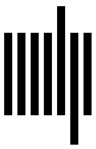
This is a section of doi:10.7551/mitpress/11805.001.0001

Data Feminism

By: Catherine D'Ignazio, Lauren F. Klein

Citation:

Data Feminism
By: Catherine D'Ignazio, Lauren F. Klein
DOI: 10.7551/mitpress/11805.001.0001
ISBN (electronic): 9780262358521
Publisher: The MIT Press
Published: 2023



The MIT Press

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

2 Collect, Analyze, Imagine, Teach

Principle: Challenge Power

Data feminism commits to challenging unequal power structures and working toward justice.

In 1971, the Detroit Geographic Expedition and Institute (DGEI) released a provocative map, *Where Commuters Run Over Black Children on the Pointes-Downtown Track*. The map (figure 2.1) uses sharp black dots to illustrate the places in the community where the children were killed. On one single street corner, there were six Black children killed by white drivers over the course of six months. On the map, the dots blot out that entire block.

The people who lived along the deadly route had long recognized the magnitude of the problem, as well as its profound impact on the lives of their friends and neighbors. But gathering data in support of this truth turned out to be a major challenge. No one was keeping detailed records of these deaths, nor was anyone making even more basic information about what had happened publicly available. "We couldn't get that information," explains Gwendolyn Warren, the Detroit-based organizer who headed the unlikely collaboration: an alliance between Black young adults from the surrounding neighborhoods and a group led by white male academic geographers from nearby universities.[1] Through the collaboration, the youth learned cutting-edge mapping techniques and, guided by Warren, leveraged their local knowledge in order to produce a series of comprehensive reports, covering topics such as the social and economic inequities among neighborhood children and proposals for new, more racially equitable school district boundaries.

Compare the DGEI map with another map of Detroit made thirty years earlier, *Residential Security Map* (figure 2.2). Both maps use straightforward cartographic techniques: an aerial view, legends and keys, and shading. But the similarities end there. The maps differ in terms of visual style, of course. But more profound is how they diverge in terms of the worldviews of their makers and the communities they seek to support. The latter map was made by the Detroit Board of Commerce, which consisted

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

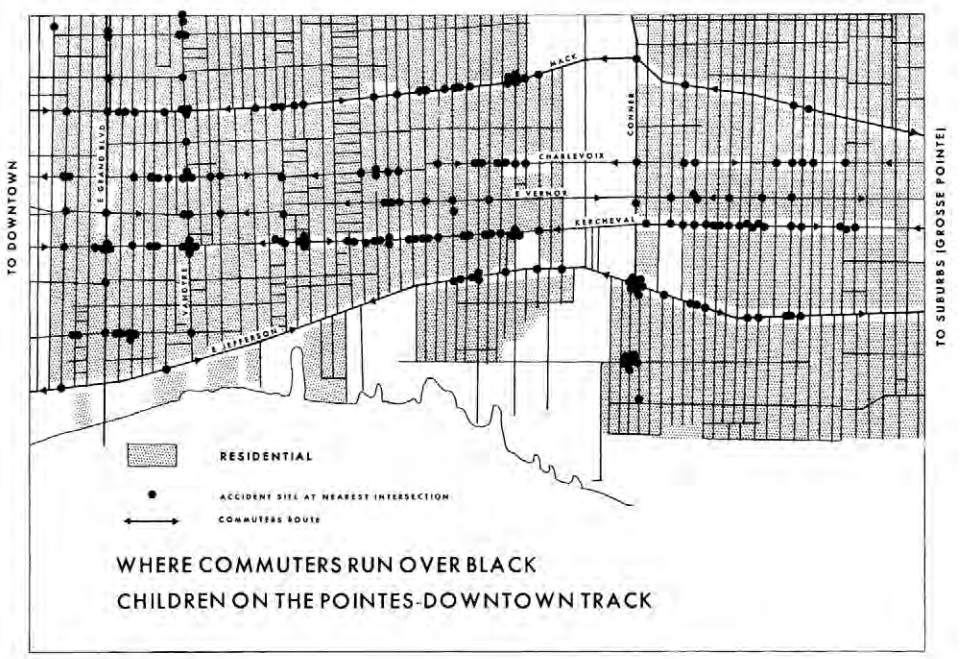


Figure 2.1

Where Commuters Run Over Black Children on the Pointes-Downtown Track (1971) is one image from a report, "Field Notes No. 3: The Geography of Children" which documented the racial inequities of Detroit children. The map was created by Gwendolyn Warren, the administrative director of the Detroit Geographic Expedition and Institute (DGEI), in a collaboration between Black young adults in Detroit and white academic geographers that lasted from 1968–1971. The group worked together to map aspects of the urban environment related to children and education. Warren also worked to set up a free school at which young adults could take college classes in geography for credit. Courtesy of Gwendolyn Warren and the Detroit Geographical Expedition and Institute.

of only white men, in collaboration with the Federal Home Loan Bank Board, which consisted mostly of white men. Far from emancipatory, this map was one of the earliest instances of the practice of *redlining*, a term used to describe how banks rated the risk of granting loans to potential homeowners on the basis of neighborhood demographics (specifically race and ethnicity), rather than individual creditworthiness.

Redlining gets its name because the practice first involved drawing literal red lines on a map. (Sometimes the areas were shaded red instead, as in the map in figure 2.2.) All of Detroit's Black neighborhoods fall into red areas on this map because housing discrimination and other forms of structural oppression predated the practice.[2] But denying home loans to the people who lived in these neighborhoods reinforced those

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

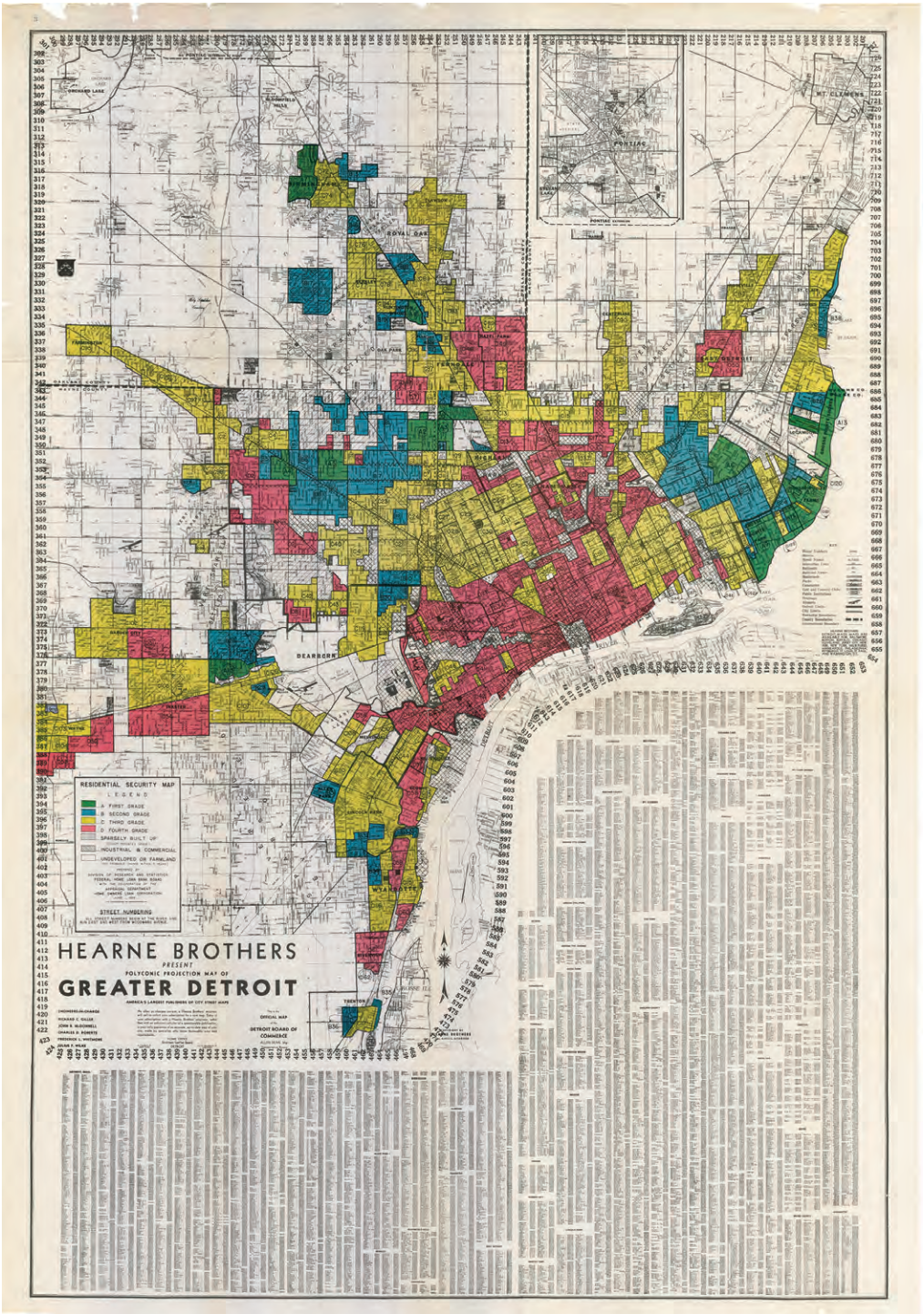


Figure 2.2
Residential Security Map, a redlining map of Detroit published in 1939. Created as a collaboration between the (all white and male) Detroit Chamber of Commerce and the (majority white and male) Federal Home Loan Bank Board, the red colors signify neighborhoods that these institutions deemed red neighborhoods, at "high risk" for bank loans. Courtesy of Robert K. Nelson, LaDale Winling, Richard Marciano, Nathan Connolly, et al., *Mapping Inequality: Redlining in New Deal America*.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

existing inequalities and, as decades of research have shown, were directly responsible for making them worse.[3]

Early twentieth-century redlining maps had an aura very similar to the "big data" approaches of today. These high-tech, scalable "solutions" were deployed across the nation, and they were one method among many that worked to ensure that wealth remained attached to the racial category of whiteness.[4] At the same time that these maps were being made, the insurance industry, for example, was implementing similar data-driven methods for granting (or denying) policies to customers based on their demographics. Zoning laws that were explicitly based on race had already been declared unconstitutional; but within neighborhoods, so-called covenants were nearly as exclusionary and completely legal.[5] This is a phenomenon that political philosopher Cedric Robinson famously termed *racial capitalism*, and it continues into the present in the form of algorithmically generated credit scores that are consistently biased and in the consolidation of "the 1 percent" through the tax code, to give only two examples of many.[6] What's more, the benefits of whiteness accrue: "Whiteness retains its value as a 'consolation prize,'" civil rights scholar Cheryl Harris explains. "It does not mean that all whites will win, but simply that they will not lose."[7]

Who makes maps and who gets mapped? The redlining map is one that secures the power of its makers: the white men on the Detroit Board of Commerce, their families, and their communities. This particular redlining map is even called *Residential Security Map*. But the title reflects more than a desire to secure property values. Rather, it reveals a broader desire to protect and preserve home ownership as a method of accumulating wealth, and therefore status and power, that was available to white people only. In far too many cases, data-driven "solutions" are still deployed in similar ways: in support of the interests of the people and institutions in positions of power, whose worldviews and value systems differ vastly from those of the communities whose data the systems rely upon.[8]

The DGEI map, by contrast, challenges this unequal distribution of data and power. It does so in three key ways. First, in the face of missing data, DGEI compiled its own counterdata. Warren describes how she developed relationships with "political people in order to use them as a means of getting information from the police department in order to find out exactly what time, where, how and who killed [each] child."[9] Second, the DGEI map plotted the data they collected with the deliberate aim of quantifying structural oppression. They intentionally and explicitly focused on the problems of "death, hunger, pain, sorrow and frustration in children," as they explain in the report.[10] Finally, the DGEI map was made by young Black people who lived in the community, under the leadership of a Black woman who was an organizer in the community, with support provided by the academic geographers.[11] The identities of these

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

makers matter, their proximity to the subject matter matters, the terms of their collaboration matter, and the leadership of the project matters.[12]

For these reasons, the DGEI provides a model of the second principle of data feminism: *challenge power*. Challenging power requires mobilizing data science to push back against existing and unequal power structures and to work toward more just and equitable futures. As we will discuss in this chapter, the goal of challenging power is closely linked to the act of examining power, the first principle of data feminism. In fact, the first step of challenging power is to examine that power. But the next step—and the reason we have chosen to dedicate two principles to the topic of power—is to take action against an unjust status quo.

Taking action can itself take many forms, and in this chapter we offer four starting points: (1) *Collect:* Compiling counterdata—in the face of missing data or institutional neglect—offers a powerful starting point as we see in the example of the DGEI, or in María Salguero's femicide maps discussed in chapter 1. (2) *Analyze:* Challenging power often requires demonstrating inequitable outcomes across groups, and new computational methods are being developed to audit opaque algorithms and hold institutions accountable. (3) *Imagine:* We cannot *only* focus on inequitable outcomes, because then we will never get to the root cause of injustice. In order to truly dismantle power, we have to imagine our end point not as "fairness," but as co-liberation. (4) *Teach:* The identities of data scientists matter, so how might we engage and empower newcomers to the field in order to shift the demographics and cultivate the next generation of data feminists?

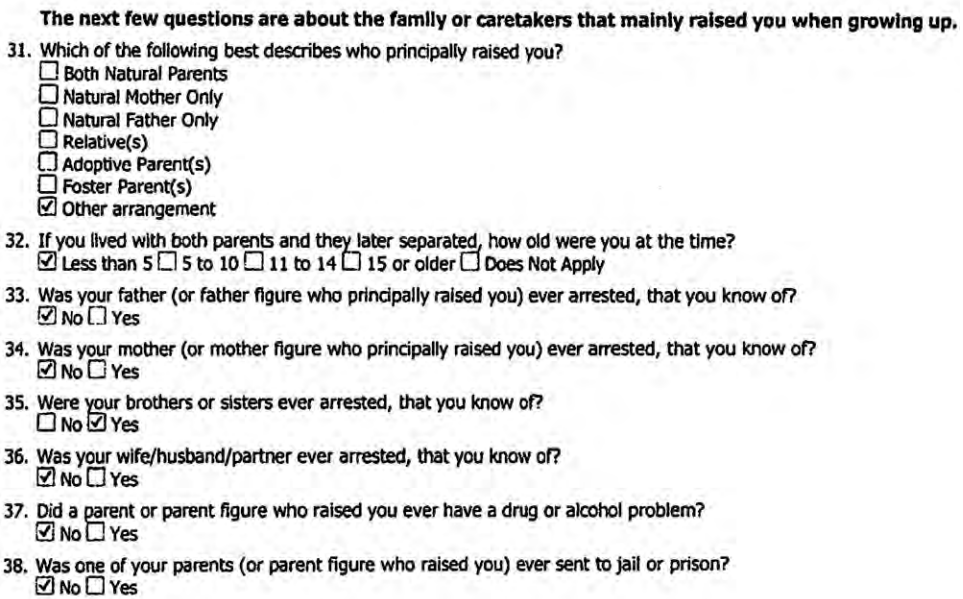
Analyze and Expose Oppression

One can make a direct comparison between yesterday's redlining maps and today's risk assessment algorithms. The latter are used in many cities in the United States today to inform judgments about the length of a particular prison sentence, the amount of bail that should be set, and even whether bail should be set in the first place. The "risk" in their name has to do with the likelihood of a person detained by the police committing a future crime. Risk assessment algorithms produce scores that influence whether a person is sent to jail or set free, effectively altering the course of their life.

But risk assessment algorithms, like redlining maps, are neither neutral nor objective. In 2016, Julia Angwin led a team at ProPublica to investigate one of the most widely used risk assessment algorithms in the United States, created by the company Northpointe (now Equivant).[13] Her team found that white defendants are more often mislabeled as low risk than Black defendants and, conversely, that Black defendants are mislabeled as high risk more often than white defendants.[14] Digging further into

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

the process, the journalists uncovered a 137-question worksheet that each detainee is required to fill out (figure 2.3). The detainee's answers feed into the software, in which they are compared with other data to determine that person's risk score. Although the questionnaire does not ask directly about race, it asks questions that, given the structural inequalities embedded in US culture, serve as proxies for race. These include questions like whether you were raised by a single mother, whether you have ever been suspended from school, or whether you have friends or family that have been arrested. In the United States, each of those questions is linked to a set of larger social, cultural, and political—and, more often than not, racial—realities. For instance, it has been demonstrated that 67 percent of Black kids grow up in single-parent households,



The next few questions are about the family or caretakers that mainly raised you when growing up.

31. Which of the following best describes who principally raised you?
☐ Both Natural Parents
☐ Natural Mother Only
☐ Natural Father Only
☐ Relative(s)
☐ Adoptive Parent(s)
☐ Foster Parent(s)
☑ Other arrangement

32. If you lived with both parents and they later separated, how old were you at the time?
☑ Less than 5 ☐ 5 to 10 ☐ 11 to 14 ☐ 15 or older ☐ Does Not Apply

33. Was your father (or father figure who principally raised you) ever arrested, that you know of?
☑ No ☐ Yes

34. Was your mother (or mother figure who principally raised you) ever arrested, that you know of?
☑ No ☐ Yes

35. Were your brothers or sisters ever arrested, that you know of?
☐ No ☑ Yes

36. Was your wife/husband/partner ever arrested, that you know of?
☑ No ☐ Yes

37. Did a parent or parent figure who raised you ever have a drug or alcohol problem?
☑ No ☐ Yes

38. Was one of your parents (or parent figure who raised you) ever sent to jail or prison?
☑ No ☐ Yes

Figure 2.3
Equivant's risk assessment algorithm is called Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) and is derived from a defendant's answers to a 137-question survey about their upbringing, personality, family, and friends, including many questions that can be considered proxies for race, such as whether they were raised by a single mother. Note that evidence of family criminality would not be admissible evidence in a court case for a crime committed by an individual, but here it is used as a factor in making important decisions about a person's freedom. Courtesy of Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner for ProPublica, 2016.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

whereas only 25 percent of white kids do.[15] Similarly, studies have shown that Black kids are punished more harshly than are white kids for the same minor infractions, starting as early as preschool.[16] So, though the algorithm's creators claim that they do not consider race, race is embedded into the data they are choosing to employ. What's more, they are using that information to further disadvantage Black people, whether because of an erroneous belief in the objectivity of their data, or because they remain unmoved by the evidence of how racism is operating through their technology.

Sociologist Ruha Benjamin has a term for these situations: the *New Jim Code*—where software code and a false sense of objectivity come together to contain and control the lives of Black people, and of other people of color.[17] In this regard, the redlining map and the Equivant risk assessment algorithm share some additional similarities. Both use aggregated data about *social groups* to make decisions about *individuals*: Should we grant a loan to this person? What's the risk that this person will reoffend? Furthermore, both use past data to predict future behavior—and to constrain it. In both cases, the past data in question (like segregated housing patterns or single parentage) are *products* of structurally unequal conditions. These unequal conditions are true across large social groups, and yet the technology uses those data as *predictive elements* that will influence one person's future. Surya Mattu, a former ProPublica reporter who worked on the story, makes this point directly: "Equivant didn't account for the fact that African Americans are more likely to be arrested by the police regardless of whether they committed a crime or not. The system makes an assumption that if you have been arrested you are probably at higher risk."[18] This is one of the challenges of using data about people as an input into a system: the data are never "raw." Data are always the product of unequal social relations—relations affected by centuries of history. As computer scientist Ben Green states, "Although most people talk about machine learning's ability to predict the future, what it really does is predict the past."[19] Effectively such "predictive" software reinforces existing demographic divisions, amplifying the social inequities that have limited certain groups for generations. The danger of the New Jim Code is that these findings are actively promoted as objective, and they track individuals and groups through their lives and limit their future potential.

But machine learning algorithms don't just predict the past; they also reflect current social inequities. A less well-known finding from the ProPublica investigation of Equivant, for example, is that it also surfaced significantly different treatment of women by the algorithm. Due to a range of factors, women tend to recidivate—to commit new crimes—less than men do. That means the risk scale for women "is such that somebody with a high risk score that's a woman is generally about the level of a medium risk score

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

for a man. So, it's actually really shocking that judges are looking at these and thinking that high risk means the same thing for a man and a woman when it doesn't," explains lead reporter Julia Angwin.[20]

Angwin decided to focus the story on race in part because of the prior work of criminologists such as Kristy Holtfreter, which had already highlighted some of these gender differentials.[21] But there was another factor at play in her editorial decision: workplace sexism faced by women reporters like Angwin herself. Angwin explains how she had always been wary of working on stories about women and gender because she wanted to avoid becoming pigeonholed as a reporter who *only* worked on stories about women and gender. But, she explains, "one of the things I woke up to during the #MeToo movement was how many decisions like that I had made over the years"—an internalized form of oppression that had discouraged her from covering those important issues. In early 2018, when we conducted this interview, Angwin was hiring for her own data journalism startup, the Markup, founded with a goal of using data-driven methods to investigate the differential harms and benefits of new technologies on society. She was encouraged to see how many job candidates of all genders were pitching stories on issues relating to gender inequality. "In the era of data and AI, the challenge is that accountability is hard to prove and hard to trace," she explains. "The challenge for journalism is to try to make as concrete as possible those linkages when we can so we can show the world what the harms are."

Angwin is pointing out a tricky issue that is unlikely to go away. The field of journalism has long prided itself on "speaking truth to power." But today, the location of that power has shifted from people and corporations to the datasets and models that they create and employ. These datasets and models require new methods of interrogation, particularly when they—like Equivant's—are proprietary. How does one report on a *black box*, as these harmful algorithms are sometimes described?[22] Much like the situation encountered by Gwendolyn Warren when she looked into the data on the Detroit children's deaths, or like María Salguero when she started logging femicides in Mexico, ProPublica found no existing studies that examined whether the risk scores were racially biased, or existing datasets they could use to point them to answers. To write the risk assessment story, ProPublica had to assemble a dataset of their own. The researchers looked at ten thousand criminal defendants from a single county in Florida and compared their recidivism risk scores with people who actually reoffended in a two-year period. After doing some initial exploratory analysis, they created their own regression model that considered race, age, criminal history, future recidivism, charge degree, and gender. They found that age, race, and gender were the strongest predictors of who received a high risk score—with Black defendants 77 percent more likely than

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

white ones to receive a higher violent recidivism score. Their analysis also included creating models to test the overall accuracy of the COMPAS model over time and an investigation of errors to see if there were racial differences in the distribution of false positives and false negatives. As it turns out, there were: the system was more likely to predict that white people would not commit additional crimes if released, when they actually did recidivate.[23]

Angwin and her coauthors used data science to challenge data science. By collecting missing data and reverse-engineering the algorithm that was judging each defendant's risk, they were able to prove systemic racial bias. This analysis method is called *auditing algorithms* and it is being increasingly used in journalism and in academic research in order show how the harms and benefits of automated systems are differentially distributed. Computational journalism researcher Nicholas Diakopoulos has proposed that work like this become formalized into an algorithm accountability beat, which would help to make the practice more widespread.[24] He and computer scientist Sorelle Friedler have asserted that algorithms need to be held "publicly accountable" for their consequences, and the press is one place where this accounting can take place.[25] By providing proof of how racism and sexism, among other oppressions, create unequal outcomes across social groups, analyzing data is a powerful strategy for challenging power and working toward justice.

The Pitfalls of Proof

Let's pause here for a feminist *who question*, as we introduced in chapter 1. Who is it, exactly, that needs to be shown the harms of such differentials of power? And what kind of proof do they require to believe that oppression is real? Women who experience instances of sexism, as Angwin did in her workplace, already know the harms of that oppressive behavior. The young adults whom Gwendolyn Warren worked with in Detroit already knew intimately that the white commuters were killing their Black neighbors and friends. They had no need to prove to their own communities that structural racism was a factor in these deaths. Rather, their goal in partnering with the DGEI was to prove the structural nature of the problem to those in positions of power. Those dominant groups and institutions were the ones that, by privileging their own social, political, and economic interests, bore much of the responsibility for the problem; and they also, because of the phenomenon we have described as a privilege hazard, were unlikely to see that such problem existed in the first place. The theory of change that motivates these efforts to use data as evidence, or "proof," is that by being made aware of the extent of the problem, those in power will be prompted to take action.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

These kinds of data-driven revelations can certainly be compelling. When the analysis appears in a high-profile newspaper or blog or TV show (in other words: a place white enough and male enough to be considered mainstream), it can indeed prompt people in power to act. The ProPublica story on risk assessment algorithms, for example, prompted a New York City council member to propose an algorithmic accountability bill. Enacted in 2018, the bill became the first legal measure to tackle algorithmic discrimination in the United States and led to the creation of a task force focused on "equity and fairness" in city algorithms.[26] Should the city implement some of the task force's recommendations, it would influence the work of software vendors, as well as legislation in other cities. This path of influence—from community problem to gathering proof to informed reporting to policy change—represents the best aspirations of speaking truth to power.[27]

While analyzing and exposing oppression in order to hold institutions accountable can be extremely useful, its efficacy comes with two caveats. Proof can just as easily become part of an endless loop if not accompanied by other tools of community engagement, political organizing, and protest. Any data-based evidence can be minimized because it is not "big" enough, not "clean" enough, or not "newsworthy" enough to justify a meaningful response from institutions that have a vested interest in maintaining the status quo.[28] As we saw in chapter 1, María Salguero's data on femicides was augmented by government commissions, reports from international agencies, and rulings of international courts. But none of those data-gathering efforts have been enough to prompt comprehensive action.

Another feminist *who question*: On whom is the burden of proof is placed? In 2015, communications researcher Candice Lanius wrote a widely shared blog post, "Fact Check: Your Demand for Statistical Proof is Racist," in which she summarizes the ample research on how those in positions of power accept anecdotal evidence from those like themselves, but demand endless statistics from minoritized groups.[29] In those cases, she argues convincingly, more data will never be enough.

Proof can also unwittingly compound the harmful narratives—whether sexist or racist or ableist or otherwise oppressive—that are already circulating in the culture, inadvertently contributing to what are known as *deficit narratives*. These narratives reduce a group or culture to its "problems," rather than portraying it with the strengths, creativity, and agency that people from those cultures possess. For example, in their book *Indigenous Statistics*, Maggie Walter and Chris Anderson describe how statistics used by settler colonial groups to describe Indigenous populations have mainly functioned as "documentation of difference, deficit, and dysfunction."[30] This can occur even when the creators have good intentions—for example, as Kimberly Seals Allers

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

notes (see chapter 1), a great deal of the media reporting on Black maternal mortality data falls into the deficit narrative category. It portrays Black women as victims and fails to amplify the efforts of the Black women who have been working on the issue for decades.

This goes for gender data as well. "What little data we collect about women tends to be either about their experience of violence or reproductive health," explains Nina Rabinovitch Blecker, who directs communications for Data2X, a nonprofit aimed at improving the quality of data related to gender in a global context.[31] The current data encourage additional deficit narratives—in which women are relentlessly and reductively portrayed as victims of violent crimes like murder, rape, or intimate partner violence. These narratives imply that the subjects of the data have no agency and need "saving" from governments, international institutions, or concerned citizens. As one step to counteract that, Blecker chose to publish an example from Uruguay that didn't focus on violence, but rather on quantifying women's unseen contributions to the economy.[32]

So, though collecting counterdata and analyzing data to provide proof of oppression remain worthy goals, it is equally important to remain aware of how the subjects of oppression are portrayed. Working with communities directly, which we talk more about in chapter 5, is the surest remedy to these harms. Indigenous researcher Maggie Walter explains that ownership of the process is key in order to stop the propagation of deficit narratives: "We [Indigenous people] must have real power in how statistics about us are done—where, when and how."[33] Key too is a sustained attention to the ways in which communities themselves are already addressing the issues. These actions are often more creative, more effective, and more culturally grounded than the actions that any outside organization would take.

Envision Equity, Imagine Co-liberation

As the examples discussed thus far in this book clearly demonstrate, one of the most dangerous outcomes of the tools of data and data science being consolidated in the hands of dominant groups is that these groups are able to obscure their politics and their goals behind their technologies. Benjamin, whose book *Race after Technology: Abolitionist Tools for the New Jim Code* (mentioned earlier), describes this phenomenon as the "imagined objectivity of data and technology" because data-driven systems like redlining and risk assessment algorithms are not really objective at all.[34] Her concept of *imagined objectivity* emphasizes the role that cultural assumptions and personal preconceptions play in upholding this false belief: one imagines (wrongly) that datasets and algorithms are less partial and less discriminatory than people and thus more

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

"objective."[35] But as we discuss in chapter 1, these data products seem objective only because the perspectives of those who produce them—elite, white men and the institutions they control—pass for the default. Assumptions about objectivity are becoming a major focus in data science and related fields as algorithm after algorithm is revealed to be sexist, racist, or otherwise flawed. What can the people who design these computational systems do to avoid these pitfalls? And what can everyone else do to help them and hold them accountable?

The quest for answers to these questions has prompted the development of a new area of research known as *data ethics*. It represents a growing interdisciplinary effort—both critical and computational—to ensure that the ethical issues brought about by our increasing reliance on data-driven systems are identified and addressed. Thus far, the major trend has been to emphasize the issue of "bias," and the values of "fairness, accountability, and transparency" in mitigating its effects.[36] This is a promising development, especially for technical fields that have not historically foregrounded ethical issues, and as funding mechanisms for research on data and ethics proliferate.[37] However, as Benjamin's concept of imagined objectivity helps to show, addressing bias in a dataset is a tiny technological Band-Aid for a much larger problem. Even the values mentioned here, which seek to address instances of bias in data-driven systems, are themselves non-neutral, as they locate the source of the bias in individual people and specific design decisions. So how might we develop a practice that results in data-driven systems that challenge power at its source?

The following chart (table 2.1) introduces an alternate set of orienting concepts for the field: these are the six ideals that we believe should guide data ethics work. These

Table 2.1
From data ethics to data justice

Concepts That Secure Power Because they locate the source of the problem in individuals or technical systems	**Concepts That Challenge Power** Because they acknowledge structural power differentials and work toward dismantling them
Ethics	Justice
Bias	Oppression
Fairness	Equity
Accountability	Co-liberation
Transparency	Reflexivity
Understanding algorithms	Understanding history, culture, and context

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

concepts all have legacies in intersectional feminist activism, collective organizing, and critical thought, and they are unabashedly explicit in how they work toward justice.

In the left-hand column, we list some of the major concepts that are currently circulating in conversations about the uses of data and algorithms in public (and private) life. These are a step forward, but they do not go far enough. On the right-hand side, we list adjacent concepts that emerge from a grounding in intersectional feminist activism and critical thought. The gap between these two columns represents a fundamental difference in view of why injustice arises and how it operates in the world. The concepts on the left are based on the assumption that injustice arises as a result of flawed individuals or small groups ("bad apples," "racist cops," "brogrammers") or flawed technical systems ("the algorithm/dataset did it"). Although flawed individuals and flawed systems certainly exist, they are not the root cause of the problems that occur again and again in data and algorithms.

What is the root cause? If you've read chapter 1, you know the answer: the matrix of domination, the matrix of domination, and the matrix of domination. The concepts on the left may do good work, but they ultimately keep the roots of the problem in place. In other words, they maintain the current structure of power, even if they don't intend to, because they let the matrix of domination off the hook. They direct data scientists' attention toward seeking technological fixes. Sometimes those fixes are necessary and important. But as technology scholars Julia Powles and Helen Nissenbaum assert, "Bias is real, but it's also a captivating diversion."[38] There is a more fundamental problem that must also be addressed: we do not all arrive in the present with equal power or privilege. Hundreds of years of history and politics and culture have brought us to the present moment. This is a reality of our lives as well as our data. A broader focus on *data justice*, rather than *data ethics* alone, can help to ensure that past inequities are not distilled into black-boxed algorithms that, like the redlining maps of the twentieth century, determine the course of people's lives in the twenty-first.

In proposing this chart, we are not suggesting that ethics have no place in data science, that bias in datasets should not be addressed, or that issues of transparency should go ignored.[39] Rather, the main point is that the concepts on the left are inadequate *on their own* to account for the root causes of structural oppression. By not taking root causes into account, they limit the range of responses possible to challenge power and work toward justice. In contrast, the concepts on the right start from the basic feminist belief that oppression is real, historic, ongoing, and worth dismantling.

Media theorist and designer Sasha Costanza-Chock proposes a restorative approach to data justice.[40] Drawing from theories of restorative justice—meaning that decisions should be made in ways that recognize and rectify any harms of the

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

past—Costanza-Chock asserts that any notion of algorithmic *fairness* must also acknowledge the systematic nature of the *unfairness* that has long been perpetrated by certain groups on others. They give the example of college admissions—a topic that always seems to be in the news, not the least because it's a major mechanism of protecting privilege.[41] A restorative approach to college admissions would entail making decisions about who gets admitted in the present on the basis of who was historically not admitted in the past—like women, who were excluded from MIT, where Constanza-Chock teaches, for decades. According to this model, a "fair" present-day entering class that accounts for history might be composed of 90 percent women and people of color.[42]

Does this approach make fairness political? Emphatically yes, because all systems are political. In fact, the appeal to avoid politics is a very familiar way for those in power to attempt to hold onto it.[43] The ability to sidestep politics is a privilege in itself—held only by those whose existence does not challenge the status quo. If you are a Black woman or a Muslim man or a transgender service member and you live in the United States today, your being in the world is political, whether or not you want it to be.[44] So rather than design algorithms that purport to be "color-blind" (since color-blindness is of course a myth), Costanza-Chock explains that we should be designing algorithms that are *just*.[45] This means shifting from the ahistorical notion of fairness to a model of *equity*.

Equity is justice of a specific flavor, and it is different than equality. Equality is measured from a starting point in the present: $t = 0$, where t equals time and 0 indicates that no time has elapsed since now. Based on this formula, the principle of equality would hold that resources and/or punishments should be doled out according to what is happening in the present moment—the time when $t = 0$. But this formula for equal treatment means that those who are ahead in the present can go further, achieve more, and stay on top, whereas those who start out behind can never catch up. Kiddada Green, executive director of the Black Mothers Breastfeeding Association, makes the case that in a country where Black babies are dying at twice the rate of white babies, equality is actually systematically unfair: "There is a level of political correctness in America that causes some people to believe that equality is the way to go. Even when equality is unfair, some say that it's the right thing to do."[46] Working toward a world in which everyone is treated equitably, not equally, means taking into account these present power differentials and distributing (or redistributing) resources accordingly. Equity is much harder to model computationally than equality—as it needs to take time, history, and differential power into account—but it is not impossible.[47]

This difficulty also underscores the point that bias (in individuals, in datasets, in statistical models, or in algorithms) is not a strong enough concept in which to anchor

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

ideas about equity and justice. In writing about the creation of New York's Welfare Management System in the early 1970s, for example, Virginia Eubanks describes: "These early big data systems were built on a specific understanding of what constitutes discrimination: personal bias."[48] The solution at the time was to remove the humans from the loop, and it remains so today: without potentially bad—in this case, racist—apples, there would be less discrimination. But this line of thinking illustrates what whiteness studies scholar Robin DiAngelo would call the *new racism*: the belief that racism is due to individual bad actors, rather than structures or systems.[49] In relation to welfare management, Eubanks emphasizes that this often meant replacing social workers, who were often women of color, and who had empathy and flexibility and listening skills, with an automated system that applies a set of rigid criteria, no matter what the circumstances.

While bias remains a serious problem, it should not be viewed as something that can be fixed after the fact. Instead, we must look to understand and design systems that address the source of the bias: structural oppression. In truth, oppression is itself an outcome, one that results from the matrix of domination. In this model, majoritized bodies are granted undeserved advantages and minoritized bodies must survive undeserved hardships. Starting from the assumption that oppression is the problem, not bias, leads to fundamentally different decisions about what to work on, who to work with, and when to stand up and say that a problem cannot and should not be solved by data and technology.[50] Why should we settle for retroactive audits of potentially flawed systems if we can design with a goal of co-liberation from the start?[51] And here, *co-liberation* doesn't mean "free the data," but rather "free the people." The people in question are not only those with less privilege, but also those with more privilege: data scientists, designers, researchers, and educators—in other words, those like ourselves—who play a role in upholding oppressive systems.

The key to co-liberation is that it requires a commitment to and a belief in mutual benefit, from members of both dominant groups and minoritized groups; that's the *co* in the term. Too often, acts of data service performed by tech companies are framed as charity work (we discuss the limits of "data for good" in chapter 5). The frame of co-liberation equalizes this exchange as a form of relationship building and demographic healing. There is a famous saying credited to aboriginal activists in Queensland, Australia, from the 1970s: "If you have come here to help me, you are wasting your time. But if you have come because your liberation is bound up with mine, then let us work together."[52]

What does this mean? As poet and community organizer Tawana Petty explains in relation to efforts around antiracism in the United States: "We need whites to firmly

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

believe that their liberation, their humanity, is also dependent upon the destruction of racism and the dismantling of white supremacy."[53] The same goes for gender: men are often not prompted to think about how unequal gender relations seep into the institutions they dominate, resulting in harm for everyone.

This goal of co-liberation motivates the Our Data Bodies (ODB) project. Led by a group of five women, including Gangadharan and Petty, who sit at the intersection of academia and organizing work, this project is a community-centered initiative focused on data collection efforts that disproportionately impact minoritized people. Working with community organizations in three US cities, the ODB project has led participatory research initiatives and educational workshops, culminating in the recently released *Digital Defense Playbook*, a set of activities, tools, and tip sheets intended to be used by and for marginalized communities to understand how data-driven technologies impact their lives.[54]

Digital Defense Playbook was born out of many years of relationship-building and research, as well as a deliberate shift. The group explains in the playbook's introduction, "We wanted to shift who gets to define problems around data collection, data privacy, and data security—from elites to impacted communities; shine a light on how communities have been confronting data-driven problems as well as how they wish to confront these problems; and forge an analysis of data and data-driven technologies from and with allied struggles."[55] In so doing, the ODB project demonstrates how co-liberation requires not only transparency of methods but also *reflexivity*: the ability to reflect on and take responsibility for one's own position within the multiple, intersecting dimensions of the matrix of domination. Along the way, the scholars and organizers involved in the project decided to shift their research agenda, which had begun as a general project about data profiling and resistance, to surveillance, in response to the problems voiced by the communities themselves.[56]

Even within big tech itself, there is evidence of an increasing sense of reflexivity among employees for their role in creating harmful data systems. Employees have pushed back against Google's work with the Department of Defense (DoD) on Project Maven, which uses AI to improve drone strike accuracy; Microsoft's decision to take $480 million from the Department of Defense to develop military applications of its augmented reality headset HoloLens; and Amazon's contract with US Immigration and Customs Enforcement (ICE) to develop its Rekognition platform for use in targeting individuals for detention and deportation at US borders.[57] This pushback has led to the cancelling of the Google and Microsoft projects, as well as political consciousness raising across the sector, which we discuss further in the book's conclusion.[58]

Designing datasets and data systems that dismantle oppression and work toward justice, equity, and co-liberation requires new tools in our collective toolbox. We have

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

some good starting points; building more understandable algorithms is a laudable, worthy research goal. And yet what we need to explain and account for are not only the inner workings of machine learning, but also the history, culture, and context that lead to discriminatory outputs in the first place. For example, it is not an isolated incident that facial analysis software couldn't "see" Joy Buolamwini's face, as we discussed in chapter 1. It is not an isolated incident that the "Lena" image used to test most image-processing algorithms was the centerfold from the November 1972 issue of *Playboy*, cropped demurely at the shoulders.[59] It is not an isolated incident that the women who worked on the ENIAC computer were not invited to the fiftieth anniversary celebration in 1995. It is not an isolated incident that Christine Darden was not promoted as quickly as her male coworkers. None of these are isolated incidents: they are connected data points and eminently measurable and predictable outcomes of the matrix of domination. But you can only detect the pattern if you know the history, culture, and context that surrounds it.

Data people, generally speaking, have choices—choices in who they work for, which projects they work on, and what values they reject.[60] Starting from the assumption that oppression is the problem, equity is the path, and co-liberation is the desired goal leads to fundamentally different projects that challenge power at their source. It also leads to different metrics of success. These extend beyond the efficiency of a database under load, the precision of a classification algorithm, or the size of a user base one year after launch. The success of a project designed with co-liberation in mind would also depend on how much trust was built between institutions and communities, how effectively those with power and resources shared their power and resources, how much learning happened in both directions, how much the people and organizations were transformed in the process, and how much inspiration for future work, together, was co-conspired. These metrics are a little more squishy than the numbers and rankings that we tend to believe are our only option, but utterly and entirely measurable nonetheless.

Teach Data Like an Intersectional Feminist

When Gwendolyn Warren and the DGEI researchers collected their data about hit and runs on Black children or scoured Detroit playgrounds to weigh and measure the broken glass they found, they were not only doing this work to make a data-driven case for change. The "institute" part of the Detroit Geographic Expedition and Institute described the educational wing of the organization that ran classes in data collection, mapping, and cartography. It came about at Warren's insistence that the academic geographers give something back to the community whose knowledge they were drawing

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

upon for their research. She recognized that while a single map or project could make a focused intervention, education would enable her community to come away with a longer-term strategy for challenging power. As it turned out, the institutional affiliations of the academic geographers enabled them to offer free, for-credit college courses, which they taught in the community for community members.

In her emphasis on education, Warren recognized its enduring role as a mechanism of both empowerment and transformation. This belief is not new; as American educational reformer Horace Mann stated famously in 1848, "Education, then, beyond all other divides of human origin, is a great equalizer of conditions of men—the balance wheel of the social machinery." But here is the thing—it really matters how we do that equalizing and who we imagine that equalizing to serve. For his part, Mann was literal about the "men": education was to be an equalizer of men, but only certain men (read: white, Anglo, Christian) and explicitly not women.[61] Warren, on the other hand, recognized that access to education—and to data science education in particular—would have to be expanded in order for it to achieve its equalizing force.

Unfortunately, Warren's transformative vision has still yet to enter the data science classroom. As was true in Mann's era, men still lead. Women faculty comprise less than a third of computer science and statistics faculty. More than 80 percent of artificial intelligence professors are men.[62] This gender imbalance, and the narrowness of vision that results, is compounded by the fact that data science is often framed as an *abstract and technical* pursuit. Steps like cleaning and wrangling data are presented as solely technical conundrums; there is less discussion of the social context, ethics, values, or politics of data.[63] This perpetuates the myth that data science about astrophysics is the same as data science about criminal justice is the same as data science about carbon emissions. This limits the transformative work that can be done. Finally, because the goal of learning data science is modeled as *individual mastery of technical concepts and skills*, communities are not engaged and conversations are restricted. Instead, teachers impart technical knowledge via lectures, and students complete assignments and quizzes individually. We might call this model of teaching "the Horace Mann Factory Model of Data Science," because it represents the exclusionary view that Mann himself advanced. But let's just call it the *Man Factory* for short.

The Man Factory is really good at producing men, mainly elite white men like the ones who already lead the classes. It's not as good at producing women data scientists, or nonbinary data scientists, or data scientists of color. For years, researchers and advocacy organizations have recognized that there are problems with this "pipeline" for technical fields; yet this research is framed around questions like "Why are there so few women computer scientists?" and "Why are women leaving computing?"[64] Note

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

that these questions imply that it is the women who have the problem, inadvertently perpetuating a deficit narrative. Feminist scholars who are studying the issue are, not surprisingly, asking very different questions, like "How can the men running the Man Factory share their power?" and "How can we structurally transform STEM education together?"[65]

One person currently modeling an answer to these questions is Laurie Rubel, the math educator behind the Local Lotto project. If you were on the city streets of Brooklyn or the Bronx in the past five years, you may have inadvertently crossed paths with one of her data science classes. You probably didn't realize it because the classes looked nothing like a traditional classroom (figure 2.4). Teenagers from the neighborhood wandered around in small groups. They were outfitted with tablets, pen and paper, cameras, and maps. They periodically took pictures on the street, walked into bodegas, chatted with passersby in Spanish or English, and entered information on their tablets.

Rubel is a leader in an area called mathematics for spatial justice, which aims to show how mathematical concepts can be taught in ways that relate to justice concerns arising from students' everyday lives, and to do so in dialogue with people in their neighborhoods and communities. The goal of Local Lotto was to develop a place-specific way of teaching concepts related to data and statistics grounded in considerations of equity.[66] Specifically, Rubel and the other organizers of Local Lotto wanted young learners to come up with a data-driven answer to the question: "Is the lottery good or bad for your neighborhood?"

In New York, as in other US states that operate lotteries, lottery ticket sales go back into the state budget—sometimes, but not always, to fund educational programs.[67] But lottery tickets are not purchased equally across all income brackets or all neighborhoods. Low-wage workers buy more tickets than their higher-earning counterparts. What's more, the revenue from ticket purchases is not allocated back to those workers or the places they live. Because of this, scholars have argued that the lottery system is a form of regressive taxation—essentially a "poverty tax"—whereby low-income neighborhoods are "taxed" more because they play more, but do not receive a proportional share of the profit.[68]

The Local Lotto curriculum was designed to expose high school learners to this instance of social inequality. They begin by talking about the lottery and the idea of probability by playing chance-based games. Then they consider jackpot games like the Sweet Millions lottery, advertised by New York State as "your best chance from the New York Lottery to win a million for just a buck." The best chance to win one million, however, turns out to be about one in four million; an entire class session is

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

Figure 2.4
A data science classroom. The Local Lotto project (2012–2015) taught local high school students statistics and data analysis rooted in neighborhood and justice concerns. Courtesy of the City Digits Project Team, including Brooklyn College, the Civic Data Design Lab at MIT and the Center for Urban Pedagogy. This work was supported by the National Science Foundation under Grant No. DRL-1222430.

devoted to a discussion about other instances of "four million" that more closely relate to the learners' lives.[69] The learners then leave the classroom with the goal of collecting data about how other people experience the lottery, which takes them back into their neighborhoods. They map stores that sell lottery tickets. They record interviews with shopkeepers and ticket buyers on their tablets and then geolocate them on their maps. They take pictures of lottery advertising. Afterward, the learners analyze their results and present them to the class. They examine choropleth maps of income levels, they make ratio tables, and they correlate state spending of lottery profits with median family income. (No surprise: there is no correlation.) Finally, they create a data-driven argument: an opinion piece supported with evidence from their statistical and spatial analyses, as well as their fieldwork (figures 2.5 and 2.6).

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

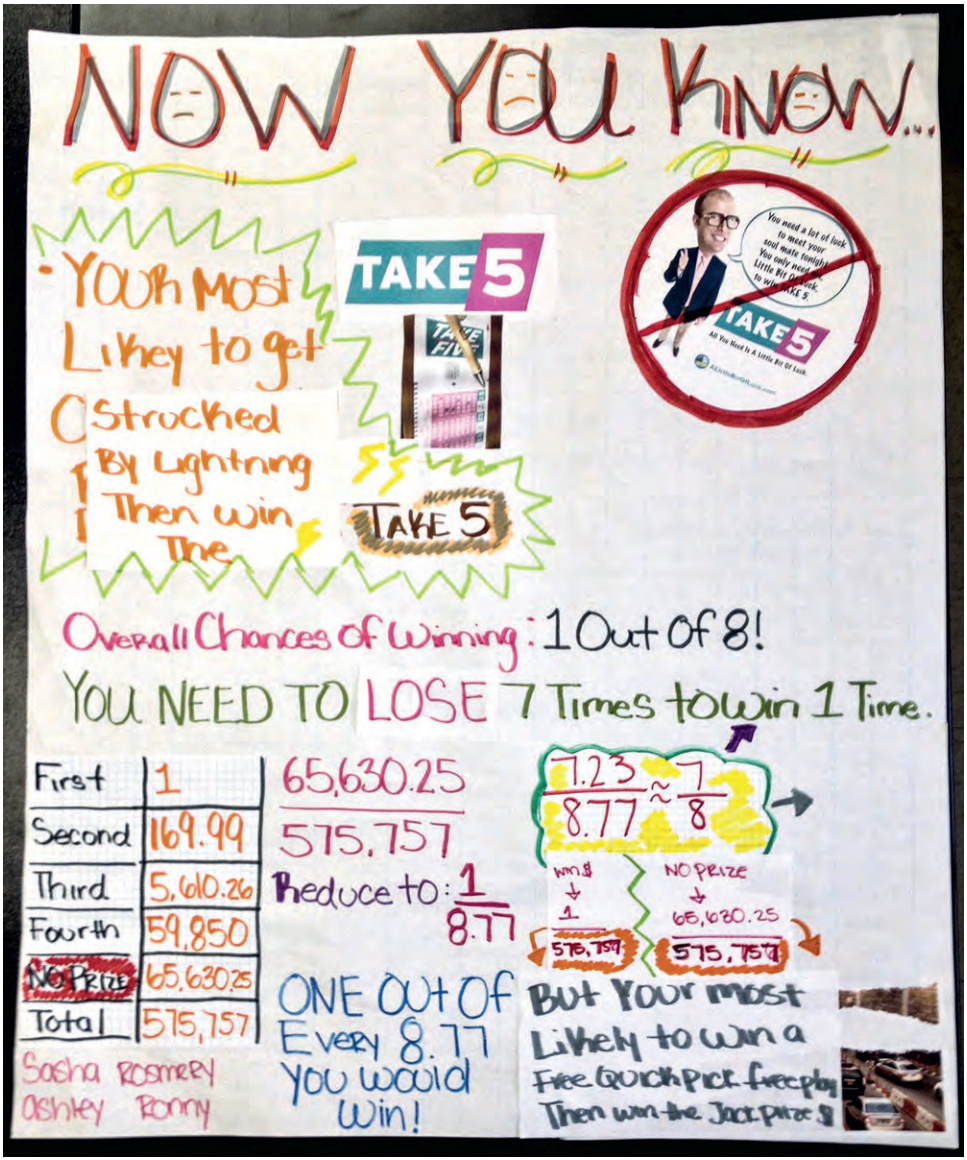


Figure 2.5

"Now You Know." A student's infographic poster responding to the New York Lottery advertising message, "Hey, you never know ... ," that explores the probability of winning anything—ranging from a million dollars to another ticket. Courtesy of the City Digits Project Team. This work was supported by the National Science Foundation under Grant No. DRL-1222430.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

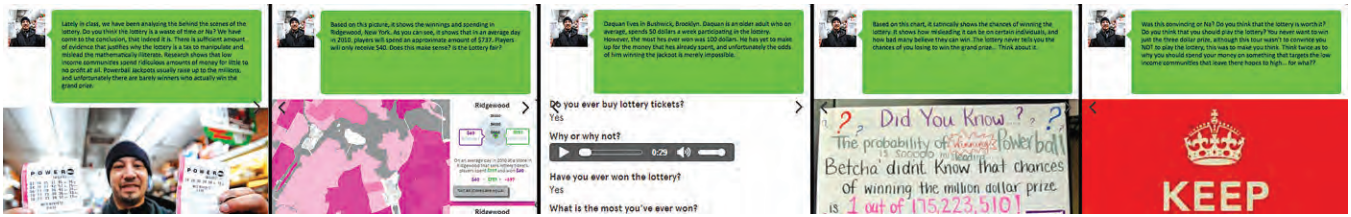


Figure 2.6
Final collaborative opinion piece asserting that the lottery is not good for the students' neighborhood and presenting evidence they collected to back up their case. See the multimedia slideshow at http://citydigits.mit.edu/locallotto#tours-tab. Courtesy of Emmanuela, Angel, Robert, and Janeva. This work was supported by the National Science Foundation under Grant No. DRL-1222430.

By formal measures, the Local Lotto approach worked: before one school's implementation of Local Lotto, only two of forty-seven learners were able to determine the correct number of possible combinations in a lottery example. Later, almost half (twenty-one of forty-seven) were successfully able to calculate the number of combinations. But perhaps more importantly, the Local Lotto approach made math and statistics relevant to the students' lives. One student shared that what he learned was "something new that could help me in my local environment, in my house actually," and that after the course, he tried to convince his mother to spend less money on the lottery by "showing her my math book and all the work." Spanish-speaking women in the class who didn't often participate in classroom discussion became essential translators during the participatory mapping module. Several students went on to teach other teachers about the curriculum, both locally and nationally.[70]

What's different about the Local Lotto approach to teaching data analysis and statistical concepts compared to the Man Factory? How is Local Lotto challenging power both inside and outside the classroom? First, it was woman-led: the project was conceived by three women leaders representing three institutions.[71] Just as with the DGEI map and school, led by Gwendolyn Warren, the identities of the creators matter. Second, rather than modeling data science as abstract and technical, Local Lotto modeled a data science that was grounded in solving ethical questions around social inequality that had relevance for learners' everyday lives: Is the lottery good or bad for your neighborhood? The project valued lived experience: the learners came in as "domain experts" in their neighborhoods. And it valued both qualitative data and quantitative data: the learners spoke with neighborhood residents and connected their beliefs, attitudes, and concerns to probability calculations. Learners used community members' voices as evidence in their final projects. Third, rather than

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

valorizing individual mastery of technical skills as the gold standard, learners worked together during every phase of the project. They used methods from art and design (like the creation of infographics and digital slideshows) to practice communicating with data.

Even as we celebrate these intentional pedagogical choices, the Local Lotto project still had its shortcomings, as the organizers noted in a 2016 paper for *Cognition and Instruction*.[72] Many of these stemmed from a basic fact: the teachers and course designers of the project were white and Asian, whereas the youth in the classes were predominantly Latinx and Black. This led to several issues. For instance, the curriculum designers had intended to focus primarily on income inequality, but they discovered that "the students consistently surfaced race." Because race and ethnicity were not part of the teaching material, the teachers felt that they did not have the experience or background to discuss them explicitly and deflected those conversations. As they write in the paper, "Youth, and in this case youth of color, have different understandings about racial boundaries; theirs are differently nuanced and scaled than affluent, white, or adult perspectives." The organizers are now taking steps to explicitly integrate discussions about race into the curriculum, as well as to include race, ethnicity, and age data in the course projects.[73]

The course designers also encountered "limited but recurring instances of resistance from students" to the project's central focus on income inequality. They attribute this resistance to the fact that the course was developed and taught by outsiders and could be seen as passing judgment on the people in their neighborhoods: that because they were not from the community, the teachers were perpetuating a deficit narrative about low-income people. This is both a sophisticated and very fair pushback from the young learners. Most people, regardless of their wealth or level of education, know they are not going to win the lottery, after all. There is an element of imaginative fantasy in purchasing a ticket. The campaign slogan, "Hey, you never know ..." appeals as much to this fantasy as it does to the reality of the odds, and this fantasy has value too. In reflecting on the unintended sense of judgment experienced by the students, the course designers determined that, in the next iteration of the course, they would work to connect students with people in the communities themselves who are actively working to address issues of income inequality.

In both its successes and its failures, as well as its commitment to iteration and trying again, Local Lotto encapsulates what it means to challenge power and privilege and work toward justice. Justice is a journey. The discomfort that comes along with this journey is par for the course. There is no such thing as mastery of feminism because those who hold positions of privilege—like those in data science, like the Local Lotto

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

course designers, and like us, the authors of this book—are constantly learning how to be better allies and accomplices across difference. In this process, what becomes most important is to "stay with the trouble," as feminist philosopher Donna Haraway would say.[74] Staying with the trouble means persisting in your work, especially when it becomes uncomfortable, unclear, or outright upsetting. One of the biggest strengths of the Local Lotto project is the courage of its creators to publicly, transparently, and reflexively interrogate themselves and their process, to detail their stumbling blocks, and to describe their commitments to doing better in the future.

Challenge Power

After examining power, the next step is to challenge it—map by map, audit by audit, community by community, and classroom by classroom. Collecting counterdata to quantify and visualize structural oppression, as Gwendolyn Warren and the DGEI did with their map, helps those who occupy positions of power understand the scope, scale, and character of the problems from which they are otherwise far removed. Analyzing biased algorithms, as Julia Angwin and ProPublica did, can show the real, material harms of automated systems, as well as build a base of evidence for political or institutional change. At the same time, it is important to remember that minoritized individuals and groups should not have to repeatedly prove that their experiences of oppression are real. And data alone do not always lead to change—especially when that change also requires dominant groups to share their resources and their power.

Those of us who use data in our work must alter some of our most basic assumptions and imagine new starting points. Shifting the frame from concepts that secure power, like fairness and accountability, to those that challenge power, like equity and co-liberation, can help to ensure that data scientists, designers, and researchers take oppression and inequality as their grounding assumption for creating computational products and systems. We must learn from—and design with—the communities we seek to support. A commitment to data justice begins with an acknowledgment of the fact that oppression is real, historic, ongoing, and worth dismantling. This commitment is one that we must teach the next generation of data scientists and data citizens, in communities and in classrooms, if we want to broaden our path toward justice.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023

MIT Press Direct

© 2020 Massachusetts Institute of Technology

This work is subject to a Creative Commons CC BY-NC-ND license. Subject to such license, all rights are reserved.

This book was set in ITC Stone Serif Std and ITC Stone Sans Std by Toppan Best-set Premedia Limited.

Library of Congress Cataloging-in-Publication Data

Names: Kanarinka, author. | Klein, Lauren F., author.
Title: Data feminism / Catherine D'Ignazio and Lauren F. Klein.
Description: Cambridge, Massachusetts : The MIT Press, [2020] | Series: <Strong> ideas series | Includes bibliographical references and index.
Identifiers: LCCN 2019036137 | ISBN 9780262044004 (hardcover)
Subjects: LCSH: Feminism. | Feminism and science. | Big data—Social aspects. | Quantitative research—Methodology—Social aspects. | Power (Social sciences)
Classification: LCC HQ1190 .K375 2020 | DDC 305.42—dc23
LC record available at https://lccn.loc.gov/2019036137

10 9 8 7 6 5 4 3 2 1

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2122496/9780262358521_cam.pdf by guest on 10 October 2023