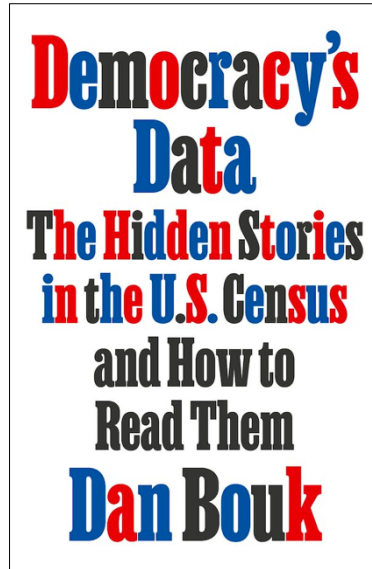


Dan Bouk, **Democracy's Data: The Hidden Stories in the U.S. Census and How to Read Them**, New York, NY: Macmillan, 2022, 362 pp., \$30.00 (hardcover), \$20.00 (paperback).

Reviewed by
Emilia Ruzicka
University of Virginia

Throughout the history of the United States, the decennial census has been a constant presence, laying the foundation for American governmental structure and determining who gets represented and how. Though it is easy to view the census as objective fact when its insights are presented as aggregated numbers in large, downloadable tables, Dan Bouk urges readers to consider the people, infrastructure, and politics behind the numbers of the 1940 census in ***Democracy's Data: The Hidden Stories in the U.S. Census and How to Read Them***. Bouk grounds his research on the idea that “our democracy is only as good as our data, and our data is only as good as our democracy” (p. 20). He begins by explaining the logistics surrounding the release of census manuscript records, which are the original forms filled out by enumerators who visited houses to count each person in the United States. The manuscript records contain the names, addresses, and other associated information of those who were surveyed. To protect the privacy of individuals, manuscript records are not released until 72 years after their corresponding census was conducted.



Bouk began his research after the 1940 manuscript records were publicized in 2012. Alongside his research assistants at Colgate University, Bouk amassed census-related records of all kinds, including manuscripts, government correspondence, congressional notes, poems, political cartoons, and newspaper clippings. Via these primary sources, Bouk in the first two chapters reconstructs the social and political moment preceding the launch of the 1940 census. By investigating the individuals included in the census design process, who Bouk dubs “the Question Men” (p. 20), chapters 0 and 1 detail the underlying motivations and biases that influenced the final form of the 1940 census. These perspectives included those of insurance company executives, labor union representatives, eugenicists, and more, all of whom were almost exclusively, Bouk pointedly notes, White men.

Chapters 2 and 3 unpack specific quirks of the 1940 census. 1940 was the first census to record every individual's name instead of only the head of the household's name and simply tallying up all other residents, which became the topic of chapter 2. Bouk focuses heavily on the interaction that happened between census enumerators and those they were counting, emphasizing that the interaction was more than just simply tallying people because “respect and honor were also up for grabs” (p. 55). Similarly, chapter 3 examines the “partner” label that was used to identify a wide variety of relationships between the head of the household and others living there, such as business partners, roommates, and potentially those in queer relationships. Bouk directs readers to ask, “How can a dataset encompass those its designers never imagined?” (p. 75).

The following two chapters examine the political biases and motivations served through the execution of the 1940 census. Bouk aptly calls chapter 4 "Counting with Friends" (p. 103) and gives a detailed explanation of when enumerators and other census officials were hired as political favors or to earn the good graces of those with power. Chapter 5 focuses on incomplete or inaccurate data in the 1940 census, unpacking the underenumeration of Black Americans, the absorption of the Mexican community into the "White" racial category, and the uncertain process of counting Indigenous Americans who belonged to sovereign tribes. Both chapters prove Bouk's argument that "the separation of politics and statistics is idle fantasy" (p. 112).

Moving into public controversies that surrounded the 1940 census, chapters 6 and 7 aim to reconstruct the political and social environment during and after the counting was conducted. Because the 1940 census was the first census to ask Americans about their wage earnings, the question of income became controversial, so much so that Bouk spends all of chapter 6 discussing it. One particular crusader, Senator Charles W. Tobey of New Hampshire, was very vocal about how invasive the income question was, stoking Americans' fears of fascism during World War II. Chapter 7 lays out the myriad ways that census data was leveraged to aid the American war effort and cause egregious harm to those of Japanese ancestry living in America. Typically, census data is considered confidential, but during WWII, the federal government passed "a provision that would ease confidentiality protections for data useful to national defense activities" (p. 215). The census data not only told the government who was a skilled laborer and who was eligible for the draft but also who it could profile as potentially "disloyal" (p. 220).

In the closing chapter, Bouk reflects on the disorder and uncertainty he found while digging through the records and media associated with the 1940 census. He laments that "the idea of naming a single 'average American' is ridiculous, but once the data was out the door, people could do all sorts of things with it, whether it was a good idea or not" (p. 241). Bouk warns of the dangers of irresponsible data aggregation and the potential for marginalized people to be systemically forgotten if they are not counted. He advocates for more detailed data storytelling that turns numbers back into the people from which they originated, humanizing the sterile tables in the census database. Lastly, Bouk expresses hope that his close reading of the 1940 census "affirm[s] each person's dignity and advocate[s] for the inherent, equal value of every individual, even or especially when the census itself did not" (p. 247).

Bouk asks many questions of both the documents he collected and his readers. He often points out which questions can be answered by the data provided by and surrounding the census and which cannot. The peppering of unanswered inquiries throughout the book creates an ongoing sense of curiosity and mystery, accurately conveying how much has yet to be uncovered in and about the survey that forms the basis of U.S. democracy.

This book is vivid, accessible, and an excellent example of an academic book written for a general audience. It includes a greater number of footnotes than might be expected from a trade publisher. However, though Bouk demonstrates some of his method of close reading census data via annotated scans of select documents he worked with at the end of each chapter, the book has less instruction about how to replicate or add to Bouk's research than the title of the book seems to suggest. There is a brief note on method at the beginning of the book, but those few pages primarily focus on what kinds of

documents Bouk used, who he was aided by, and from what perspective he approaches his research. As a result, this book is best suited for readers who are more interested in the outcomes of Bouk's census research as opposed to those who want to understand his method in detail.

As a researcher, Bouk situates himself at the intersection of many fields. First, he identifies himself as a "cultural historian" (p. 129). In doing so, he places *Democracy's Data* among other census histories such as *The American Census: A Social History* (Anderson, 2015) and "Challenges to the Confidentiality of U.S. Federal Statistics, 1910–1965" (Anderson & Seltzer, 2007), published in the *Journal of Official Statistics*. Additionally, Bouk draws heavily on the work of data and technology scholars, particularly those who are concerned with the dangers of or participate in big data and data aggregation. Some examples are Rudder's (2014) *Dataclysm: Who We Are (When We Think No One's Looking)* and O'Neil's (2016) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Lastly, Bouk integrates the research of scholars studying critical race, gender, and feminist perspectives in data. These works include *Data Feminism* (D'Ignazio & Klein, 2020) and *Lines of Descent: W.E.B. Du Bois and the Emergence of Identity* (Appiah, 2014).

Though no single book can ever encompass the vast stories contained within the data of the U.S. decennial census, Bouk effectively covers some of the most pressing and interesting issues of the 1940 census. His close reading of census manuscripts and other primary source documents paired with critical perspectives from secondary scholarly works create a captivating narrative accessible to academics and casual readers alike.

References

- Anderson, M. J. (2015). *The American census: A social history* (2nd ed.). New Haven, CT: Yale University Press.
- Anderson, M. J., & Seltzer, W. (2007). Challenges to the confidentiality of U.S. federal statistics, 1910–1965. *Journal of Official Statistics*, 23(1), 1–34.
- Appiah, K. A. (2014). *Lines of descent: W. E. B. Du Bois and the emergence of identity*. Cambridge, MA: Harvard University Press. doi:10.4159/9780674419346
- D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. Cambridge, MA: MIT Press.
- O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy* (1st ed.). New York, NY: Crown.
- Rudder, C. (2014). *Dataclysm: Who we are when we think no one's looking* (1st ed.). New York, NY: Crown.