2024 STATE OF OUR DATA REPORT

UNC CHARLOTTE URBAN INSTITUTE
+ CHARLOTTE REGIONAL DATA TRUST

CONTENTS

ABOUT **THE INSTITUTE + DATA TRUST**

This year, the Charlotte Urban Institute (Institute) celebrates its 55th year and the Charlotte Regional Data Trust (Data Trust) celebrates its 20th, each with a rich history of equipping changemakers in our communities and region with data and information that kindles community transformation and makes a difference in people's lives.

Established by the state legislature in 1969 to serve our urbanizing region, the Institute is our region's applied research and community outreach center. We engage expertise across a diverse set of disciplines and life experiences to curate data and conduct actionable research and policy analysis to convene and equip changemakers.

The Data Trust, formed in 2004, is a community–university partnership that links administrative data across service and organizational silos to provide information our community can act on. The Data Trust is a 501(c)(3) and a university-affiliated entity governed by a board of community and university stakeholders and staffed by the Institute. In addition to their fiduciary responsibilities, the Board guides the strategic priorities of the Institute. Together, the Institute + Data Trust provide information and infrastructure our community can use.

The Charlotte Urban Institute and the Charlotte Regional Data Trust equip changemakers in our communities with data and information that ***kindles community transformation and makes a difference in people's lives.***

OUR **BOARD + STAFF**

Charlotte Regional Data Trust Board of Directors

Jenny Ward, Board President
Bank of America

Jamaal Kinard, Board Vice President & Co-Chair
Community Data Advisory Committee – Lakeview Neighborhood Alliance

Raquel Lynch, Past President
Goodwill Industries of the Southern Piedmont

Jeremy Alajajian, Board Treasurer
UNC Charlotte Urban Institute

Jesh Humphrey, Board Secretary
UNC Charlotte Division of Institutional Integrity

Andrew Bowen, City of Charlotte

Casey Calloway, 26th Judicial District Court

Sherri Chisholm, Leading On Opportunity

Brian Collier, The Gambrell Foundation

Kathryn Firmin-Sellers, United Way of Greater Charlotte

Geraldine Gardner, Centralina Regional Council

Tim Gibbons, Mecklenburg County Department of Social Services

Doug Hague, Chair, Data & Technology Advisory Committee – UNC Charlotte, Corporate Engagement

Melvin Herring, Co-Chair, Community Data Advisory Committee, Johnson C. Smith University

Kristen Howard, Novant Health

Carlene Mayfield, Advocate Health

Aisling Nagel, Charlotte-Mecklenburg Schools

Ann Oshel, Alliance Health

MaryAnn Priester, Chair, Data and Research Oversight Committee, Mecklenburg County, Community Support Services

Federico Rios, Foundation For The Carolinas

Beth Thompson, Charlotte-Mecklenburg Schools

Jennifer Troyer, UNC Charlotte Division of Academic Affairs

Raynard Washington, Mecklenburg County Health Department

Institute + Data Trust Staff

Jeremy Alajajian, Business Officer

Fatimah Alnahash, Research Data Engineer

Bridget Anderson, Senior Research Associate

Kathy Bywaletz, Office Manager

Alberto Centurion, TIMS Project Leader

Asha Ellison, Director of Research Translation and Engagement

Shawn Frazer, Business Services Coordinator

Mia Gaddy, Research Associate

Angelique Gaines, Research Associate

Mecca Howe, Research Associate

Jenny Hutchison, Research Associate

Sydney Idzikowski, Associate Director of the Charlotte Regional Data Trust

Elaine Jacobs, Operations Manager

Bill McCoy, Director Emeritus/ Consultant

Andrew McGomery, TIMS Director

Eric Moore, Senior Research Associate

Liz Morrell, Director of Public Policy Research

Jenny Niu, Research Associate

Nick Occhipinti, Data Scientist

Fatima Rodriguez-Caba, Research Associate

Stephanie Rollinson, TIMS Research Assistant

Lori Thomas, Executive Director

Kailas Venkitasubramanian, Director of Research Analytics

Khou Xiong, Director of Community Research Services

Katie Zager, Research Associate

Graduate Research Assistants

Faria Amin
Koya Brown
Bec De Luna
Alish Jain
Sarah Paul
Kadae Phimia
Mariah Quinn

Report Contact

Lori Thomas, Executive Director

Authorship & Suggested Citation

Thomas, M.L., Idzikowski, S., Occhipinti, N., & Venkitasubramanian, K. (2024). State of Our Data. Charlotte, NC: UNC Charlotte Urban Institute + Charlotte Regional Data Trust. https://ui.charlotte.edu/state-of-our-data/.

LETTER FROM **THE BOARD PRESIDENT + EXECUTIVE DIRECTOR**

In 2021, the Charlotte Regional Data Trust's Board of Directors took a step we had never before taken in our then 17-year history – we started a fundraising campaign to put data to use for our community. We believed in the promise of the Data Trust to connect our community across data silos, but we also knew that we lacked critical data infrastructure to scale our impact.

The following year, we took a couple more big steps – we changed our name from the Institute for Social Capital to the Charlotte Regional Data Trust and we voted to become an advisory board for the UNC Charlotte Urban Institute. These actions underscored our commitment to serve as a trusted partner for community data and information, and to further align the resources and efforts of two UNC Charlotte community-facing centers.

We release this first State of Our Data report as we celebrate the completion of our fundraising and the launch of our new web portal created through that community investment. The report will serve as an annual check-in with our Board and committee members, our data partners and data users, and other community and university stakeholders about the available infrastructure, the people in our data, and how our data have been used in the past year.

> **This report will serve as an annual check-in with our Board and committee members, our data partners and data users, and other community and university stakeholders about the available infrastructure, the people in our data, and how our data have been used in the past year.**

We're already using the information from this report to inform recruitment for our Community Data Advisory Committee (CDAC), to ensure that the people in our data are represented in our Board governance. We intend to further use this and subsequent reports to inform our stakeholders, guide our planning, and to track our impact.

We are grateful to the many people it takes to build and sustain an enduring community data infrastructure – our Board, our committees, our donors, our staff, our data and research partners, and, especially, the people in the data. The purpose of this infrastructure isn't data for data's sake, but data to ensure that everyone in our region can thrive. Thank you.

Jenny Ward

Jennifer Ward, President
Charlotte Regional Data Trust Board of Directors

Lori Thomas

Dr. Lori Thomas, Executive Director
Charlotte Urban Institute + Regional Data Trust

INTRODUCTION

Equipping changemakers requires infrastructure – the components, systems, and practices necessary to consistently, effectively, and ethically put data to use for community benefit. One dataset, one dashboard, or one report can provide valuable insights but more is required to guide the ongoing community and organizational decision-making necessary to ensure everyone in our community thrives. The work requires an enduring data and research infrastructure that is accessible to multiple and diverse community stakeholders.

The Institute + Data Trust are committed to providing the critical building blocks of evidence-based and ethical decision-making so everyone in our region can thrive. The State of *Our* Data Report provides our communities, partners, and stakeholders with an accessible summary of this enduring infrastructure, and particularly the data we hold in trust and why it matters.

The Institute + Data Trust are committed to providing the critical building blocks of evidence-based and ethical decision-making so everyone in our region can thrive.

What do we mean by Data Trust?

In legal terms, a trust is an arrangement where a person or an entity holds an asset on behalf of a designated beneficiary. At the Institute + Data Trust, we hold data for our partners, for community benefit – especially for those who are represented in the data. The "Our" in the State of Our Data is purposeful and inclusive. We recognize that the integrated data we hold does not "belong" to the Institute + Data Trust.

Trust is also the essential work of this community asset – to build trust among our stakeholders. As is often said, "Data moves at the speed of trust" (Robert Wood Johnson Foundation, 2015). And we prioritize building and maintaining trust with all of our stakeholders – from the individuals and communities whose data we hold, to the organizations and institutions that collect and share the data, to the researchers, analysts, and people with lived expertise who help make sense of its patterns and anomalies.

Why does trust matter?

Trust is the foundation for our most basic, fruitful human interactions. And trusted relationships are required to build collaborations that last and help individuals and communities meet common goals and commitments.

We are releasing this first annual report during a time of profound distrust in our society and among our communities. Much of the growing distrust finds fuel in misinformation and disinformation and the explosion of data that is used to support dueling arguments. People from a range of political and ideological perspectives are wary of, or simply do not believe in, the objectivity of data or the neutrality of some of our key institutions like public universities or governments that routinely embrace research and data-driven solutions (Brady & Kent, 2022; Saad, 2023).

The lack of trust is also exacerbated by the lived experiences of how data and research have worked in people's lives. It is hard to trust the numbers when the data have regularly driven decisions that don't benefit your family or families like yours. Redlining, with origins in a 1933 doctoral dissertation (Hoyt, 2000; Rothstein, 2017), was supported by population data and financial risk models and, yet, resulted in racially segregated housing, and in residential and business lending practices that continue to impact outcomes as diverse as health and generational wealth (e.g., Krieger et al., 2020). The closure of 108 rural North Carolina hospitals since 2005 (Sheps Center, 2024) and the persistence of health professional shortage areas in both our rural and urban communities (HRSA, n.d.) speak louder than the data-driven financial models that are cited in their closures. When your family and community's livelihood and well-being are regularly threatened by policies and business practices that are supported by data and research, it's hard to trust the organizations championing the numbers.

It's also hard to trust the numbers when data and research regularly describe your community in incomplete and primarily deficit-focused ways. If the map of your neighborhood or community consistently illustrates what is wrong, without a more complete context that includes community assets and the histories and policies that have cumulatively resulted in negative outcomes, it may be difficult to trust the research or those responsible for it.

To nurture the relationships and collaboration needed to build communities and a region that works for everyone, we have to take these concerns seriously and work to build trust among the many who never had, or have lost trust in data and research.

At the Institute + Data Trust, we hold data for our partners, for community benefit – especially for those who are represented in the data. The *"Our"* in the State of *Our* Data is purposeful and inclusive. We recognize that the integrated data we hold does not "belong" to the Institute + Data Trust.

Trust is also the essential work of this community asset – to build trust among our stakeholders.

And we prioritize building and maintaining trust with all of our stakeholders – from the individuals and communities whose data we hold, to the organizations and institutions that collect and share the data, to the researchers, analysts, and people with lived expertise who help make sense of its patterns and anomalies.

How do we build trust?

For each of our stakeholders, we are responsible for pursuing and ensuring trustworthiness. Here's how we work to achieve it:

- **For the people represented, and overrepresented in our data,** we recognize that data points are people and we treat the data with the same care and respect we would want information about us to be treated. We structure ways for people to participate in decisions about how their data are used, and we work to ensure that research about communities can inform and be used by those communities. We also create accessible and affordable capacity-building opportunities to understand research and use data.

- **For our data partners,** we recognize the legal and ethical responsibility to protect data privacy and to use data to improve programs for the people they serve. We know organizations face significant challenges to develop and maintain the legal and technical capacity to effectively access and use data. We work on the backend to sustain critical community data infrastructure and involve organizations in decisions about the data they share.

- **For our data users,** we recognize that impactful research and analysis move faster when our processes and communication are clear and timely. We also recognize that research is more accurate and effective when researchers have access to robust metadata, relevant history and context. We work continuously to enhance the information that is available to those who use our data resources.

Purpose of the report

The State of *Our* Data report is an annual description of the state of this community data infrastructure, the data the Institute + Data Trust hold, and the ways our stakeholders have used the data in the previous year. This annual report is shared with the Data Trust's Board of Directors and our partners, and is available to the public on our website.

In addition, we will use annual information about who is in our data to inform representation on our board and its committees. We will also use this information to guide our communication and dissemination strategies. And, the report will help us monitor the extent of data coverage across key quality of life domains and across our region.

We anticipate that, as we engage our current and new stakeholders, this report will evolve. We will be able to report on new partners, new research questions, and new uses of the data. And we will be able to report on the growing impact of our shared community resources.

Finally, this report is one of the many ways we practice transparency in our work and honor the investment of UNC Charlotte, our donors, and particularly our communities, whose lives are represented in our data and whose quality of life we seek to ensure. Trust requires transparency.

OUR COMMUNITY DATA

INFRASTRUCTURE


TRANSPORTATION
ARTS & CULTURE
CIVIC ENGAGEMENT & BELONGING
ECONOMY & WORK
EDUCATION
ENVIRONMENT
HEALTH, MENTAL HEALTH & WELLBEING
HOUSING & HOMELESSNESS
SAFETY & JUSTICE
A CITY & REGION WHERE EVERYONE THRIVES
COMMUNITY DATA INFRASTRUCTURE
PEOPLE & PARTNERS

OUR **COMMUNITY DATA INFRASTRUCTURE**

At the Institute + Data Trust, we equip changemakers with the information they need to create a city and region where everyone thrives. Through our community data infrastructure, we provide access to information about interconnected domains such as education, the economy, and housing that determine people's quality of life.

Our infrastructure includes the data tools we provide like our integrated administrative data system, the Charlotte-Mecklenburg Quality of Life Explorer, and the emerging Regional Explorer. It also includes the services, technology, policies, and practices that ensure access, security, ethical use, and impact.

And crucially, our community data infrastructure is inextricably linked to the people and partners who contribute data, expertise, and context, join us in research and dialogue, and put the tools, services, and information to work for our communities and region. These relationships not only support our work, they are also what makes our infrastructure notable. On the following pages, we briefly highlight each component of the ecosystem.

Interconnected Domains

The Institute + Data Trust work across nine interconnected domains that influence quality of life and thriving at individual, organizational, and community levels. Issues and opportunities in these domains rarely exist in isolation. Instead, they are dynamically connected and often require multiple forms of data, diverse research methods, and a range of expertise to understand. We continuously seek new data sources and conduct applied research to help us and our communities better grasp these arenas and the ways that different people in our region may experience them.

Data Infrastructure

Our community data infrastructure supports understanding and improving these nine domains for everyone's benefit, particularly those who have not historically shared in the prosperity of our Charlotte region. Key tools of our infrastructure include:

- The Data Trust's **integrated administrative data system** links individual-level data across organizational silos. Administrative data are data collected for operational purposes as people engage with services and agencies. Even though the data are not collected for research purposes, it can provide rich information to better understand complex social problems and evaluate solutions. At the Data Trust, we repurpose administrative data for research, evaluation and planning. Data can be accessed through a secure data request process on our web portal.

- The **Charlotte-Mecklenburg Quality of Life Explorer** (Quality of Life Explorer) is a publicly accessible data tool provided in partnership with the City of Charlotte, Mecklenburg County, and the towns of Cornelius, Davidson, Huntersville, Matthews, Mint Hill and Pineville. Users can access data on more than 80 variables for 462 neighborhoods in Charlotte-Mecklenburg. Data for the Quality of Life Explorer is drawn from a variety of federal, state and local sources, and includes variables that generally correspond to our nine infrastructure domains. The tool allows users to create custom geographies (i.e., school zones, business districts, light rail corridors) or use the designated neighborhood boundaries. Data are updated annually and can be accessed on the Quality of Life Explorer dashboard and on the Data Trust's web portal.

- The emerging **Regional Explorer** will create a census tract-based tool for the 14-county Charlotte metropolitan region. The application will have a similar look and feel to the Quality of Life Explorer, but will add functionality to compare neighborhoods, cities and counties across time, and conduct exploratory analysis via a companion dashboard. In a later phase, the Regional Explorer will integrate existing research reports and stories by the Institute or our partners to provide additional context on quantitative data points available through the tool. The Regional Explorer will offer a unique, integrated lens on how the Charlotte region is changing across time. Data from the tool will be available through the dashboard and on the Data Trust's web portal.

- The Institute has historically conducted **public opinion surveys** and, after a pandemic pause, we are resuming this service that focuses on residents' quality of life, perceived well-being, and other topics important to our communities and partners. Leveraging a representative sample of the community, we analyze the survey results to identify patterns and trends on various topics ranging from public safety to education, health, transportation, and housing. Survey data will complement local and regional indicators and will become an additional longitudinal data source available on the Data Trust web portal.

- **The Impact, Data, and Evaluation Academy (IDEA)** is a community capacity-building initiative that provides access to training and skill-building so people can put data and information to use in their organizations and communities. A community-driven certificate program on the foundations of research, evaluation, and data management is currently in development and will be piloted in 2025. This program will be open to all but is specifically intended to build on the assets of small and grassroots organizations and address the barriers they face in research and evaluation.

- The **Institute's data, research, & policy services** provide support alongside community and university partners on numerous local and regional projects and initiatives. Our staff researchers bring expertise in disciplines ranging from public health, social work, urban planning, and political science to communications, computing, and anthropology. In addition to data application and research, we are committed to convening meaningful conversations, as well as translating our research and technical findings into information that's easier to understand and apply.

In addition to the technology, policies and practices that support them, these components offer a community data infrastructure that we support, maintain, and grow for broader community use and application.

People & Partnerships

The strength of our infrastructure is dependent on our relationships – the people that volunteer to govern and guide the work, the partners that share data, the people who use it, question it, and push the Institute + Data Trust to make our infrastructure more applicable. There are many hands that shape our work and we highlight key stakeholder groups below.

- **Board of Directors and Committee Members.** The Data Trust Board maintains responsibility over the data held by the Data Trust and use of the data is managed by the Board's Data and Research Oversight Committee (DAROC). A suite of legal agreements and data governance practices comply with federal and state privacy laws and help ensure the integrity and security of the data. Our Data and Technical Advisory Committee (DTAC) monitors our technical infrastructure, and particularly the completion and implementation of the new web portal and cloud-based data system. A new Community Data Advisory Committee (CDAC) will work to ensure that people who are represented in the data participate in decisions related to its use.

- **Data Partners.** Our integrated administrative data partners sign legal agreements to share individual-level data with the Data Trust and include three types of partners. Our standing partners deposit data at least annually into our integrated data system. Our project-specific partners do not have a regular deposit schedule but they have standing data-sharing agreements and their data may be requested for use. Our legacy partners have signed data-sharing agreements in the past, but have not yet signed updated data agreements that allow participation in our new data portal.

  The Quality of Life Explorer is also dependent on data partnerships, with ongoing data sharing among the three lead organizations – the City of Charlotte, Mecklenburg County, and the Institute + Data Trust. In addition, we work with the towns of Cornelius, Davidson, Huntersville, Matthews, Mint Hill, and Pineville to access their specific municipal data for the Quality of Life Explorer. More information about data partners is available on the Data Trust website.

- **People in the Data.** The administrative data and Explorer data represent people and their experiences. We are working to include and partner with people and groups represented in our data through our Community Data Advisory Committee and by partnering with small and grassroots organizations for data sharing and mutual capacity building.

- **Data Users.** Community, faculty, and staff researchers turn the data available through the Data Trust's web portal and Explorer tools into information our community can use. Those projects most recently approved for use of administrative data are listed later in the report.

- **UNC Charlotte Partners.** The Institute is a unit of UNC Charlotte's Division of Academic Affairs and the Data Trust is a 501(c)(3), affiliated entity of the University. We rely on partnerships with other University units and divisions to support key functions of the Institute + Data Trust. The divisions of Academic Affairs, Advancement, Legal Affairs, OneIT, and Research are key internal partners in our ongoing operations, their engagement spelled out in part in the Data Trust's cooperation agreement with UNC Charlotte. Our sister unit, urbanCORE (Community Oriented Research & Engagement), as well as multiple academic departments are key partners that inform use of and engagement with the Institute + Data Trust.

- **Community Investors.** Since 2019, a number of public and private funders have contributed to our data infrastructure improvements, including seed funding in 2019 from The Gambrell Foundation. The Gambrell gift allowed us to bring in a third party to assess our data and technical infrastructure and make recommendations that led to the Board decision to launch a fundraising campaign to improve our infrastructure. Pledges from Bank of America and UNC Charlotte kicked off the campaign and we successfully completed it with a gift from an anonymous corporate foundation, allowing us to fund student data stewards through our capacity-building initiative, the Impact, Data, and Evaluation Academy (IDEA). Our community investors include:

  - Advocate Health
  - Bank of America
  - City of Charlotte
  - The Duke Endowment
  - Foundation For The Carolinas
  - The Gambrell Foundation
  - The Knight Foundation
  - Mecklenburg County
  - Trane Technologies
  - United Way of Greater Charlotte
  - UNC Charlotte
  - UNC Charlotte School of Data Science

OUR DATA


Tree Canopy, 2012
COUNTY
45.1%
NPA Distribution, 2012
Adopt-a-Street Participation
Residential Recycling
Residential Solid Waste
Residential Solid Waste Diversion
Tree Canopy
Commuters Driving Alone
Energy Consumption - Electricity
2012
County Commission

This section describes two key data sources at the Institute + Data Trust: (1) the spatial variables used for the Quality of Life Explorer and the emerging Regional Explorer, and (2) annual deposits of administrative data into the Data Trust's integrated data system. Additional details on data, methods, and analyses are available in Appendix A.

Explorer Data

The Charlotte-Mecklenburg Quality of Life Explorer and Regional Explorer use data from multiple sources to map variables on quality of life domains to our neighborhoods and communities. These data help us understand our own spaces and places, but also help us see patterns – areas where we are thriving together, or areas where we have work to do to ensure that everyone can share in the prosperity of our city and region.

The focus of our communities' Explorer data is on place. The Quality of Life Explorer focuses on our neighborhoods in Charlotte-Mecklenburg, called Neighborhood Profile Areas (NPAs) in the Explorer. NPA boundaries were determined using data from community partners and in conversations with neighborhoods conducted by the City, County, and UNC Charlotte in 2011. While neighborhood boundaries may change over time, we generally do not adjust neighborhood boundaries, which make it possible to observe patterns in the same place over time.

While neighborhoods are the focus for the Quality of Life Explorer, the Regional Explorer will display data at the larger census-tract level across the Institute's 14-county region. Eventually, and as regional neighbors express interest, the Institute can work with municipalities, county governments, and their residents to create more granular tools like the Charlotte-Mecklenburg Quality of Life Explorer. In the meantime, the Regional Explorer will offer an integrated lens on how quality of life in our region is changing across place and time. Additional details on the Explorer tools and the methods we use are available in Appendix A

OUR COMMUNITY
DATA SOURCES

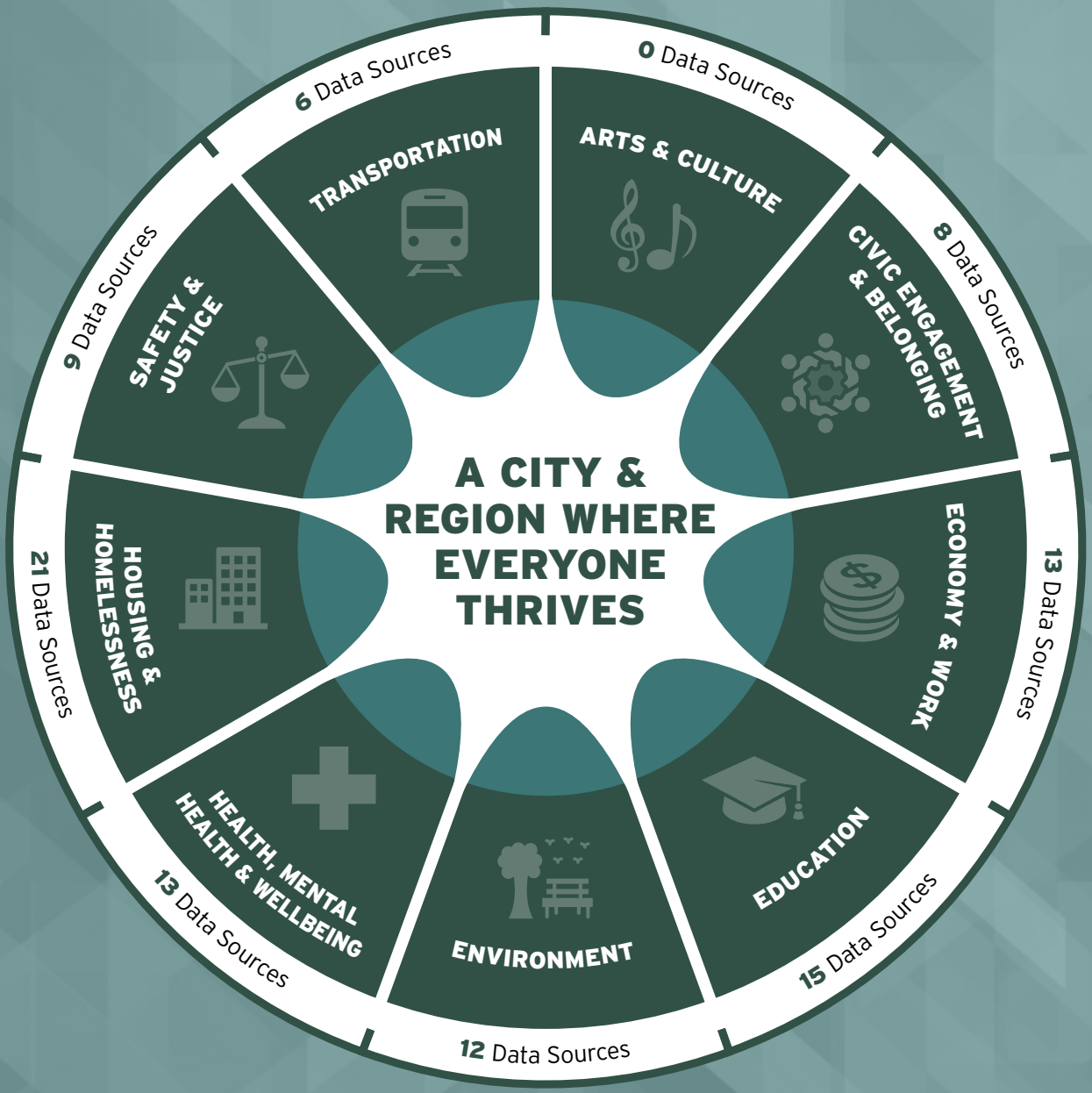


FIGURE 1 – Our Community Data Sources

Our Data Sources by Arena

DATA TYPE LEGEND ● Explorer Data ● Integrated Administrative Data

Data partners in all caps deposit data annually to the Data Trust or are in the process of beginning annual deposits. Other partners deposit data as requested or are legacy partners.

CIVIC ENGAGEMENT & BELONGING

- City of Charlotte
- Mecklenburg County Board of Elections
- Town of Cornelius
- Town of Davidson
- Town of Huntersville
- Town of Matthews
- Town of Mint Hill
- Town of Pineville

ECONOMY & WORK

- American Community Survey
- Bureau of Labor Statistics
- City of Charlotte
- City of Charlotte Neighborhood & Business Services
- CRISIS ASSISTANCE MINISTRY
- Federal Deposit Insurance Corporation
- GOODWILL INDUSTRIES OF THE SOUTHERN PIEDMONT
- LAKEVIEW NEIGHBORHOOD ALLIANCE
- Longitudinal Employer Household Dynamics
- MECKLENBURG COUNTY DEPARTMENT OF SOCIAL SERVICES – COMMUNITY RESOURCES
- Mecklenburg County Register of Deeds
- Mecklenburg County Tax Parcels
- National Credit Union Administration

EDUCATION

- American Community Survey
- Bethlehem Center
- Cabarrus County Schools
- Central Piedmont Community College
- Charlotte Mecklenburg library
- CHARLOTTE-MECKLENBURG SCHOOLS
- City of Charlotte
- Communities in Schools
- Freedom School Partners
- Girl Scouts Hornet's Nest Council
- Kannapolis City Schools
- SCHERMCO FOUNDATION
- SMART START OF MECKLENBURG COUNTY
- University of North Carolina at Charlotte
- YMCA

ENVIRONMENT

- CharMeck 311
- Charlotte Water
- Charlotte land use planning
- City of Charlotte
- City of Charlotte Solid Waste Services
- Duke Energy
- Mecklenburg County Land Use & Environmental Services Agency
- Mecklenburg County Parks and Recreation
- Piedmont Natural Gas
- U.S Department of Agriculture
- U.S Geological Survey
- US Environment Protection Agency

HEALTH, MENTAL HEALTH & WELLBEING

- Advocate Health
- Cardinal Innovations
- Center for Disease Control
- Lakeview Neighborhood Alliance
- MECKLENBURG COUNTY DEPARTMENT OF SOCIAL SERVICES – CHILD, FAMILY AND ADULT SERVICES
- MECKLENBURG COUNTY DEPARTMENT OF SOCIAL SERVICES – COMMUNITY RESOURCES
- Mecklenburg County Health and Mental Health Department
- Mecklenburg County Parks and Recreation
- North Carolina Board of Pharmacy
- Novant Health
- SMART START OF MECKLENBURG COUNTY
- U.S Department of Agriculture
- YMCA

HOUSING & HOMELESSNESS

- National Housing Preservation Database
- American Community Survey
- Charlotte Family Housing
- Charlotte land use planning
- City of Charlotte

HOUSING & HOMELESSNESS *(continued)*

- CITY OF CHARLOTTE HOUSING AND NEIGHBORHOOD SERVICES
- City of Charlotte Neighborhood & Business Services
- CRISIS ASSISTANCE MINISTRY
- DreamKey Partners
- HABITAT FOR HUMANITY OF THE CHARLOTTE REGION
- HEAL CHARLOTTE
- Hearts United For Good
- HOMELESS MANAGEMENT INFORMATION SYSTEM (30+ AGENCIES)
- Inlivian
- Mecklenburg County Code Enforcement
- Mecklenburg County Register of Deeds
- Mecklenburg County Tax Parcels
- North Carolina Housing Finance Agency
- THE HOUSING COLLABORATIVE
- The Relatives
- U.S. Department of Housing and Urban Development

SAFETY & JUSTICE

- CHARLOTTE-MECKLENBURG POLICE DEPARTMENT
- City of Charlotte Fire Department
- Cornelius Police Department
- Huntersville Police Department
- Matthews Police Department
- MECKLENBURG COUNTY SHERIFF'S DEPARTMENT
- Mecklenburg E911
- Mint Hill Police Department
- Pineville Police Department

TRANSPORTATION

- American Community Survey
- Charlotte Area Transit System
- Charlotte Department of Transportation
- Charlotte land use planning
- City of Charlotte
- NC Department of Transportation

Administrative Data

The Data Trust's integrated administrative data system uses data that are collected for an organization's operational purposes such as managing program enrollment or tracking services provided. While collected for another purpose, administrative data can be repurposed for research, evaluation, and planning.

Integrating administrative data addresses the central challenge of organizational and information silos, particularly across the often fragmented health and human services landscape. Community challenges and opportunities span organizations and sectors, but the data that would help us better understand issues and create solutions are trapped in data silos. Integrated administrative data systems allow us to span our information silos to more comprehensively understand how people move within and across organizations and sectors.

The focus of the Data Trust's integrated administrative data system is on people. Because the Data Trust's administrative data is about people and, because the information that organizations collect on people's lives can be sensitive and private, there is an additional layer of legal responsibility and ongoing governing processes that accompany our work. We manage the deposit, use, and privacy of these data with a suite of legal agreements and a decision-making process that all partners agree to before joining the Data Trust (See the Data Trust website for a description of how the Data Trust works).

The administrative data described in the chart to the right includes the data holdings of our standing partners, both individual and collaborative partnerships. We use abbreviations and colors to refer to data partners in the graphics.[1]

In the next section, we describe the count of unduplicated individuals in our data holdings, the extent to which individuals are present in the data of multiple partner organizations, and individual's demographic characteristics. We describe an annual count of 2023, our most recent year, and we also provide longitudinal information since 2016.[2] See Appendix A for a more in depth description about partner years and methods included in analyses.

Standing Data Partner Names, Abbreviations, and Colors

NAME	ABBREV. & COLOR
Charlotte-Mecklenburg Homeless Management Information System (34 partner organizations)	HMIS
Charlotte-Mecklenburg Schools	CMS
Crisis Assistance Ministry	Crisis
Goodwill Industries of the Southern Piedmont	GW
Habitat for Humanity of the Charlotte Region – Homeowner	Habitat – Homeowner
Habitat for Humanity of the Charlotte Region – Critical Home Repair	Habitat – CHR
Mecklenburg County Department of Social Services – Economic Services	DSS – ES
Mecklenburg County Department of Social Services – Abuse and Neglect Investigations	DSS – ANI
Mecklenburg County Department of Social Services – Family In Home Services	DSS – FIHS
Mecklenburg County Department of Social Services – Children in Custody	DSS – CIC
Mecklenburg County Sheriff's Office – Arrest Processing Center[3]	MCSO – Arrest
Mecklenburg County Sheriff's Office – Booking	MCSO – Booking

How many unique people are reflected in our administrative data?

In 2023, administrative data on 370,434 unduplicated individuals were included in the Data Trust, approximately 32% of the Mecklenburg County population (American Community Survey 1-Year Estimates, 2023). Since 2016, a total of 767,325 unique individual's administrative data have been reflected in the Data Trust, while overall annual unduplicated counts of individuals remain fairly

[1]Note that of our standing partners, Heal Charlotte, City of Charlotte Housing and Neighborhood Services, Charlotte-Mecklenburg Police Department, and those listed as "New Partners in Process" above are either still in the data onboarding process or we didn't have data at the time of the data pull for this project and will be reflected in a future report. In addition, Habitat for Humanity for the Charlotte Region, Mecklenburg County Sheriff's Office, and Mecklenburg County Department of Social Services (DSS) share more than one distinct dataset with the Data Trust. Depending on the analysis, they may be described as one organization or as a specific dataset.

[2]Note: This dataset is sent to us by Mecklenburg County Sheriff Office (MCSO) and includes data from the Arrest Processing Center, which includes arrest data from all arresting entities in Mecklenburg County.

[3]When possible, data partner datasets are provided for dates 1/1/2016-12/31/2023. For Charlotte-Mecklenburg Schools, the 2022/2023 school year is the most recent school year used in this report. The Homeless Management Information System Data is through 12/31/2022, the MCSO Arrest data goes through 4/30/2022, and Habitat for Humanity of the Charlotte Region Homeowner and Critical Home Repair go through 2/29/2022. Goodwill Industries of the Southern Piedmont data begins 1/1/2020. For more information See Appendix A.

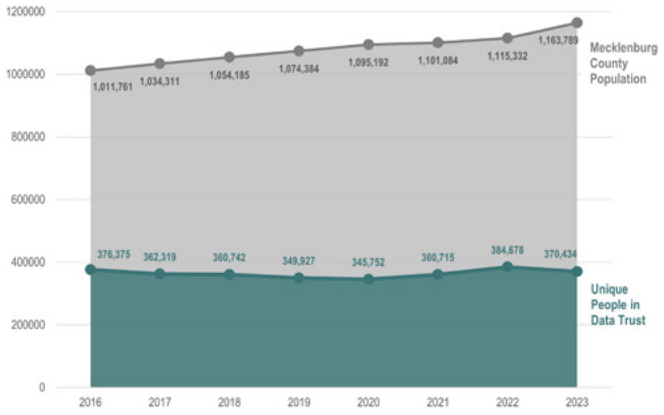
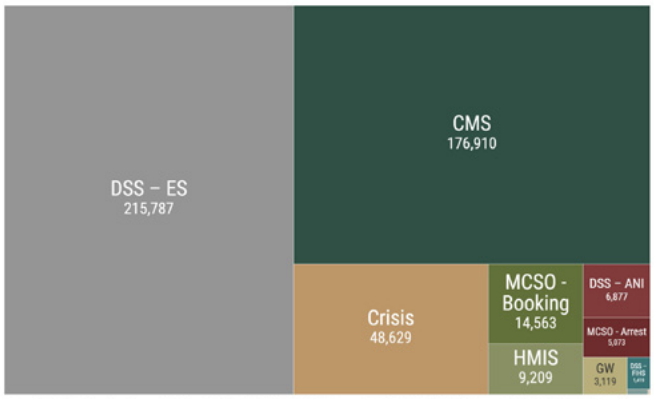


FIGURE 2 – Since 2016, the number of unique people in the Data Trust has remained relatively stable, representing about 35% of the Mecklenburg County population.



Note: CMS 22/23 school year; HMIS 1/1/22-12/31/22; MCSO-Arrest 1/1/22-4/30/22; Habitat 1/1/22-2/29/22; Habitat for Humanity Homeowner & CHR served 19 people between 1/1/22-2/29/22. DSS-ANI represents 6 months of 2023.

FIGURE 3 – DSS – ES, CMS, and Crisis represent the largest data partners (n= 481,808).

steady, ranging between 345,752 in 2020 to 384,678 in 2022, or 37.2% to 34.4% of the annual Mecklenburg County population. Figure 2 below describes the annual unduplicated count of individual data in the Data Trust compared to the annual Mecklenburg County population.

Figure 3 displays the total number of unique individuals by data partner. Unless otherwise noted, the numbers are for the entire year of 2023. Data partners vary widely by how many people they serve and the size of their datasets. The largest datasets are Mecklenburg County Department of Social Services – Economic Services, Charlotte-Mecklenburg Schools, and Crisis Assistance Ministry.

What are the characteristics of the people in the Data Trust's administrative data?

The administrative data held by the Data Trust includes information our partners collect about the characteristics of the people they serve. Whether or not specific data are collected, however, isn't always consistent. These variations are typical in administrative data and can be driven by government regulations, funding requirements, and/or simply the evolution of an organization's data system. Figure 4 below describes the types of characteristics that are collected by each of the Data Trust's standing data partners in their most recent deposit.

	Race*	Ethnicity*	Under 18*	Over 18*	Geography*	Gender*	Education Level	Disability	Income	Immigration	Military/Vet Status
CMS	✔	✔	✔		✔	✔					
Crisis	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
DSS-ANI	✔	✔	✔	✔		✔					
DSS-CIC	✔	✔	✔	✔	✔	✔					
DSS-FIHS	✔	✔	✔	✔		✔					
DSS-ES	✔	✔	✔	✔	✔	✔					
GW	✔	✔	✔	✔		✔			✔		
Habitat – CHR	✔	✔		✔				✔	✔		✔
Habitat – Homeowner	✔	✔		✔	✔	✔			✔	✔	
HMIS	✔	✔	✔	✔	✔	✔		✔	✔		✔
MCSO – Arrest	✔	✔	✔	✔		✔					
MCSO – Booking	✔	✔	✔	✔	✔	✔					

FIGURE 4 – There is variation in the characteristics collected by data partners.

*Characteristics described further on next page.

As the figure on the previous page shows, there is variation in what data partners collect. This section focuses on variables that are collected the most consistently across data partners: race, ethnicity, gender, age, and geography.

Normalization: The data collected about people's characteristics may vary from partner to partner. The Data Trust normalizes these data, meaning we make methodological decisions to standardize the data so they can be grouped and compared. We normalize the names of the racial, ethnic, and gender groups because there is variation across data partners. The Data Trust acknowledges the need to standardize data as well as the risk of doing so, particularly when it can obscure and oversimplify identities. More information about normalization is available in Appendix A.

Representation: In several places, this report describes proportionate representation and disproportionate overrepresentation or underrepresentation. This refers to how the data we hold compares to data on the overall population of the county or the particular area we are describing. Understanding representation is important in the use and interpretation of Data Trust data.

Race & Ethnicity

In 2023, the majority of individuals (49.6%, n=183,609) whose administrative data are in the Data Trust identified as Black or African American (see Figure 5). White individuals represented 23.9 % of individuals (n=88,874). Those identifying as Native Hawaiian or Pacific Islander or Native Alaskan or American Indian each made up less than 1% of the Data Trust's administrative data (0.1%, n=306; 0.2%, n=712, respectively). The multi-race category, which we discuss in more detail below, included 12.6% (n=47,075) of individuals, while individuals identifying as Hispanic or Latino were 7.3% (n=27,228) of the most recent Data Trust data. Those identifying as Asian were 4.8% (n=17,793) of the data.

Since 2016, the proportions of racial and ethnic groups represented in Data Trust data have been relatively similar. There has been a slight increase over time in the percentage of Hispanic or Latino people and a slight decrease in White people, similar to larger population trends in Mecklenburg County over the same period (Figure 6).

Compared to the Mecklenburg County population, people who identify as Black or African American are overrepresented in Data Trust data (see Figure 7). In Mecklenburg County in 2022, the latest available data on the county population,

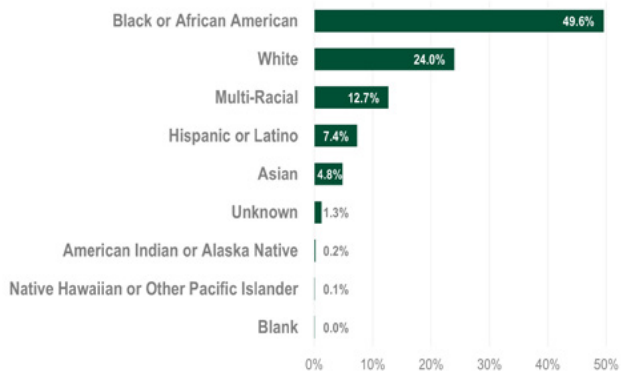


FIGURE 5 – In 2023, the majority of people in the Data Trust identified as Black or African American (n=370,434).

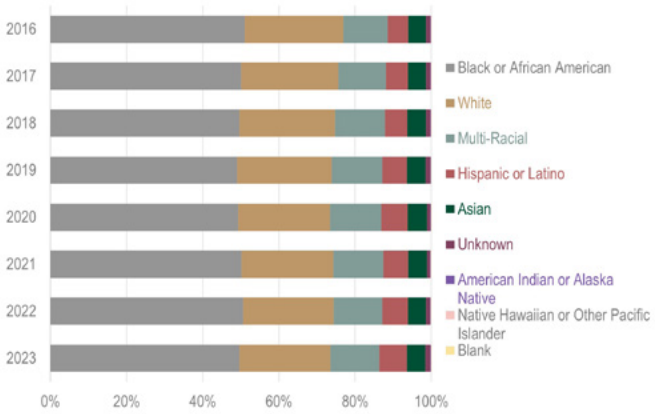


FIGURE 6 – Since 2016, the share of racial and ethnic identities has been fairly steady, with a slight increase of Hispanic or Latino people.

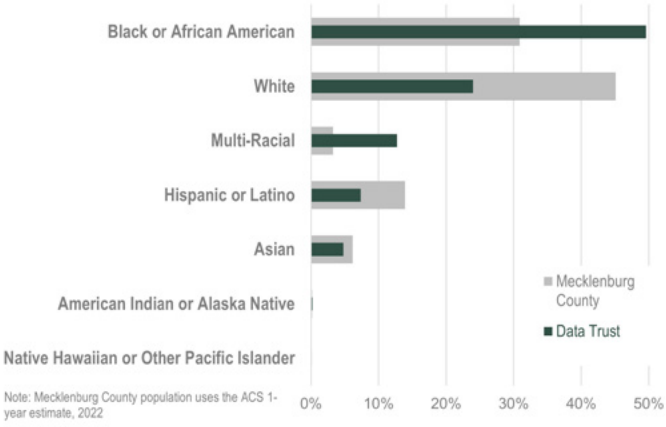


FIGURE 7 – The 2023 population of people in the Data Trust differs from the overall population of Mecklenburg County. People who identify as Black or African American or Multi-Racial are overrepresented (n=370,434).

Black or African American individuals were 30.9% of the population. White individuals are underrepresented in Data Trust data in comparison to the Mecklenburg County population, as are Hispanic or Latino individuals. There is a slightly higher percentage of people who identify as American Indian or Alaska Native and Native Hawaiian or Pacific Islander than the Mecklenburg County population. There is also a slightly lower percentage of people who identify as multi-racial. This may be due to the fact that, in the Data Trust data, we included a count for all racial categories a person identified within any dataset. We may be capturing more detail than the US Census is able to report.

Among people who are in more than one dataset (n=288,626), the overrepresentation is even more pronounced (see Figure 8). People who identify as Black or African American, and who are in more than one dataset, represent 60.9% (n=175,669) of the population compared to 48.2% (n=369,946) of the overall population of people in the Data Trust. Similarly, the percentage of people who identify with two or more races, and are in more than one dataset, is double the percentage of the overall population of people in the Data Trust (20.8%; n=60,071 and 9.3%; n=71,145, respectively). All other racial and ethnic identities of people who are in more than one dataset represent smaller percentages than in the overall population of the Data Trust.

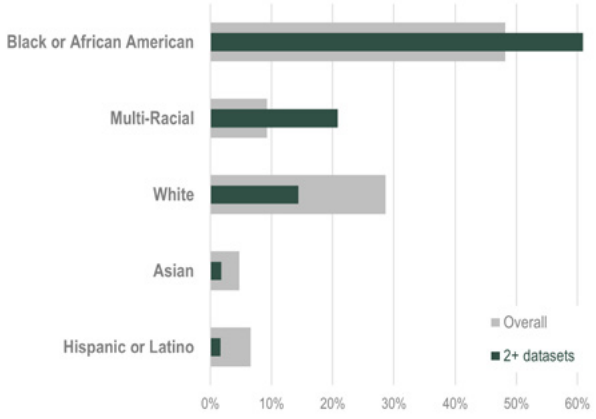


FIGURE 8 – For people who were in more than one dataset, the representation of people who are Black or African American or Multi-Racial is even more pronounced (Overall n=766,679; 2+ datasets n=288,626).

Understanding Data Overrepresentation and Underrepresentation

There are a number of possible reasons for overrepresentation and underrepresentation in Data Trust data. One reason may be the implications of eligibility criteria. Many of the datasets at the Data Trust are associated with economic hardship. For example, people who apply for Mecklenburg County Department of Social Services – Economic Services (our largest dataset) are seeking financial support. Undocumented individuals, however, are not eligible for many government programs and this may deter mixed-status households (households that include both documented and undocumented immigrants) from enrolling in services (e.g., Bovell-Ammon et al., 2019; Vargas & Pirog, 2016). This could lead to underrepresentation of Hispanic or Latino people in the Data Trust data despite their size and growth in the broader Mecklenburg County population.

While eligibility criteria may account for the above underrepresentation, explanations for disproportionate representation in administrative data may be less straightforward and more systemic in nature. For example, historic policies and practices like redlining and subsidized housing support explicitly restricted equal access to housing and homeownership, and has been linked to a negative impact on wealth generation for families of color (e.g., Scott, 2024). These early and subsequent policies have contributed to the wealth gap between White and Black households impacting who has household or network wealth to prevent or delay use of public safety net programs, partly explaining why people who identify as Black or African American are overrepresented in the Data Trust data.

It is also important to consider who is more likely to be surveilled in our society and for what reasons. Even in an age of surveillance through our phones and wearables, we are more likely to collect sensitive data during people's most difficult times and disproportionately so from families of color and families who are poor (e.g., Eubanks, 2018; Fong, 2020). Understanding the extent and implications of disproportionate representation in our data can help us make ethical decisions about data use and guard against research that furthers harm.

One benefit of an integrated administrative data system is that it allows us to combine racial identities across datasets, so where one partner has used a multi-racial category and another has allowed a person to list all their racial or ethnic identities, we begin to see the complexity reflected in a seemingly simple category.

Here, we describe the racial or ethnic category for each person whose administrative data is in the Data Trust. The charts combine all racial and ethnic identities for every unique person, across all datasets from 2016-2023. The visualizations describe how major racial and ethnic categories are linked to an array of diverse racial and ethnic identities. Figure 9, called a chord diagram, shows the racial and ethnic identities for all unique people represented in the Data Trust, and the makeup of multi-racial and ethnic identities.

How to read the chord diagram: Each segment of the chart displays a racial or ethnic category. The size of each segment shows how large that racial or ethnic group is based on how many unique people identify with that category. The different colored lines in each segment link to the additional categories people identified. Thicker lines represent more people. For example, under the Hispanic or Latino segment, about half of the people who identified as Hispanic or Latino also identified as White. And under the Black segment, a small group also identified as Hispanic or Latino.

Black or African-American is the largest group in the Data Trust's integrated data, followed by White, Hispanic or Latino, and then Asian. The chord diagram demonstrates, however, that these categories are not mutually exclusive and that diversity is not necessarily captured in our typical categories of race and ethnicity. The Sankey diagrams below (Figures 10-16) provide a more detailed picture of each racial or and ethnic identity and the extent to which other identities are part of each category in Data Trust Data.

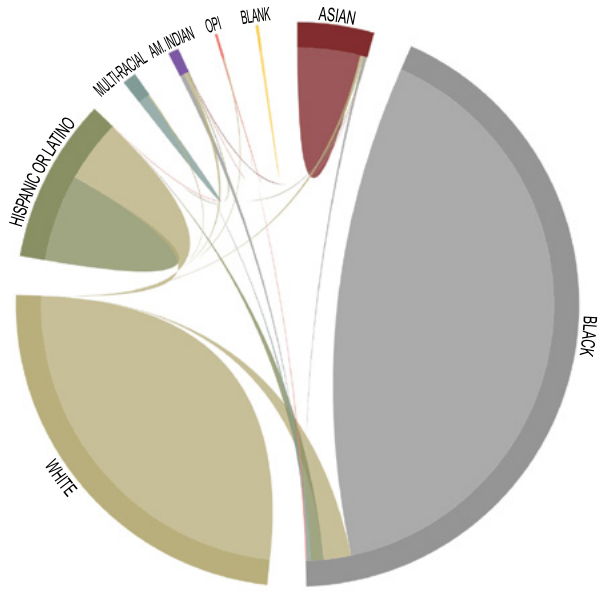


FIGURE 9 – The racial and ethnic identities represented by the people in the Data Trust are complex and multifaceted (n=767,325).

How to read the Sankey diagrams: Each chart displays a racial or ethnic category. The left side of the chart is the main racial or ethnic category, and the right side shows which other races are linked to that main racial or ethnic category. Thicker lines indicate more people. Small percentages are not shown. See Appendix A for data tables.

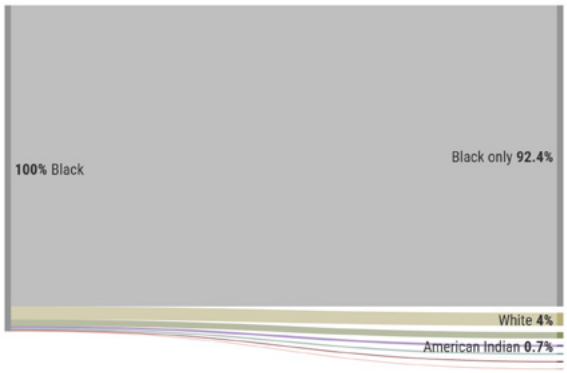


FIGURE 10 – Sankey of Black or African American (n=400,472).

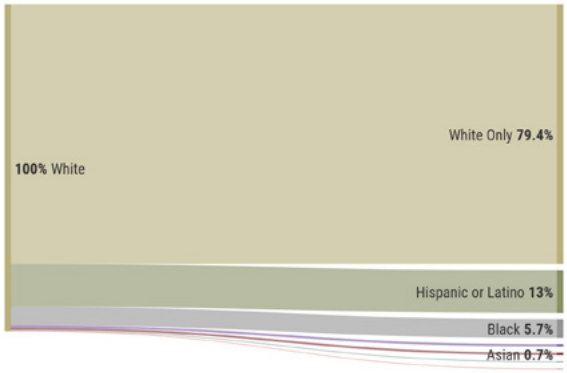


FIGURE 11 – Sankey of White (n=277,116).

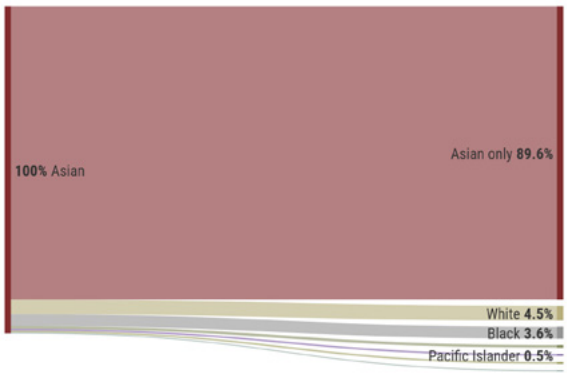


FIGURE 12 – Sankey of Asian (n=40,754).

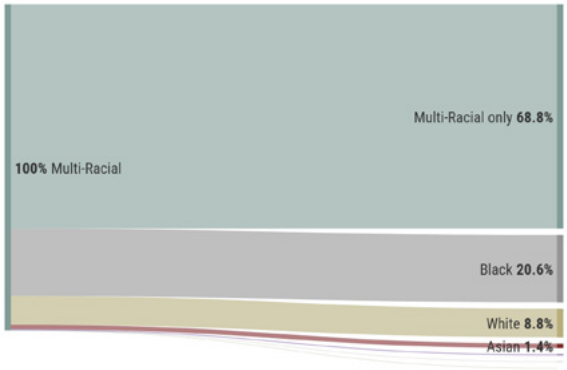


FIGURE 13 – Sankey of Multi-Racial (n=9,573).

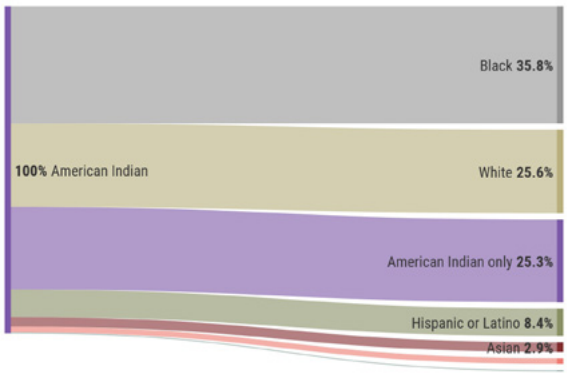


FIGURE 14 – Sankey of American Indian or Alaska Native (n=7,533).

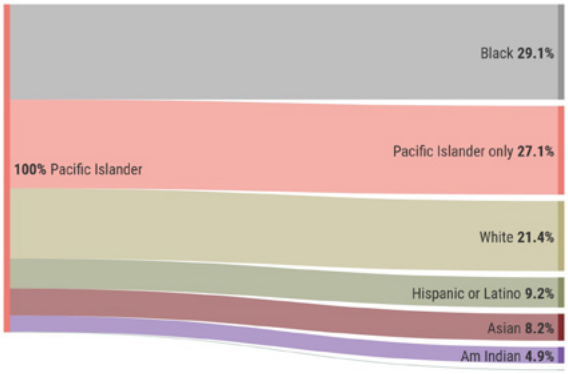


FIGURE 15 – Sankey of Native Hawaiian or Other Pacific Islander (n=2,699).

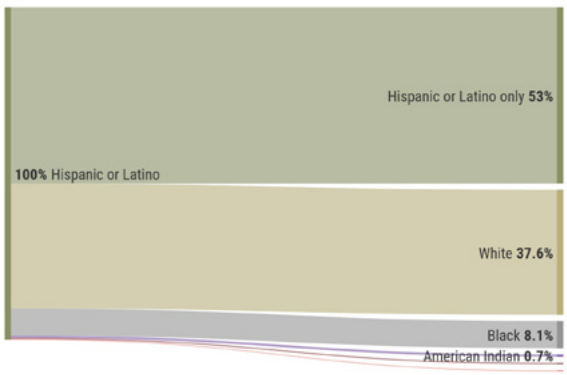


FIGURE 16 – Sankey of Hispanic or Latino (n=2,699).

The integration of administrative data allows us to see a more nuanced description of race and ethnicity among our partners. Both the chord and Sankey diagrams remind us of the complexity of the categories we often default to and particularly challenge our framing of issues around simple racial binaries like Black or White. The data also suggest care around the use of catch all categories like "multi-racial." More information on methods regarding the above chord and Sankey diagrams is available in Appendix A.

Gender

In 2023, people who identified as female (52.5%, n=194,327) made up a slightly larger share of people represented in the Data Trust than those who identified as male (46.8%, n=173,486, see Figure 17). This is representative of the population in Mecklenburg County. In 2023, females represented 51.7 % of the population and males made up 48.3 % (ACS 1-Year Demographic and Housing Estimates, 2023. Table DP05).

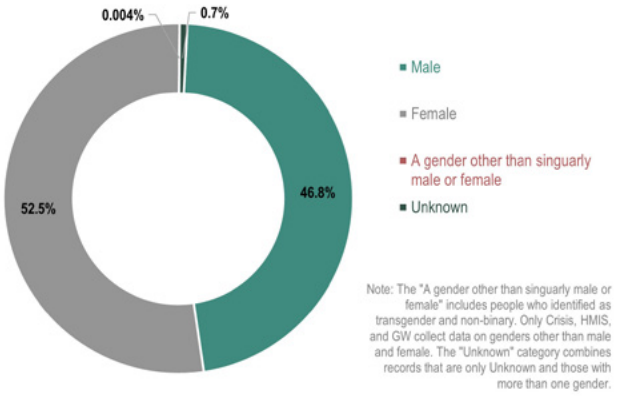


FIGURE 17 – The genders of people in the Data Trust are representative of Mecklenburg County. A very small number of people identify with a gender other than singularly male or female (n=370,434).

A small number of people represented in the Data Trust (0.004%, n=16) identified as a gender other than singularly male or female, including transgender and non-binary. Only some data partners collect data on genders other than male or female: Charlotte-Mecklenburg Homeless Management Information System (since 2014), Crisis Assistance Ministry (since 2012), and Goodwill Industries of the Southern Piedmont (since 2019). This means that we are likely undercounting the number of people who identify with genders other than male or female. The US Census does not collect gender data other than male or female. Since 2016, the share of females, males, and people identifying as genders other than male and female represented in the Data Trust has remained stable.

Age

In the Data Trust data, people are represented across different phases of the life course – from infancy to old age. Historically, the Data Trust focused primarily on children and families, but subsequent data partnerships extended to adults. For some data partners, like Mecklenburg County Department of Social Services – Abuse and Neglect Investigations and Charlotte-Mecklenburg Schools (CMS), only children are represented; while some data partners, like Goodwill Industries and Mecklenburg County Sheriff's Office, only collect information about adults. Other data partners, like the Homeless Management Information System and Crisis Assistance Ministry, collect information about entire households, so both children and adults are represented.

Figure 18 displays the median age and age ranges for all data partners between 2016 and 2023 as of their most recent entry. Age was calculated as of the most recent entry. Except for Habitat for Humanity's Critical Home Repair, the median age across all other datasets is under 40 years.

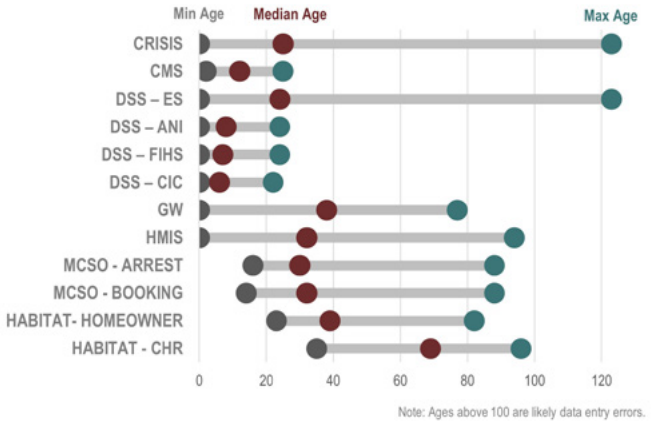


FIGURE 18 – People across the life cycle are represented in the Data Trust.

Where do people in the Data Trust data live?

Address information in some partner datasets allow us to understand where people who are included in Data Trust data live across Mecklenburg County. Of the 370,434 unique individuals in Data Trust data in 2023, 87.8% (n=325,367) had one or more available addresses. The map below (Figure 19) describes the neighborhood profile areas that correspond with the address of those in three Data Trust datasets that also include addresses: Charlotte-Mecklenburg Schools (CMS), Mecklenburg County DSS – Economic Services (DSS-ES), and Crisis Assistance Ministry (Crisis) . The darker the red, the higher the percentage of people represented in the Data Trust.

As the map portrays, individuals whose data are in the Data Trust are present in all of our NPAs. But as the shading suggests, there are areas with a greater concentration

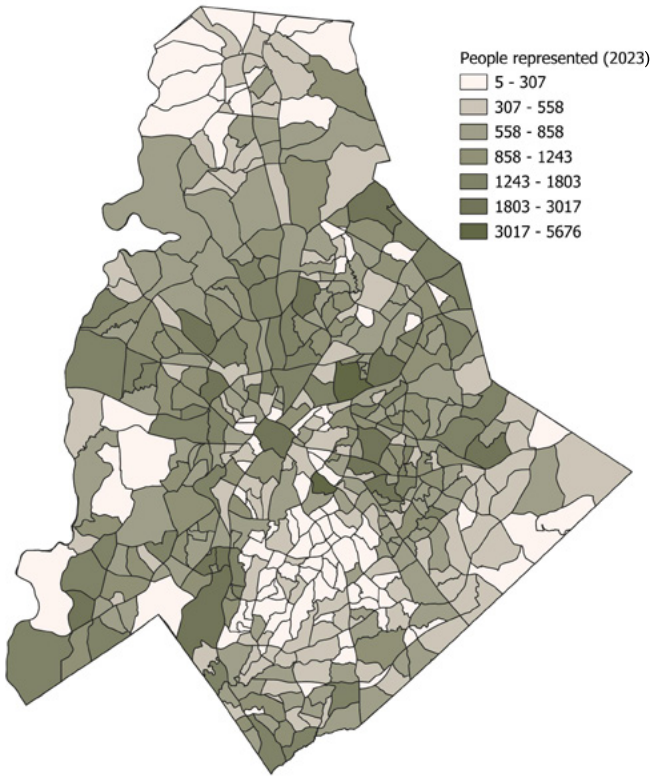


FIGURE 19 – People in the Data Trust are represented in every NPA, with the lowest numbers in south and southeast parts of the county.

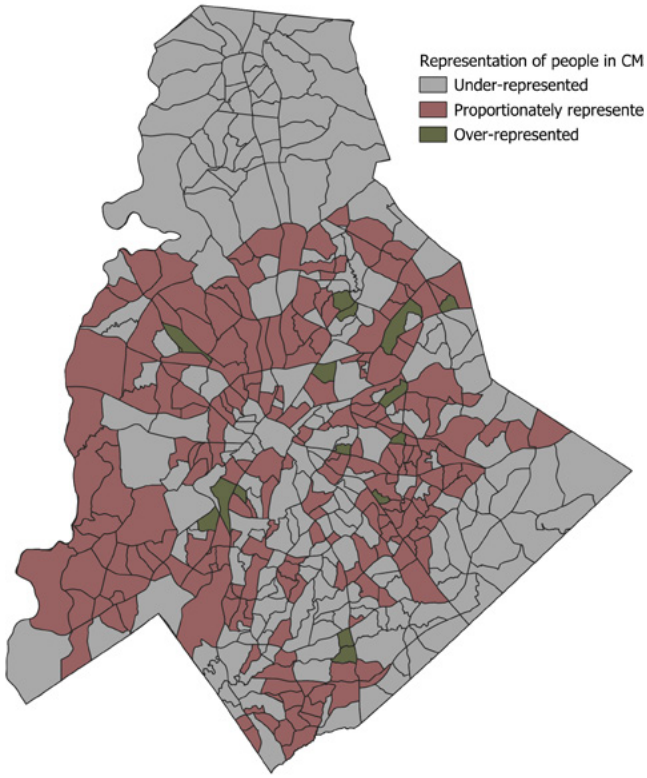


FIGURE 20 – Representation of where CMS students live is fairly evenly spread across the county.

of people from Data Trust data. The maps below help us better understand those areas of concentration. Each map describes a partner dataset compared to the overall share of the Mecklenburg County population in each of the neighborhoods.

For example, if an NPA has 5% of Mecklenburg County's population and also has 5% of the total people in one of the datasets, we estimate that they are proportionally represented. But if a neighborhood has 1% of the County's population but 5% of the total people in one of the datasets, we estimate that they are disproportionately overrepresented. And if a neighborhood has 5% of the population but 1% of the dataset, we estimate that they are disproportionately underrepresented. Additional information on our estimation method is available in Appendix C.

Figure 20 is a map of the CMS dataset for 2023. In the red NPAs, the number of students in the CMS dataset is approximately the same as that of the neighborhood. About 47% (n=219) of the total NPAs are proportionally represented in the data. In the gray areas, CMS students are underrepresented and in the green areas, students are overrepresented. Enrollment in charter schools, private schools and homeschooled students may explain the disproportionate representation.

In the maps for the Crisis dataset (Figure 21) and DSS-ES (Figure 22), only around 20% (n=91) and 22% (n=103) of the total NPAs are proportionately represented, respectively. In both maps, there is a defined crescent of green overrepresentation to the west, north, and east of the uptown area, and more pronounced areas of gray underrepresentation than in the CMS dataset.

Comparing the three maps, students in CMS are more evenly distributed across the County, however, patterns of underrepresentation in the north and southeast of the County remain discernable. Both Crisis and DSS-ES datasets show clear patterns of overrepresentation in the crescent NPAs and underrepresentation in the north and southeast, unsurprising given that both Crisis and DSS-ES serve individuals in a financial crisis and the households with the lowest incomes are concentrated in the crescent NPAs and higher incomes are concentrated in the north and southeast of the county. These patterns also correspond to the concentration of racial and ethnic groups in the NPAs, where patterns of overrepresentation have a higher concentration of our county's Black population and areas of underrepresentation have a higher concentration of the county's White population (see Appendix A for the correlations and more detail about the methodology).

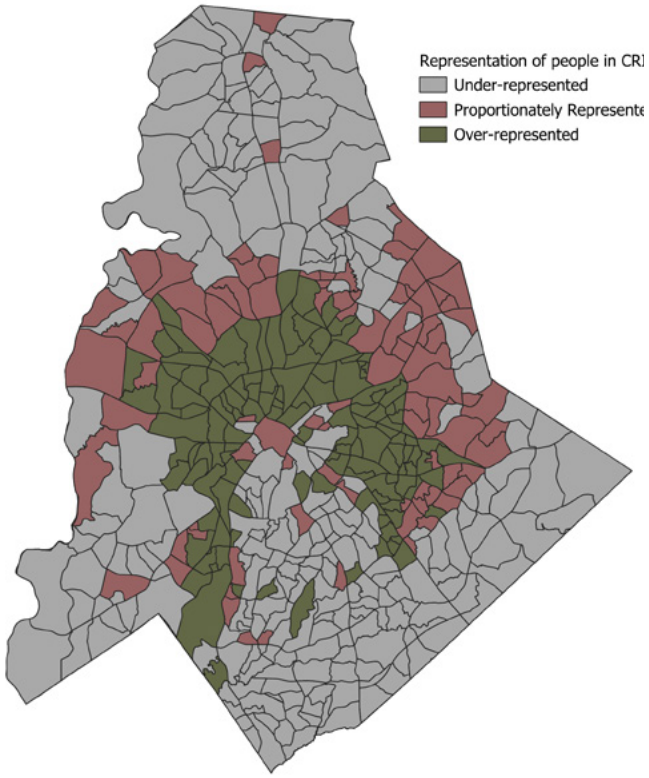


FIGURE 21 – People who used Crisis are more heavily concentrated within City of Charlotte limits and are overrepresented in the "crecent."

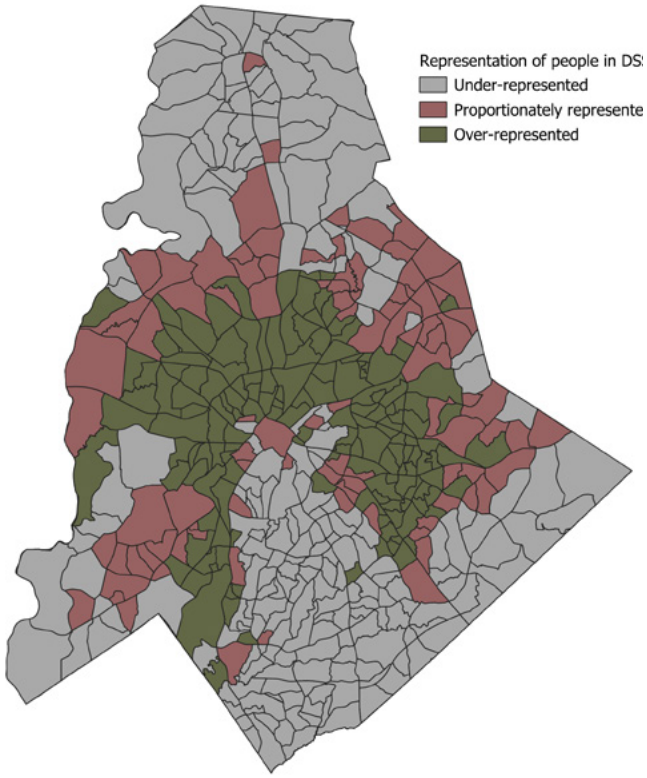


FIGURE 22 – People who used DSS-ES primarily live in the "crecent" and are overrepresented in those areas. Primarily live across Mecklenburg.

What and who is missing in Data Trust data?

The characteristics of the data described above demonstrate the complexity of the data and the people represented in the Data Trust. And yet, significant information is still missing from the available data. Having a better sense of who and what is missing in our data can guide future data partnerships and research focus.

This inaugural report only includes a description of race/ethnicity, gender, and age data. We hope to include characteristics like income, education level, disability, and veteran status in future reports. These characteristics were not included in this report because comparatively few partners collect the information and decisions about standardizing these categories should be informed in collaboration with our data partners.

In addition to the characteristics we didn't include in this report, the available data on the characteristics we did include may not be sufficiently comprehensive. For example, we should assume that we are underestimating the number of people who identify with a gender other than singularly male or female, since only a few partners collect expanded information on gender identity. In addition, the available data is limited to a single ethnicity, Hispanic or Latino. And similarly, we do not have enough information about the country of origin of people in the data. We know that the Charlotte region has a rich and diverse immigrant population, and country-of-origin data would help us better understand these groups.

Since many of our data partners are serving people who are in some form of economic crisis or a crisis that's closely associated with lower income households (homelessness for instance), the Data Trust is missing data on people with higher incomes, who are likely not to experience some of the same hardships. Among our partners who regularly deposit data, Charlotte Mecklenburg Schools is one of the few that serve a range of income levels. If higher incomes are a desired outcome for our community, including additional partners whose data include individuals across income levels will be important.

Additionally, the data may omit people who do not meet eligibility requirements for programs and services. This includes people who are not eligible for programs based on immigration status or because their income is just above the eligibility threshold. The Data Trust data may, therefore, not capture those in our community who still struggle, but don't qualify for services, including those who could most benefit from interventions that may prevent a person from further hardships.

Finally, we are missing important data partners that could help us better understand who is thriving and who is left behind and why. While many areas such as homelessness and education are well represented in our current data, gaps remain in these and other domains. Arts & culture program data, early childhood data and older adult data, housing instability and evictions data, workforce data, and data from our broader metropolitan region would strengthen this community resource and engage more partners and researchers in its use and application. This annual report can serve as an ongoing assessment of who and what is missing in our data and the success of strategic efforts to create a more comprehensive and representative resource.

How are people reflected across partners?

A key benefit of the Data Trust's administrative data system is the ability to understand the extent to which people served by one partner are also served by other partners. Operationally, it helps organizations understand if there are gaps in their services, if they are duplicating services with another provider, or if their programs are reaching their target audience who may be best represented in another partner's data. From both an operations and research perspective, it is helpful to understand if involvement in one organization may impact outcomes in another. For example, if a student is in the child welfare system, how might that involvement may impact school attendance or performance? Or if a person who is experiencing homelessness is housed, are they less likely to be arrested or use the emergency room?

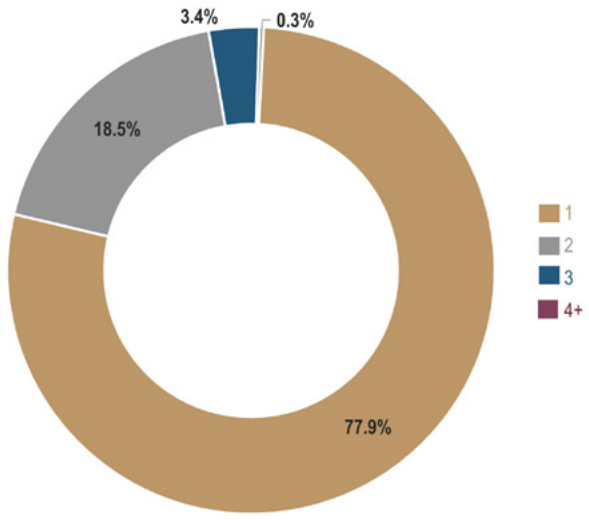


FIGURE 23 – In 2023, most people were in 1 or 2 data partners Only about 4% of people overlapped with 3 or more data partners (n=370,434).

Overlap across partners

In 2023, the majority of individuals whose administrative data were in the Data Trust's integrated data system were reflected in one partner's data (77.9%, n=288,510) and 18.5% (n=72,889) were reflected in two partners' data. Comparatively few individuals were reflected in three or more partners' data with only 3.4% (n=14,928) reflected in three, and less than 1% reflected in four or more (0.23%,n=1,013).[4]

Cumulatively, since 2016, 65.3% (n=500,847) of individuals whose administrative data is in the Data Trust were in one partner's data and 21.8% (n=167,519) were reflected in two partners' data. Importantly, the nearly 13% who are in three or more partners' data number nearly 100,000 people (12.9%, n=98,959). About 4% (3.8%, n=29,180) were reflected in four or more partners' data.

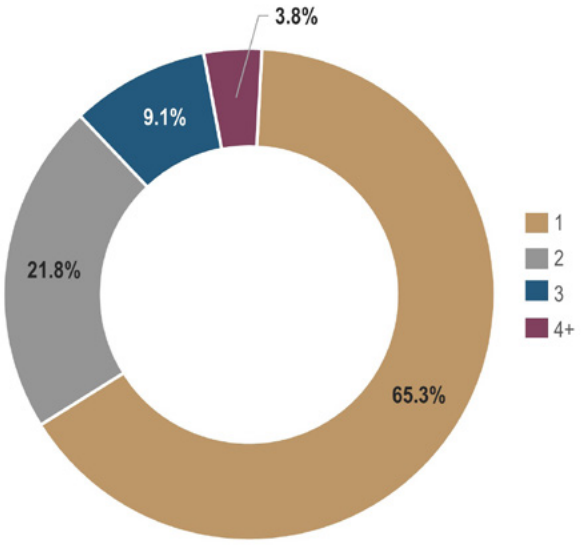


FIGURE 24 – Since 2016, roughly 1/5 of people in the Data Trust were in 2 data partners. Nearly 100,000 people overlapped with 3 or more data partners (n=767,325).

Both annually and cumulatively since 2016, the majority of individuals in the data were only in one partner's data, with only a small minority reflected in three or more partners. This distribution represents a rich area for further research and understanding. For example, what factors are associated with individuals' involvement with multiple organizations or with only one?

Notably, over time, we expect to see greater overlap between partners than we would see annually because individuals may be reflected in multiple services as they age. For example, a child or adolescent receiving child welfare services may be reflected as an adult in Goodwill or Charlotte-Mecklenburg Homeless Management Information System (HMIS). These longitudinal patterns are also important patterns for further research. What factors in childhood predict whether the individual is more or less likely to become homeless as an adult? Administrative data along with Explorer data can be used to examine these and related questions.

Overlap across datasets

Since several partners have more than one distinct dataset, it is also helpful to understand the extent to which each dataset overlaps with other datasets. Figure 25 visualizes the overlap across all datasets from 2016 through 2023 and includes only the individuals that overlap across two or more datasets (n=288,626). Again, thicker lines represent more people. The figure demonstrates the variety of ways that people are linked across datasets, presenting opportunities to understand the nature of the overlaps and more comprehensively develop and consider solutions that span organizations.

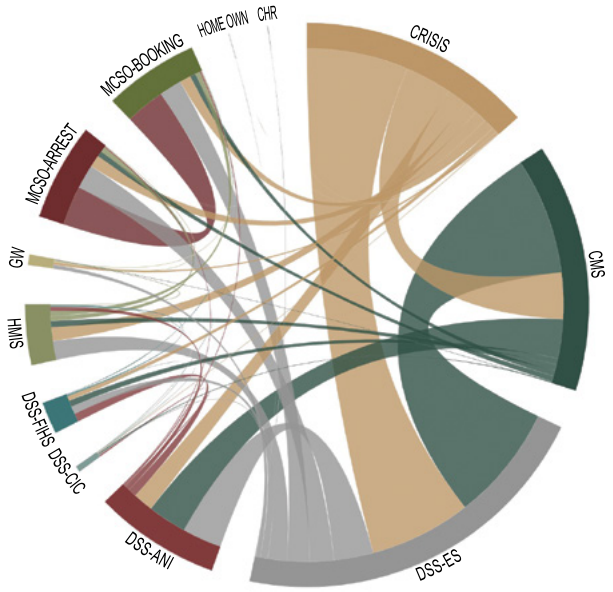


FIGURE 25 – Every data partner serves people who are also served by other partners (n=288,626).

The figures on the next page describe, by partner, the extent to which the individuals in each partner dataset were reflected in that of other partners between 2016 and 2023. The figures, called tree maps, visualize the percentage of individuals in one dataset that overlap with the others. The title of each tree map indicates the source dataset that is described. For example, the first tree map shows how people in the HMIS dataset overlapped with other partner datasets – 63% of people in HMIS were also in Mecklenburg County Department of Social Services – Economic Services, 45.3% were also in Crisis Assistance Ministry's dataset, and 20.% were also in Charlotte-Mecklenburg School's dataset.

[4]Note, for the purposes of the analyses visualized in Figures 23 and 24, we combined the three child welfare datasets (DSS-ANI, DSS-FIHS, and DSS-CIC) and two MCSO datasets (MCSO-Arrest and MCSO Booking), since the datasets within child welfare and criminal justice reflect many of the same people. For figures 25-37, we separated the three child welfare and two criminal justice databases since the distinctions between the datasets or the programs they represent may be meaningful to partners.

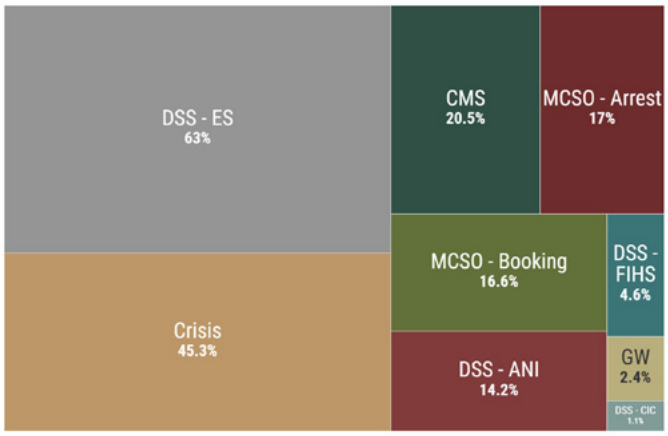


FIGURE 26 – HMIS overlap with other datasets (n=36,089).

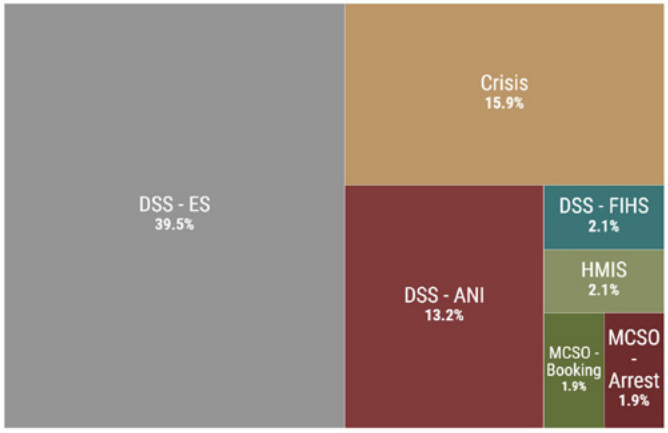


FIGURE 27 – CMS overlap with other datasets (n=355,314).

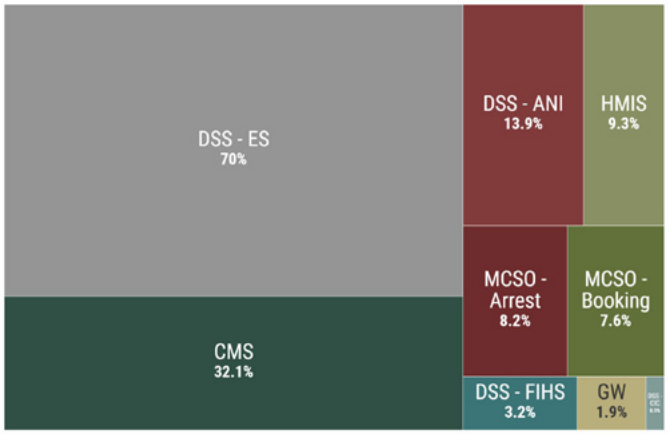


FIGURE 28 – Crisis overlap with other datasets (n=175,434).

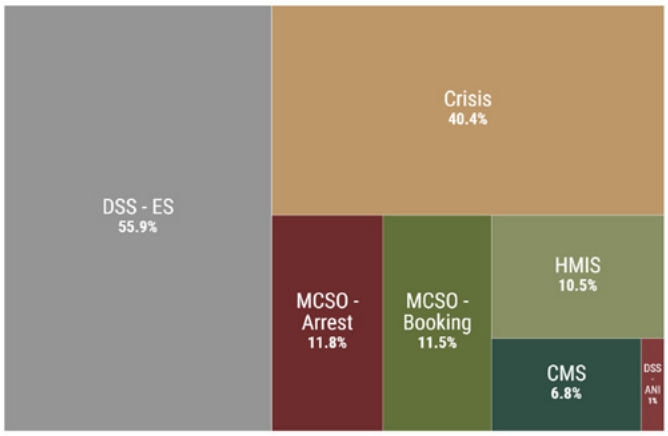


FIGURE 29 – GW overlap with other datasets (n=8,352).

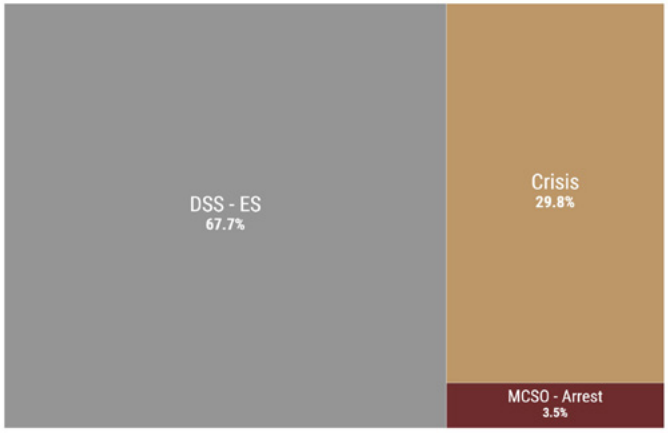


FIGURE 30 – Habitat – Homeowner overlap with other datasets (n=282).

FIGURE 31 – Habitat – CHR overlap with other datasets (n=515).

The figures above allow us to see how the systems and services in our community are linked by the people they serve. Every data partner serves people who are also served by other partners. Further analysis to understand these connections allows us not only to break down data silos, but better understand, coordinate, and improve service delivery.

The absence of expected overlap can also be telling. In one of the first integrated data reports on unhoused students completed in partnership with Mecklenburg County Community Support Services in 2017, the overlap of data showed that 38% of the students in emergency shelter were not enrolled in the McKinney-Vento programs at Charlotte Mecklenburg Schools (CMS) designed to support students who were unhoused or housing insecure. Seeing

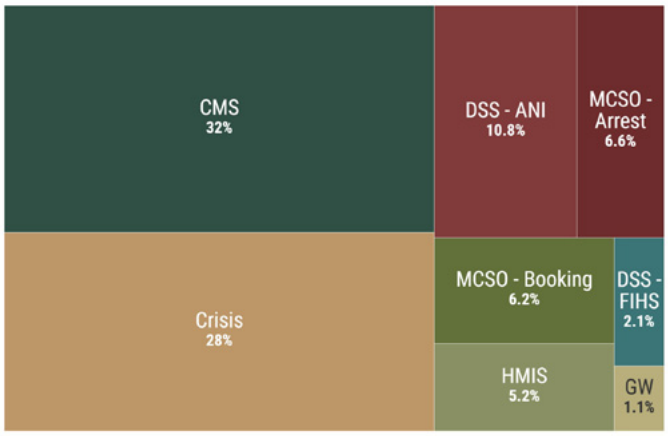


FIGURE 32 – DSS – ES overlap with other datasets (n=438,886).

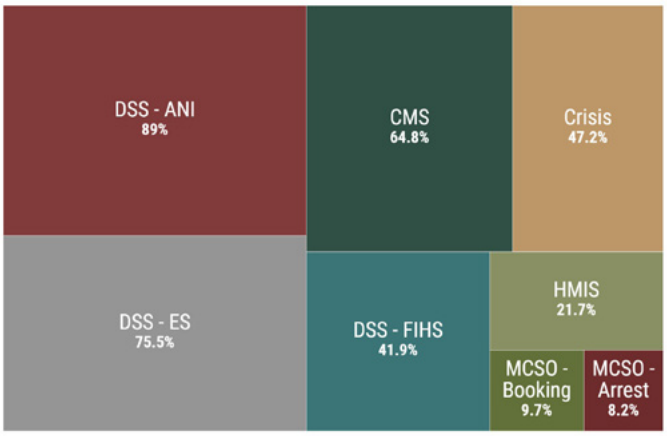


FIGURE 35 – DSS – CIC overlap with other datasets (n=1,881).

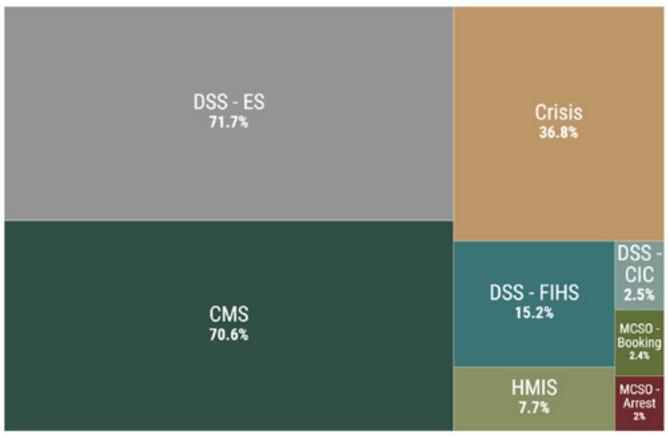


FIGURE 33 – DSS – ANI overlap with other datasets (n=66,407).

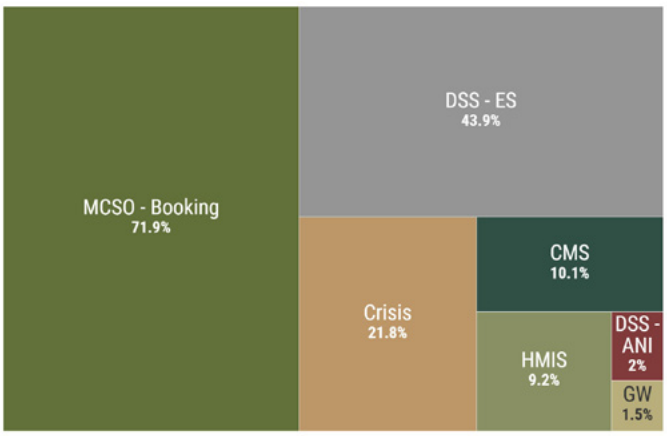


FIGURE 36 – MCSO – Arrest overlap with other datasets (n=66,317).

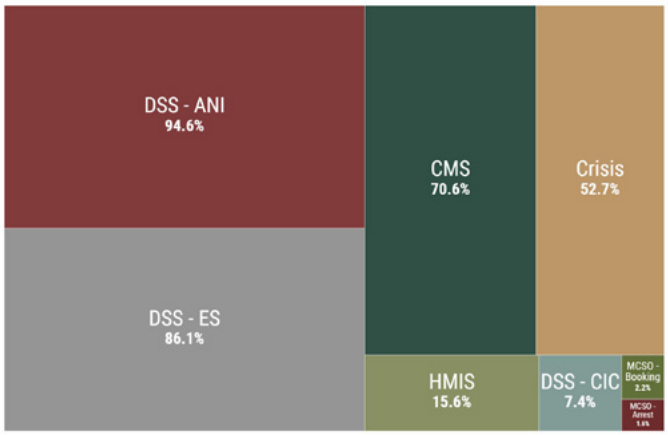


FIGURE 34 – DSS – FIHS overlap with other datasets (n=10,649).

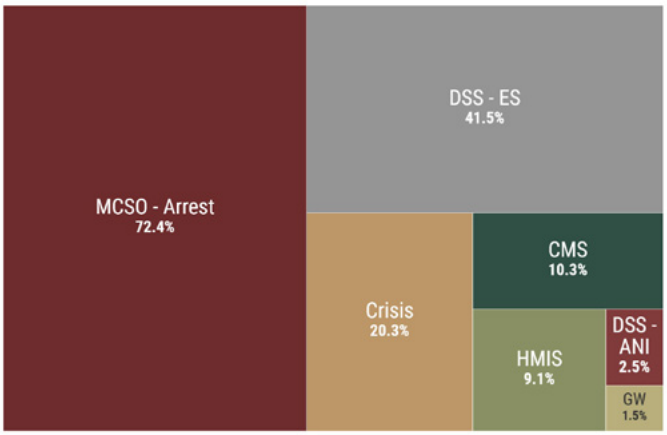


FIGURE 37 – MCSO – Booking overlap with other datasets.

this unexpected gap in program enrollment allowed the Salvation Army Center of Hope and CMS to work together to ensure that all students residing in emergency shelter were also enrolled in the services available to them at CMS. The various ways data overlap or fail to do so are rich opportunities for additional research.

USE AND IMPACT

The Institute + Data Trust maintains a community data infrastructure so that the data can be used for community benefit, specifically through research, evaluation and planning. As a resource to both the community and the University, we engage researchers who are based in our local government and nonprofits, as well as academic researchers doing community-engaged research and advancing knowledge in the nine domains we address. This section summarizes the ways in which Data Trust data were used in 2023 and through October 31, 2024. Additional details are available in Appendix B.

Data License Requests

People interested in using linked data from the Data Trust must submit a Data License Request (DLR). DLRs are reviewed and decided upon by the Data and Research Oversight Committee (DAROC), comprised of data partner representatives. Note: DLRs may be submitted during one calendar year and approved in the next calendar year and it may take several years for a project to be completed.

In 2023, a total of eight DLRs were submitted and approved, the highest number submitted and approved in five years. The project names of these DLRs are below. More information can be found in Appendix B:

- **Quality of Life Explorer 2023 Update** – Project Lead: Katie Zager, UNC Charlotte Urban Institute
- **System contact and exposure to firearm related violence in Mecklenburg County** – Project Lead: Dr. Jennifer Langhinrichsen- Rohling, UNC Charlotte, Department of Psychological Science
- **Understanding educational outcomes of Habitat Charlotte Region families** – Project Lead: Dr. Eric Moore, UNC Charlotte Urban Institute
- **Understanding proportionality and representation in Data Trust datasets** – Project Lead: Sydney Idzikowski, UNC Charlotte Urban Institute
- **Examining educational outcomes for children involved in foster care** – Project Lead: Dr. Robert Pinka, Council for Children's Rights
- **Integrating healthcare and social service data to evaluate the effect of COVID-19 on cancer disparities for underserved patients** – Project Lead: Dr. Carlene Mayfield, Advocate National Center for Health Equity, part of Atrium Health
- **Student Emergency Fund** – An exploratory analysis of community connections – Project Lead: Dr. Ellissa Brooks Nelson, UNC Charlotte Student Affairs
- **Transition age foster care youth and housing instability** – Project Lead: Bridget Anderson, UNC Charlotte Urban Institute

Much of the research conducted through these requests is longitudinal. Through the studies, researchers seek to understand how experiences with systems over time impact outcomes from gun violence to homelessness. Two of the DLRs take a deep dive into the educational and housing outcomes, respectively, of children involved with foster care in Mecklenburg County with a goal to inform policy, practice, and collaboration to better serve this population. Other studies shed light on what COVID-19 meant for the care of cancer patients and tell us more about students who applied for financial assistance during and after the pandemic.

The majority of the DLRs (n=5) requested five or more datasets. One DLR requested two datasets, and two DLRs requested three datasets. Figure 38 shows how many times a dataset was requested. Charlotte-Mecklenburg Schools remains the most requested dataset. In 2023, child welfare data from DSS-Abuse and Neglect Investigations, DSS – Children in Custody, and DSS – Economic Services were also requested in the majority of DLRs. Two DLRs linked a dataset that is external to the Data Trust. The majority of the DLRs in 2023 were from UNC Charlotte faculty and staff (see Figure 39). We also received requests from Atrium Health and Council for Children's Rights.

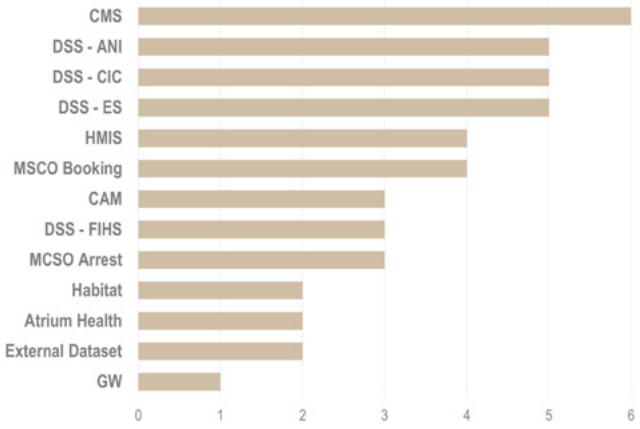


FIGURE 38 – CMS, DSS – ANI, & DSS – ES were the most requested datasets in 2023.

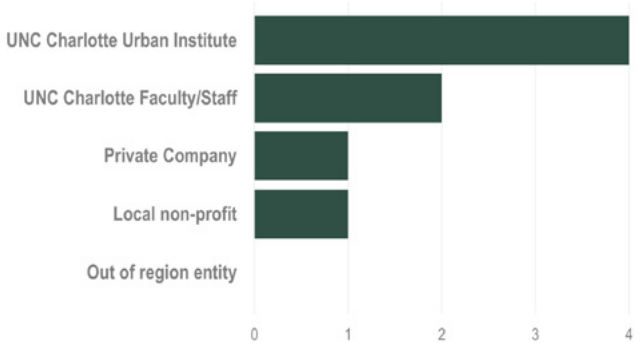


FIGURE 39 – In 2023, most DLRs were from UNC Charlotte faculty and staff.

For 2024, a total of four DLRs have been submitted and approved through October 31. The projects include:

- **Quality of Life Explorer 2023 Update** – Project Lead: Katie Zager, UNC Charlotte Urban Institute
- **Youth data intervention initiative** – Project Lead: Dr. Mikaela Rabinowitx, National Institute on Criminal Justice Reform
- **Integrating healthcare and educational data to evaluate the effectiveness of school-based virtual care and teletherapy on healthcare utilization and academic outcomes** – Project Lead: Dr. Carlene Mayfield, Advocate National Center for Health Equity, part of Atrium Health
- **CAPE: Assessing equity in computer science and computing education in CMS schools** – Project Lead: Dr. Xiaoxia Newton, UNC Charlotte College of Education (and Gambrell Faculty Fellow)

Final Products

Four projects were completed in 2023 and through October 31, 2024. See Appendix B for more information. As available, links to the reports are provided. They were:

- **Incarceration and Homelessness Integrated Data Report (May 2023):** Part of the State of Housing Instability and Homelessness report series, the study examined the ways in which people experiencing homelessness interact with the Mecklenburg County jail. They found that nearly seven percent of people experienced homelessness in the two years before or after incarceration; that people with a prior history of homelessness were 22 times more likely to have another episode of homelessness after their release from incarceration; and Black individuals are overrepresented in both the jail and homelessness services. The full report can be found here.
- **Girl Scouts Program Evaluation (October 2023):** This study helped Girl Scouts Hornets Nest Council better understand the extent to which participation in the Girl Scout Leadership Experience affects the behavioral and academic outcomes of its participants. The results of the analysis indicate that being a member of the Girl Scouts had a positive effect on math proficiency, science proficiency, and school attendance.
- **Quality of Life Updates (March 2023):** The Quality of Life team submits an annual DLR to update data on the Quality of Life Explorer. The variables from the Data Trust used on the Quality of Life Explorer are recipients of Food and Nutrition Services from DSS – Economic Services and EOG/EOC, attendance, neighborhood school attendance, and graduation rate from Charlotte-Mecklenburg Schools. The Quality of Life Explorer can be accessed here.
- **Understanding Educational Outcomes of Habitat Charlotte Region Families (August 2024):** This study was an update to a previous study that examined the educational outcomes of children who live in Habitat homes. The study explored differences in attendance, test proficiency, and suspensions compared to two groups of similar students.

Future reports will include a description of the use of Explorer data, as well as descriptions of how research has catalyzed impact in Charlotte and our region. As this report continues in subsequent years, we expect that reporting our impact will become a key focus of the overall report.

CONCLUSION

The Institute + Data Trust's inaugural State of Our Data report summarizes and describes our data ecosystem, our data, and its use. The report itself is intended to be a tool, focusing not only on what we have completed and developed over the past year, but also on who is missing, what needs to be improved, where there are new opportunities for collaboration, and critically, what is our impact. This inaugural report also reminds us:

Our data are people. Whether describing people or the places they live, the data we hold, curate, and use for research and decision-making reflect human lives. This brings added responsibility, both in terms of protecting the privacy of data, but also in using it so it benefits the people reflected.

Simple categories hold immense complexity. The categories that we often take for granted to describe who we and others are hold a great deal of nuance and complexity. Categories we use to describe characteristics like race and ethnicity help us standardize and compare data, but they can also oversimplify or erase important aspects of people's identities and experiences. A community data infrastructure minds this tension and a resource like the Data Trust's integrated data system has the capacity to help us better understand and recognize the level of detail and the people behind the categories.

Understanding representation is essential. When compared to Mecklenburg County, people who identify as Black or African American, or who are captured in multi-racial categories are overrepresented in Data Trust data. People in our datasets are also disproportionately located in lower income "crescent" neighborhoods in Charlotte-Mecklenburg. Understanding disproportionate representation matters for research methods, analysis, and interpretation. It also matters strategically for the Data Trust, particularly in the recruitment of new data partners, as well as board and committee members.

Understanding disproportionate overrepresentation in Data Trust data also points to the importance of understanding data in context and linking administrative data with other forms of data, like broader population data from the Quality of Life Explorer as well as more qualitative, contextual, and historical data to build a fuller picture of the people in our community and their experiences. If we relied only on integrated administrative data, much of it collected during very difficult points in people's lives, we would only count the ways something is going wrong in people's lives. There is a bigger story that additional forms of data allow us to tell.

CONCLUSION

Organizations and services in our community are connected by the people they serve. Whether intentional or not, services in our community are linked by the people receiving them. Seeing those connections across programs and partners allows us to ask questions about them to better coordinate services, observe gaps, and better serve those in our community that are engaged in multiple organizations and systems.

A community data infrastructure requires ongoing commitment. The report details the extent of our partnerships and the wide variety of data required to develop and maintain an enduring community data infrastructure. Accessible people and place data across multiple, interconnected domains like housing, health, and transportation requires more than one leader or one organization or one sector. The community infrastructure stewarded by the Institute + Data Trust and our partners is a community effort and a community commitment.

This year, the State of Our Data report is a beginning. Subsequent annual reports will allow us to reflect on the further development of our community data ecosystem, the people represented in our partner data, and the ways our data are used for their benefit and the benefit of our communities.

We invite you to join us as a partner in this work – *use the QR Code to let us know if you are interested in becoming a data partner or learning more about using our community data.*

REFERENCES

Bovell-Ammon, A., Ettinger de Cuba, S., Coleman, S., Ahmad, N., Black, M. M., Frank, D. A., ... & Cutts, D. B. (2019). Trends in food insecurity and SNAP participation among immigrant families of US-born young children. *Children, 6*(4), 55.

Brady, H. E., & Kent, T. B. (2022). Fifty years of declining confidence & increasing polarization in trust in American institutions. *Daedalus*, *151*(4), 43-66.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor.* St. Martin's Press.

Fong, K. (2020). Getting eyes in the home: Child protective services investigations and state surveillance of family life. *American Sociological Review, 85*(4), 610-638.

Health Resources & Services Administration (HRSA). (n.d.). *Find shortage areas*. U.S. Department of Health & Human Services. https://data.hrsa.gov/tools/shortage-area.

Hendriks, F., Kienhues, D., & Bromme, R. (2016). Trust in science and the science of trust. *Trust and communication in a digitized world: Models and concepts of trust research*, 143-159.

Hoyt, H. (2000). *One hundred years of land values in Chicago: The relationship of the growth of Chicago to the rise of its land values, 1830-1933.* Beard Books.

Kauh TJ, Read JG, Scheitler AJ. The Critical Role of Racial/Ethnic Data Disaggregation for Health Equity. *Popul Res Policy Rev. 2021;40*(1):1-7. doi: 10.1007/s11113-020-09631-6. Epub 2021 Jan 8. PMID: 33437108; PMCID: PMC7791160.

Krieger, N., Wright, E., Chen, J. T., Waterman, P. D., Huntley, E. R., & Arcaya, M. (2020). Cancer stage at diagnosis, historical redlining, and current neighborhood characteristics: Breast, cervical, lung, and colorectal cancers, Massachusetts, 2001–2015. *American Journal of Epidemiology, 189*(10), 1065-1075.

Robert Wood Johnson Foundation. (2015). *Data for Health: Learning What Works - A report from the Data for Health Advisory Committee.* Virginia Tech. https://www.rwjf.org/content/rwjf-web/us/en/about-rwjf/newsroom/2015/04/report-highlights-publics-hopes-fears-about-using-data-to-improve-health

Rothstein, R. (2017). *The color of law: A forgotten history of how our government segregated America.* Liveright Publishing.

Saad, L. (2023) Historically Low Faith in U.S. Institutions Continues, *Gallup News*, July 6, 2023, https://news.gallup.com/poll/508169/historically-low-faith-institutions-continues.aspx.

Scott, R. H. (2024). The Evolution of Redlining in the United States Housing Market. *Journal of Economic Issues, 58*(2), 588–597. https://doi.org/10.1080/00213624.2024.2344442

Vargas, E. D., & Pirog, M. A. (2016). Mixed status families and WIC uptake: The effects of risk of deportation on program use. *Social science quarterly, 97*(3), 555-572.

URBAN INSTITUTE
REGIONAL DATA TRUST

ui.charlotte.edu / 704.687.1210
urbaninstitute@charlotte.edu

REPORT METHODS & DATA TABLES

Key Methodological Notes

The data included in this report is from an approved data license request. The Data Trust team submitted a data license request that was reviewed and approved by the Data and Research Oversight Committee. Data from the request will also be used to inform recruitment to the Community Data Advisory Committee, which led to the data license request.

The years referenced in this report are calendar years. Many of our data partners share data based on the calendar year.

Data Generation Process:

The first part of the data generation process was to identify individuals represented in the Data Trust and to create a unique id for every person. A person could have multiple data partner ids so the deduplication process assigns IDs related to the same person to a globally unique ID. To accomplish this we used a similar process as the record linkage process described below. The normal record linkage process is used to identify the same person in two different datasets, while with deduplication we searched for common people within the records of the same dataset. Performing the deduplication process allowed us to get a more accurate number of the amount of unique people in the dataset.

Record Linkage Process:

The record linkage process involved connecting individuals present in different datasets. Since a common unique identifier, such as a Social Security number or email address, was not provided, the linkage primarily relied on names and dates of birth to identify records representing the same person. Exact name matching proved challenging due to variations such as typographical errors, shortened versions of first names (e.g., Sam vs. Samuel), or hyphenated last names where one dataset included only part of the last name while another had the full hyphenated version.

Names, dates of birth, race, and gender were collected from datasets meeting the specified criteria. Names were cleaned by converting them to lowercase, removing whitespace and special characters, and separating multi-part last names into individual records. For example, a record with the last name "Smith-Jones" was cleaned into "smith jones" and split into three records: one with "smith," one with "jones," and one retaining "smith jones." Additionally, Soundex codes—a phonetic encoding method—were generated for all names. These codes facilitated the matching process by accounting for slight spelling differences, as they could produce the same code for similar-sounding names.

Race and gender categories were also standardized to address inconsistencies across data partners. For instance, one partner might categorize gender as "Male" and "Female," while another used "M" and "F." These categories were mapped to a common naming convention for consistent comparison. If individuals had multiple races, each race was treated as a separate record, similar to the approach for multi-part last names. Consequently, a single person could be associated with multiple records. A unique key was generated for the cleaned dataset.

Once the data was cleaned, the next step involved identifying potential matching candidates. For large datasets, comparing all possible record combinations would have been computationally expensive due to the sheer volume. To narrow down the list of potential matches, subsets of fields were used: combinations of birth date with the Soundex code for first names and birth date with the Soundex code for last names. Birth dates had to match exactly, while the Soundex codes employed a Sorted Neighborhood Index to identify the closest matches, even when minor variations existed.

The resulting candidate list contained pairs of unique keys from both datasets to compare additional fields. These fields included first, middle, and last names, which were evaluated using a string similarity method that generated a score. Matches were identified if the score exceeded a set threshold. Soundex codes for names, as well as birth date, race, and gender fields, required exact matches. Since birth dates were already exact matches from the candidate selection step, all potential matches shared this attribute. The standardization of race and gender fields ensured accurate comparisons.

After comparing records, a table of binary fields was generated, with columns indicating whether each field matched (1) or did not match (0). A weighted match score was calculated, prioritizing name fields, and ranged from 0 to 1, with 1 representing a perfect match. The table was also analyzed using a classification model to categorize records as either "No Match" (0) or "Match" (1). Final matches were determined by a match score of 0.85 or higher and a classification as "Match."

In the matched dataset, checks were performed to ensure that an ID from one dataset did not match multiple IDs from the other. In cases of multiple matches, the record with the highest match score was retained, and others were excluded. If scores were identical, the records were treated as representing the same individual and assigned the same ISC_ID. The final dataset included matches and non-matches, with an ISC_ID assigned to each unique person. Matched individuals shared a common ISC_ID across both datasets. This final dataset was saved to a file and linked to subsequent outputs.

Standing Data Partner Names, Abbreviations, and Colors

NAME	ABBREV. & COLOR
Charlotte-Mecklenburg Homeless Management Information System (34 partner organizations)	HMIS
Charlotte-Mecklenburg Schools	CMS
Crisis Assistance Ministry	Crisis
Goodwill Industries of the Southern Piedmont	GW
Habitat for Humanity of the Charlotte Region – Homeowner	Habitat – Homeowner
Habitat for Humanity of the Charlotte Region – Critical Home Repair	Habitat – CHR
Mecklenburg County Department of Social Services – Economic Services	DSS – ES
Mecklenburg County Department of Social Services – Abuse and Neglect Investigations	DSS – ANI
Mecklenburg County Department of Social Services – Family In Home Services	DSS – FIHS
Mecklenburg County Department of Social Services – Children in Custody	DSS – CIC
Mecklenburg County Sheriff's Office – Arrest Processing Center[3]	MCSO – Arrest
Mecklenburg County Sheriff's Office – Booking	MCSO – Booking

Data Partner Legend – Abbreviations and Colors

How many unique people are reflected in our administrative data?

COUNT OF UNIQUE INDIVIDUALS OVERALL

Using the year fields to filter the dataset for records between 2016 and 2023 a dataset is created for each dataset that will contain the ISC_ID and a record for each year they are in the dataset. The datasets are aggregated together into one dataset with a unique record for each ISC_ID and year the person is in the datasets. The data is then grouped by year and a count of every ISC_ID per year is calculated and reported below.

To obtain the Mecklenburg County population, we used the American Community Survey 1-year estimates.

Figure 2

Year	2016	2017	2018	2019	2020	2021	2022	2023
Unique people in Data Trust	376,375	362,319	360,742	349,927	345,752	360,715	384,678	370,434
Mecklenburg County Population	1,011,761	1,034,311	1,054,185	1,074,384	1,095,192	1,101,084	1,115,332	1,163,789
% of Mecklenburg Population	37.2	35.03	34.22	32.57	31.57	32.76	34.49	31.83

COUNT OF UNIQUE INDIVIDUALS BY DATA PARTNER

Using the year field to filter the dataset for records between 2016 and 2023 a dataset is created for each data partner dataset that will contain the ISC_ID and a record for each year they are in the dataset. The data is combined into one main dataset where all data partners are present. The data partner and year fields are used to group the data and get the count of the number of ISC_IDs represented in each data partner dataset for every year between 2016 and 2023. The most recent year of available data is returned in the table below.

Figure 3

Data Partner	Number of unique individuals 2023 (or most recent data available)
Crisis	48,629
DSS – ES	215,787
DSS – ANI	6,877
DSS – FIHS	1,419
DSS – CIC	203
CMS	176,910
GW	3,119
HMIS	9,209
MCSO-Arrests	5,073
MCSO - Booking	14,563
Habitat - Homeownership & CHR	19
Note: CMS 22/23 school year; HMIS 1/1/22-12/31/22; MCSO Arrest 1/1/22-4/30/22; Habitat 1/1/22-2/29/22; Habitat for Humanity Homeowner & CHR served 19 people between 1/1/22-2/29/22	

Note: after the publication of the report on 12/6/24, the Data Trust learned that the DSS-ANI data deposit contained duplicate records across files, instead of distinct records for the entire 2023 calendar year. This means that not all people who were in the DSS-ANI dataset were reflected in the analyses. This will be resolved in future reports.

What are the characteristics of the people in the Data Trust's administrative data?

Here we explore the race, ethnicity, gender, and age characteristics of the people in the Data Trust.

Normalization: The data collected about people's characteristics may vary from partner to partner. The Data Trust ***normalizes*** these data, meaning we make methodological decisions to standardize the data so they can be grouped and compared. Namely, we normalize the names of the racial, ethnic, and gender groups because there is variation across data partners. The Data Trust acknowledges the need to standardize data as well as the risk of doing so, particularly when it can obscure and oversimplify identities.

Representation: In several places, this report describes proportionate representation and disproportionate overrepresentation or underrepresentation. This refers to how the data we hold compares to data on the overall population of the county or the particular area we are describing. Understanding representation is important in the use and interpretation of Data Trust data.

RACE

The data in the Race_Ethnicity_Gender table counts all races the individual has ever identified with in that category. Every racial/ethnicity category present in this dataset is then joined with the dataset used to store the year each ISC_ID is present in a data partner dataset. The joined datasets are then grouped by year and aggregated to count each ISC_ID. This process is repeated for each racial/ethnicity category present in Figure 5. The percentages are calculated by dividing the count of each racial/ethnicity group by the number of total ISC_IDs for each year. The category Unknown is created when the only information about an ISC_ID Racial/Ethnic category are values such as Not Specified, Unable to Determine, Unreported or Unknown. The category Blank is created when no category is provided at all for an ISC_ID. The count and percentages represented in Figure 5 are from 2023. Since people can be part of multiple Racial/Ethnic categories the percentages will not add to 100%.

Figure 5

Racial/Ethnic Category	Percent	Count
Black or African American	49.6%	183609
White	24.0%	88874
Multi-Racial	12.7%	47075
Hispanic or Latino	7.4%	27228
Asian	4.8%	17793
Unknown	1.3%	4767
American Indian or Alaska Native	0.2%	712
Native Hawaiian or Other Pacific Islander	0.1%	306
Blank	0.0%	70

Figure 6

For the purposes of this chart and section, all people with more than one race are in the multi-racial section.

	2016	2017	2018	2019	2020	2021	2022	2023
Black or African American	192,012	181,250	178,745	171,627	170,376	181,062	194,760	183,609
White	97,296	92,548	90,973	86,569	83,652	86,980	91,338	88,874
Multi-Racial	43,913	45,310	46,962	46,607	46,169	47,417	49,265	47,075
Hispanic or Latino	20,390	20,829	212,95	22,622	24,059	23,427	25,550	27,228
Asian	17,421	17,374	17,588	17,426	17,510	17,804	18,531	17,793
Unknown	4,242	4,055	4,141	4,089	3,048	2,802	3,767	4,767
American Indian or Alaska Native	719	617	617	537	528	600	741	712
Native Hawaiian or Other Pacific Islander	240	249	283	269	270	255	306	306
Blank	142	87	138	181	140	368	420	70

Figure 7

Race/Ethnic Category	Mecklenburg County	Data Trust
Native Hawaiian or Other Pacific Islander	0.0%	0.1%
American Indian or Alaska Native	0.2%	0.2%
Asian	6.2%	4.8%
Hispanic or Latino	13.9%	7.4%
Multi-Racial	3.3%	12.7%
White	45.1%	24.0%
Black or African American	30.9%	49.6%

Note: Mecklenburg County population uses the American Community Survey 1 year estimates, 2022. The Data Trust percentages are from 2023.

To avoid counting people multiple times when they have two or more races, the categories with the exception of "Two or more races" represent people only having one Race/Ethnic category. The Overall Count was calculated using the Race_Ethnicity_Gender table that represents all ISC_IDs. To calculate the 2+ Datasets Count the "Dataset Matches By ISC_ID" dataset which captures ISC_IDs that are linked to more than one dataset was joined to the Race Ethnicity Gender table in order to only count ISC_IDs that have overlap with multiple datasets.

Figure 8

Race/Ethnic Category	2+ Datasets Count	Overall Count
Black or African American	175,669	369,946
White	41,511	220,005
Asian	5,326	36,522
American Indian or Alaska Native	154	1,903
Native Hawaiian or Pacific Islander	85	731
Two or more races	60,071	71,145
Hispanic or Latino	4,825	50,826
Unknown	860	14,814
Blank	125	1,433

UNDERSTANDING MULTIPLE RACIAL AND ETHNIC IDENTITIES

The data was calculated by going through each racial/ethnic category and generating counts based on two conditions, the first was how many ISC_IDs only represent the one Racial/Ethnic category. For example in the first record of Figure 9 the Target racial/ethnic identity and other racial/ethnic identity values are both Asian which means this is the amount of ISC_IDs are only represented as Asian and no other racial/ethnic identities. The second condition was how many ISC_IDs represent multiple racial/ethnic identities category combinations. Records 2-6 in Figure 9 represent the number of people represented as Asian and a different racial/ethnic identity.

Figure 9

The chart was created using the counts of people and their identified racial/ethnic groups. Each race/ethnicity is represented by a segment of the circle. The size of the segment is proportional to the total count of people of a given race/ethnicity. People who identified with more than one race or ethnicity are represented using chords that connect the segments, with their thickness representing the total count of a given pair of ethnicity. Due to size limitations, some race/ethnic names needed to be changed. A crosswalk of the different names are below:

- Black refers to the Black or African American group
- OPI refers to the Native Hawaiian or Other Pacific Islander group
- Am. Indian refers to the American Indian or Alaska Native group

Target Race/Ethnic Identity	Other Race/Ethnic Identity	Count
Asian	Asian	36,522
Asian	Black	1,475
Asian	White	1,837
Asian	Hispanic or Latino	343
Asian	Multi-Racial	134
Asian	OPI	222
Black	Black	369,946
Black	White	15,842
Black	Hispanic or Latino	7,785
Black	Multi-Racial	1,947
Black	OPI	784
White	White	220,005
Hispanic or Latino	White	36,096
Hispanic or Latino	Hispanic or Latino	50,826
Hispanic or Latino	Multi-Racial	<10
Hispanic or Latino	OPI	247
Multi-Racial	White	831
Multi-Racial	Multi-Racial	6,496
Multi-Racial	OPI	<10
Am.Ind	Asian	221
Am.Ind	Black	2,693
Am.Ind	White	1,928
Am.Ind	Hispanic or Latino	631

Am.Ind	Multi-Racial	25
Am.Ind	OPI	1903
Am.Ind	OPI	132
OPI	White	577
OPI	OPI	731
Blank	Blank	1,433

The Sankey visualizations below are created for every race and will show how races are linked to each other. Similar to the chord diagram, each Sankey chart takes the total count of people in a given race/ethnicity and shows the relative proportion of people who identify with a single race and those who identify with other race/ethnicity categories. Larger 'flows' in a given race/ethnicity's chart to other races/ethnicities represent higher incidence of multi-racial identification.

Figures 10-16

Target Racial/Ethnic Category	Other Racial/Ethnic Category	Percent
Black	Black only	92.37%
Black	White	3.96%
Black	Hispanic or Latino	1.94%
Black	Multi-Racial	0.49%
Black	Pacific Islander	0.20%
Black	American Indian	0.67%
Black	Asian	0.37%
Pacific Islander	White	21.38%
Pacific Islander	Pacific Islander only	27.08%
Pacific Islander	Multi-Racial	0.22%
Pacific Islander	Asian	8.23%
Pacific Islander	Hispanic or Latino	9.15%
Pacific Islander	Black	29.05%
Pacific Islander	Am Indian	4.89%
White	White Only	79.38%
White	Asian	0.66%
White	Black	5.72%
White	Hispanic or Latino	13.03%
White	Multi-Racial	0.30%

White	American Indian	0.70%
White	Pacific Islander	0.21%
Multi-Racial	Asian	1.42%
Multi-Racial	Black	20.62%
Multi-Racial	Hispanic or Latino	0.05%
Multi-Racial	White	8.80%
Multi-Racial	Multi-Racial only	68.79%
Multi-Racial	Pacific Islander	0.06%
Multi-Racial	American Indian	0.26%
American Indian	Asian	2.93%
American Indian	Black	35.75%
American Indian	White	25.59%
American Indian	Hispanic or Latino	8.38%
American Indian	Multi-Racial	0.33%
American Indian	American Indian only	25.27%
American Indian	Pacific Islander	1.75%
Asian	Asian only	89.62%
Asian	Black	3.62%
Asian	White	4.51%
Asian	Hispanic or Latino	0.84%
Asian	Multi-Racial	0.33%
Asian	Pacific Islander	0.54%
Asian	American Indian	0.54%
Hispanic or Latino	Asian	0.36%
Hispanic or Latino	Black	8.14%
Hispanic or Latino	White	37.58%
Hispanic or Latino	Hispanic or Latino only	53%
Hispanic or Latino	American Indian	0.66%
Hispanic or Latino	Pacific Islander	0.26%

GENDER

This output will get the number of people that are associated with only one type of gender and calculate the percentage. For the Unknown category this combines the records that are only Unknown and records with more than one gender. The counts will add up to the total number of people in all the datasets.

Figure 17

Year	Male	Female	A gender other than singularly male or female	Unknown	Blank
2023	173,486	194,327	16	2,605	0

AGE

Age was calculated as of the most recent date of entry for all individuals in every data partner dataset between 2016-2023. In future reports, we will calculate age as of the most recent entry of the most recent year.

Figure 18

Data Partner	Min Age	Median Age	Max Age
CRISIS	0	25	123
CMS	2	12	25
DSS – ES	0	24	123
DSS – ANI	0	8	24
DSS – FIHS	0	7	24
DSS – CIC	0	6	22
GW	0	38	77
HMIS	0	32	94
MCSO - ARREST	16	30	88
MCSO - BOOKING	14	32	88
HABITAT- HOMEOWNER	23	39	82
HABITAT - CHR	35	69	96
Note: Ages above 100 are likely data entry errors.			

Where do people in the Data Trust live?

To understand which parts of Charlotte/Mecklenburg County are represented by the Data Trust data, we geocoded the addresses in the datasets and aggregated them to neighborhood profile areas

(NPAs). Location data is not available with all partners. The address data used in this section is from CMS, DSS – ES, and Crisis.

For each of these datasets, we first computed the total count of people in each NPA. To analyze if a given NPA is proportionately represented, we used the 2020 population distribution map of Mecklenburg County from the Quality of Life Explorer. Proportion of the total population was calculated for each NPA as the reference. Then, the proportion of NPA counts to the total count was calculated for each dataset and compared with the reference proportion. For example, if a NPA has 1% of the total population and the same NPA has 1% of the total people in Crisis for a given year (or set of years) we classify the NPA as proportionately represented.

Given the wide variation in population concentrations in NPAs as well as potential missing value and address quality issues, the maps represent only a rough estimate of representation. Thus, we give a margin of error of 30% and classify any NPA within 30% of the population proportion as proportionately represented. Any NPA above 30% is over represented and any NPA below 30% is considered underrepresented.

Detailed data tables are available on request.

How are people reflected across partners?

For the purposes of the analyses visualized in Figures 23 and 24, we combined the three child welfare datasets (DSS-ANI, DSS-FIHS, and DSS-CIC) and two MCSO datasets (MCSO-Arrest and MCSO Booking), since the datasets within child welfare and criminal justice reflect many of the same people. For Figures 25-37, we separated the three child welfare and two criminal justice databases since the distinctions between the datasets or the programs they represent may be meaningful to partners.

Figure 23

To calculate this data the dataset developed by the Charlotte Regional Data Trust to track the years each ISC_ID is represented in the data partner datasets was used and joined with the Dataset Matches by ISC_ID table, which contains ISC_IDs that overlap with more than one dataset. This resulted in a dataset containing all the ISC_IDs and a binary (0 or 1) field for each data partner dataset to indicate if they are found in the dataset. The dataset was then grouped by the ISC_ID and year and then aggregated to get the count of the number of data partner datasets each ISC_ID was found in for every year.

Year	Number of datasets	Number of unique people

2023	1	288,510
	2	68,497
	3	12,414
	4+	1,013

Figure 24

To calculate this data, the Dataset Matches by ISC_ID dataset was outer joined with the Race_Ethnicity_Gender table which resulted in a dataset with one record for each ISC_ID (767,325) and binary (0 or 1) fields for each data partner dataset to indicate if they are present in each dataset. The total number of datasets was calculated for every ISC_ID based on summing all the binary fields. A count was then performed on the Number of Datasets field to get a count of the number of unique people grouped by the number of datasets they are linked to.

Year	Number of datasets	Number of unique people
2016-2023	1	500,847
	2	167,519
	3	69,779
	4	24,816
	5+	4,364

OVERLAP BY DATA PARTNERS

The output calculates the overlap between a target data partner and the other data partners between 2016-2023. The target counts and percentages represent the number of people that are in the target dataset. The other datasets represent how many people in their dataset match also match with the people in the target dataset. For example, with Crisis Assistance Ministry, there were 175,434 people total and 76.2% of the people in the Crisis Assistance Ministry dataset matched with at least one other dataset. When compared to DSS - Economic Services, there are 122,792 people that match in both the Crisis Assistance Ministry dataset and DSS - Economic Services datasets. The percent overlap for DSS - Economic Service is calculated by dividing the 122,792 people to the 175,434 total people in the Crisis Assistance Ministry dataset.

Figures 25 - 37

Target Data Partner	Linked Data Partner	Count Overlap
CRISIS	DSS-ES	122,792
CRISIS	CMS	56,325
CRISIS	DSS-ANI	24,436
CRISIS	HMIS	16,348
CRISIS	MCSO-ARREST	14,453
CRISIS	MCSO-BOOKING	13,378
CRISIS	DSS-FIHS	5,609
CRISIS	GW	3,371
CRISIS	DSS-CIC	888
CRISIS	HABITAT-CHR	152
CRISIS	HABITAT-HOMEOWNER	84
CMS	DSS-ES	140,397
CMS	CRISIS	56,325
CMS	DSS-ANI	46,893
CMS	DSS-FIHS	7,513
CMS	HMIS	7,382
CMS	MCSO-BOOKING	6,757
CMS	MCSO-ARREST	6,721
CMS	DSS-CIC	1,219
CMS	GW	569
CMS	HABITAT-HOMEOWNER	0
CMS	HABITAT-CHR	0
DSS-ES	CMS	140,397
DSS-ES	CRISIS	122,792
DSS-ES	DSS-ANI	47,618
DSS-ES	MCSO-ARREST	29,106
DSS-ES	MCSO-BOOKING	27,298
DSS-ES	HMIS	22,720

DSS-ES	DSS-FIHS	9,172
DSS-ES	GW	4,666
DSS-ES	DSS-CIC	1,420
DSS-ES	HABITAT-CHR	273
DSS-ES	HABITAT-HOMEOWNER	191
DSS-ANI	DSS-ES	47,618
DSS-ANI	CMS	46,893
DSS-ANI	CRISIS	24,436
DSS-ANI	DSS-FIHS	10,071
DSS-ANI	HMIS	5,126
DSS-ANI	DSS-CIC	1,675
DSS-ANI	MCSO-BOOKING	1,615
DSS-ANI	MCSO-ARREST	1,342
DSS-ANI	GW	86
DSS-ANI	HABITAT-HOMEOWNER	0
DSS-ANI	HABITAT-CHR	0
DSS-CIC	DSS-ANI	1,675
DSS-CIC	DSS-ES	1,420
DSS-CIC	CMS	1,219
DSS-CIC	CRISIS	888
DSS-CIC	DSS-FIHS	789
DSS-CIC	HMIS	409
DSS-CIC	MCSO-BOOKING	183
DSS-CIC	MCSO-ARREST	154
DSS-CIC	GW	<10
DSS-CIC	HABITAT-HOMEOWNER	0
DSS-CIC	HABITAT-CHR	0
DSS-FIHS	DSS-ANI	10,071
DSS-FIHS	DSS-ES	9,172

DSS-FIHS	CMS	7,513
DSS-FIHS	CRISIS	5,609
DSS-FIHS	HMIS	1,658
DSS-FIHS	DSS-CIC	789
DSS-FIHS	MCSO-BOOKING	235
DSS-FIHS	MCSO-ARREST	169
DSS-FIHS	GW	18
DSS-FIHS	HABITAT-HOMEOWNER	0
DSS-FIHS	HABITAT-CHR	0
HMIS	DSS-ES	22,720
HMIS	CRISIS	16,348
HMIS	CMS	7,382
HMIS	MCSO-ARREST	6,130
HMIS	MCSO-BOOKING	6,004
HMIS	DSS-ANI	5,126
HMIS	DSS-FIHS	1,658
HMIS	GW	873
HMIS	DSS-CIC	409
HMIS	HABITAT-HOMEOWNER	<10
HMIS	HABITAT-CHR	0
GW	DSS-ES	4,666
GW	CRISIS	3,371
GW	MCSO-ARREST	986
GW	MCSO-BOOKING	961
GW	HMIS	873
GW	CMS	569

GW	DSS-ANI	86
GW	DSS-FIHS	18
GW	DSS-CIC	<10
GW	HABITAT-HOMEOWNER	<10
GW	HABITAT-CHR	<10
MCSO-ARREST	MCSO-BOOKING	47,671
MCSO-ARREST	DSS-ES	29,106
MCSO-ARREST	CRISIS	14,453
MCSO-ARREST	CMS	6,721
MCSO-ARREST	HMIS	6,130
MCSO-ARREST	DSS-ANI	1,342
MCSO-ARREST	GW	986
MCSO-ARREST	DSS-FIHS	169
MCSO-ARREST	DSS-CIC	154
MCSO-ARREST	HABITAT-HOMEOWNER	10
MCSO-ARREST	HABITAT-CHR	<10
MCSO-BOOKING	MCSO-ARREST	47,671
MCSO-BOOKING	DSS-ES	27,298
MCSO-BOOKING	CRISIS	13,378
MCSO-BOOKING	CMS	6,757
MCSO-BOOKING	HMIS	6,004

MCSO-BOOKING	DSS-ANI	1,615
MCSO-BOOKING	GW	961
MCSO-BOOKING	DSS-FIHS	235
MCSO-BOOKING	DSS-CIC	183
MCSO-BOOKING	HABITAT-HOMEOWNER	<10
MCSO-BOOKING	HABITAT-CHR	<10
HABITAT-HOMEOWNER	DSS-ES	191
HABITAT-HOMEOWNER	CRISIS	84
HABITAT-HOMEOWNER	MCSO-ARREST	10
HABITAT-HOMEOWNER	MCSO-BOOKING	<10
HABITAT-HOMEOWNER	GW	<10
HABITAT-HOMEOWNER	HMIS	<10
HABITAT-HOMEOWNER	CMS	0
HABITAT-HOMEOWNER	DSS-ANI	0
HABITAT-HOMEOWNER	DSS-FIHS	0
HABITAT-HOMEOWNER	DSS-CIC	0
HABITAT-HOMEOWNER	HABITAT-CHR	0
HABITAT-CHR	DSS-ES	273
HABITAT-CHR	CRISIS	152
HABITAT-CHR	GW	<10

HABITAT-CHR	MCSO-BOOKING	<10
HABITAT-CHR	MCSO-ARREST	<10
HABITAT-CHR	CMS	0
HABITAT-CHR	DSS-ANI	0
HABITAT-CHR	DSS-FIHS	0
HABITAT-CHR	DSS-CIC	0
HABITAT-CHR	HMIS	0
HABITAT-CHR	HABITAT-HOMEOWNER	0