Volume 14, No. 6
November 2016
ISSN: 1545-679X

INFORMATION SYSTEMS EDUCATION JOURNAL

In this issue:

The **Information Systems Education Journal** (ISEDJ) is a double-blind peer-reviewed academic journal published reviewed published by **ISCAP,** Information Systems and Computing Academic Professionals. The first year of publication was 2003.

ISEDJ is published online (http://isedj.org). Our sister publication, the Proceedings of EDSIGCon (http://www.edsigcon.org) features all papers, panels, workshops, and presentations from the conference.

The journal acceptance review process involves a minimum of three double-blind peer reviews, where both the reviewer is not aware of the identities of the authors and the authors are not aware of the identities of the reviewers. The initial reviews happen before the conference. At that point papers are divided into award papers (top 15%), other journal papers (top 30%), unsettled papers, and non-journal papers. The unsettled papers are subjected to a second round of blind peer review to establish whether they will be accepted to the journal or not. Those papers that are deemed of sufficient quality are accepted for publication in the ISEDJ journal. Currently the target acceptance rate for the journal is under 40%.

Information Systems Education Journal is pleased to be listed in the 1st Edition of Cabell's Directory of Publishing Opportunities in Educational Technology and Library Science, in both the electronic and printed editions. Questions should be addressed to the editor at editor@isedj.org or the publisher at publisher@isedj.org. Special thanks to members of AITP-EDSIG who perform the editorial and review processes for ISEDJ.

INFORMATION SYSTEMS EDUCATION JOURNAL

Editors

ISEDJ Editorial Board

Progression of a Data Visualization Assignment

Joni K. Adkins
jadkins@nwmissouri.edu
Mathematics, Computer Science, & Information Systems Department
Northwest Missouri State University
Maryville, MO 64468, USA

Abstract

The growing popularity of data visualization due to increased amounts of data and easier-to-use software tools creates an information literacy skill gap for students. Students in an Information Technology Management graduate course were exposed to data visualization not only through their textbook reading but also through a data visualization assignment. Through a review of current literature and practice, student feedback, and instructor experience, the data visualization assignment has been updated multiple times in an effort to best serve the students' learning needs. This paper provides a brief literature review on why data visualization is growing and why students should learn about it and then outlines the data visualization assignment through each iteration. Lessons learned from student feedback on a student survey and suggestions for future assignments are also included.

Keywords: data visualization, data viz, Tableau, interactive data visualization, story, assignment

1. INTRODUCTION

Data visualization (or data viz) is a broad term referring to both the visual representation of data and the study of the presentation of data in a visual way (Turban, Volonino, & Wood, 2013). Data viz can also be defined as "the presentation of information in graphical or pictorial form, such as dashboards, interactive reports, and interactive presentations" (Brands, 2014, p. 56). Data visualization is becoming more popular as companies and organizations have access to more data and better software tools to handle the data. The popularity and prevalence of data visualization compels information systems instructors to add data visualization instruction and assignments to course work to expose students to this information literacy skill. This paper discusses a data visualization assignment as it has progressed through four semesters.

2. DATA VIZ POPULARITY

"Gartner predicts that analytics will reach 50 percent of potential users by 2014. By 2020, that figure will be 75 percent, and we will be in a world where systems of record, systems of differentiation and systems of innovation are enabling IT, business and individuals to analyze data in a much denser fashion than before" ("Gartner Says Business Analytics," 2013, para. 3). Trends that are helping data visualization growth include visualization software that non-programmers can use, improved aesthetics in today's software, and growth in self-service and mobile applications (Turban et al., 2013). In addition, the existence of cheap data, cloud computing, and availability of software tools such as Hadoop are other forces that feeding big data and data visualization growth ("Visualizations Make Big Data Meaningful," 2014).

Data visualizations can replace cognitive calculations and improve understanding as people are able to see trends and patterns in the data (Heer, Bostock, & Ogievetsky, 2010). Individuals who use data visualization tools state that it would take approximately nine hours longer to detect trends and patterns if they did not use data visualization (SAP News, 2014). The visual forms and patterns allow users to link the pictures to the information and may bring "aha" moments faster than examining rows in a spreadsheet (Brands, 2014). Since working memory is limited, a

meaningful data visualization can provide a clear and straightforward way to communicate a message (Evergreen & Metzner, 2013).

Traditional pie and bar charts have been used for years in companies, but now there is a desire for more interesting and interactive displays of data ("Visualizations Make Big Data Meaningful," 2014). Many organizations are going beyond basic Excel functionality for their charts and graphs and using interactive data visualizations where users can interact, filter, display more or less detail, or select certain items for display (Janvrin, Raschke, & Dilla, 2014). In fact, interactive data visualization is experiencing more growth than traditional business intelligence tools with tools like Tableau, QlikTech, and Spotfire ("Gartner Says Business Analytics," 2013).

Another reason for the growth of data visualization is that it can be applied to any field, not just statistics or business (McCandless, 2014). Research on data visualization will show plenty of examples of data visualization used in health care, business, science, government, and many other areas.

3. NEED FOR INSTRUCTION

One of the steps in deploying visualizations is making sure the users have the correct skill set (Polsky, 2013). Kelly (2015) emphasizes that college graduates need skills in data visualization to be effective in today's business world. Womack (2014) agrees that data visualization is now a core information literacy skill. Many of today's college students have not yet been exposed to data visualization and the proper techniques in generating visual displays.

Current data visualization tools allow students to easily create visualizations which may or may not be appropriate for the data. Students should learn to evaluate the various types of visualizations available and justify whether it is a proper representation to use (Womack, 2014). Determining an appropriate data visualization is a challenge for creators (Heer et al., 2010). Creating a data visualization requires the user to determine the questions to ask, identify the correct data, and then select an effective visual display to represent the data (Heer et al., 2010). Data visualization principles include "striving for clarity, avoiding clutter, and emphasizing the most relevant data" (Womack, 2014, p. 15). Students can easily get caught up in making data visualizations that are attractive but might not represent the data correctly. For example, some students use a line graph on data that does not have any logical connection. Instead a bar chart is generally a better choice. In addition, default settings in software programs may overcomplicate the graph by including gridlines, details, extra labels, or distracting colors (Evergreen & Metzner, 2013).

Data visualizations can be very pleasing to the eye, but students should remember that data visualizations are a tool to be used to better understand the data (Womack, 2014). Since most college students have not been exposed to graphic or data visualization principles, adding a component on data visualization can be beneficial to the student and potential employers.

4. ASSIGNMENT PROGRESSION

Information Technology Management is a graduate course for MBA and M.S. Applied Computer Science students. Students in this course are exposed to data visualization in their textbook so it was logical for them to practice the concepts with an assignment. This course is taught online, blended, and in a traditional format. Approximately 350 students take this course each year. The course does not have any pre-requisites so students can take it at any time during their graduate studies. The data viz assignment is approximately 10 percent of their overall grade.

The first time the data visualization assignment was attempted was during a summer blended class. The class meetings were reserved for spreadsheet activities, case presentations, library instruction, and exams so there was not time to include data visualization during a class session. The students were to watch the TED talk by David McCandless titled "The Beauty of Data." This video does an excellent job of showing students the power of data visualization. Then the students were to read the article "A Tour Through the Visualization Zoo" by Jeffrey Heer, Michael Bostock, and Vadim Ogievetsky. The article explains different types of charts to use for different kinds of data. Next the students found a data visualization to share with others. The students then participated in a threaded discussion when they discussed the video, posted the data visualization they found and discussed what type it was, and then suggested how a data visualization could be used with a company they had previously studied in a case assignment. The

assignment went okay, but we wanted them to practice creating a data visualization in the next iteration.

The assignment was significantly changed for the fall course offering. The fall course was a 15-week traditional course meeting twice a week. The students had approximately three weeks to work on the visualization assignment. The assignment was done in groups of 4-5 students where one group member located data and installed Tableau desktop on their notebook computer. Tableau software "helps people to see and understand data" ("Tableau Software," 2015). Tableau is a free data visualization program for students. Students can load Excel files into Tableau where they can create the data visualizations. Each student in a group created one chart in Tableau. As a group, they created a story and a dashboard with their Tableau charts. In addition, the student group was to make an infographic using Piktochart, another free tool. While data visualization is a singular view of data, an infographic is a collection of visualizations that tell a story (McCandless, 2014). The student teams then did a presentation in class to display their Tableau charts and Piktochart infographic. Students were surveyed at the end of the class, and they shared that the most frustrating part of the assignment was that the software and files were on one person's computer so that one student seemed to do the majority of the work. Also since the students were responsible for finding their own data, the team member who found the data set tended to have more knowledge and interest in the data set. Most student teams used data sets available from their jobs on campus.

The instructor used the feedback from the fall assignment and made the spring data visualization assignment an individual assignment to remove the barrier of sharing one computer. Tableau was selected as the software, and an infographic was not required. A scoring guide was created where the students typed their responses to the assignment and submitted it as part of the grade. The first due date of the assignment required students to find a data set, clean it up (if necessary), watch at least 5 videos on Tableau and include the links and note which video was most helpful on their sheet, and write three questions they wanted to answer with the data set. The second part of the assignment was to create three charts to answer the three questions. The students had to justify and provide an explanation why they chose the data visualization they did. The students also created a story and a dashboard. The latest version of the assignment sheet is shown in Appendix A.

Instead of class presentations, students met individually with the instructor for a 10-minute demonstration. Interaction is an important part of data visualization (Kosara & Mackinlay, 2013). The demonstration allowed the students to review their story and then the instructor could ask questions so the students could show how they could interact with the data. The student brought their notebook computer with Tableau open and their charts, story, and dashboard to the instructor's office where they sat together and went through the assignment sheet. The downside to interaction is it can interfere with a story (Kosara & Mackinlay, 2013). When the instructor asked a question during the story, sometimes the student would add a filter or make a change that would impact the story. A better technique is to wait for the end of the story to begin interaction (Kosara & Mackinlay, 2013).

At the end of the meeting, each student was asked what the most challenging part of the assignment was. The most common response was finding the data set. The students had been provided a list of web sites where they could get data sets, but they often had to search to find a data set in an Excel format.

The next class offering was the 6-week blended summer class. We made some changes to allow the new data visualization assignment to fit into the course schedule. The day the assignment was initiated, the students watched the TED video in class so they all were introduced to data visualizations. Each instructor provided an Excel data set to each course section so they did not have to find their own data set. The data sets came from data.gov, and students were provided the URL to the data set so they could learn more about the data. Some of the students in the spring course had used the filter feature to make their data visualizations more interactive; therefore, using a filter was added as a requirement for the summer class. The other parts of the assignment remained similar to the spring assignment as there were again two due dates to force the students to install the software, watch videos, and determine their questions in the first week of the assignment. Again students signed up for a 10-minute demonstration with the instructor. At the end of the summer data visualization assignment, a survey was made available to the students so the instructors could

see where the students were starting with their data visualization experience and to learn what should be kept and/or changed about the assignment.

5. RESULTS

259 students enrolled in Information Technology Management in summer 2015 participated in the voluntary survey. They were awarded two extra credit points for their participation.

The first question asked about their prior experience with data visualization. Forty-eight percent had no previous experience with data visualization. Approximately 45 percent of the participants had visualization experience with Excel charts and graphs. Other products that students also indicated some experience with were Tableau (11 percent), Weka (9 percent), SAS or R (3 percent), Qlik (2 percent), Business Objects, Cognos, Hyperion, and Microsoft SSRS tool (all less than one percent). Students were asked to rank their prior experience with data visualization tools and the average score was 3.05 on a 7-point scale. Overall, these students had limited or no experience with data visualization.

The participants were asked if watching the TED data visualization video was helpful or not helpful to their understanding of data visualization. Ninety-eight percent indicated that watching the video was helpful. One student commented that the TED video showed creative ways to visualize data, it created enthusiasm in me."
Participants indicated the likelihood of using Tableau again by with 1 being extremely unlikely and 7 extremely likely. The average score was 5.94 on the 7-point scale. Participants were more likely to recommend Tableau to others with an average score of 6.27 on a 7-point scale.

Participants were then asked if they preferred a 10-minute demonstration with the instructor or a class presentation to showcase their data visualization. Sixty-six percent preferred the demonstration while 34 percent would rather give a presentation in front of the class. One participant commented that "the one-on-one session was really helpful in knowing the viewpoint of people when we represent something on Tableau. I learned how to think in the viewer's perspective." Another participant shared "The most rewarding part of assignment is presenting the viz created. That 10 minutes where we presented our visualization is when we had to convince why the visualization suited the data."

The data set was given to the summer class sections so they had to use the data set provided. The participants were asked if they would rather find their own data set. Only 12 percent indicated they would prefer to find their own data set for the assignment. The last two questions were open-ended and allowed participants to write about the rewarding and challenging parts of the assignment. The responses were read and analyzed to find common themes.

The most common challenge was forming the questions to ask about the data as participants indicated "to frame the questions from a large set of data was really challenging." There are several possible reasons for this including the data set is not familiar to the students, and students were required to write the questions before they did the data visualizations. Some wanted to change their questions later after they started working on the data visualizations. Another challenge that participants included was learning Tableau as there were no instructor-provided materials. One participant commented "As I am new to this software, there was steep learning curve at the beginning of the assignment. But after some hours, I got used to it. Now I love it." They were to use web resources from Tableau and YouTube to learn how to use Tableau. There are lots of good learning materials available online. The third top challenge was selecting an appropriate data visualization for the data displayed. The fourth challenge was understanding the data set. Other challenges listed were analyzing the data, cleaning up the data, telling the story, and using filters.

The most mentioned rewarding part of the data visualization assignment was working with Tableau. As indicated previously, students had to learn Tableau on their own, but after learning it, many saw learning the new software as a positive. The second positive was the general comment of learning to visualize data. The third most common rewarding point was learning something new. One student commented that the most rewarding part was "self-learning of the tool, which gave me a confidence to learn any type of tool in the coming future" while another wrote "this assignment gave me confidence that I can work on my own with the help of online information and succeed in any tasks which I have no idea." Other rewarding points of the assignment were the good output, displaying data

in different formats, the dashboard, the story, and using filtering. In general, participants seemed to be proud of their progress made in data visualization and in their final Tableau dashboard and story.

6. LESSONS MOVING FORWARD

One lesson learned is that the data visualization assignment should remain as part of the class. Given the limited exposure students have to data visualization prior to the class and the need for students to be familiar with data visualization techniques (Kelly, 2015; Womack, 2014), the data visualization assignment does provide an introduction to the emerging data visualization topic.

The students in the summer class were almost all students in the applied computer science master's program so their career goal is usually software development. This assignment was introduced as one where they would have to learn the software on their own, just as they will have to do throughout their career. Given that a number of participants commented that their confidence in learning new software tools increased with the assignment, there are no plans to make or provide additional tutorials or videos to students. The students should be able to use existing materials to learn about data visualization.

The use of the TED video on data visualization was justified since nearly all students found the video beneficial. Also the decision to change the assignment to a demonstration instead of a class presentation was confirmed. The time commitment for an instructor is generally greater than a class presentation but does give the chance to meet one on one with students and provide immediate feedback on their assignment. Since data visualization is about engaging with the data for increased understanding (Evergreen & Metzner, 2013), having the chance to talk one on one with the student lets them showcase their story and engage with the instructor. While a small percentage of students wanted to find their own data, most were satisfied with having the data set provided to them. Participant feedback verified all of these assignment conditions.

Since forming the questions was the most challenging part of the assignment for many students, further examination is required. One solution could be to have the questions be part of the second part of the assignment so students could essentially change the question after they have created their data visualizations. Another solution would be to spend some time in class reviewing the data and having the students generate multiple questions and then determine which one could be answered with the data. This process could help them develop a better story for the data. A drawback could be that several students would end up with the same questions. A third option would be to have a preliminary assignment with a small data set to give students some practice forming questions. The data sets provided were not related to the course or the information technology industry. Perhaps if the data were related to a student's interest or area of study, he or she would be able to generate better questions.

In the first iteration of this assignment, students were required to read an article that discussed different kinds of data visualizations and when they were appropriate to use. The assignment sheet the students currently fill out requires them to discuss this but they were not required to read any articles about selecting an acceptable data viz. Students could again be required to read about various types of data visualizations. This could help them overcome the challenge of determining the correct data viz.

7. CONCLUSION

The need for better data skills continues as "86 percent of business decision makers say they believe all employees in a company will eventually need to be 'data geeks,' meaning they will require skills to analyze company data and make decisions based on that analysis" (SAP News, 2014, para. 1). Data visualization is a current topic that deserves attention in an information systems class as graduates will be expected to interact with and understand data as a basic element of information literacy (Womack, 2014).

Information systems educators realize that the content and tools in our classroom will constantly change. As a result, the assignments used to teach and assess students must also adapt to meet the changing needs of students and their future employers. The data viz assignment will continue to evolve as new tools are developed and student skill levels increase. Feedback learned from the participants in the summer study will impact the assignment in future semesters. Teaching data visualization in a completely online class will begin in the Fall 2015 term. The Blackboard Collaborate software

currently used for group presentations will be used for the student demonstration. The instructor will be able to talk with the student online and the student will be able to share their desktop so they can display their data visualizations.

Educators and students realize that practice helps develop knowledge and skills so this experience with data visualization creation will help develop skill and wisdom in making good visualizations (Womack, 2014). As one participant noted, "Learning new techniques and the usage of Tableau in various ways was a rewarding experience." While the assignment does not attempt to cover all aspects of data visualization, it does provide an introductory experience so the students can begin practicing this information literacy skill.

8. REFERENCES

Brands, K. M. (2014). Data Visualization and Discovery. *Strategic Finance*, *96*(12), 56–57.

Evergreen, S., & Metzner, C. (2013). Design Principles for Data Visualization in Evaluation: Design Principles for Data Visualization. *New Directions for Evaluation*, *2013*(140), 5–20. http://doi.org/10.1002/ev.20071

Gartner Says Business Analytics Will Be Central for Business Reinvention. (2013). Retrieved July 6, 2015, from http://www.gartner.com/newsroom/id/2510815

Heer, J., Bostock, M., & Ogievetsky, V. (2010). A Tour Through the Visualization Zoo. *Communications of the ACM*, *53*(6), 59–67. http://doi.org/10.1145/1743546.1743567

Janvrin, D. J., Raschke, R. L., & Dilla, W. N. (2014). Making sense of complex data using interactive data visualization. *Journal of Accounting Education*, *32*(4), 31–48. http://doi.org/10.1016/j.jaccedu.2014.09.003

Kelly, S. (2015). Teaching Infographics: Visually Communicating Data for the Business World. *Business Education Forum*, *59*(3), 35–37.

Kosara, R., & Mackinlay, J. (2013). Storytelling: The Next Step for Visualization. *Computer*, *46*(5), 44–50. http://doi.org/10.1109/MC.2013.36

McCandless, D. (2014). Who Doesn't Like a Good Data Visualisation? *Creative Review*, *34*(1), 41–44.

Polsky, A. (2013). Data Visualization: 7 Considerations for Visualization Deployment. International Institute for Analytics. Retrieved from http://resources.idgenterprise.com/original/AST-0116272_DataViz_7_considerations.pdf

SAP News. (2014). SAP-Sponsored Survey Finds Business Decision Makers Struggle to Unlock the Power of Big Data. Retrieved July 6, 2015, from http://www.news-sap.com/sap-sponsored-survey-finds-business-decision-makers-struggle-unlock-power-big-data/

Tableau Software: We help people see and understand data. (2015). Retrieved August 20, 2015, from http://mission.tableau.com

Turban, E., Volonino, L., & Wood, G. R. (2013). *Information Technology for Management: Advancing Sustainable, Profitable Business Growth* (9 edition). Hoboken, NJ: Wiley.

Visualizations Make Big Data Meaningful. (2014). *Communications of the ACM*, *57*(6), 19–21. http://doi.org/10.1145/2601074

Womack, R. (2014). Data Visualization and Information Literacy. IASSIS Quarterly, 38(1), 12–17.

Appendix A

Data Viz Assignment – Summer 2015

To get started, download Tableau, choose and watch your training videos, and get familiar with your data set.

Questions	Type your answer in this column. The shaded boxes do not need to be filled in.	Due date	Points
What training videos did you watch? Include the URLs to the videos. These can be from Tableau or from YouTube. You should watch at least 5 training videos. Which one was the most helpful?		5/28	4
What data set did you use? What "cleaning up" or changes did you have to perform?		5/28	4
What questions are you trying to answer/display in your visualizations? Write these as specific questions.		5/28	4
Chart 1: What type of viz did you create? Why did you select the data and the viz that you did?		6/3	4
Chart 2: What type of viz did you create? Why did you select the data and the viz that you did?		6/3	4
Chart 3: What type of viz did you create? Why did you select the data and the viz that you did?		6/3	4
Create your story		6/3	4
Create your dashboard		6/3	4
Demonstration to instructor. Be prepared to talk about your data and your choices and demonstrate your filter(s).		6/3	8
Total points			40