
Invisible in Data, Excluded from Research: A Literature Review of Sexual Orientation and Gender Identity Data

Prepared by
DataHaven

Anyuyue Feichu Ai, Research Assistant

May 2024

Introduction

The lesbian, gay, bisexual, transgender, and queer (LGBTQ+) community comprises some of the most vulnerable populations in the United States, yet they are underrepresented in research. The social crises facing the LGBTQ+ community remain understudied. Despite persistent efforts of advocacy groups across the country, data on LGBTQ+ people are sparse and, when available, hard to access. Invisible in data and excluded from the research process, LGBTQ+ people are rendered missing from important conversations. The lack of data collected about LGBTQ+ people and issues materially impacts policy decisions and the (re)distribution of resources. Disparities go ignored, and entire communities are vanished. Comprehensive data are crucial to quantify the challenges facing the LGBTQ+ community, from health outcomes to economic precarity to discrimination. Data help policymakers and activists to advocate for LGBTQ+ rights and to triage the issues facing their communities.

Data collection efforts inclusive of LGBTQ+ demographics are urgently needed at this moment of American history. With the rise of anti-LGBTQ+ animus — alongside waves of far-right extremism and white supremacy¹ — anti-LGBTQ+ public demonstrations, threats of violence, and more have risen to the highest levels recorded.² Hostile legislation has gained momentum, as seen through the surge in anti-trans bills that restrict trans people’s full participation in public life. In 2023, a record-breaking 600 anti-trans bills were introduced across 49 states, a nearly ten-fold increase from the number introduced in 2020.³ All of this culminates in deadly violence that disproportionately takes the lives of Black trans women. Thirteen percent of the trans community is estimated to be Black, yet nearly three-quarters of known victims of anti-trans homicides are Black trans women.⁴ The LGBTQ+ community faces interwoven crises of homophobia, transphobia, anti-Blackness, racism, ableism, and more. Action is swiftly needed.

Data are useful tools for LGBTQ+ advocacy. Sexual orientation and gender identity (SOGI) data are vital in ascertaining how homophobia and transphobia move along other axes of power and identity, like race, class, and disability. Although public and private entities collect a wealth of data, demographic questions on LGBTQ+ people are routinely left out. For instance, the

¹ Bennett, G., Norris, C., & Hastings, D. (2023, August 31). *How the rise of anti-LGBTQ+ hate and violence is impacting the community*. PBS.

<https://www.pbs.org/newshour/show/how-the-rise-of-anti-lgbtq-hate-and-violence-is-impacting-the-community>

² Jones, S., & Kishi, R. (2023, October 26). *UPDATE | Fact Sheet: Anti-LGBT+ Mobilization on the Rise in the United States*. ACLED. <https://acleddata.com/2022/11/23/update-fact-sheet-anti-lgbt-mobilization-in-the-united-states/>

³ Trans Legislation Tracker. (2024). *2023 anti-trans legislation*. Trans Legislation Tracker. <https://translegislation.com/bills/2023>

⁴ Mandler, C. (2022, October 14). *U.S. Murders of trans people nearly doubled over past 4 years, and Black trans women are most at risk, report finds*. CBS News.

<https://www.cbsnews.com/news/transgender-community-murder-rates-everytown-for-gun-safety-report/>

American Community Survey by the U. S. Census Bureau is a comprehensive survey that has been running since 2005, but it only began to collect data on same-sex couples in 2019.⁵ Furthermore, the Census — the gold standard of federal data collection — has failed to ask any SOGI questions despite continual pressure from LGBTQ+ organizations.⁶ Data invisibility of LGBTQ+ people is not a coincidence. It is manufactured.

Even when surveys attempt to collect information on gender, they will often ask about sex assigned at birth but not gender identity, two different but often conflated subjects. Furthermore, most surveys handle topics of sex and gender in a binary way, which erases nonbinary, two-spirit, and intersex respondents. Specific and distinct categories of LGBTQ+ identity are often combined, making disparities within the community unclear.

This literature review aims to address the ongoing erasure of LGBTQ+ people in datasets by comparing public and private databases and measuring how often SOGI data are collected and which specific SOGI questions are asked. We examine whether data were collected based on the following categories: sexual orientation (lesbian, gay, bisexual, queer, etc.); transgender identity; nonbinary identity; two-spirit identity; and intersex status. The goals of this document are to visualize data invisibility and assist researchers and activists in accessing datasets on LGBTQ+ populations.

⁵ *Same-Sex Couples*. United States Census Bureau. (2022, August 5).
<https://www.census.gov/topics/families/same-sex-couples.html>

⁶ Velte, K. (2019). Straightwashing the Census. *Boston College Law Review*, 61(69), 1–55.
<https://doi.org/10.2139/ssrn.3358453>, 2.

Methods

We searched for datasets that collected data on the national level or on the Connecticut state level. We were interested in datasets that covered socioeconomic indicators such as harm and victimization; health; COVID-19; income and employment; housing; race, ethnicity, Indigeneity, and migration; disability; youth; aging; and polling.

To narrow our search for databases to include in this literature review, the following criteria were used:

- Must include microdata and/or raw statistics generated by research studies
- Must be publicly available or available through request
- Must be collected or updated within the past ten years
- Must be collected by a federal or state authority or by a reputable private organization (i.e. receives funding from the government or is sponsored by a university)

The following federal and university catalogs aided in the search for databases:

- ["Data Sources"](#) Webpage of Healthy People 2030
- ["Compendium of International Data Sources for Parents with Disabilities | 2023"](#) Webpage of Brandeis University's National Research Center for Parents with Disabilities
- ["Publicly Available Sources of Data for Health & Social Determinants of Health"](#) Webpage of the University of Pittsburgh's Health Sciences Library System
- ["LGBTQ+ Data and Collections"](#) Webpage of Johns Hopkins's Sheridan Libraries

Ultimately, we assessed 96 datasets in this literature review.

Clarification of Terminology

For this project, we collected data based on sexual orientation, transgender identity, nonbinary identity, two-spirit identity, and intersex status.

We distinguish between transgender and nonbinary because not all nonbinary people identify as transgender. Therefore, surveys that merely ask whether someone is transgender, or whether someone is a "transgender man" or "transgender woman," miss nonbinary people who do not identify as transgender. The Human Rights Campaign provides the following explanation:

"Transgender, or trans, is an umbrella term for people whose gender identity is different from the sex assigned to them at birth... Non-binary is an identity embraced by some people who do not identify exclusively as a man or a woman. Non-binary people may identify as being both a man and a woman, somewhere in between or as falling completely outside of these categories. While many also identify as transgender, not all non-binary people do. Non-binary can also be used as an umbrella term encompassing identities such as agender, bigender, genderqueer or gender fluid."⁷

We also distinguish two-spirit from the categories of transgender, nonbinary, and queer. Two-spirit is a pan-Native American term that describes a traditional third gender (or more). While such gender roles are understood differently across Indigenous Nations, two-spirit is an umbrella term that bridges Western and Indigenous conceptions of gender and sexuality. Two-spirit is distinct from transgender in that it is rooted in Indigenous understandings. Geo Soctomah Neptune, a Passamaquoddy two-spirit basketmaker, artist, and educator, offers clarity:

"You may recognize [two-spirit] as the '2' that sometimes appears on the end of 'LGBTQ2.' There are many definitions and understandings of two-spirit, and each is nation specific. The term was intentionally introduced by Native people with the goal of finding common ground and helping educate about traditional teachings in a contemporary context. The Two Spirit Society of Denver offers the following definition: 'Two-spirit refers to another gender role believed to be common among most, if not all, first peoples of Turtle Island (North America), one that had a proper and accepted place

⁷ Human Rights Campaign. (n.d.). *Transgender and Non-Binary People FAQ*. Human Rights Campaign. <https://www.hrc.org/resources/transgender-and-non-binary-faq>

within Native societies. This acceptance was rooted in the spiritual teachings that say all life is sacred."⁸

Additionally, we collected data on intersex status. Intersex and trans people have overlapping needs and issues, but intersex ought not be confused with the identities of transgender and nonbinary. An intersex person is someone whose physical sex characteristics (reproductive anatomy, hormones, chromosome patterns) fall outside the strict male/female sex binary. Intersex people can be of any gender. The Gender and Sexuality Campus Center of University of Wisconsin-Madison offers the following explanation:

"Intersex describes people's bodies that do not fit invented categories of 'male sex' and 'female sex.' There is no single 'intersex body'; the term encompasses a wide variety of conditions that do not have anything in common except that they are deemed 'abnormal' by the medical community. What makes intersex people similar is their experiences of medicalization, not anatomy. Generally speaking, intersex is not an identity category. While some intersex people do reclaim "intersex" as part of their identity, most regard it as a medical condition or just a unique physical quality. Intersex people may be straight, cis, or any of the identities in the LGBTQ umbrella."⁹

Finally, we use the term Latine (pronounced lah-TEE-ney) to describe people of Latin American descent.¹⁰ While Latino and Latina are commonly used, these terms are gendered and thus do not fully capture identities beyond the gender binary. Latinx has emerged as a gender-neutral alternative, but the term has been widely criticized for being difficult to pronounce in Spanish.¹¹ Latine, on the other hand, is a gender-neutral term that is easy to pronounce in Spanish and fits within Spanish language conventions.¹² We thus adopt the term Latine throughout our report.¹³

⁸ Neptune, G. S. (2018, December 11). *What Does "Two-Spirit" Mean? | InQuery | them*. YouTube. <https://www.youtube.com/watch?v=A4lBibGzUnE>, 0:33 - 1:13

⁹ (N.d.). (rep.). *Frequently asked questions: Intersex identities*.

¹⁰ Dictionary.com. (2022, September 26). *Latine vs. Latinx: How And Why They're Used*. Dictionary.com. <https://www.dictionary.com/e/latine-vs-latinx/#difference-between-latine-and-latinx>

¹¹ Villegas, P. (2023, April 2). *How 'Latinx' united — and divided — a community seeking to redefine itself*. The Washington Post. <https://www.washingtonpost.com/nation/2023/04/02/latinx-meaning-controversy-culture/>

¹² Lopez, R. S. (2022, August 19). *Switching to "Latine" from "Latinx."* Out Boulder County. <https://www.outboulder.org/blog/switching-to-latine-from-latinx>

¹³ DataHaven typically uses Latino in its reporting based on feedback from various stakeholders, but because this report focuses specifically on gendered concepts and other identities, Latine is used here.

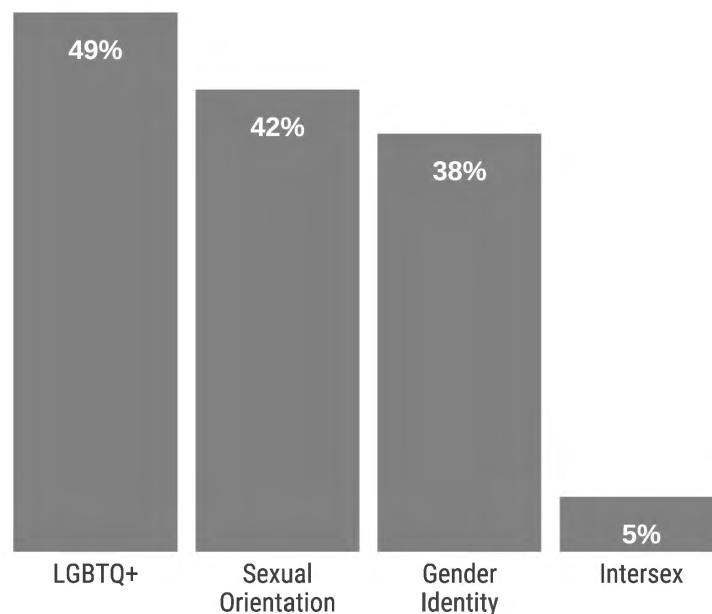
Findings

The criteria for a database to be considered to collect SOGI data were strict. LGBTQ+ identities had to be explicit and clear. For example, we considered a survey to collect two-spirit data only if the survey directly and clearly asked about two-spirit identity. Surveys that asked for gender and listed "male," "female," and "other" did not count as collecting nonbinary and intersex data because the "other" category was ambiguous, and "male" and "female" do not describe gender but rather sex. Both nonbinary and intersex individuals might select "other" since one can reasonably interpret the wording as asking for either gender or sex. This made the resulting aggregated data for "other" respondents unhelpful. Surveys that listed "other" under questions about sex, however, were considered as collecting intersex data.

Representation of LGBTQ+ people in data is limited. Out of the 96 databases assessed in this literature review, only 47 collected SOGI data meeting our criteria. Sexual orientation was the most collected demographic for the LGBTQ+ community, with 40 databases containing that data. Gender identity (transgender, nonbinary, two-spirit) was collected by 36 databases. Intersex status was collected by only five databases.

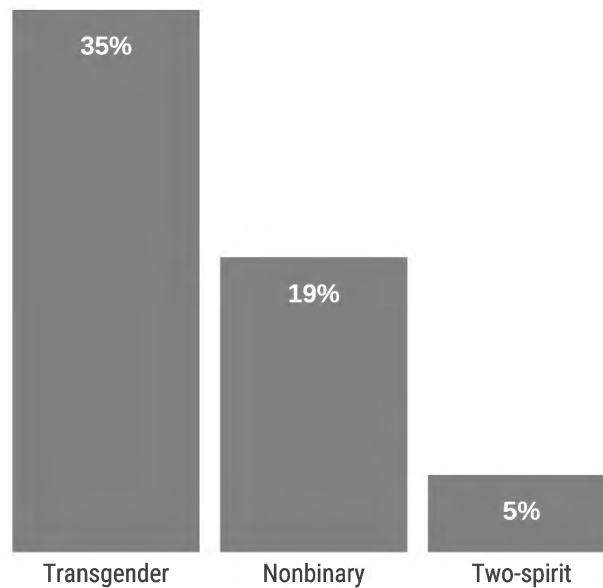
Figure 1. Sexual orientation and gender identity (SOGI) data are often absent in prominent public and private databases that have been published or updated within the past ten years.

Out of 96 databases reviewed, 49 percent contain any SOGI data at all, 42 percent collect sexual orientation, 38 percent collect gender identity, and 5 percent collect intersex status.



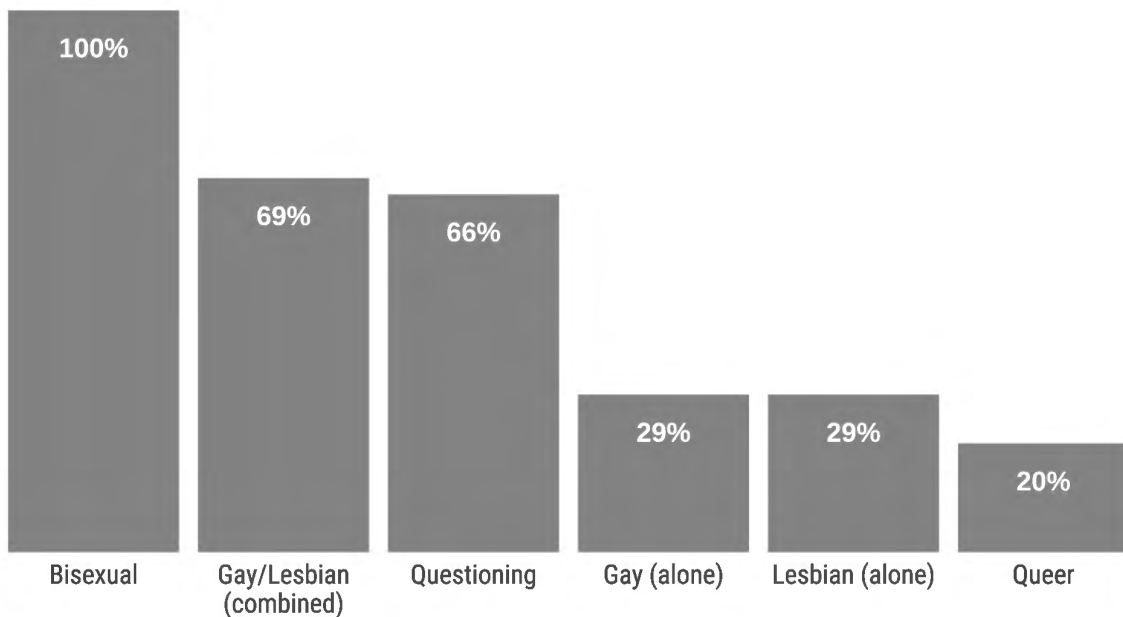
Databases frequently did not include gender identity data. While transgender identity was occasionally collected, data on specific trans identities like nonbinary and two-spirit were even less common.

Figure 2. Data on transgender, nonbinary, and two-spirit people are not commonly collected by prominent public and private databases that have been published or updated within the past ten years. Out of 96 databases reviewed, 35 percent collect transgender identity, 19 percent collect nonbinary identity, and 5 percent collect two-spirit identity.



Sexual orientation data were disaggregated into specific identities in 35 of the 40 databases that asked for sexual orientation. However, few options were provided. Bisexual was the most common identity collected, followed by gay/lesbian (combined), then questioning. The options "not sure," "unsure," and "don't know" were considered the questioning identity. Gay (alone), lesbian (alone), and queer were less often listed.

Figure 3. When collected, sexual orientation data are often disaggregated into specific identities by prominent public and private databases that have been published or updated within the past ten years. Out of 35 databases reviewed that collect sexual orientation data, 100 percent collect bisexual, 69 percent collect gay/lesbian (combined), 66 percent collect questioning, 29 percent collect gay (alone), 29 percent collect lesbian (alone), and 20 percent collect queer.



Future Directions

Current practices of data collection are limited for LGBTQ+ populations, as researchers often do not ask for SOGI data. When LGBTQ+ people are represented in data, researchers fail to capture the rich diversity of the community.

For instance, data on LGBTQ+ adolescents are rare, as most datasets are limited to adult respondents only despite young people identifying as LGBTQ+ at increasingly higher rates than older populations.¹⁴ This is especially concerning given the vulnerability and unique challenges, such as bullying, that youth in the community face.¹⁵

Current research practices also do not adequately represent the fluidity of LGBTQ+ identities. Sexuality and gender are diverse and expansive, and LGBTQ+ people may be of multiple identities. A lesbian might also identify as asexual, and a bi-gender person may identify as both man and woman. Surveys often limit responses to SOGI questions to one selection, which fail to capture the fluidity of queerness and transness.

Although collecting SOGI data is a crucial first step, the research process as a whole ought to be reimagined. It is not enough to simply include questions and metrics on LGBTQ+ identity. Datasets still face the problem of insufficient sample sizes. Surveys often lack a sufficient number of LGBTQ+ respondents, making analyses difficult or impossible. Survey providers ought to make the deliberate and active recruitment of LGBTQ+ people a priority. To address data invisibility, researchers must commit to comprehensive and faithful representations of the LGBTQ+ community.

Furthermore, SOGI data ought to be handled responsibly. In a climate that is actively hostile to LGBTQ+ people, the privacy of LGBTQ+ respondents should be treated seriously. Data have been increasingly weaponized to surveil trans, nonbinary, and two-spirit people in particular. In 2023, Missouri Attorney General Andrew Bailey barraged a Kansas City hospital with 54 requests for the health data of trans patients.¹⁶ Widely criticized as unlawful and transphobic, Bailey's investigations also extended to Planned Parenthood, from which he demanded health records

¹⁴ Chavez, N. (2024, January 25). *Gen Z adults identify as LGBTQ at much higher rates than older Americans, report shows*. CNN. <https://www.cnn.com/2024/01/25/us/gen-z-adults-lgbtq-identity-reaj/index.html>

¹⁵ Green, A., & Price-Feeney, M. (2020, August 18). *LGBTQ Youth Face Unique Challenges Amidst COVID-19*. Harvard Medical School Center for Primary Care. <https://info.primarycare.hms.harvard.edu/perspectives/articles/lgbtq-youth-challenges-covid-19>

¹⁶ The Associated Press. (2023, April 15). *Hospital sues Missouri's top prosecutor over trans care data*. AP News. <https://apnews.com/article/transgender-care-missouri-attorney-general-hospital-lawsuit-217e78b46cfd50432c97e3cc759f9717>

of trans youth.¹⁷ Similarly, Texas Attorney General Ken Paxton sought data from the Texas Department of Public Safety on people who changed their gender on their state driver's license. Although Paxton's office was unable to secure such data, the Texas Department of Public Safety identified over 16,000 Texans in 2022.¹⁸ Amidst the rise in digital surveillance, the identities of LGBTQ+ respondents ought to be protected. Data can be readily co-opted by hostile actors. Researchers must respect the privacy of LGBTQ+ people, including their right to refuse participation in studies.

As researchers begin to prioritize responsible methods of SOGI data collection, it is important that queerness and transness not be studied in isolation. Depictions of LGBTQ+ populations are incomplete if they are not intersectional. Calls for more LGBTQ+ representation within research are inextricable from ongoing efforts to dismantle systemic racism, sexism, ableism, and more. SOGI data ought to be analyzed in tandem with these and other identities. Racism is a crisis within and outside the LGBTQ+ community, with LGBTQ+ people of color faring worse across a broad range of social, economic, and health outcomes than White LGBTQ+ people.¹⁹ Combating the invisibility of queer and trans people of color must entail an honest reckoning with the Whiteness of the LGBTQ+ community and the centering of Black LGBTQ+ voices and lives. The ongoing whitewashing of the LGBTQ+ history extends to representations in data. Given the compounding effects of homophobia/transphobia and racism (and other systems of oppression), researchers ought to adopt an intersectional approach that centers the lives of those most marginalized, especially Black, Latine, and Indigenous LGBTQ+ people.

¹⁷ Ballentine, S. (2023, March 31). *Missouri Planned Parenthood sues state AG over transgender care investigation*. PBS.

<https://www.pbs.org/newshour/nation/missouri-planned-parenthood-sues-state-ag-over-transgender-care-investigation>

¹⁸ Kladzyk, R. (2023, June 28). *Policing Gender: How Surveillance Tech Aids Enforcement of Anti-Trans Laws*. Project on Government Oversight.

<https://www.pogo.org/investigations/policing-gender-how-surveillance-tech-aids-enforcement-of-anti-trans-laws>

¹⁹ Wilson, B. D. M., Bouton, L. J. A., & Mallory, C. (2022, January). *Racial Differences Among LGBT Adults in the US*. The Williams Institute. <https://williamsinstitute.law.ucla.edu/publications/racial-differences-lgbt/>

Moving With and Beyond Data

Data are useful in making the struggles of the LGBTQ+ community legible to the state and broader population, but at the same time, we must resist the "datafication of injustice," wherein researchers may inadvertently abstract real-world concerns into the realm of the theoretical.²⁰ Data can flatten people with complex lives into mere numbers to be observed, diminishing their humanity. Reflecting on anti-Blackness in the field of statistics, sociologist Anthony Ryan Hatch cautions against the "hyper-fixation on racist numerology: rates, numbers, percentages, risk estimations, abstractions, instrumentalizations." Indeed, the preoccupation with quantitative evidence "feels like we're not even talking about fleshy human beings."²¹ Numbers cannot encapsulate the depth of people's lives and struggles. Narrating the lives of LGBTQ+ people cannot be done through statistics alone.

LGBTQ+ people, especially queer and trans people of color, have long mobilized against the violences of the cisheteropatriarchy. Their lived experiences capture worlds beyond data. The demand for data may sometimes come at the expense of this rich knowledge because "[f]or many, injustice is not invisible and waiting to be documented: it is an intimate fact of daily life."²² Researchers Roderic Crooks and Morgan Currie elaborate on this tension between data and community knowledge:

"Creating or working with data demands resources, including, in many cases, the labor of community members themselves. Such a demand to produce data directs resources toward clerical work, public relations, and scientific communication rather than the redress of structural inequality. Ironically, the privileging of arguments when they are made in the form of data devalues the varied forms of knowledge held within communities, a value that contradicts the ethos of community work."²³

Data are tools for fomenting social change. The collection of data itself ought not be the end goal of research. The imperative to gather more data misses how data alone are insufficient. Knowledge production must be paired with direct action to meaningfully transform society. Hatch elaborates on these confines of data collection in context of racial disparities in health research:

²⁰ Benjamin, R. (2019). *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity, 13.

²¹ Hatch, A. R. (2020, May 22). *Two Meditations in Coronatime*. The Science, Knowledge, and Technology Section of the American Sociological Association. <https://asaskat.com/2020/05/22/two-meditations-in-coronatime/>

²² Crooks, R., & Currie, M. (2021). Numbers will not save us: Agonistic Data Practices. *The Information Society*, 37(4), 201–213. <https://doi.org/10.1080/01972243.2021.1920081>, 206.

²³ Crooks, R., & Currie, M. (2021), 206.

"To borrow from Janet Jackson: what has the data done for us, lately?... [there exists] the assumption that collecting data on racial health disparities in the COVID-19 pandemic will lead to the reduction or elimination of those disparities, an assumption that keeps scientists in an endless search for more and more refined measurements of racism's harms, while the political and economic systems that comprise the fundamental causes of those harms are given a pass until all the data are counted."²⁴

Data are necessary but insufficient for achieving LGBTQ+ rights. Researchers ought to thoughtfully reflect on the purpose of their efforts. Data cannot move social institutions alone. Researchers and advocates must actively put data findings into practice to achieve durable social transformation.

²⁴ Hatch, A. R. (2022). The data will not save us: Afropessimism and racial antimatter in the COVID-19 pandemic. *Big Data & Society*, 9(1), 1–13. <https://doi.org/10.1177/20539517211067948>, 2.

Conclusion

We find that LGBTQ+ people are often missing from data collection efforts. While there is some representation of queer sexualities in data, the invisibility of trans, nonbinary, two-spirit, and intersex people continues in a way that mirrors how trans, nonbinary, two-spirit, and intersex people are erased in society at large. With the rising moral panic over queerness and transness, the heightened visibility of queer and trans people has largely been contained to negative rhetoric and propaganda.

The pattern of data invisibility is real and concerning, especially for two-spirit and intersex people. Data are not neutral but rather embedded within social relations. The ongoing erasure of LGBTQ+ people within data must be contextualized within the broader efforts to remove LGBTQ+ people from public life. As such, the struggles for LGBTQ+ rights extend to the realm of data and tech, too. Data and tech have yet to truly grapple with queerness and transness that meet the moment of now.

We urge researchers to prioritize responsible and reciprocal processes of data collection. Just data practices are not only about who is studied in the research, but also about who is conducting the research. LGBTQ+ people ought to be actively recruited as participants and ought to be the researchers carrying out the studies in the first place. This means centering LGBTQ+ voices throughout the entire research process. Knowledge production is incomplete without Black, Latine, and Indigenous LGBTQ+ people, and especially Black trans women.

Just as producing data analyses is important, so too is acting on findings. Data are some of the many tools that researchers and activists can use to fight for LGBTQ+ equality. Ultimately, data ought to be mobilized in service of social movements that seek to end the marginalization of LGBTQ+ people. Research ought to be paired with direct action to best achieve social equity.

Acknowledgements

Many thanks to Camille Seaberry and Kelly Davila for their generous contributions to the editing of this report.

Appendices

Table 1. Databases on General Information

Database	Last Updated	Demographic Information Collected				
		Sexual Orientation	Transgender	Nonbinary	Two-spirit	Intersex
2015 U.S. Transgender Survey	2019	✓	✓	✓	✓	✓
Household Pulse Survey	2024	✓	✓	✓		
TransPop Study	2021	✓	✓	✓		
2020 Census	2023					
American Community Survey	2024					
CDC/ATSDR Social Vulnerability Index	2022					

Table 2. Databases on Harm and Victimization

Database	Last Updated	Demographic Information Collected				
		Sexual Orientation	Transgender	Nonbinary	Two-spirit	Intersex
Fatal Encounters	2021		✓			
Fatal Force Database	2024			✓		
National Crime Victimization Survey (NCVS)	2024	✓				
National Intimate Partner and Sexual Violence Survey (NISVS)	2020	✓	✓			
National Violent Death Reporting System (NVDRS)	2022	✓	✓			
Police-Public Contact Survey (PPCS)	2022	✓				
Prison Policy Initiative Data Toolbox	2024	✓	✓			
School Crime Supplement (SCS)	2024	✓				
Survey of Prison Inmates (SPI)	2024	✓	✓			
UCR Hate Crime Statistics	2019	✓	✓			
Annual Probation Survey	2021					
Annual Parole Survey	2021					
Annual Survey of Jails (ASJ)	2023					
Capital Punishment (NPS-8)	2023					

Census of Jails (COJ)	2022					
Census of State and Federal Adult Correctional Facilities (CCF, Formerly CSFACF)	2022					
Mapping Police Violence	2024					
National Prisoner Statistics (NPS)	2024					
Survey of Sexual Victimization (SSV)	2021					
The Counted	2016					
UCR Federal Crime Data	2019					
UCR Human Trafficking	2019					
UCR Law Enforcement Officers Killed and Assaulted (LEOKA)	2019					
UCR National Incident-Based Reporting System (NIBRS)	2019					
U.S. Criminal Justice Data	2024					

Table 3. Databases on Health

Database	Last Updated	Demographic Information Collected				
		Sexual Orientation	Transgender	Nonbinary	Two-spirit	Intersex
Behavioral Risk Factor Surveillance System (BRFSS)	2022	✓	✓	✓		
California Health Interview Survey (CHIS)	2022	✓	✓	✓		
Health Center Patient Survey (HCPS)	2022	✓	✓			
Health Information National Trends Survey (HINTS)	2023	✓				
Monitoring the Future	2024					✓
Mpox Ongoing 2022 Global Outbreak Cases and Data	2024		✓			
National Health Interview Survey	2024	✓	✓	✓		
National HIV Behavioral Surveillance (NHBS)	2022	✓	✓			✓
National Survey on Drug Use and Health (NSDUH)	2022	✓				
Population Assessment of Tobacco and Health (PATH) Study	2024	✓	✓			
Pregnancy Risk Assessment Monitoring System (PRAMS)	2024	✓				
Research and Development Survey (RANDS)	2023	✓	✓		✓	
CDC WONDER	2024					

Current Population Survey (CPS)	2023					
Division of Nutrition, Physical Activity, and Obesity (DNPAO)	2023					
Fatality Analysis Reporting System (FARS)	2024					
Health, United States	2023					
Medicaid and Children’s Health Insurance Program (CHIP) Enrollment Data	2024					
Medical Expenditure Panel Survey (MEPS)	2024					
Medicare Current Beneficiary Survey (MCBS)	2023					
Mental Health Client-Level Data (MH-CLD)	2024					
National Health and Nutrition Examination Survey (NHANES)	2023					
National Immunization Survey	2024					
National Survey of Family Growth	2019					
National Survey on Drug Use and Health	2024					
National Vital Statistics System	2024					
National Post-acute and Long-term Care Study	2022					
Treatment Episode Data Set: Admissions (TEDS-A)	2023					
Treatment Episode Data Set: Discharges (TEDS-D)	2023					

Table 4. Databases on COVID-19

Database	Last Updated	Demographic Information Collected				
		Sexual Orientation	Transgender	Nonbinary	Two-spirit	Intersex
COVID-19 Case Surveillance Public Use Data	2024					✓
Connecticut Department of Public Health Respiratory Viral Disease Surveillance	2024					
COVID Data Tracker	2024					
U.S. COVID-19 Self-Test Data	2024					
Vaccine Safety Datalink (VSD)	2023					

Table 5. Databases on Income and Employment

Database	Last Updated	Demographic Information Collected				
		Sexual Orientation	Transgender	Nonbinary	Two-spirit	Intersex
Current Employment Statistics	2024					
Survey of Income and Program Participation (SIPP)	2022					

Table 6. Databases on Housing

Database	Last Updated	Demographic Information Collected				
		Sexual Orientation	Transgender	Nonbinary	Two-spirit	Intersex
Annual Homeless Assessment Report	2023		✓	✓		
Connecticut Coalition to End Homelessness	2021	✓	✓	✓		
American Housing Survey	2021					
FHEO Filed Title VIII Cases	2024					

Table 7. Databases on Race, Ethnicity, Indigeneity, and Migration

Database	Last Updated	Demographic Information Collected				
		Sexual Orientation	Transgender	Nonbinary	Two-spirit	Intersex
American Indian Adult Tobacco Survey	2018	✓	✓		✓	
Democracy Remixed (Black and Latino Youth: The Future of American Politics Memo Series)	2014	✓	✓			
GenForward Survey	2023	✓	✓	✓	✓	
MMIP (Murdered and Missing Indigenous People) Database	2023		✓		✓	
Annual Survey of Jails in Indian Country (SJIC)	2022					
Black Women’s Health Study	2024					
HHS Unaccompanied Children Program	2024					
Jackson Heart Study	2024					

Table 8. Databases on Disability

Database	Last Updated	Demographic Information Collected				
		Sexual Orientation	Transgender	Nonbinary	Two-spirit	Intersex
National Core Indicators - Intellectual and Developmental Disabilities (NCI-IDD)	2018					

Table 9. Databases on Youth

Database	Last Updated	Demographic Information Collected				
		Sexual Orientation	Transgender	Nonbinary	Two-spirit	Intersex
Adolescent Behaviors and Experiences Survey (ABES)	2022	✓				
Future of Families and Child Wellbeing Study (FFCWS)	2024	✓		✓		
Growing Up Today Study (GUTS)	2015	✓				
Healthy Minds Study (HMS)	2023	✓	✓	✓		✓
National College Health Assessment (NCHA)	2023	✓	✓	✓		✓
National Longitudinal Study of Adolescent to Adult Health (Add Health)	2024	✓	✓	✓		
National Survey of Youth in Custody (NSYC)	2018	✓	✓			
National Youth Tobacco Survey (NYTS)	2022	✓	✓			
Youth Risk Behavior Surveillance System (YRBSS)	2023	✓	✓			

Table 10. Databases on Aging

Database	Last Updated	Demographic Information Collected				
		Sexual Orientation	Transgender	Nonbinary	Two-spirit	Intersex
Generations: A Study of the Life and Health of LGB People in a Changing Society, United States, 2016-2019	2023	✓	✓	✓		
Health and Retirement Study	2022	✓				
National Health and Aging Trends Study (NHATS)	2024					

Table 11. Databases on Polling

Database	Last Updated	Demographic Information Collected				
		Sexual Orientation	Transgender	Nonbinary	Two-spirit	Intersex
Gallup	2023	✓	✓	✓		
General Social Survey	2022	✓	✓			
Harvard Opinion Research Program	2022	✓	✓	✓		
Pew Research Center	2024	✓	✓	✓		

Table 12. Databases that Disaggregate Sexual Orientation Data

Database	Last Updated	Sexual Orientations Collected					
		Bisexual	Gay/Lesbian (combined)	Questioning	Gay (alone)	Lesbian (alone)	Queer
2015 U.S. Transgender Survey	2019	✓			✓	✓	✓
Adolescent Behaviors and Experiences Survey (ABES)	2022	✓	✓	✓			
American Indian Adult Tobacco Survey	2018	✓	✓	✓			
Behavioral Risk Factor Surveillance System (BRFSS)	2022	✓	✓	✓			
California Health Interview Survey (CHIS)	2022	✓	✓	✓			
Connecticut Coalition to End Homelessness	2021	✓		✓	✓	✓	
Democracy Remixed (Black and Latino Youth: The Future of American Politics Memo Series)	2014	✓			✓	✓	
Future of Families and Child Wellbeing Study (FFCWS)	2024	✓	✓				
Gallup	2023	✓					
General Social Survey	2017	✓	✓	✓			
Generations: A Study of the Life and Health of LGB People in a Changing Society, United States, 2016-2019	2023	✓			✓	✓	✓

GenForward Survey	2023	✓	✓	✓			✓
Growing Up Today Study (GUTS)	2015	✓	✓	✓			
Harvard Opinion Research Program	2022	✓	✓				✓
Health and Retirement Study	2022	✓	✓				
Health Center Patient Survey (HCPS)	2022	✓	✓	✓			
Health Information National Trends Survey (HINTS)	2023	✓	✓				
Healthy Minds Study (HMS)	2023	✓		✓	✓	✓	✓
Household Pulse Survey	2024	✓	✓	✓			
National College Health Assessment (NCHA)	2023	✓		✓	✓	✓	✓
National Crime Victimization Survey (NCVS)	2024	✓		✓	✓	✓	
National Health Interview Survey	2024	✓	✓	✓			
National Intimate Partner and Sexual Violence Survey (NISVS)	2020	✓			✓	✓	
National Longitudinal Study of Adolescent to Adult Health (Add Health)	2024	✓	✓	✓			
National Survey of Youth in Custody (NSYC)	2018	✓	✓	✓			
National Survey on Drug Use and Health (NSDUH)	2022	✓	✓	✓			

National Youth Tobacco Survey (NYTS)	2022	✓	✓	✓			
Pew Research Center	2024	✓	✓	✓			
Population Assessment of Tobacco and Health (PATH) Study	2024	✓	✓	✓			
Prison Policy Initiative Data Toolbox	2024	✓	✓				
Research and Development Survey (RANDS)	2023	✓	✓	✓			
Survey of Prison Inmates (SPI)	2024	✓	✓	✓			
TransPop Study	2021	✓			✓	✓	✓
UCR Hate Crime Statistics	2019	✓			✓	✓	
Youth Risk Behavior Surveillance System (YRBSS)	2023	✓	✓	✓			