

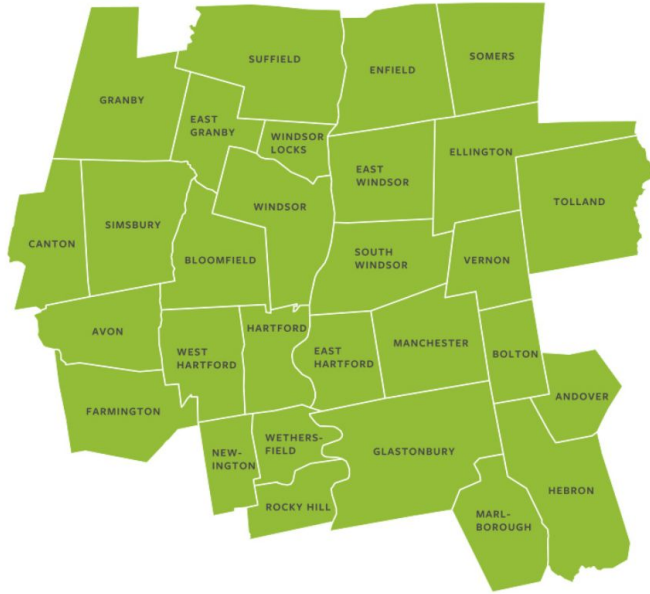
Sexual Orientation and Gender Identity (SOGI) Data Accessibility and Availability

A Connecticut Case Study

Presented by Shannon Carter and Camille Seaberry, DataHaven

Introduction

Motivation



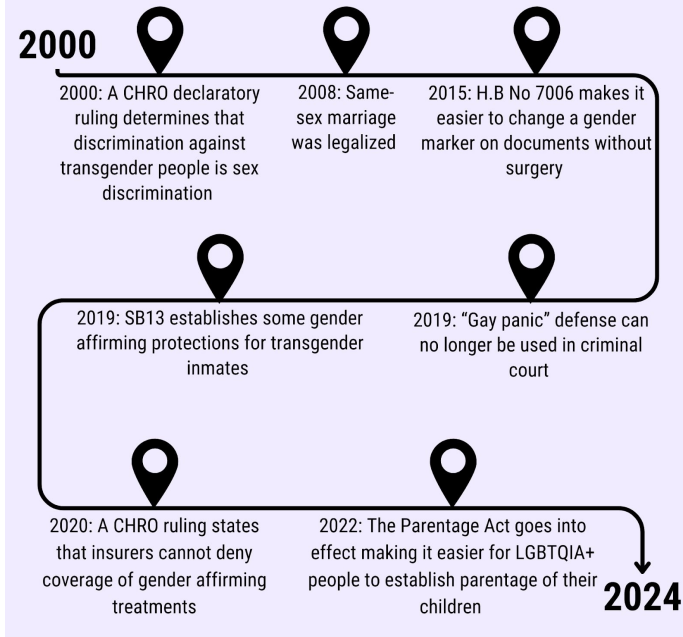
- Partnership with the Hartford Foundation for Public Giving
- Initiated as part of their Equality Fund
- Needed high quality, hyper local information about the queer community
- Drive funding and program decision making

Process

Started with a review of

- Publicly available, disaggregated datasets
- State, federal and local policy
- Relevant academic literature

21st Century Connecticut LGBTQIA+ Policy Landscape Sample



Constructing data

Where are LGBTQ people in data?

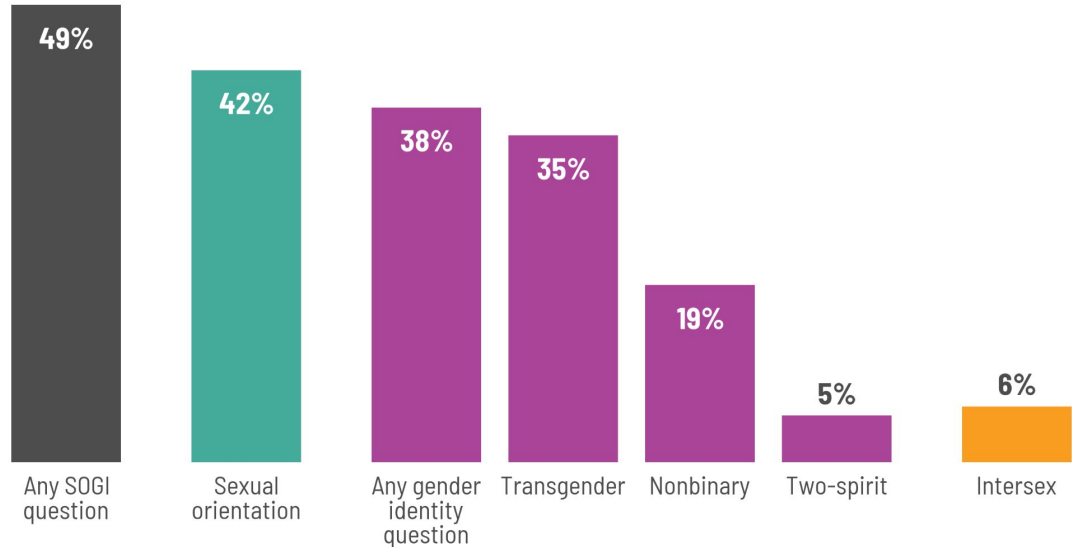
Landscape: LGBTQ people in data

In our review of 96 major datasets, 49% asked a question about SOGI

- Phrasing and depth of questions varies

About half of major datasets studied include any question on sexual orientation or gender identity

Prevalence of SOGI questions in major datasets (N = 96)



Source: Ai (2024), DataHaven.
Note: Questions are not mutually exclusive

Qualitative research

- When LGBTQ people aren't in standard data sources, qualitative research helps us create our data ourselves
- Requires trust and rapport
 - Valid reasons to not trust researchers---points to need for LGBTQ people as researchers for our community

	N	Share
Number of employees		
Less than 10	8	42%
10-49	7	37%
50 and greater	4	21%
Budget		
Under \$500k	8	42%
\$500k to \$5 million	7	37%
\$5+ million	4	21%
Organization type		
Non-profit	14	74%
Governmental agency	3	16%
Community volunteers	1	5%
For profit	1	5%

Qualitative research: results

Emergent themes

- Connecticut's national role
- Complicated relationships with data
- Information seeking processes
- Underfunding
- Discrimination & oppression

"And my trans friend just moved back to Connecticut. And was like, 'I need a name change and what are the good orgs'...I had like five on hand ...I was like, 'Okay, so we're doing something here' because they said Google wouldn't really tell them."

“The goal for the [organization] is to keep the kids alive, period...almost all either attempt, or they’re survivors of multiple attempts to suicide...In fact, when I sit with those parents they just want to get them through the teenage years, through adolescence, they want to keep them alive. And it's very helpful to me that there are parents like that, but number one, is to help them to continue to love their children. And we support the children and what they need, so they'll stay alive, and to see far enough ahead, that there's a hope that there's joy, and there's a full life waiting for them.”

Quantitative research

- Small sample sizes, inconsistent definitions hinder quantitative research
- DataHaven's Community Wellbeing Survey includes a question on gender identity
 - Our phrasing has improved over past few years
 - Several rounds of the survey before this became clear priority
 - Small sample in a small state means we still can't ask in as much depth as we'd like
 - Other orgs can design surveys that meet this need, but it takes resources
- Other national surveys intentionally target small populations, e.g. US Trans Survey

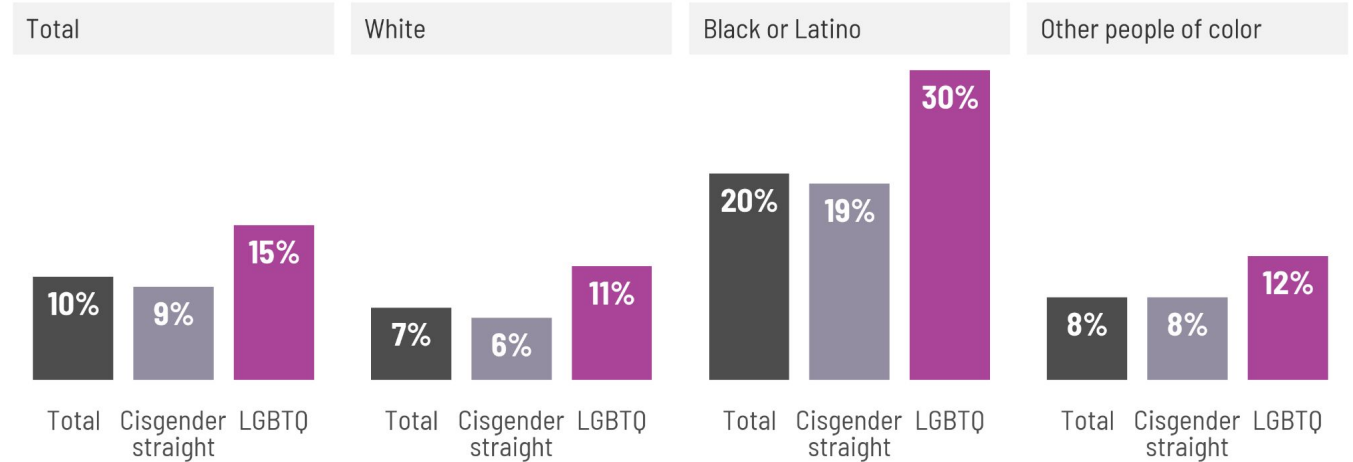
Quantitative research: results

Household Pulse Survey: large enough sample to get into some intersections

e.g. vulnerabilities like food insecurity hit LGBTQ, Black & Latino adults; hit doubly hard if in both groups

Black and Latino LGBTQ adults have elevated rates of food insecurity

Share of Connecticut adults reporting having not had enough food, past 7 days, 2021-2023



Source: DataHaven analysis (2024) of US Census Bureau Household Pulse Survey

Quantitative research: results

DataHaven Community Wellbeing Survey: smaller survey, but responsive to local needs

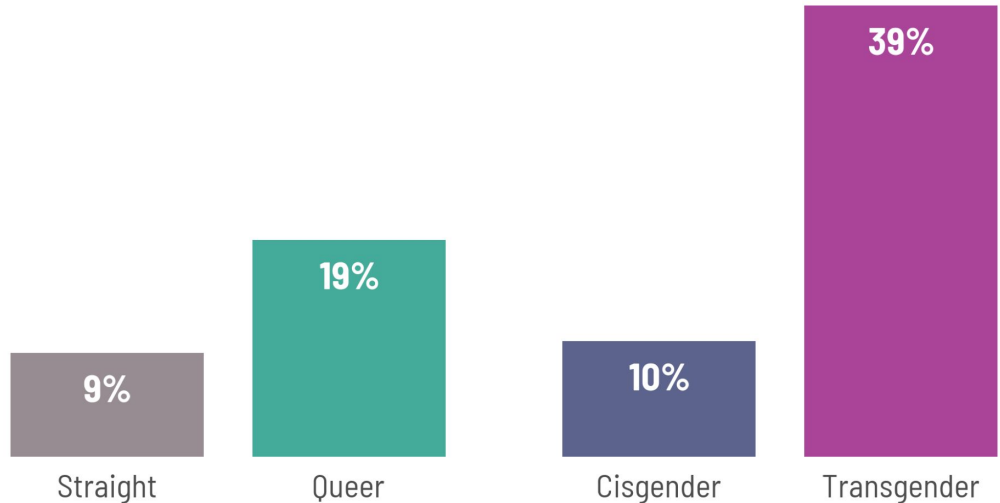
Could get better data by facilitating local LGBTQ-focused orgs implementing own surveys

Transgender adults commonly miss medical care

Share of adults missing needed medical care in past year
By sexual orientation and gender identity, Connecticut, 2018-2022

By sexual orientation

By gender identity



Source: DataHaven analysis (2024) of DataHaven Community Wellbeing Survey

Bigger picture questions

Definitions & questions

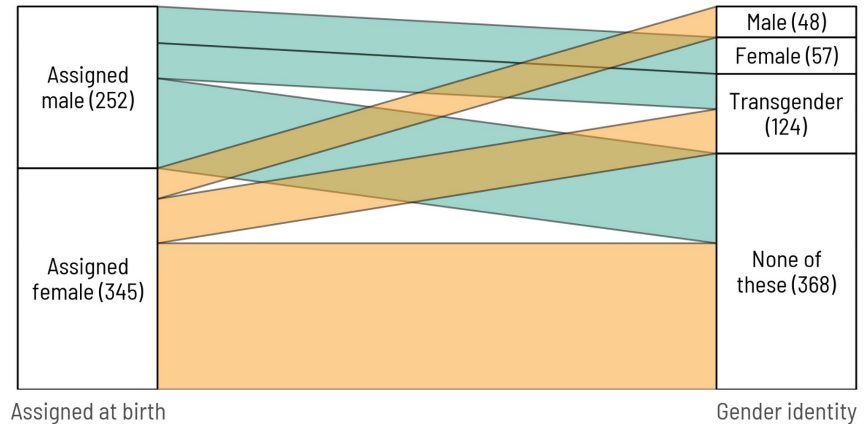
- Definitions and phrasing impact how people answer questions, or even whether they do
- Outsiders might not use language that reflects community
- Need consistency to compare across surveys, over time, between generations
- Figure out what you need to know and why

Definitions & questions: example from Household Pulse

- HPS asks 2 questions for gender:
 - Gender assigned at birth (male / female)
 - Current gender identity (male / female / transgender / none of these)
- How does a trans woman answer this?
- Our first analysis missed 18% of trans adults *because of this phrasing*

Only 21% of transgender adults respond with 'transgender' as a gender identity

Distribution of assigned gender versus gender identity (N = 597)
Household Pulse Survey, transgender Connecticut adults, 2021-2023



Source: DataHaven analysis (2024) of US Census Bureau Household Pulse Survey

Access to & inclusion in data

- More research *about* LGBTQ community done *by* LGBTQ community *for* LGBTQ community
 - Can learn from other marginalized / over-researched communities that have pushed back on research institutions
 - We can be attentive to needs of our community that others probably cannot, e.g. catching phrasing of Household Pulse question
- Who does more diverse data serve? Are we just a DEI checkbox?

Data as a risk

- Is it safe to be included? Who does it serve? Do benefits to LGBTQ people outweigh risks?
 - Plenty of examples of weaponization of data on vulnerable people, esp. trans people
- How much should techniques follow political & social currents?
 - Might be safe here & now but need to preempt policy changes
 - Where do data destruction, obfuscation fit into your data pipeline?

Data as a hurdle

Demanding data of vulnerable populations can impede their work caring for their own community

- Respect qualitative data, anecdotal evidence, lived experience of community
- Find proxy measures that are easier to gather---might not be perfect but good enough
- Funders / decisionmakers: is all that data a *need*, or just a *want*?
- e.g. do we need to measure trans people's access to affirming medical care *yet again*, or can we just start doing what we know works?

Ask honestly: are people dying or in harm's way while you're collecting data on them?

Final thoughts

Takeaways: what next?

- Quantitative data can help us identify strengths and needs, and qualitative data can help us put that data into context
- We need more, better data, but our communities need access to it
- When data we need doesn't exist, we can create it ourselves
- Do no harm approach: LGBTQ people's safety & health cannot be the price to pay for better data
- Trust wealth of knowledge the community has about itself