

MIT Open Access Articles

*This is a supplemental file for an item in DSpace@MIT*

**Item title:** Data Visualization Literacy: A Feminist Starting Point

**Link back to the item:** <https://hdl.handle.net/1721.1/125276>



Chapter Title: Data visualization literacy: A feminist starting point

Chapter Author(s): Catherine D'Ignazio and Rahul Bhargava

Book Title: Data Visualization in Society

Book Editor(s): Martin Engebretsen, Helen Kennedy

Published by: Amsterdam University Press. (2020)

Stable URL: <https://www.jstor.org/stable/j.ctvzgb8c7.19>

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



This book is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0). To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc-nd/4.0/>.



*Amsterdam University Press* is collaborating with JSTOR to digitize, preserve and extend access to *Data Visualization in Society*

# 13. Data visualization literacy: A feminist starting point

*Catherine D'Ignazio and Rahul Bhargava*

## Abstract

We assert that visual-numeric literacy, indeed all data literacy, must take as its starting point that the human relations and impacts currently produced and reproduced through data are unequal. Likewise, white men remain overrepresented in data-related fields, even as other STEM (Science, Technology, Engineering and Medicine) fields have managed to narrow their gender gap. To address these inequalities, we introduce teaching methods that are grounded in feminist theory, process, and design. Through three case studies, we examine what feminism may have to offer visualization literacy, with the goals of cultivating self-efficacy for women and underrepresented groups to work with data, and creating learning spaces where, as Philip et al. (2016) state, 'groups influence, resist, and transform everyday and formal processes of power that impact their lives'.

**Keywords:** Data literacy; Feminism; Community; Inequality; The arts

## Introduction

There is a growing body of literature arguing that working with data is a key modern skill (Letouzé et al., 2015; Wolff et al., 2016). And yet, while highly valued as a precursor to evidence-driven insight, data are expensive—to collect, maintain, and mobilize. Corporations, governments, and elite universities are the primary institutions which have the resources to undertake this work. Within those institutions, white men remain overrepresented in data-related fields such as computer science, engineering, and artificial intelligence, even as other STEM fields like biology have managed to narrow their gender gap (Corbett & Hill, 2015; Neuhauser, 2015; West, Whittaker,

---

Engebretsen, M. and H. Kennedy (eds.), *Data Visualization in Society*. Amsterdam: Amsterdam University Press, 2020

DOI 10.5117/9789463722902\_CH13

& Crawford, 2019). This has resulted in a growing literature around bias in data collection (Angwin et al., 2016), algorithmic decision-making (O'Neil, 2016), and machine learning training sets (Buolamwini & Gebru, 2018). Acknowledging these basic inequalities in the ecosystem—that data and skills to work with them are in the possession of groups that are already privileged in society—lays the groundwork for how educators can start to discuss data literacy more broadly.

We assert that visual-numeric literacy, indeed all data literacy, must take as its starting point that the human relations and impacts currently produced and reproduced through data are unequal. Thus, educators are faced with a choice. They may either 'integrate the younger generation into the logic of the present system' or teach learners how to 'participate in the transformation of their world' (Freire, 1968, p. 16) through data-driven inquiry. The vast majority of data science programmes, trainings, and tools choose the former. This choice may not be nefarious or intentional, but rather because alternatives may not be readily apparent. This chapter explores an emancipatory approach to data visualization literacy based in feminist scholarship and pedagogy in an attempt to chart an alternate course.

A body of work that owes its emergence to the women's suffrage movements of the nineteenth century, feminist theory encompasses a range of ideas about how identity is constructed, how power is assigned, and how knowledge is generated, as well as how a range of intersectional forces such as gender, race, class, and ability combine to influence the experience of being in the world. It is important to note that while feminist scholarship uses gender as a starting point for considerations of social inequality, a feminist approach is not only about cis and trans women, nor only about gender. Deployed as a tool for critical inquiry, feminist thinking seeks to situate knowledge in specific human bodies and to 'unmask universalism' (Davis, 2008)—to show how things that appear to be neutral or objective are in fact biased towards the bodies that hold power—typically male, white, abled, heterosexual, and well-educated. For example, the quintessential feminist critique of data visualization is Donna Haraway's characterization of it as 'the gaze from nowhere' (1989, p. 581). Because the view is not situated in a body or a perspective, it has the aura of neutrality. But, of course, the view from nowhere is always the view from somewhere—more often than not it is the view from a dominant location of power over people whose views and knowledge are suppressed and silenced (Collins, 2009, p. 251; Eubanks, 2018; Noble, 2018; Walter & Andersen, 2013).

For this chapter, we draw specifically on prior work (D'Ignazio & Klein, 2016) that connects feminist theory to the design of data visualizations. Our

goal is to demonstrate the relevance of feminist concerns with gender, social difference, and power in relationship to the teaching and learning of data visualization. Data visualization is sometimes taught with the idea that data are neutral and objective; visualizations are methods for depicting those data; and the right method of depiction can be found by understanding the basics of human visual perception and cognition—which are sometimes imagined to be universally the same across contexts, culture, and history (Kennedy et al., 2016a). Instead, we wish to craft an alternate starting point that acknowledges the social and political context in which data are collected and communicated, cultivate self-efficacy in women, people of colour, and other minoritized groups to work with data and visualization, and focus learners' attention on what happens in the world as a result of an act of data communication.

The editors of this volume ask, *'What does literacy mean when it comes to data visualization, and how can visual-numeric literacy be enhanced?'* (this volume). We assert that because visualizations are outputs of a process, visual-numeric literacy is part of a larger process of data literacy, which itself draws on other approaches such as statistical literacy, numeracy, and critical information literacy. In earlier work, we proposed that data literacy *'includes the ability to read, work with, analyze and argue with data as part of a larger inquiry process'* (D'Ignazio & Bhargava, 2016, p. 1).

While this definition makes it sound like it is an individual ability, integrating feminist thinking opens up questions as to the nature of literacy itself. Is data literacy a set of autonomous skills acquired by an individual? Or, following bell hooks's notion of an 'open learning community' (1994) and proponents of new literacy studies (Street, 1994), is data literacy a set of social practices, learned and practised in and through a social context such as an organization or community? Or, following feminist computer scientist Lynette Kvasny (2006), is teaching about data a site of ideological transmission, a place where, if we are not careful, we risk reinscribing structural oppression? From our teaching practices we would answer these questions 'yes' and 'yes' and 'yes'.

From this complicating ground, then, data literacy cannot begin with technical skills like making and interpreting bar charts and network diagrams. It necessitates a starting point grounded in higher-order critical thinking and making skills that connect data back to the social and political reality from which they were produced.

But what might this look like in practice? In the remainder of this chapter, we outline three short cases of data visualization learning that come from our practice as educators and then analyse them in relation to the design

principles laid out by one of us, Catherine, and Lauren Klein (2016), in our paper 'Feminist Data Visualization'. In that paper, we created six preliminary principles of feminist data visualization, drawing from work in feminist science and technology studies, feminist human-computer interaction, feminist digital humanities, and critical cartography & GIS. The principles are: (1) Rethink binaries, (2) Embrace pluralism, (3) Examine power and aspire to empowerment, (4) Consider context, (5) Legitimize affect and embodiment, and (6) Make labour visible. Due to space considerations as well as the exploratory nature of this work, we focus on analysing (2), (4), and (5) as they relate most directly to data visualization literacy.

### Three cases of data visualization learning

#### The Groundwork Somerville data mural

Our first case study focuses on the process of working with a community group—Groundwork Somerville (GW)—and local youth to design and paint a data-driven story as a community mural. This example of a 'data mural' documents one approach to an action-oriented, community-situated activity that builds various data literacies.

GW focuses on empowering participants to improve environmental, economic, and social well-being, specifically through nature-focused programmes. One of their main programmes involves youth to create, plant, and maintain gardens as well as sell the produce that results. Immigrants and low-income families are the main beneficiaries. Additionally, many of the vegetables planted are chosen to reflect the immigrant makeup of the community. GW was interested in working with us in order to reinforce their goals of youth development, to beautify an urban garden, and to tell a story about their impact.

The collaboration followed a process which moved from identifying data, finding a story, collaboratively designing a visual to tell that story, painting the mural, and finally hosting an unveiling event (see Bhargava et al., 2016 for details). GW shared qualitative and quantitative data with our team, and we worked together to narrow in on data to include in a multi-page handout for the youth. In terms of demographics, there were six young women of colour, seven young men of colour, two young white women and two young white men on the GW team. With the data in hand, we hosted a brainstorming session with youth to analyse these data and generate a story and visuals to tell it. Based on these handouts, the participants identified



Figure 13.1. Some of the sketches students created. Photos by Emily Bhargava. Printed with permission.

a number of facts and quotes in order to tell the GW story. Inspired by some data visualizations we shared with them, the youth then sketched visuals for telling the story. This invitation to sketch concretized the data and helped them bridge into a narrative structure more readily. The youth were responsible for the visual language of the narrative, which was a key pedagogical goal of ours.

The resulting narrative arc told a story about the GW 'winning formula' and how it was benefiting the community through 'together livin' better'. The visual designs were stitched together by Emily Bhargava into a consistent mural design. Painted on the large metal fence behind one of the converted lots, roughly 80 feet long and 10 feet tall, it showcases the GW impact story at the site of one of the reclaimed urban farms; literally telling a story about the space, in the space itself.

At the unveiling, viewers and participants alike commented on the impacts. One attendee said, 'What strikes me is that you've managed to tell a story with an equation and very simple images'. Others commented on the visual encodings and symbolic language—'The bike sticks in my mind'. Validating our goal of increasing data literacy with the youth participants, one commented that 'I learned that by pictures you can also send out a message', and another said, 'I learned how to take data and make a story'.

This example highlights that people who don't 'speak' data or self-identify as 'geeks' or 'techies' can be effectively involved in data analysis and storytelling by focusing on an arts-based, socially-oriented invitation. Our goals centred on building the confidence of youth to engage with data and enhancing the built environment in an impactful way. The choice of a mural as the medium leveraged the long history of public murals being used to comment on and change the public discourse about a topic. The



Figure 13.2. The data mural. Photo by Rahul Bhargava. Printed with permission.

participatory, youth-driven process offers an example of how to engage people in a collaborative meaning-making process to amplify their understanding of how data can communicate. In this case, data analysis and visualization are methods for connecting more deeply to the community, not modelled as end points in themselves.

### **‘Asking questions’ with WTFcsv**

For the second case of data visualization learning, we introduce an activity called ‘Asking questions’ from the DataBasic.io suite of tools and activities which we built. DataBasic.io consists of simple, web-based tools for beginners that introduce concepts of working with data ranging from quantitative text analysis to network analysis. For the purposes of this case, we focus on the tool WTFcsv and its accompanying learning activity ‘Asking questions’.

WTFcsv helps learners analyse a comma-separated-values (CSV) file to look for potential data-driven stories to tell. The software analyses each column from a spreadsheet file uploaded by the learner and returns a data visualization that summarizes the patterns in each column (Figure 13.3).

Newcomers often approach data thinking of it as consisting only of numbers. Two of the primary learning goals for WTFcsv are that 1) learners understand that data have many types, including numbers, categories, text, and dates, and 2) that different types of visualizations are appropriate for summarizing different types of data. For example, temporal data are shown as a line-chart histogram on a time-series axis. Numeric data are shown as a histogram, with buckets created linearly. Text data are shown as a column chart if there are only a few types (categorical data), or a word cloud if there are many entries (i.e. open text).

‘Asking questions’ is the learning activity that accompanies the WTFcsv tool, based on Tactical Technology Collective’s notion of ‘asking your data some questions’ (Tactical Tech, 2014). While newcomers to spreadsheet analysis often attribute some wizardry to the data analysis process, this activity tries to introduce them to a simple, inquiry-based process for getting acquainted with a dataset.





Figure 13.3. The WTFcsv results screen. Printed with permission.

Learners break into small groups, choose one of the sample datasets, examine WTFcsv's summary visualizations, and brainstorm questions that they want to ask the data.<sup>1</sup> Facilitators encourage learners to use the visualizations to generate many types of questions, including context questions ('What's the source of these data? Why did they collect it? Who uses it?'), ethical questions ('Is it OK to publish people's full names? How did sex end up as a binary variable?'), quality questions ('Are these data complete? How were they acquired?'), data formatting questions ('What does the "Parch" column mean?'), as well as data analysis questions ('Did women survive at a higher rate than men?'). After 10 minutes of brainstorming questions and sources for connected data, learners share their most interesting question back to the group. The debrief conversation focuses on how rich questions often involve multiple data sources, the variety of questions that can come from one dataset, and the critical process of recognizing any bias in identified questions.

We have run more than 30 workshops with WTFcsv. In an evaluation of the WTFcsv activity, learners responded well to these choices. One participant commented that the activity 'helps you from the beginning to understand the possibilities of your spreadsheet'. The fact that they framed 'possibilities' as plural is meaningful, in the sense that it is important for newcomers to understand the role of *exploratory data visualization*—the way in which visual aggregation can serve to provoke important questions and next steps towards the formulation of a knowledge claim. Another commented that the tool was 'different because usually there are just text and numbers, not lots of images and graphs and the ability to look at them all right away'. This fulfils our primary goal that learners understand that data visualization can play an important role in the exploration and meaning-making process.

### 'ConvinceMe' with the Data Culture Project

The third case study is an activity we developed to enable people to practice the skill of making arguments with data to convince people to take action, called 'ConvinceMe'. Many definitions of data literacy focus on the ability to read and write with data; fewer include the idea of arguing with data as a core skill. We believe that this is critical for putting data into action in the real world. Without this, data end up divorced from the fact that

1 Current options for English-language speakers include Titanic Passengers, UFO Sightings, and Dogs of NYC. When viewed in other languages, the tool offers different culturally and geographically appropriate examples.

they often represent real people, and are used to make decisions about them. This activity is the final one in our Data Culture Project (see <http://datacultureproject.org>), a lightweight self-service curriculum available for free to any organization that wishes to work on building an inclusive data culture.

The activity itself is simple to run, and relies on the creativity of the participants. It begins with a room of at least 10 people. Facilitators share a printed data visualization about some topic. We use one about water conservation, which argues that choosing to consume beef has an extraordinarily high 'water cost'. After talking through the graphic, we ask the group to identify 3-5 key stakeholder groups that can influence that topic. For example, in the case of water conservation, stakeholders might be farmers, policymakers, a shopping family member, or a restaurant owner. Volunteers are solicited to role-play those stakeholders, and invited to stand in the front of the group. The rest of the group is asked to make a short data-driven argument to specific stakeholders, asking them to change some behaviour. For example, this could consist of telling a 'shopping parent' stakeholder about the high water cost of meat, and invite them to try a vegetarian diet by joining the 'meatless Monday' movement.

The primary learning goal for ConvinceMe is for learners to practice making data communication decisions in a situated environment, with a specific audience in mind. For newcomers, data often appear neutral and abstract, but to practice making arguments with data you have to re-concretize them. The role-playing stakeholders embody those being asked to take action—they physically step forward if they are convinced by the arguments and step backwards if they are not. Another learning goal for ConvinceMe is to situate data visualizations and data-driven arguments as tools for advocacy and social transformation. The goal for the activity is not to create a picture, but rather to move a particular stakeholder towards a desired action. The act of inviting participants to make a persuasive argument with data breaks down the narrative of data as neutral.

A group of 25 non-profit organizations participated in the first cohort of the Data Culture Project in Fall 2017 and ran the ConvinceMe activity, with positive feedback. One noted that their arguments 'used a lot of shame and guilt', leading them to reconsider how they frame their calls to action. A mid-sized non-profit valued most the 'importance of identifying stakeholders and trying to understand their interests'. The act of practising the arguments with real people impacted their approaches to picking stories to tell, and how to tell them. Another group began rethinking their whole approach to data storytelling, and decided to do this activity 'before anyone

even collects any data', so they could 'think up front about who should be convinced about what'.

## Learning and feminist data visualization

While mainstream data visualization teaching often starts by instruction in chart types, or learning a software tool, or learning about human perception, we argue that we need to begin with a wider lens before zooming into technical specifics. In this section, we will consider the three cases of data visualization learning described above in relationship to three design principles of feminist data visualization outlined by D'Ignazio and Klein (2016)—'Consider context', 'Legitimize affect & embodiment', and 'Embrace pluralism'—in order to explore their implications for a feminist starting point for visual numeric literacy.

### Consider context

One of the central tenets of feminist epistemology is that knowledge is 'situated' (Haraway, 1989, p. 581). What this means is that context matters—What are the social, cultural, historical, and material conditions in which knowledge is produced? What are the identities of the humans making the knowledge? A feminist perspective advocates for connecting datasets and data visualizations back to their context, to better understand their limitations and ethical obligations, and, ultimately, the ways in which power and privilege may obscure truths.

Situating datasets and data visualizations for learners is a particular challenge, since the conventions of both spreadsheets and precise graphics make them appear objective (Kennedy et al., 2016a), particularly for non-technical newcomers. The 'Asking questions' activity illustrates one way to situate data and visualizations. Instead of trying to 'find stories', the position of asking questions helps model for learners a process of inquiry and exploration where meaning is not something to be 'found' hidden in the dataset, but rather produced through an iterative process of investigation that involves many bits of information that are not included in the dataset itself. And encouraging many types of questions, including questions about trust in sources, missing data, and data formatting, helps learners start to connect the data back to the institutional and historical context where they were collected, emphasizing that those things also matter deeply to any meaning that comes from patterns observed in the data. The fact that learners use

visualizations to ask those questions matters as well—this demonstrates the value of visualization not as the definitive, objective word on a subject, but rather as a helpful exploratory step in a process of building meaning.

ConvinceMe functions in a similar way to draw connections back to situated human experience, and grounds data visualization in acts of communication between different types of stakeholders. While data visualization research within technically oriented disciplines often focuses on time to task metrics, such as how well an individual can decode the meaning of a particular chart, there is very little research on how data visualizations or data-driven arguments help move groups from different positions and/or different cultures to action (or not). If you know that your data visualization needs to move farmers to use less water, then you will make different decisions about what data to highlight and what format to use than if you need to convince parents. ConvinceMe creates a lightweight, bounded playground in which people can begin to understand visualization as acts of situated communication.

### **Legitimize affect & embodiment**

This principle of feminist data visualization derives from the argument by feminist thinkers that experiences that derive from sensation and emotion have been systematically devalued over quantitative methods of knowing. Patricia Hill Collins notes in her articulation of Black feminist epistemology that in an ideal knowledge situation ‘neither emotion nor ethics is subordinated to reason’ (2009, p. 266). There has been work on the rhetorical function of data visualization in narratives (Hullman & Diakopoulos, 2011) and persuasion (Pandey et al., 2014). But the role of emotion in data visualization has been understudied, with the exception of work by Kennedy, Hill, Aiello, and Allen (2016) and several chapters in this book (Gray, this volume; Hill, this volume; Simpson, this volume). In contrast to learning about data which involves abstraction and distance from the subject matter (or from the subjects themselves), how might we acknowledge embodied and affective experiences in the data visualization learning process? How might an intimate, emotional connection to data be considered an asset to the analysis, visualization, and learning process?

The GW data mural legitimates affect and embodied experience in several ways. First, the data that the youth analysed and used to tell a story are data about themselves and their organization. Data visualization techniques are often discussed as though the subject matter of the data is interchangeable and neutral (Kennedy, Hill, Allen & Kirk 2016). Many teaching examples use

so-called 'classic' datasets like *mtcars* (Kosara, 2018), but, as D'Ignazio has stated in prior work, 'Cars. Who cares?' (2017, p. 8), people will learn better and more deeply from data that they have an experiential understanding of and an emotional connection to (Kennedy, Hill, Aiello & Allen, 2016). The GW project began with data that was about something intimate and emotionally connected to the youth and their community. Likewise, the project ended with a data visualization that was emotionally connected to them—the input and the output were deeply situated.

Both the GW data mural and the ConvinceMe activity also use embodied, arts-based ways of knowing for learner engagement and think beyond the screen in terms of data visualization. In the case of GW, the youth not only came up with the iconography but also painted it onto a giant mural in the garden. The act of assembling the data visualization was itself an embodied, social act, undertaken in community. And ConvinceMe uses performance and narrative to construct a humorous social situation where peers have to use a data graphic to convince each other to shift their behaviour. Data visualization is important to each of these cases, but a 2D screen-based graphic is not modelled as the end point. Significantly, ConvinceMe seeks to value personal testimony in addition to data-driven graphics. Indeed, the graphic is the jumping off point, but needs the embodied personal testimony (the speaker) and custom-tailoring to an audience (the speaking situation) in order to move them to action. Legitimizing affect and embodiment may mean seeking the appropriate form for the appropriate audience, as well as modelling in learning activities how visualizations fit in to an embedded advocacy or community-building process.

### Embrace pluralism

The design principle 'Embrace pluralism' comes from feminist scholarship's long history of challenging claims of objectivity, neutrality, and universalism, emphasizing instead how knowledge is always constructed within the context of a specific subject position as well as within a community of knowers (Harding, 1991). Black feminist scholars like Patricia Hill Collins have demonstrated how discourses of objectivity systematically exclude the voices of women and people of colour, among others, with the burden of oppression most heavily borne by those who sit at the intersections of the 'matrix of domination' such as Black women (Collins, 2008). A key contribution of this line of feminist thinking has been to recognize how a multiplicity of voices, rather than one single loud, magical, or technical voice, often results in a more complete picture of the issue at hand.

Embracing pluralism in data visualization learning can counter the media narrative that sometimes constructs data scientists as ‘unicorns’ or ‘wizards’—solo technical geniuses who can use advanced analytical or artistic techniques to tame large datasets into insights. Learning activities for newcomers often have to counteract these preconceptions by modelling an alternative meaning-making process that is social and dialogue-based rather than individual and technical. While skill development is necessary for data visualization, we would argue that starting with those things reinforces the naive notion that data are about solo technical mastery. We choose to model a process that communicates that answers are best found in dialogue with a community of knowers, who approach a topic area with many perspectives.

It is significant that all three of the cases discussed above model a process of valuing different voices and producing knowledge through dialogue in group. Rather than students being positioned as individual learners in front of computer screens, the learning situation is social. In the case of the data mural, youth worked in small groups to contextualize the GW numbers and in one large group to paint the mural. In the case of the WTFcsv activity, learners work in groups of three to brainstorm questions—people learn from their peers’ questions. And ConvinceMe is an activity conducted in a larger group which intentionally centres the idea of multiple stakeholder voices as well as formulating appeals to those particular standpoints. You ‘win’ by making a data-driven appeal to one of those stakeholders, convincing them to move towards the speaker. While the first two activities model a social, pluralistic process for deriving meaning from data and their visualization, this last activity embraces pluralism on the reception side of data communication, helping learners understand that different audiences may be moved by different narrative and visual arguments.

## Conclusion

As stated at the beginning, one of the reasons that a feminist approach to data is useful and necessary is because of the power differentials and inequalities in the data ecosystem. Resources to collect, store, and analyse data are not distributed equally, nor are the technical skills to work with data. The cases we have discussed, along with the feminist design principles that guide them, represent a starting point for visual-numeric literacy. This is an area for further research and evaluation: does a feminist-informed learning programme lead to increased self-efficacy around data and its visualization for more women, people of colour, and other minoritized groups?

Relatedly, while many of our learning activities *model a feminist process*, they do not explicitly tackle issues of power, structural inequality, and bias in the *content of what is taught*. What might learning activities for the same audience (adult, non-technical newcomers) look like that specifically address concerns of gender and racial bias, the political economy of data, and so on? In order to integrate these conversations into the learning situation, Philip, Olivares-Pasillas, and Rocha (2016, p. 365) argue that we need to consider cultivating racial literacy and gender literacy side-by-side with data literacy. They write, 'Spaces must be facilitated for students to engage with the structural and ideological contexts of data visualizations if these tools are to authentically engage them in democratic deliberations—spaces where they grapple with how groups influence, resist, and transform everyday and formal processes of power that impact their lives'.

## References

- Angwin, J., Larson, J., Mattu, S., Kirchner, L. (2016, May 23). Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. *ProPublica*. Retrieved from <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Bhargava, R., Kadouaki, R., Bhargava, E., Castro, G., & D'Ignazio, C. (2016). Data murals: Using the arts to build data literacy. *Journal of Community Informatics*, 12(3). Retrieved from <http://ci-journal.org/index.php/ciej/article/view/1276>
- Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research: Proceedings of the 1st Conference on Fairness, Accountability, and Transparency*, 81, 77-91.
- Collins, P. H. (2009). *Black feminist thought: Knowledge, consciousness, and the politics of empowerment* (2nd ed.). New York: Routledge.
- Corbett, C., & Hill, C. (2015). *Solving the equation: The variables for women's success in engineering and computing*. American Association of University Women. Retrieved September 18, 2019 from <https://www.aauw.org/research/solving-the-equation/>
- D'Ignazio, C. (2017). Creative data literacy: Bridging the gap between the data-haves and data-have nots. *Information Design Journal*, 23(1), 6-18. <https://doi.org/10.1075/idj.23.1.03dig>
- D'Ignazio, C., & Bhargava, R. (2016). DataBasic: Design principles, tools and activities for data literacy learners. *The Journal of Community Informatics*, 12(3), 83-107. Retrieved from <http://ci-journal.org/index.php/ciej/article/view/1294>



- D'Ignazio, C., & Klein, L. (2016, October). *Feminist Data Visualization*. Paper presented at and published in the workshop proceedings from the Workshop on Visualization for the Digital Humanities at IEEE VIS Conference, Baltimore, Maryland.
- Davis, K. (2008). Intersectionality as buzzword: A sociology of science perspective on what makes a feminist theory successful. *Feminist Theory*, 9(1), 67-85. <https://doi.org/10.1177/1464700108086364>
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. New York: St. Martin's Press.
- Freire, P. (2015). *Pedagogy of the oppressed* (50th Anniv.). New York: Bloomsbury Academic.
- Haraway, D. J. (1988). Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist Studies*, 14(3), 575-599.
- Harding, S. G. (1991). *Whose science? Whose knowledge? Thinking from women's lives*. Ithaca: Cornell University Press.
- Hooks, B. (1994). *Teaching to transgress: Education as the practice of freedom*. New York & London: Routledge.
- Hullman, J., & Diakopoulos, N. (2011). Visualization rhetoric: Framing effects in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics*, 17(12), 2231-2240.
- Kennedy, H., Hill, R. L., Aiello, G., & Allen, W. (2016). The work that visualisation conventions do. *Information, Communication and Society*, 19(6), 715-735. <https://doi.org/10.1080/1369118X.2016.1153126>
- Kennedy, H., Hill, R. L., Allen, W., & Kirk, A. (2016). Engaging with (big) data visualizations: Factors that affect engagement and resulting new definitions of effectiveness. *First Monday*, 21(11). <https://doi.org/10.5210/fm.v21i11.6389>
- Kvasny, L. (2006). Social reproduction and its applicability for community informatics. *The Journal of Community Informatics*, 2(2). Retrieved from <http://ci-journal.org/index.php/ciej/article/view/342>
- Kosara, R. (2018, March 21). How to get excited about standard datasets [Blog post]. Retrieved from [eagereyes.org](http://eagereyes.org).
- Letouzé, E., Bhargava, R., Deahl, E., Noonan, A., Sangokoya, D., & Shoup, N. (2015). *Beyond data literacy: Reinventing community engagement and empowerment in the age of data*. Retrieved from <http://datapopalliance.org/item/beyond-data-literacy-reinventing-community-engagement-and-empowerment-in-the-age-of-data/>
- Neuhaus, A. (2015, June 29). 2015 STEM Index Shows Gender, Racial Gaps Widen. *US News & World Report*. Retrieved from <https://www.usnews.com/news/stem-index/articles/2015/06/29/gender-racial-gaps-widen-in-stem-fields>
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York: New York University Press.

- O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. New York: Crown.
- Pandey, A. V., Manivannan, A., Nov, O., Satterthwaite, M., Bertini, E. (2014). The persuasive power of data visualization. *IEEE Transactions on Visualization and Computer Graphics*, 20(12), pp. 2211-2220.
- Philip, T. M., Olivares-Pasillas, M. C., & Rocha, J. (2016). Becoming racially literate about data and data-literate about race: Data visualizations in the classroom as a site of racial-ideological micro-contestations. *Cognition and Instruction*, 34(4), 361-388. <https://doi.org/10.1080/07370008.2016.1210418>
- Street, B. (2003). What's 'new' in New Literacy Studies?: Critical approaches to literacy in theory and practice. *Current Issues in Comparative Education*, 5(2), 77-91.
- Tygel, A. F., & Kirsch, R. (2016). Contributions of Paulo Freire for a critical data literacy: A popular education approach. *The Journal of Community Informatics*, 12(3), 108-121.
- Walter, M., Andersen, C. (2013) *Indigenous statistics: A quantitative research methodology*. Walnut Creek, CA: Left Coast Press.
- West, S. M., Whittaker, M., & Crawford, K. (2019). *Discriminating systems: Gender, race and power in AI*. AI Now Institute. Retrieved from <https://ainowinstitute.org/discriminatingystems.html>.
- Wolff, A., Gooch, D., Caverio Montaner, J., Rashid, U., & Kortuem, G. (2016). Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics*, 12(3), 9-26. Retrieved from <http://ci-journal.net/index.php/ciej/article/view/1286>

## About the authors

**Catherine D'Ignazio** is Assistant Professor of Civic Media and Data Visualization at Emerson College, a senior fellow at the Emerson Engagement Lab and a research affiliate at the MIT Media Lab. [catherine\\_dignazio@emerson.edu](mailto:catherine_dignazio@emerson.edu)

**Rahul Bhargava** is a Research Scientist at the MIT Center for Civic Media and MIT Media Lab. [rahulb@media.mit.edu](mailto:rahulb@media.mit.edu)