Teaching Data Visualization in /nFirst-Year Courses

Position paper

Ryan Clement

22 July 2019

Introduction

Data science and data visualization have become subjects of intense interest in academia, with many institutions of higher education launching or developing programs in these areas in recent years (National Academies of Sciences 2018; Tate 2017). While some of these programs offer introductory experiences in data science that satisfy a general education requirement, often these programs are tied to majors, minors, or certificate programs. Other avenues of exposure to working with data are regularly tied to other areas of study, such as majors in the sciences and social sciences. These disciplines typically don't cover topics such as data visualization in introductory courses, often setting them aside as more advanced topics for research methods courses or upper-level seminars. This approach can lead to wide variability in students' ability to produce and critically examine data visualization if they have not been a part of these programs.

As the Data Services Librarian at Middlebury College, I work primarily with undergraduate students on finding, acquiring, and visualizing data. Like many higher-ed institutions, Middlebury requires all students to take a first-year course that introduces them to academic writing, reading, and critical thinking. Several faculty at Middlebury, across a variety of disciplines, have begun including data visualization work and critique in their first-year courses, treating this another form of "writing." Other faculty teach first-year courses where the inclusion of data visualization topics would be well-suited, but they have chosen not to include these topics for a variety of reasons. Students who later want to work on projects where they need to visualize data or critically examine visualizations in publications can be at a disadvantage if they have not been exposed to these topics. Working to include some level of data visualization instruction in required first-year courses should lead to a broader group of students who are better able to create and critically examine data visualizations.

Research so far

I have begun studying the current practices, challenges, and needs of faculty teaching first-year courses at Middlebury College, in order to develop best practices and recommendations for teaching data visualization in these courses. I began by identifying faculty who have taught a 'required first-year course' (at Middlebury, these courses are known as 'First Year Seminars') in the past four years (2015-2019). I then narrowed this group of 193 faculty down to 25 who mentioned any of the following in their course descriptions:

- data
- statistics
- mapping
- visualization
- traditionally "data-centric" departments (economics, ecology, computer science, geography, etc.)

I then began reaching out to these faculty to ask them to participate in one-hour, semi-structured interviews, as well as asking them to provide learning artifacts (syllabi, assignments, rubrics, etc.) if they were comfortable doing so. As of July 2019, six faculty have completed interviews. Participants have come from these departments: economics, geography, and geology. In the interviews, participants are asked open-ended questions about:

- their own past learning about data visualization
- teaching data visualization across all courses
- teaching data visualization in first-year courses
- the library's role in supporting data visualization

Despite the small size of the participant group so far, several trends have begun to emerge from the interviews. These are discussed below. A major, albeit unsurprising, theme that has emerged from these interviews is that *teaching first-year students about data visualization is challenging, but important.* As one participant said, 'learning to write for college is not just about using words, but all of the ways you can convey information to your audience.'

Implications for library visualization instruction

Several trends with implications for library visualization instruction have begun to emerge from the interviews completed so far. These include:

1. The need for data with which to work that is complex, real-world data, but that also does not need extensive cleaning. There are opportunities here for librarians to engage in collection development, curation, and access.
2. The need for students in first-year courses to learn about data visualization techniques in ways that are robust, but do not have the high learning curve

of tools like R, Python, Stata, ArcGIS, etc.

This issue could be solved in two ways. On one hand, librarians can work on finding and learning these 'entry-level' tools and educating faculty about them. On the other hand, librarians can also work to provide more workshops geared toward novices that emphasize powerful, open-source tools like QGIS, Python, R, etc. This second approach would ease some of the burden faculty feel to keep up with and learn 'all of these new tools' themselves, while still giving librarians the opportunity to support the powerful, open tools we tend to favor.

3. The need to build students' confidence. Almost all of the faculty participants spoke of overestimating students' comfort with technology. Sometimes this overestimation actually comes from not recognizing the diversity of background that students bring to our institutions. While some students may come with extensive technology, research, and library experience, students on the other side of the digital divide rarely have this experience. This overestimation may also come from a mistaken reliance on the existence of the 'digital native' – several participants expressed surprise at the level of technical knowledge that even very experienced students had. The myth of the digital native has been sufficiently debunked (Margaryan, Littlejohn, and Vojt 2011; Sorrentino 2018; Kirschner and De Bruyckere 2017), but often this research has not made its way into faculty practice.

Future work

There is still a need to continue pursuing other faculty participants here at Middlebury. Summer research schedules have made it difficult to schedule interviews, and some re-education on the aims of the project is necessary for other faculty (several faculty did not understand that they would be valuable participants *because* they had not addressed data visualization in their courses). Additionally, it would be useful to extend this research to both another peer institution (another small, liberal arts college) as well as a larger research institution with a robust first-year program.

References

Kirschner, P.A., and P. De Bruyckere. 2017. "The Myths of the Digital Native and the Multitasker." *Teaching and Teacher Education* 67: 135–42. https://doi.org/10.1016/j.tate.2017.06.001.

Margaryan, A., A. Littlejohn, and G. Vojt. 2011. "Are Digital Natives a Myth or Reality? University Students' Use of Digital Technologies." *Computers and Education* 56 (2): 429–40. https://doi.org/10.1016/j.compedu.2010.09.004.

National Academies of Sciences, and Medicine, Engineering. 2018. "Data Science Education." In *Data Science for Undergraduates: Opportunities and Options*. Washington (DC): National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK532765/.

Sorrentino, P. 2018. "The Mystery of the Digital Natives' Existence: Questioning the Validity of the Prenskian Metaphor." *First Monday* 23 (10). https://doi.org/10.5210/fm.v23i10.9434.

Tate, Emily. 2017. "Data Analytics Programs Taking Off at Colleges." *Inside Higher Ed*. https://www.insidehighered.com/digital-learning/article/2017/03/15/data-analytics-programs-taking-colleges.