Usability of Data Visualization Activity Worksheets in the Context of a Critical Data Visualization Workshop: Findings from a Usability Survey

Dr. Vetria Byrd PhD, Purdue University-Main Campus, West Lafayette (College of Engineering)

Dr. Vetria Byrd is an assistant professor in the Department of Computer Graphics Technology in the Polytechnic Institute at Purdue University in West Lafayette, Indiana. Dr. Byrd is the founder and organizer of the biennial Broadening Participation in Visualization (BPViz) Workshop. Dr. Byrd has given numerous invited talks on visualization and has been featured in HPC Wire online magazine (2014), and numerous workshops nationally and internationally. Dr. Byrd received her graduate and undergraduate degrees at the University of Alabama at Birmingham, in Birmingham, Alabama which include: Ph.D. in Computer and Information Sciences, Master's degrees in Computer Science and Biomedical Engineering and a Bachelor's degree in Computer Science. Dr. Byrd's research interests include: data visualization, data visualization capacity building, high performance visualization, big data, and collaborative visualization. Dr. Byrd's research utilizes data visualization as a catalyst for communication, as a conduit for collaboration, as a pathway to STEM and as a mechanism for broadening participation and inclusion.

Dr. Kendall Roark, Purdue University at West Lafayette

Kendall Roark is an applied cultural anthropologist who engages in ethnographic fieldwork and anthrodesign projects in Canada and the United States. Her research and teaching interests focus on participatory and speculative design, queer and feminist technoscience studies, and data ethics. Dr. Roark is the cofounder and faculty lead for the Critical Data Studies Collective at Purdue University.

Brent T. Ladd, Center for Science of Information, Purdue University

Brent Ladd serves as Director of Education (and Interim Director of Diversity) for the Center for Science of Information NSF Science and Technology Center based at Purdue University. His education outreach activities and research focus on interdisciplinary collaboration among students and young scholars and evaluating learning outcomes for science-based training programs.

Usability of Data Visualization Activity Worksheets in the Context of a Critical Data Visualization Workshop: Findings from a Usability Survey

Abstract

This evidence-based practice paper presents outcomes on workshop participants' perception of the usability of data visualization worksheets utilized in a critical data visualization workshop. Training students to think critically about data is a process that takes practice. Thinking critically about data requires understanding the data, its provenance and context, and the best way to represent insights contained in the data. A 5-day workshop on teaching critical and ethical approaches to data visualization was hosted by the Center for Science of Information at Purdue University in June 2019. The motivation for the workshop was to provide an introduction to data science through the lens of critical data visualization. In addition to instruction and hands-on data visualization and data manipulation exercises supported by activity worksheets, workshop participants were exposed to small group discussion and active-learning focused team science and data visualization ethics modules. This three-prong approach provides the context in which the data visualization activity worksheets were utilized.

This paper focuses solely upon the results of a usability survey of data visualization activity worksheets to introduce the data visualization process and reports students' perception of the worksheets to introduce the data visualization process to novice users. Overall participant perception of the worksheet method was positive: 61% thought the method was easy to use, felt most people would learn to use the method quickly, felt confident using the worksheet method after using the worksheets for multiple visualization challenges and indicated they would like to use the method frequently when applying the data visualization process. This work is funded by The National Science Foundation (Award #CCF-0939370). Next steps involve integrating team science and data visualization ethics activities into future iterations of the data visualization activity sheets and usability survey.

Keywords: Collaboration, data visualization, workshop, data visualization activity worksheets, usability

Introduction

This work reports participants' perception of a data visualization activity worksheet method designed to introduce data visualization to participants of a workshop on critical and ethical approaches to data visualization during the Summer of 2019. The activity worksheet method consists of seven worksheets, one for each of the 7 stages of the data visualization process. The activity worksheet approach and subsequent worksheet usability survey is informed by evidence-based approach to data visualization education. We contend that this unique approach facilitates critical thinking about the data visualization process [1]. The method utilized, worksheets and a usability survey, was developed by Byrd [2] and was designed to introduce the data visualization process to novice visualizers in a short period of time,

Background and Context of the Workshop

With a goal of training the next generation of interdisciplinary scholars, an annual engaged learning summer workshop was designed to introduce diverse cohorts of students to data science techniques while providing positive interdisciplinary research team experiences [3]. The

motivation for the workshop in 2019 was to provide an introduction to data science through a focus on critical and ethical approaches to data visualization. In addition to instruction and hands-on data visualization exercises supported by visualization activity worksheets, workshop participants were exposed to small group discussion and active-learning modules focused on team science (one half-day equivalent) and data visualization ethics (two half-day equivalents). While this three-prong approach provides the context in which the data visualization activity worksheets were utilized, the focus of this paper is the attempt to empirically test a novel process [2] of learning data visualization through specially designed activity worksheets utilized throughout the hands-on data visualization component of the workshop. In an upcoming iteration of the workshop, the authors plan to test the feasibility of integrating data visualization ethics and team science components into the design of the data visualization activity worksheet.

Data Visualization Activity Worksheets

The design process is a non-trivial process which requires more than one week to master. In the interest of introducing a complex process in a short period of time, a series of data visualization activity worksheets were used to facilitate the process. The data visualization worksheet approach has been used in many platforms for teaching the design process for data visualization [4] - [5]. The goal of these methods are to develop data visualization systems. It was beyond the scope and unrealistic to expect a visualization system as a deliverable for this workshop. The Data Visualization Activity Worksheet Method, utilized in the workshop is designed for novice users who are new to data visualization. The aim of the workshop was to give participants a strong foundation in visualization principles, design and ethics so that they would continue to use these skills after the workshop in their academic and professional practices. The scope of this paper focuses on students' perception of the usability of the data visualization activity worksheets.

The Data Visualization Worksheet method consists of multiple worksheets – each worksheet represents one stage of the data visualization process. The worksheets allow for the introduction of data visualization in a progressive way; the results from one worksheet serves as input for the next worksheet. There are specific tasks associated with each stage of the visualization process. The worksheets are designed to help users capture key components needed for each stage. Figure 1 shows examples of the worksheets for three of seven stages of the visualization process: acquire, parse and mine. Goals, and objectives are included in each worksheet to inform the user of the expected outcomes for each worksheet. The data visualization process is an iterative, non-linear, dynamic process, that can be overwhelming for a novice visualizer. The worksheets foster a progressive path of completion with guidance along the way regarding what to look for and what key elements students should focus on. Outcomes of each worksheet are pieces to a larger puzzle that, once compete, will provide insight into what the underlying data represents. The worksheets also help students to identify skills they need to strengthen (for example, students with little or no programming background, tend to find the parsing worksheet challenging) or call to their attention aspects of the process they would like to explore more to build their data visualization capacity skills.

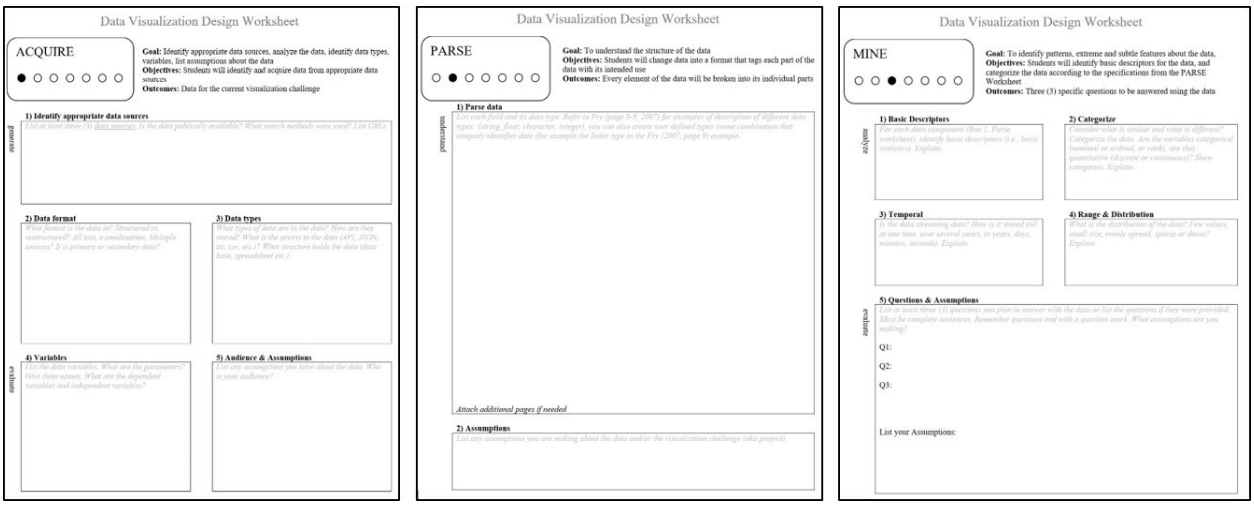

Figure 1 Data Visualization activity worksheets for three stages of the data visualization process. Each worksheet captures information specific to each stage of the process. Layout adapted from [4-5].

Usability Survey for Data Visualization Activity Worksheets

To assess the usability of the Data Visualization Activity Worksheet Method the System Usability Scale (SUS) reported by [6] was adapted. SUS is a reliable, low-cost usability scale that can be used for global assessments of systems usability [6]. To prevent response biases caused by respondents not having to think about each statement, the SUS model alternates positive and negative items to ensure that participants make some effort to think whether they agree or disagree with a statement. The SUS Usability Scale and questions were found to be appropriate for assessing the usability of the data visualization activity method and were adapted in the work. The SUS questions were revised by replacing, where needed, the word "system" with the word "method," and the word "functions" with "various parts."

After completing all of the worksheet activities, participants were asked to complete the survey on their perceived usability of the Data Visualization Activity Worksheet Method. Participants were asked to indicate the level of agreement with 11-statements (Table 1) using a five-point Likert scale: strongly agree (5), agree (4), neutral (3), disagree (2), and strongly disagree (1).

Table 1 Usability statements.

Statement Number	Please rate your agreement with the following statements
1	I think that I would like to use the Data Visualization Activity Worksheet Method frequently for applying the data visualization process.
2	I found the method unnecessarily complex.
3	I thought the method was easy to use.
4	I think that I would need the support of a technical person to be able to use the Data Visualization Activity Worksheet Method.
5	I found the various parts in the method were well integrated.

6	I thought there was too much inconsistency in the Data Visualization Activity Worksheet Method.
7	I would imagine that most people would learn to use the Data Visualization Activity Worksheet method very quickly.
8	I found the Data Visualization Activity Worksheet method very cumbersome to use.
9	I felt very confident using the method the first time.
10	After using the worksheets for multiple visualization challenges, I felt very confident using the method.
11	I need to learn a lot of things before I could get going with the Data Visualization Activity Worksheet method.

Workshop Overview

The workshop enables students from all backgrounds, domains, and academic majors access to learn the process of visualizing data in a scientific and ethically sound manner – becoming "Agents of Insight" in the world of data science.

Participants of the workshop learned through hands-on exercises in interdisciplinary teams and worked individually to reinforce concepts covered in short lectures by working through the key stages of visualizing data using Tableau and Excel software tools. Participants concurrently examined case studies and engaged in hands-on creative data and visualization activities informed by Critical Data Studies, critical geography and data journalism in order to develop frameworks for ethical and just data visualization practices that can be applied to their own research [7] – [12].

Workshop Demographics

Student diversity in the traditional STEM areas of computer science and engineering is low [13] as evident by the National Science Foundation's charge to focus on increasing female participants as a core diversity goal [14]. In addition to achieving gender balance, diverse student backgrounds, experience, and institutional breadth was a priority for recruitment for the workshop. Students were recruited in light of the overall learning objectives and diversity goals of the training. The workshop attracted a cross-section of participants who are interested in learning data visualization and data science skills in general, as well as in experiencing interdisciplinary team collaborations. Funding was provided for students to travel and attend the workshop. Twenty-five attendees participated in the week-long workshop.

Workshop Elements

Pre-workshop tasks included completing several online tutorials, readings, and instructions for Tableau software, watching the "Introduction to Data Visualization" video [15], and a "Data Viz and Misrepresentation" video. Elements of the 5-day workshop included: lectures, hands-on activities, a studio component, readings and videos. The introduction of the data visualization process described by [2], [16] was adapted from a 16-week format to accommodate a shorter introductory period. Hands-on activities followed lectures to reinforce lecture content. The

workshop agenda, shown in Table 2, details workshop activities designed to introduce the data visualization process using the widely recognizable Walmart Recruiting – Store sales forecasting data [17] as the training dataset.

Table 2 Workshop Agenda and Activity.

Day	Lecture Topics	Activity
1	Introduction to the acquire, parse and mine stages of the data visualization process, using training data	Acquire, parse and mine the training data; teams meet to brainstorm about ideas and topics for team projects
2	AM: Revisit mine stage, introduction to the filter and represent stages of the data visualization process; introduce data visualization check list. PM: Data and Platform Bias; acquire group project data	Look for trends in the training data, pose questions to answer, generate first representation of the data. teams meet to discuss and prepare proposal for group project, teams research topic, look for and acquire preliminary data, generate basic visuals;
3	Introduce the importance of feedback in the data vis process. PM: Intro to Critical DataViz/Design Values; Intro to Critical DataViz/Epistemic Challenges	Peer and self-critique using the data visualization checklist as a rubric; teams come to consensus on group topic, and prepare proposal.
4	Visualizing principles: color, perception, best practices; Designing for impact; create project related visualizations	Review and combine feedback and self-critique and refine visualizations; teams continue to finalize proposal.
5	Team presentations of preliminary findings and interests..	Teams decide if they will submit an optional, full proposal to continue to develop the project after the workshop concludes.

Participants also completed examples and practiced material covered during the lecture on their own after daily events of the workshop concluded each day.

Group Projects

The workshop cohort involved 6 teams, with 3-5 members per team. Members of teams were selected to comprise broad interdisciplinary perspectives, with students and postdocs from multiple institutions, gender and racial diversity, as well as a mix of graduate, undergraduate, and postdocs, as these multiple levels of diversity combined have shown consistent positive outcomes [17]. During the pre-workshop period students read about and were asked to prepare to engage in best practices for successful team science [18].

During the workshop, teams met and worked on their projects in the afternoons and evenings following morning training sessions. Team members spent a great deal of time together discussing their projects and the various approaches and potential methods. The space created for this experience emphasized the creative wisdom that each student brought to the process. They were not only allowed, but encouraged, to explore new questions, ways of thinking, and use of software tools. The workshop week culminated in team project presentations where each team presented their overarching problem/topic. For the final project, students combined all their

acquired skills to create innovative, effective visualizations which tell a story about a data set of their choice [19]. This included the methods they have used to analyze data, results gained thus far, and any plans for future collaborations to continue the project. The teams fielded questions from the audience which provided additional insights.

Results

A total of 23 participants completed the usability survey. Statistical indicators (mean, median, and mode) calculated from quantified Likert scale responses are shown in Table 3 for each of the eleven statements from the usability survey instrument.

Interdisciplinary Team Context. The interdisciplinary team project context allowed students to experience bridging across disciplines, identifying and coordinating team member strengths, while learning to think about problems broadly. A second survey instrument was used to capture participant responses about the team-based experience. Participant results ($n=22$) from their overall team-based experience in the workshop showed that 87% reported "Good" or "Excellent" outcomes from their team collaborations and 91% reported "Good" or "Excellent" results related to team presentations and deliverables. Ninety-one percent reported "Strongly Agree" or "Agree" with the statement, "I gained an improved understanding and experience in approaching a data problem within an interdisciplinary team," the remaining 9% "Disagreed" with the statement. These results show that the positive outcomes of the team-based experience had no negative impact on the use of the data visualization worksheets.

Table 3 Statistical indicators (* indicates the value was rounded up).

Statement		**Mean**	**Median**	**Mode**
S1	I think that I would like to use the Data Visualization Activity Worksheet Method frequently for applying the data visualization process.	3.6*	4	4
S2	I found the method unnecessarily complex.	3.0*	3	2
S3	I thought the method was easy to use.	3.5	4	4
S4	I think that I would need the support of a technical person to be able to use the Data Visualization Activity Worksheet Method.	2.3*	2	2
S5	I found the various parts in the method were well integrated.	3.8	4	4
S6	I thought there was too much inconsistency in the Data Visualization Activity Worksheet method.	2	2	2
S7	I would imagine that most people would learn to use the Data Visualization Activity Worksheet method very quickly.	3.6	4	4
S8	I found the Data Visualization Activity Worksheet method very cumbersome to use.	3.0*	3	4
S9	I felt very confident using the method the first time.	3.3	3	4

S10	After using the worksheets for multiple visualization challenges, I felt very confident using the method.	3.6	4	4
S11	I need to learn a lot of things before I could get going with the Data Visualization Activity Worksheet method.	2.4*	2	2

Discussion

The usability survey of the Data Visualization Activity Worksheet was modeled the System Usability Scale (SUS) created by [6]. The approach provided by [6] was adapted to capture feedback from users regarding the Data Visualization Activity Worksheet method (henceforth referred to as the DVAW-SUS). The DVAW-SUS assessment tool was refined to include follow-up questions if respondents' answers were considered more negative than positive. Responses were categorized as positive, neutral or negative.

The SUS Scale was originally developed by Brooke (1996) as a "quick and dirty" survey scale that would allow the usability practitioner to quickly and easily assess the usability of a given product or service. The survey provides a single score on a scale between 0 and 100 where higher scores indicate better usability [20]. Because the statements alternate between positive and negative, care must be taken when scoring the survey. See Brooke (1996) for details of scoring the SUS. For this work the questions from the SUS survey were modified for current project; however, a 5 point Likert scale, ranging from strongly disagree to strongly agree, was used to capture level of agreement from participants. Questions from the SUS Scale were adapted in the context of viewing the data visualization activity worksheet method as a system; however, results from responses are presented as the collection of responses and not in terms of each individual perception of the system.

Responses were grouped into usability and learnability subscales. Ninety-two responses for each subscale were coded and used to calculate the mean. Because the statements alternate between the positive and negative, care is taken when calculating the mean [20] for each subscale. The latest research suggests that negatively worded items should not be used because they measure different constructs [22]. To address this issue, items from even numbered statements were reverse scored before calculating the mean score. The first subscale reflects responses related to the participants' perception of the usability of the method. For this subscale, the mean score was calculated from responses to statements 2, 3, 9 and 10. The mean value for the usability subscale is 3.43, indicating on average, participants were neutral. The mode and median values for the subscale is 4, indicating most of the responses tended to be more positive (agree). Participants did not find the method to be unnecessarily complex, thought the method was easy to use, felt very confident using the method for the first time and after using the worksheet method for multiple visualization challenges.

The second subscale reflects responses related to the participants' ability to learn to use the method. For this scale, the mean score was calculated from responses to statements 4, 6, 7, and 8. The mean value for the learnability subscale is 3.6 (rounded up), indicating on average,

participants found the method easy to learn and use. The mode and median values for the subscale is 4, indicating most of the responses tended to be more positive (agree). Participants did not feel technical support is needed to be able to use the method, there is consistency across the seven data visualization worksheets, found the method to be easy to use and indicated most people would learn to the use method very quickly.

Mean values from statements 1, 5 and 11 (Table 3) support the authors' aim to create a method that participants would frequently use when applying the data visualization process, find the various parts in the method well integrated, and would find the method intuitive, not requiring knowledge about a lot of things before being able to implement the method. The design data visualization worksheet approach has been used in many platforms for teaching the design process for data visualization [4] - [5]. These methods assume some knowledge of data preprocessing and focus on developing data visualization systems. The approach used in this workshop includes a deeper look at the preprocessing stages of the data visualization process.

Limitations

Although the overall response to the data visualization activity worksheets was positive, the following limitations are acknowledged: (1) a lot of new material was covered in a short period of time, (2) basically, a 16-week data visualization course was scaled down to 1-week which equated to 4 days of lecture/hands-on activities, (3) due to the interdisciplinary nature of the workshop, participants were from various disciplines and academic backgrounds, (4) participants arrived with little knowledge of the data visualization process, and (5) students were asked to quickly get to know and work with their team members, while also navigating potential projects and the relevant data needed to start a project. The data visualization worksheet method was introduced to help mitigate the intricate process of visualizing data.

Conclusions & Outcomes

Teaching critical and ethical approaches to data visualization within the context of interdisciplinary teams is a powerful and effective model, as it infuses the learning process with diversity of experience, background, knowledge, and skills that are shared and integrated. This fosters a broader, but also deeper, insight into the data, creative and cogent questions of the data, and potential pathways of analysis and visualization. Overall the response to the novel data visualization activity worksheet process was positive; however, there is room for refinement of the method. The data visualization process, much like the engineering process, is an intricate process with intrinsic characteristics that require critical thinking, and problem solving skills. A method that facilitates the practice of transferrable skills will benefit students beyond the classroom into research practices and daily tasks. Six of the participating students continued their collaborations after the workshop and were involved in project results presented at a regional STEM conference in January of 2020, and are currently preparing to present at a national conference in the summer of 2020.

Future work will include refining the data visualization activity worksheets to incorporate suggestions from participants on how to make the method more intuitive and easier to use, more fully integrating team science and data visualization ethics into the data visualization activity worksheets and developing additional strategies for capturing impact of the critical and ethical data visualization process either in the usability survey or by other means. Co-authors discussed how the final visualization project itself might be analyzed to gain insight. We also discussed how impact might be captured longitudinally over the twelve-month period following the workshop.

References

[1] J. A. Koenig, "Assessing 21st Century Skills: Summary of a Workshop," 2015. [Online]. Available: https://www.learntechlib.org/p/159080/. [Accessed January 21, 2020].

[2] V. Byrd, "Introducing Data Visualization: A Hands-on Approach for Undergraduates," in *Proceedings of E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, Las Vegas, NV, USA, November 4-7, 2018, Association for the Advancement of Computing in Education (AACE), pp. 730-73, 2018.

[3] B. T. Ladd, "The Information Frontiers Program: Expanding Student Capacity for Crossing Domain and Institutional Borders," in *Association for Interdisciplinary Studies 41st Annual Conference*, Oct. 23-26, 2019, University van Amsterdam, The Netherlands, 2019.

[4] J. C. Roberts, C. Headleand, and P. D Ritsos, "Sketching Designs Using the Five Design-Sheet Methodology." *IEEE Transactions on Visualization and Computer Graphics,* vol. 22, no. 1, pp. 419-28, 2016.

[5] S. McKenna, L. Alexander, and M. Meyer. "Worksheets for Guiding Novices through the Visualization Design Process." ArXiv.org, p. 17+, 2017. [Online]. Available: https://arxiv.org/pdf/1709.05723.pdf. [Accessed January 30, 2020].

[6] J. Brooke, "SUS-A quick and dirty usability scale," *Usability Evaluation in Industry*, 189(194), 4-7, 1996.

[7] A. Cairo, "Ethical Infographics," *IRE Journal*, vol. 37, no. 2, pp. 25–27, 2014.

[8] C. D'Ignazio, "Creative data literacy: Bridging the gap between the data-haves and data-have nots," *Information Design Journal*, vol. 23, no. 1, pp. 6-18, 2017.

[9] C. Dalton, and J. Thatcher, "What does a critical data studies look like, and why should we care? Seven points for a critical approach to 'big data'" *Society + Space*, vol. 29, May 12, 2014.

[10] L. Floridi, and M. Taddeo, "What is data ethics?" *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, p. 374, 2016.

[11] K. Hepworth, "Big data visualization: promises & pitfalls." *Communication Design Quarterly Review*, vol. 4, no. 4, pp. 7-19, 2017. https://dl.acm.org/citation.cfm?id=3071090

[12] K. Tanyoung, and C. DiSalvo, "Speculative visualization: A new rhetoric for communicating public concerns," in *Proceedings of the Design Research Society (DRS) international conference design & complexity*. 2010. http://www.drs2010.umontreal.ca/data/PDF/066.pdf

[13] B. T. Ladd, and R. E. Brown, "Broader Impacts of the Information Frontiers Integrated Education and Diversity Program*," National Alliance for Broader Impacts Summit*, May 1, 2019.

[14] Building the Future Investing in Innovation and Discovery: NSF Strategic Plan 2018-2022.

[Online] https://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf18045. Last accessed January 30, 2020.
[15] Blue Waters Data Visualization Seminar Series, 2016. [Online]. Available: https://bluewaters.ncsa.illinois.edu/webinars/visualization/introduction. [Accessed January 30, 2020].
[16] V. Byrd, "Using Bloom's Taxonomy to Support Data Visualization Capacity Skills," in *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education, New Orleans, Louisiana, November, 2019*, Association for the Advancement of Computing in Education (AACE), 2019, pp. 1039-1053.
[17] Kaggle.com. URL: https://www.kaggle.com/c/walmart-recruiting-store-sales-forecasting Last accessed March 13, 2020.
[18] B. T. Ladd, "Case Study of Interdisciplinary Student Research Teams: Factors, Outcomes, and Lessons Learned," Science of Team Science National Conference, Galveston, TX, May 23, 2018.
[19] B. T. Ladd, "Best Practices Guide for Formation of Interdisciplinary Science Teams," 2019.
[20] J. Beyer, H. Strobelt, M. Oppermann, L. Deslauriers, and H. Pfister, "Teaching visualization for large and diverse classes on campus and online," in Pedagogy of Data Visualization Workshop at IEEE VIS 2016. Baltimore, MD, USA, Oct 23, 2016 – Oct 28, 2016.
[21] A. Bangor, P.T. Kortum, and J. T. Miller. (2008). An empirical evaluation of the system usability scale. Intl. Journal of Human–Computer Interaction, 24(6), 574-594.
[22] J. A. Krosnick. (2018) Questionnaire Design. In: Vannette D., Krosnick J. (eds) The Palgrave Handbook of Survey Research. Palgrave Macmillan, Cham.