

COUNTERDATA
COUNTERDATA
COUNTERDATA

BY
SEYI OLOJO

COUNTERDATA
COUNTERDATA
COUNTERDATA

COUNTERDATA

By Seyi Olojo

Introduction

Communities, activists, and organizers have long used data to narrate their everyday lives. Data, often situated as a tool for truth-telling, is at the foundation of empirical research and investigative reporting, but data has also long been a tool for storytelling and grassroots resistance. For historically marginalized communities, one way of controlling the narratives about them is through data activism.¹ This includes collecting their own data to either contest state-led projects of legibility or to collect missing data.² *Counterdata* is data that is collected to contest a dominant institution or ideology. While the practice of counterdata emerged from communities, it has been conceptualized and theorized within geographic scholarship and has gradually found its way into emerging research fields, such as critical data studies and human-computer interaction.

Counterdata can be defined as data produced as a means for enabling disadvantaged communities to reclaim political power. Key components of counterdata include (1) the correction of misrepresentative data, (2) the control of data collection and production, and (3) the strategic use of data to benefit the political and social emancipation of communities. It's important to note that counterdata is also reflective of a relationship between the state and its subjects, therefore it is inherently reactionary. Later, I'll discuss the possibilities of data activism beyond reactionary purposes. Counterdata is

- 1 Stefania Milan and Miren Gutierrez, "Technopolitics in the age of Big Data," *Networks, Movements and Technopolitics in Latin America: Critical Analysis and Current Challenges*, eds. Francisco Sierra Caballero and Tommaso Gravante (Switzerland: Palgrave Macmillan, 2018), 95–109, and Catherine D'Ignazio, "Chapter 1 — A Short Genealogy of Femicide and Data Activism," in *Counting Femicide: Data Feminism in Action*, <https://mitpressonpubpub.mitpress.mit.edu/pub/cf-chap1>.
- 2 Alessandra Jungs de Almeida, Lauren Klein, and Catherine D'Ignazio, "Missing Data," *Keywords of the Datafied State*, eds. Jenna Burrell, Ranjit Singh, and Patrick Davison (Data & Society, 2024).

not simply limited to the question of who collects the data. This term is rooted in practice and aims to describe the work that activists do in using data to realize a collective consciousness about a political issue. As a result, not only does counterdata reflect a history of data-driven grassroots methods of resistance; but it also names a shift in the way data is understood and conceptualized within statistical reporting.

Counterdata only begins to describe the ways communities and organizations subvert official datasets, given that contestation of official data has long been a tactic used by grassroots organizations. At present, state control is both mobilized and reinforced through data-driven technologies. In the United States, communities of color are continuously surveilled through federally funded data collection efforts and historically, this surveillance has led to significant structural discrimination. Well-documented examples include gerrymandering to promote voter suppression, the over policing of predominantly African American neighborhoods, and unequal distribution of federal funding for public schools. Additionally, technological failures caused by the incorrect use of data continue to harm communities that have historically been victimized by bureaucratic institutions. For example, a 2019 *Washington Post* article summarized the findings of a study conducted by The National Institute of Standards and Technology (NIST) that found a statistically significantly high rate of false positives within the racially biased facial recognition technology used by law enforcement.³ Ultimately, the justification for this surveillance is embedded within state projects for resource allocation and management. As impacted communities realize the explanatory power of data, numerous grassroots initiatives have collected counterdata to contest numbers or official statistics produced by institutions. This keyword essay attempts to summarize the emergence of the term within critical data studies while specifically historicizing it within the

3 Drew Harwell, "Federal study confirms racial bias of many facial-recognition systems, casts doubt on their expanding use," *Washington Post* (2019).

context of radical Black methodologies and grassroots organizing, which persists today. In addition to this keyword entry, Massaro et al.⁴ provide a case study analysis of counterdata production against recidivism within a Pennsylvania correctional facility.

Black Methods: Rewriting Narratives With Data

Ida B. Wells' investigative work provides an early and influential example of counterdata. Wells, an investigative journalist, and researcher, collected statistics on lynching in the late 19th century. She famously collected and curated her own datasets in an effort to show the public the alarming and often unjustified rate at which Black men were being lynched in America. Her collection of data disaggregated by race dared to fill the missing information on the landscape of lynching in the United States.⁵ In this process, she was able to debunk myths surrounding the nature of Black life that were often used to falsely justify killings.

Wells was not only contesting the state's misreporting of lynching deaths, nor the pervasively dehumanizing narratives of Black life; she was also responding to a history of quantification by acknowledging its deep racist and eugenic origins.⁶ Wells' anti-lynching pamphlets became famous for changing the narrative of lynching, which at the time was considered a justified punishment for people charged with rape. Black men overrepresented the number of men who died annually from lynching with alleged charges of rape or attempted rape. Famously, the *Chicago Tribune* reported lynching statistics every year, containing information on the alleged offense, and the race of the accused. The *Tribune's* report failed to highlight that the lynchings were racially motivated, rather than the result of accusations of rape.

4 Jungs de Almeida, Klein, and D'Ignazio, "Missing Data."

5 Ida B. Wells-Barnett, "Lynching and the Excuse for It," *The Independent* 53, no. 2737 (1901): 1133–1136.

6 Anne M. Brubaker, "Who Counts? Urgent Lessons from Ida B. Wells's Radical Statistics," *American Quarterly* 74, no. 2 (2022): 265–93.

Wells knew that the *Tribune's* data was incorrect. They failed to aggregate the data by race, which would have illustrated the disproportionate amount of Black people that made up a majority of lynching deaths. In turn, Wells published *The Red Record*, which used her own empirical investigation and the *Tribune's* data to correct its shortcomings. She provided percentages of Black people killed in proportion to the overall reported to be killed within the *Tribune*. She also specifically provided aggregates according to the reported offense. This extra step of *descriptive statistics*, adding analysis to data sources from white media outlets, was used by Wells to correct a false narrative through a quantitative approach, understood at the time to be irrefutable.

Another foundational example can be found in the work of the pioneering sociologist, Dr. W. E. B. DuBois. Insistent on presenting a narrative of Black life that empirically captured the nuances of Black communities, DuBois was in constant search of data. As a result, the data visualizations he created for the 1900 Paris exhibition were the world's very first infographics on Black life.⁷ Additionally, DuBois's *Philadelphia Negro*⁸ was amongst the first in-depth sociological studies of African-American life in an urban environment. His survey data and analysis eventually led to the release of a special report that interrogated the then-normative and limiting theories of race in America. Both Wells and DuBois understood that collecting data, often missing data, to contest official statistics would open up new possibilities for how society could come to know Black life. Perhaps the racist ideologies promoting a monolithic representation of Blackness could be corrected. Nonetheless, the struggle to correct false narratives persists, but correction serves as one of the primary motives behind the counter production of data. Currently, data remains at the center of liberation and political mobilization for Black communities. In addition to the task of correcting false statistical reporting, community-based organizations have prioritized fighting for the control of their data.

7 Witney Battle-Baptiste and Britt Rusert, *W.E.B. DuBois's Data Portraits: Visualizing Black America*. (New York: Princeton Architectural Press, 2018).

8 W. E. B. DuBois, *The Philadelphia Negro: A Social Study* (Philadelphia: University of Pennsylvania Press, 1996).

Today, organizations like Data for Black Lives, COVID Black, and Campaign Zero exist as the custodians of the most reliable databases for COVID-19 and police violence reporting for Black communities. Data for Black Lives, which rose in popularity during the pandemic, was one of the first organizations to specifically track COVID-19 data amongst Black communities. Created in reaction to the lack of accurate reporting on the impacts of COVID-19 on Black communities, Data for Black Lives created a network of scientists at various city hubs in the country to collect and report on local COVID-19 data. Another organization that uses data to correct narratives on Black life is Campaign Zero. Its Police Data Transparency Index reports on police activity and police misconduct across each state. Additionally, COVID Black, an organization that was created during the pandemic, gathers and publishes data on Black health. In partnership with academic institutions, healthcare organizations, and Black communities, COVID Black produces Black health data that helps organizations report more accurate statistics on the state of Black health. These organizations have strategically situated themselves as reputable data sources that counter the authority of institutions. By doing this, they maintain data ownership. Reminiscent of Well's original vision of better subverting racist justifications for the lynchings, these organizations maintain and produce data sources to hold policymakers accountable for the epidemic that is premature death within Black communities.

Counter-mapping

The conceptual work of the prefix counter — has shown up in many disciplines. Counterdata has clear relationships to concepts like counterpublics, or even theories of the counterculture, both of which discuss how groups

organize in opposition to central power. However, I believe a key aspect of counterdata — specifically the role of data — can be understood by examining the history of counter-mapping. Counter-mapping is a concept developed in critical geography, reflecting on the historical role of maps in the power of the state. Critical geographers Craig Dalton and Jim Thatcher coined the term “counterdata actions” in reflection of previous scholarship in counter-mapping.⁹ They go on to cite Nancy Peluso’s seminal article describing the Indigenous processes of forest mapping in Kalimantan, Indonesia as divergent from dominant mapping practices.¹⁰ Here, we first see counter-mapping defined. Peluso recalls how the Indonesian state’s techniques for rendering communities within limited logistics of representation, so that customary natural resource allocation practices may remain intact. Counter-mapping, in her essay, makes visible the fact that mapping has always been highly political. We see this in the idea of *terra nullius*, uninhabited land, that promoted Western frontierism. New maps were constructed to affirm the privatization of indigenous lands, therefore creating the grounds for counter-mapping to take deliberate departures away from dominant institutions and methods for legibility.

Even though maps have historically been used for capital accumulation and to promote state legibility, critical geographers are imagining new ways of mapmaking. Within situated geographic practices, we’ve seen a break from traditional modulations of data structures to embrace technical relations that work better for the communities the data aims to serve.¹¹ Examples of this can be seen within scholarship depicting counter-mapping practices and its bottom-up approach to resisting exploitative relationships between people and capital accumulation.¹² Additionally, counter-mapping helps situate a theory of counterdata which is reminiscent of Indigenous histories of resisting dispossession. They are explicit in asserting

- 9 Craig M Dalton, Linnet Taylor, and Jim Thatcher, “Critical Data Studies: A Dialog on Data and Space,” *Big Data & Society* 3, no. 1 (June 2016): <https://doi.org/10.1177/2053951716648346>.
- 10 Nancy Lee Peluso, “Whose Woods Are These? Counter-Mapping Forest Territories in Kalimantan, Indonesia,” *Antipode* 27, no. 4 (1995): 383–406, <https://doi.org/10.1111/j.1467-8330.1995.tb00286.x>.
- 11 Clancy Wilmott, “Small moments in spatial big data: Calculability, authority and interoperability in everyday mobile mapping,” *Big Data & Society* 3, no. 2 (2016), <https://doi.org/10.1177/2053951716661364>.
- 12 Craig M. Dalton and Tim Stallmann, “Counter-Mapping Data Science,” *Canadian Geographies / Géographies Canadiennes* 62, no. 1 (2018): 93–101, <https://doi.org/10.1111/cag.12398>.

that counterdata actions are already happening, making clear that this concept was developed from grassroots approaches to subvert state control. Furthermore, counterdata can be produced for a multiplicity of reasons, but it can be seen as data practices that are in accord with the political and epistemological visions of the communities they represent. Even though counterdata production adheres to the logic of quantification, a long-preferred method of empiricism, communities are able to prioritize which numbers are most meaningful to them. Counterdata centers marginalized communities as the main stakeholders with the technical skills to create artifacts that tell narratives that are most authentic to their lived experience.

Counterdata: Defined by Case Studies

Even though counterdata can be connected back to other theories of data activism, it is important to note that counterdata production has always been a dynamic global grassroots practice. Counterdata has always been produced to return agency to oppressed communities. Given its origins in grassroots organizing and social movements, most scholarly writing on counterdata, much of it emerging in the past decade, theorizes the concept using case studies. This notably includes work following femicide data collection in Latin America¹³ and data collection surrounding the epidemic of missing and murdered indigenous women and girls. These examples of data collection as a response to misogyny-driven killings and abuse highlight the important role that counterdata plays in disrupting the power relations that missing data¹⁴ maintains. In her work on the mapping of femicide accounts in Uruguay, activist-scholar Helena Suárez Val¹⁵ describes the affordances of collecting this data, “In this sense, digital records and cartographies of femicide are a form of research-creation where data about violence becomes

13 Harini Suresh, Rajiv Movva, Amelia Lee Dogan, Rahul Bhargava, Isadora Cruxen, Angeles Martinez Cuba, Guilia Taurino, Wonyoung So, and Catherine D'Ignazio, “Towards Intersectional Feminist and Participatory ML: A Case Study in Supporting Femicide Counterdata Collection,” in *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency, FAccT '22* (New York, NY, USA: Association for Computing Machinery, 2022), 667–78, <https://doi.org/10.1145/3531146.3533132>; Catherine D'Ignazio, Helena Suárez Val, Silvana Fumega, Harini Suresh, Isadora Cruxên, Wonyoung So, Ángeles Martínez, Mariel García-Montes, “Femicide & Machine Learning: Detecting Gender-based Violence to Strengthen Civil Sector Activism,” Mechanism Design for Social Good Workshop, (August 2020). <http://www.kanarinka.com/wp-content/uploads/2021/01/DIgnazio-et-al.-2020-Femicide-Machine-Learning-Detecting-Gender-ba.pdf>; Helena Suárez Val, “Datos discordantes. Información pública sobre femicidio en Uruguay,” *Mundos Plurales-Revista Latinoamericana De Políticas Y Acción Pública* 7, no. 1 (2020): 53–78.; Helena Suárez Val, “Datos discordantes. Información pública sobre femicidio en Uruguay,.” *Mundos Plurales-Revista Latinoamericana De Políticas Y Acción Pública* 7, no. 1 (2020): 53–78; Sarah Meagan Upton, “Moving Beyond Awareness: Ni Una Más and Approaches to the Problem of Femicide in Ciudad Juárez,” (MA Thesis, Georgetown University, 2010).

14 Jungs de Almeida, Klein, and D'Ignazio, “Missing Data.”

15 Helena Suárez Val, “Affect Amplifiers: Feminist Activists and Digital Cartographies of Femicide,” in *Networked Feminisms: Activist Assemblies and Digital Practices*, eds. Shana Macdonald, Brianna I. Wiens, Michelle MacArthur, and Milena Radzikowska (Lanham: Lexington Books, 2021): 163–187.

public displays of feminist activists' emotional and affective — and political — responses to femicide.”¹⁶

Suárez Val has been tracking femicide in Uruguay since 2015. In her essay, “Affect Amplifiers: Feminist Activist and Digital Cartographies of Femicide,” she argues that feminist rearticulations of femicide data create the conditions for social change. The act of accounting for these deaths, through recordings, mappings, and visualizations, invokes feminist affective politics that effectively contest the rampant nature of femicide. Val situates her work in the feminist affective cultures of Uruguay, recalling local histories of women advocating against violence against women in the 90s. The maps that Val presents in this paper center the emotions of feminist activists on the front lines of reporting and data collection. Data collection isn't simply the act of sourcing data, it holds the weight of the lives lost. Val describes one map in particular where data points were illustrated through the use of tear shapes. The atrocity of these deaths is communicated beyond the epistemological boundaries of numbers.

Another example of counterdata can be found in the work of Amanda Meng and Carl DiSalvo.¹⁷ In partnership with the Westside Atlanta Land Trust (WALT), they observed how this community was able to collect its own data to advance its advocacy for a community land trust. The official data, a survey of the built environment, was conducted in 2014 to assess property vacancies in the English Avenue and Vine City neighborhoods. The survey data was supposed to help local organizations combat gentrification, but after reviewing the survey data, they quickly determined that the data significantly undercounted abandoned property. WALT, one of many groups advocating for the development of a community land trust in Fulton County, was able to spot this issue because of their own lived experience of their

¹⁶ Suárez Val, “Affect Amplifiers,” 1.

¹⁷ Amanda Meng and Carl DiSalvo, “Grassroots Resource Mobilization Through Counter-data Action,” *Big Data & Society* 5, no. 2 (2018): <https://doi.org/10.1177/2053951718796862>.

neighborhood. This local knowledge propelled a collective resistance to the official data. For them, data collection not only created a new artifact but encouraged a collective consciousness amongst the people. Their data strategically contained information, such as back tax amounts for each property. Properties that had back taxes were flagged as potential sites for the community land trust. The organized resistance against the county's original survey created an opportunity for community members to think about datasets that would be of better use to their communities and that would effectively subvert state power.

Morgan Currie, Britt S. Paris, Irene Pasquetto, and Jennifer Pierre¹⁸ illustrate an example of counterdata production in the reporting of Police Officer Involved Homicides (POIHs) in Los Angeles County. Drawing from Dalton and Thatcher's seminal paper, Currie et al. provide another case study approach to defining the term. They define counterdata actions as "acts of resistance to politically dominant datasets." In this case, we see data produced in ways that center communities as the primary designers of the data. In the absence of official POIH data, Currie et al. observe how communities reinterpret previously existing datasets that fail to accurately capture the landscape of law enforcement homicides in their communities. They describe findings from a community hackathon — showing that participants were able to leverage their technical skills in spotting inaccuracies in federal datasets while also creatively thinking of ways to index and process community-derived datasets. In addition to making community-derived data more accessible, participants resisted the boundaries of quantitative data collection. They researched the social media profiles of victims to shed light on the lives they lived and the communities that memorialized them. Here, we see counterdata as not just a quantitative approach to official data, but as sites for qualitative and interpretivist methodologies. Not only are authors

¹⁸ Morgan Currie, Britt S. Paris, Irene Pasquetto, and Jennifer Pierre, "The Conundrum of Police Officer-involved Homicides: Counter-data in Los Angeles County," *Big Data & Society* 3, no. 2 (December 2016), <https://doi.org/10.1177/2053951716663566>.

dealing with a missing data issue (no official database of POIH data) the collection of their counterdata challenges policing narratives in LA. In this case, we see that counterdata actions allow for access to information. This includes challenging the inadequate ways in which data has been made available. Qualitative data collection often exists as a form of counterdata within research environments that fetishize quantitative methods. The above examples assert the importance of including local knowledge and narratives in datasets that are able to better represent communities. Qualitative methodologies can often shed light on the incomplete and incorrectly labeled information contained within official datasets.

Limitations and Conclusion

Counterdata actions can be instructive in how researchers think about data as a representative tool. State and institutional powers have almost exclusively reduced data into ledgers that are made to produce a surveillant order to everyday life.¹⁹ The history of statistics is embedded within this goal of building the nation-state²⁰ and promoting eugenic ideals.²¹ Counterdata has consistently purported that qualitative approaches to accounting for people not only produce more accurate data but provide opportunities to contest subjectivity. Inherently resistant to the objectives of scale and generalizability for hegemonic use, might it be possible to see qualitative data collection and ethnographic practices as political resistance?

Counterdata is corrective, which means that it will always be produced in response to incorrect official data. It can illuminate hierarchies of power embedded within datasets and furthermore, politicize previously depoliticized datasets. This always situates the hegemonic institutions as political

¹⁹ James C. Scott, *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed* (New Haven: Yale University Press, 2020).

²⁰ Mara Loveman, "Nation-state Building, 'Race,' and the Production of Official Statistics: Brazil in Comparative Perspective," (Ph.D. Diss., University of California, Los Angeles, 2001).

²¹ Ian Hacking, "Trial by Number; Karl Pearson's Chi-Square Test", *Science* '84, no. 5 (November 1984): 69–71.

actors and data as tools for resistance. But what can data outside the context of resistance afford communities? Indigenous data practices have often pushed back against the limitations of counterdata in favor of data practices that primarily serve their own communities.²² Nonetheless, the rhetoric around counterdata is still constrained by the construct of the statistical dataset — an inherently quantitative tool.

Finally, this essay asserts that the work of defining counterdata originated from grassroots organizing and activism. This term evolved in the hands of scholars attentive to the varied ways communities are creating datasets within their own terms. Considering this, the concept can also assist in how we rhetorically talk about data structures that exist beyond strict quantitative approaches. We've learned that the flattening of people, often victims of hegemonic violence, begets statistics that are often incorrect or even nonexistent. Corrective actions against this often implore the lived experiences and rich narratives of communities. Counterdata positions memory, place, and storytelling as viable resistance strategies against state control and legibility. Even though counterdata cannot be the sole means by which the subjugated emancipate themselves, counterdata production brings visibility to important methodologies that can hopefully continue promoting the narratives communities create for themselves.

22 Annita Hetoehotohke'e Lucchesi, "Mapping Violence against Indigenous Women and Girls: Beyond Colonizing Data and Mapping Practices," *ACME: An International Journal of Critical Geographies* 21, no. 4 (May 2022): 389–98; Stephanie Russo Carroll, Marisa Duarte, and Max Liboiron, "Indigenous Data Sovereignty," in *Keywords of the Datafied State*, eds. Jenna Burrell, Ranjit Singh, and Patrick Davison (Data & Society, 2024).