

EXPLORING DATA DETECTIVE PRACTICES AS A CLASS ACTIVITY

By Ruishan Wu, Zezhong Wang, Haidan Liu, Krithika Balasubramanyam, Kanak Gautam, Aham Gupta, Shannon McAllister and Sheelagh Carpendale

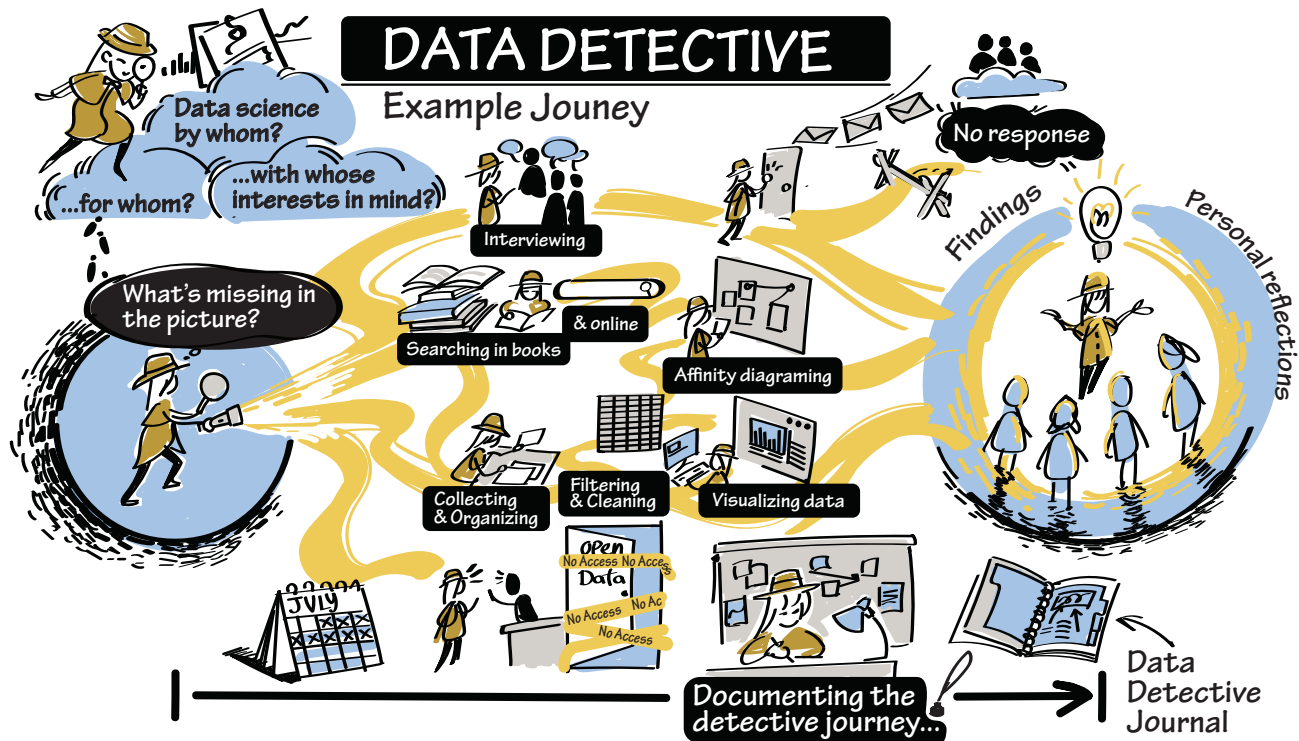


Figure 1. Example journey of Data Detective: beginning with defining a critical problem or question, identifying gaps (“What’s missing in the picture?”), searching for data, and confronting barriers—missing, partial, or deliberately obscured datasets. Each step provided unique insights into the relationship between data, society, and power dynamics. (Illustration by ©Zezhong Wang and Ruishan Wu)

WE reflect on our experiences arising from a recent computer science graduate class about data feminism, during which we explored the idea of being data detectives. In this report, we explain what we mean by *Data Detective* as an active approach where we, as individuals, could approach the underlying questions, as suggested by D’Ignazio and Klein “Data science by whom? Data science for whom? Data science with whose interests in mind?” [1]. By connecting individually through personal reflection, data literacy, and critical engagement, our goal is to inform and inspire those who are interested in integrating similar methods into their classes.

As our society continues to evolve, more and more of the information we need is stored as data, and many of these repositories are growing and becoming what we refer to as *Big Data*. In this process, data becomes more challenging and less accessible to us as individuals. We, as visualization researchers, work on the creation

of visualizations as at least part of the solution to this problem. However, much of our data is still not visualized, and even when it is, individuals still often find it challenging to understand. How do we cope with this? How do we teach our students to cope with this continually expanding problem? In our data feminism class, we introduced concepts such as visual variables, physicalizations, assumptions about knowledge development (e.g., Positivism and Interpretivism), along with reflection and discussion on reading the book *Data feminism* [1]. We then explored developing an active practice through which we would document our investigations of both qualitative and quantitative data under various themes. We now term this active practice as being a *Data Detective*.

Our concept of *Data Detective* is modeled on detective work in a more general sense, where a person uses coherent, time-based, record-keeping of their activities to gain a better understanding of that which they initially do not know but want to under-

stand. Thus, to act as a *Data Detective* is to discover and conduct purposeful, documented, and reflective actions needed to gain access to the desired data. This detective work ideally results in access to the desired data, an understanding of the data, and the detective process involved.

The term *Data Detective* appears in various contexts, making it important to clarify our specific approach. Unlike children’s books [2] that suggest counting objects (like red cars versus white cars), or Harford’s statistical literacy guide with its ten rules for making sense of statistics [3], or visualization workshops for children by providing a gamified sense of accomplishment [5]. Our approach also differs from Inselberg’s multidimensional data detective work [4], which focuses on analyzing existing visualizations, and from data activism approaches that emphasize community engagement.

At its core, being a data detective is more

than systematic data analysis; it emphasizes attentiveness to meaning. Rather than merely focusing on technical proficiency or computational methods, we approached data detective work as an interpretive practice. The aim was not simply to analyze data but to explore its broader significance—examining what data reveals and what it conceals, who it benefits, and whose perspectives it marginalizes.

We visualized our investigative approaches as journeys: beginning by defining a critical problem or question, identifying gaps (“What’s missing in the picture?”), searching for data, and confronting barriers—missing, partial, or deliberately obscured datasets. Each step provided unique insights into the relationships between data, society, and power dynamics.

EXAMPLES

Throughout the semester, students undertook diverse projects with strong societal relevance, including topics such as gender bias in politics, barriers faced by women in entrepreneurship, the functions and ideology of pockets constrained by historical gender roles, and gender representation within STEM academia.

One student examining women’s representation in political institutions vividly illustrated the practical challenges of data detective work. Initial exploration quickly highlighted systemic data gaps as key datasets were fragmented or unavailable. The student navigated through a frustrating landscape marked by opaque official sources, partial records, and silences. Despite challenges, this data detective journey offered significant emotional and intellectual rewards. The student discovered patterns of marginalization, for instance, women are frequently relegated to peripheral roles rather than core decision-making positions. Each painstakingly gathered dataset provided clarity about structural inequalities. Ultimately, the effort became a tangible act of resistance against invisibility and marginalization.

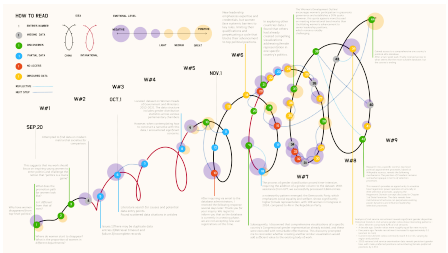


Figure 2. Data Detective journey created by ©Ruishan Wu.

Another student explored the challenges women encounter in achieving tenure in Canadian academia. Initially optimistic, the student encountered considerable barriers, including incomplete or outdated datasets and inconsistent categorization across institutions. Interviews became essential to fill these gaps, highlighting how data detective work can require alternative methods beyond computational data collection. The journey revealed systemic biases: women disproportionately assigned tasks correlated with lower job satisfaction and hindered career progression.

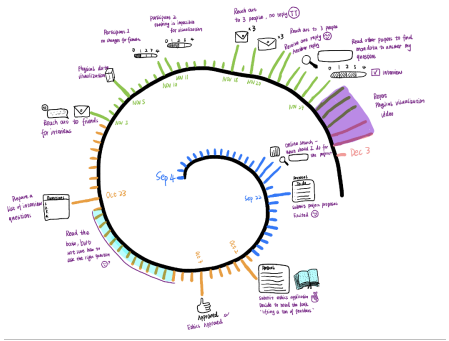


Figure 3. Data Detective journey created by ©Haidan Liu.

WORKING WITH BOTH BIG DATA & PERSONAL DATA

As we moved through the process, we found ourselves blending two approaches to data visualization that are often kept separate: working with big data and working with personal data. Big data showed up in the external datasets we chose to investigate, such as government records, institutional statistics, or public health databases. These are the kinds of large-scale, structured data commonly associated with the term *big data*. On the other hand, personal data and visualization came into play as we reflected on our own experiences navigating these data landscapes. By documenting our paths through note-taking, diagramming, and visualizing our steps, we deepened our understanding of the datasets themselves and uncovered what was missing, what was hard to access, and where our questions should lead next.

Central to our pedagogy was encouraging students to critically reflect on their data practices. We structured reflective exercises to surface the implicit power dynamics in data collection and usage. Students were prompted regularly to question: *Whose data are we using? Who collected it, and for whose benefit? Who*

controls access, and how does that affect analysis? This reflexivity deepened our critical engagement, enabling us to overcome technical challenges and interpret the implications of their findings from the data. We suggest one possible pathway to actively take on the role of being a Data Detective:

- Initially clarify what one is looking for - this is before one has the data.
- Develop a timeline starting from the current moment, which will track the process by which one gains or loses access to the data.
- Choose a currently promising direction to find more information (could be: ask a person, search on the web, go to an institution, etc.)
- Collect and reflect on the information collected, filling in one’s timeline, with data, facts, responses, including emotional and frustration level responses.

Actively conducting Data Detective projects in our class, where we used personal visualization of our detective process to teach us about both institutional and personal data, whilst revealing many factors about our society. Each data point gathered and each visualization created represents a small act of making the invisible visible, contributing to more equitable and inclusive understandings of our complex social world.

REFERENCES

- [1] C. D’ignazio and L. F. Klein. *Data feminism*. MIT press, 2023.
- [2] C. Donelson and J. Morrow. *The Data Detective at the Carnival*. Technics Publications, Piscataway, NJ, Aug. 2024. Illustrated by Mariam Trejo.
- [3] T. Harford. *The data detective: Ten easy rules to make sense of statistics*. Penguin, 2022.
- [4] A. Inselberg. Multidimensional detective. In *Proceedings of VIZ’97: Visualization Conference, Information Visualization Symposium and Parallel Rendering Symposium*, pp. 100–107. IEEE, 1997.
- [5] L. Lotteraner, R. Schuster, and J. Staudner. Engaging young minds: Learnings from data visualization workshops with children and teenagers. 2025.