasp of lecture material, hands-on skill development, and cumulative learning in the form of formal exams and concept application. Out-of-classroom discussion boards and homework assignments complement in-class learning. Students are required to submit finished visualizations and full Tableau workbooks, including data extracts, for instructor review.

In addition to Tableau, students use the ubiquitous workplace tool Microsoft Excel, as well as industry data-wrangling tools like Alteryx to prepare data prior to bringing it into the Tableau environment for analysis and visualization. The intent is that through this integrated approach students learn the fundamentals of *Data Visualization* for business use as well as achieve critical tool-based learning that add additional value for student marketability. This approach enables students to compete for jobs that require general *Data Visualization* skills and knowledge as well as hands-on software expertise as found in our analysis of the labor and industry demands discussed previously. In addition, tools like d3.js and more sophisticated visualization algorithms are also covered. While the integration of the specific tool Tableau is based on analysis of job postings, integration of other visualization tools could also be incorporated so that students are exposed to other commonly used tools.

Using Storytelling as Capstone

Most visualization courses culminate in a project. To incorporate the qualitative skill requirements that were found in the job postings, courses should include data storytelling as a key learning concept. In our course, whether comprised of individual visualizations or dashboards that include techniques like filtering, highlights, and annotations, *Data Visualizations* are used to build visual data narratives. Students present their narratives as course capstone projects, and are assessed on two fronts: 1) visualization methods and capabilities, and 2) the ability to clearly communicate a data "story" to a targeted audience. The intent for this project is to pair technical knowledge and with baseline skill demand, including communication, research, writing, and creativity. This type of approach can and should also be emphasized in a traditional visualization programming course so that students become familiar with the concept of data communication.

CONCLUSION

As *Data Visualization* continues to move out of the lab and into the workplace, it is important to understand what skills are necessary so that educators can respond appropriately in ways in which *Data Visualization as a Skill* is taught to both professional and science-based students. While we may want all students to be able to create sophisticated visualizations from scratch, we have to understand the demands of the workforce and create classes that speak to a wider range of students. Tools like Burning Glass and EMSI are becoming increasingly effective at allowing educators to feel the pulse of the labor market and to incorporate this into courses. In this viewpoint, we have demonstrated how to use these new tools to create a labor-market-oriented visualization course. This methodology can be used for many other types of courses and curriculums to help understand which technologies are used in industry, and to prepare students for various careers. Educators need to understand not only what the labor market is actively seeking when it asks for *Data Visualization* as a skill, but also what this means and how to best prepare students with the fundamental skills and knowledge they need to be successful. Additionally, questions such as how *Data Visualization* applies in the context of other analytic skills, which tools are used for which industry, and where employment opportunities are can all be investigated. Knowing about the rich sources of data available today and using it for curriculum development and to help inform advising can provide an enriched classroom experience that is more attuned to the labor market.

SIDEBAR: RESOURCES FOR TEACHING VISUALIZATION IN PROFESSIONAL PROGRAMS

With labor demand for data visualization as a skill steadily increasing (Figure 4), keeping pace with market demands in addition to the tools and competencies desired by industry and balancing these within the confines of professional and science-related course in higher education can be a challenging task. However, there are several resources available that educators can utilize, particularly for courses focused on preparing professional analytics students. These resources include numerous textbooks targeted for business schools teaching visualization courses. Popular texts today include *The Visual Display of Quantitative Information* [Tufte, 2001], *Visual Data Storytelling with Tableau* [Ryan, 2018], *Data Points: Visualization That Means Something* [Yau, 2013], and *Show Me the Numbers: Designing Tables and Graphs to Enlighten* [Few, 2012], and some information visualization textbooks such as *Visualization Analysis and Design* [Munzner, 2014] and *Information Visualization Perception for Design* [Ware, 2013], among others. Further, in addition to their Tableau for Teaching free software licensing program, Tableau also has educator-designed curricula suitable for graduate-level students that can be incorporated into classroom pedagogy for educator use.

SIDEBAR REFERENCES

1. [Tufte, 2001] Tufte, E. (2012). *The Visual Display of Quantitative Information 2nd Edition* (2nd ed.). Cheshire, CT: Graphics Press.
2. [Ryan, 2018] Ryan, L. (2018). *Visual Data Storytelling with Tableau.* Boston, MA: Pearson Education.
3. [Yau, 2013] Yau, N. *Data Points: Visualization That Means Something.* Indianapolis, IN: John Wiley and Sons, Inc.
4. [Few, 2012] Few, S. (2012). *Show Me the Numbers: Designing Tables and Graphs to Enlighten* (2nd Ed.). Burlingame, CA: Analytics Press.
5. [Munzer, 2014] Munzer, T. (2014). *Visualization Analysis and Design. A K Peters Visualization Series*, CRC Press.
6. [Ware, 2013] Ware, C. *Information Visualization Perception for Design*. Elsevier.

REFERENCES

7. [Borkin *et al* 2013] Borkin, M., Vo, A., Bylinski, Z., Isola, P., Sunkavalli, S., Oliva, A., Pfister, H. (2013). Proceedings from IEEE INFOVIS 2013: *What Makes a Visualization Memorable?*. Atlanta, GA: IEEE.
8. [Burning Glass 2018] Labor Insight, Burning Glass Technologies, http://burningglass.com [last access: 2018]
9. [D3 2018] D3, Data Driven Documents, https://d3js.org/ [last access: 2018]
10. [Davenport and Patil 2012] Davenport, T. H., & Patil, D. J. (2012). Data scientist. *Harvard business review*, *90*(5), 70-76.
11. [Dykes 2016] Dykes, B. (2016). Data storytelling: The essential data science skill everyone needs. Forbes.
12. [EMSI 2018] EMSI, https://economicsmodeling.com, 2018
13. [Kandel *et al* 2012] Kandel, S., Paepcke, A., Hellerstein, J. M., & Heer, J. (2012). Enterprise data analysis and visualization: An interview study. *IEEE Transactions on Visualization and Computer Graphics*, *18*(12), 2917-2926.[Keshif 2018] Keshif Online, https://keshif.me/demo/VisTools [last access: 2018]
14. [Kart *et al* 2013] Kart, L., Heudecker, N., & Buytendijk, F. (2013). Survey analysis: big data adoption in 2013 shows substance behind the hype. *Gartner Report GG0255160.*
15. [LaValle *et al* 2011] LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT sloan management review*, *52*(2), 21.
16. [Meeks 2017]. https://medium.com/@Elijah_Meeks/2017-data-visualization-survey-results-40688830b9f2
17. [Ryan and Snow 2016] Ryan, L., & Snow, N. (2016). Supporting student success with intuitive, approachable data visualization. In K. A. Flores, K. D. Kirstein, C. E. Schieber, & S. G. Olswang (Eds.), *Supporting the Success of Adult and Online Students: Proven Practices in Higher Education*. San Bernardino, CA: CreateSpace Independent Publishing.
18. [Tableau 2018] Tableau for Teaching, http://www.tableau.com/academic/teaching [last access: 2018]
19. [Velez et al 2009] A Cognition-based framework for the development of visualization literacy, Ph.D. Thesis, Rutgers University.
20. [VTK 2018] The Visualization Tookit (VTK), http://www.vtk.org [last access 2018]

21. [Wixon *et al* 2014] Wixom, B., Ariyachandra, T., Douglas, D. E., Goul, M., Gupta, B., Iyer, L. S., ... & Turetken, O. (2014). The current state of business intelligence in academia: The arrival of big data. *CAIS*, *34*, 1.

AUTHORS' BIOS

Lindy Ryan is an award-winning Professor in the Dept. of Information Management and Business analytics at Montclair State University. She also teaches in the Masters of Professional Science program at Rutgers University and is an active researcher at the Rutgers Discovery Informatics Institute. She is the author of *The Visual Imperative* (2016) and *Visual Data Storytelling with Tableau* (2018). Contact her at lindy.ryan@rutgers.edu.

Deborah Silver Deborah Silver is a Professor in the Dept. of Electrical and Computer Engineering at Rutgers University and the Executive Director of the Professional Science Master's Program. Contact her at silver@jove.rutgers.edu.

Robert S Laramee is an Associate Professor of Data Visualization at Swansea University, Wales in the Department of Computer Science. Contact him at rlaramee@gmail.com.

David Ebert is a Professor of Electrical and Computer Engineering at Purdue University, a University Faculty Scholar, a Fellow of the IEEE, and Director of the Visual Analytics for Command Control and Interoperability Center, the Visualization Science team of the Department of Homeland Security's Command Control and Interoperability Center of Excellence. Contact him at ebertd@purdue.edu.

Contact department editor Theresa-Marie Rhyne at theresamarierhyne@gmail.com