GUIDE

Using Data to Assess Fair Housing and Improve Access to Opportunity

A Guidebook for Community Organizations

Leah Hendey and Mychal Cohen
August 2017

URBAN INSTITUTE · ELEVATE · THE · DEBATE

ABOUT THE URBAN INSTITUTE

The nonprofit Urban Institute is dedicated to elevating the debate on social and economic policy. For nearly five decades, Urban scholars have conducted research and offered evidence-based solutions that improve lives and strengthen communities across a rapidly urbanizing world. Their objective research helps expand opportunities for all, reduce hardship among the most vulnerable, and strengthen the effectiveness of the public sector.

Contents

Acknowledgments

Part I: Data in Action

The places where we live shape our lives, our health, and our prospects for economic mobility. However, for many groups, including people of color, people with disabilities, and other protected classes, neighborhood and housing choices are too often constrained by private discrimination and public policies that concentrate poverty, prevent investments from flowing into distressed communities, or make housing prohibitively expensive in neighborhoods with quality schools and access to good jobs. The US Congress adopted the Fair Housing Act in 1968[1] not only to remedy past discrimination but also to provide tools to create more inclusive and prosperous communities and ensure that all American families have access to neighborhoods of opportunity (see box 1).

Five decades after the passage of the Fair Housing Act, and even after amendments,[2] however, the United States still faces significant challenges to creating inclusive communities. Researchers' use of data has been essential in telling this story. For example, researchers have used data to demonstrate the persistence of segregation and growing concentration of poverty across cities and regions (Logan and Stults 2011; Pendall et al. 2011). Mounting evidence shows that racial and economic segregation brings steep costs not only for individuals living in segregated neighborhoods but also for regions as a whole (Acs et al. 2017; Carr and Kutty 2008; Li, Campbell, and Fernandez 2013; Turner and Rawlings 2009; Wachter and Ding 2016). Beyond describing and quantifying the problem, data can also be used to spark dialogue among stakeholders and allow community leaders to design effective solutions (Greene and Pettit 2016; McCormick 2017).

This guide is an introduction to the use of secondary data for organizations working to advance opportunity and fair housing in their communities. Community organizations may be familiar with using their own primary data, such as client or property management systems, to improve local housing choice and access to opportunity for neighborhood residents. Secondary data from sources like the US Census Bureau and other national and local government agencies can be used to describe how the community is doing, the characteristics and needs of its people, and how the community compares to neighboring communities, its city, or and its region overall. Understanding and analyzing secondary data can allow community groups not only to improve their own program planning, but also to enhance their ability to advocate on behalf of their clients and effectively engage in fair housing and other public planning processes, such as those leading to the development of an assessment of fair housing (AFH) or consolidated plan. The guide will also be useful for local government agencies leading these planning processes, implementing programs, and analyzing local conditions.

The Fair Housing Act and the Affirmatively Furthering Fair Housing Rule

The Fair Housing Act was passed in 1968 to protect people when they are renting, buying, or seeking financing for housing from discrimination based on race, color, religion, sex, or national origin (the act was later amended to include protections for people with disabilities and families with children). In addition to prohibiting explicit and intentional discrimination in housing and lending markets, the Fair Housing Act protects individuals from policies, practices, and programs that have a disparate impact on protected groups. It also requires the federal government and local governments that receive federal housing and community development funding to take affirmative steps to remove barriers to housing choice and reduce disparities between neighborhoods. The latter provision has become known as the "affirmatively furthering fair housing" (AFFH) requirement.[a]

In 2015, the US Department of Housing and Urban Development (HUD) published regulations that interpret the AFFH requirement and provide guidance on compliance for state and local agencies that receive federal housing funds. The AFFH regulation requires grantees to use data to complete an AFH. This assessment entails identifying and documenting patterns of segregation, concentrated poverty, and access to opportunity, as well as the key local and regional factors that drive these patterns. As part of the assessment, local jurisdictions and housing authorities must engage community residents and the public to understand and address these barriers and prioritize goals. To assist in this process, HUD provides data that are available across jurisdictions, a mapping tool, and instructions and guidance. It also expects grantees to supplement uniform, federal data with local data and community knowledge to accurately describe barriers to fair housing and access to opportunity.

Note: For more on disparate impact, see HUD, Implementation of the Fair Housing Act's Discriminatory Effects Standard, 78 Fed. Reg. 32, (February 13, 2014) (Codified at 24 C.F.R. 100), https://portal.hud.gov/hudportal/documents/huddoc?id=discriminatoryeffectrule.pdf.
[a] Alan Greenblatt, "Will New Housing Rules Really Reduce Racial Segregation?" *Governing.* July 16, 2015, http://www.governing.com/topics/urban/gov-hud-housing-supreme-court-discrimination.html.

Structure of the Guide

The remainder of part I introduces how community organizations can use secondary data to improve access to opportunity and support fair housing and provides examples from three communities. It also describes types of secondary data and the resources available, including online tools and platforms and potential institutional partners, to help organizations navigate and find the right information.

Part II of this guide delivers more detailed information about data sources across many domains related to fair housing and access to opportunity, including demographics and segregation, housing,

land use, disability, education, employment, environment, health, and public safety. Under each topic appropriate sources of national and local data are described along with potential indicators. Finally, appendices catalog and cross reference all the data sources mentioned for organizations, communities, and agencies using the guide during the AFH process.

It is beyond the scope of this guide to cover formal primary data collection methods, such as conducting household and property condition surveys, structured interviews, or focus groups. These methods may be needed depending on the questions your organization has, but generally secondary data sources should be assessed before beginning primary data collection, which takes significant time and resources and can burden participants. This guide also does not cover sources of information gathered informally that are valuable for community organizations. Neighborhood residents and organizations can share important experiences that can give context to quantitative data, fill in gaps when data are unavailable, and reveal problems or solutions that are not captured in the numbers. When possible, quantitative data and local knowledge should be paired to give a more complete picture of neighborhood conditions.

Using Data to Advance Opportunity

Community-based organizations can use secondary data in policy advocacy and planning processes and to describe local and regional conditions.[3] We offer suggestions for how to think strategically about using data and concrete examples of how community organizations have incorporated secondary data in their advocacy efforts to improve outcomes for their communities. Part I closes by defining general types of secondary data that organizations may encounter, identifying trade-offs to consider in choosing data, and outlining resources that can help organizations access and translate data.

How Can Data Help?

Organizations can use secondary data to better understand local and regional conditions in each stage of the public policy process: problem identification, planning and formulation, implementation, and evaluation. Analyzing data can demonstrate patterns of segregation in a city and region or inequities in outcomes. Looking at data concurrently, across a range of resident characteristics and community conditions like education, health, housing, income, and crime, helps show the interactions among these factors, better reflecting how residents experience their communities and helping them define their goals for quality housing and schools, safe streets, and access to good jobs. By using data to depict local

conditions, organizations can begin to build a shared understanding of their community with other local institutions, residents, and government.

Examining data across multiple domains can also help break down information silos and can encourage organizations and agencies operating in different spaces to work more collaboratively to advance opportunity. For example, an analysis by the University of Pittsburgh showed that one of the biggest factors influencing absenteeism for students in the public school system of Pittsburgh, PA, was changing schools during the school year (Deitrick et al. 2016). School officials realized that housing instability, a likely cause of at least some school switches, cannot be thought of only as a housing or income issue, but also as an issue with implications for education.

Secondary data analysis in conjunction with the community is important for any strategic planning process for local agencies or for organizations looking to improve their own programming. Agencies or local universities may share analyses of local conditions as written reports or online community profiles. Community discussions of these products with a variety of stakeholders on what the data capture (or the nuances they fail to capture) about people's experiences within their communities are critical sources of information. In these discussions, organizations and residents can use the data to test assumptions about the opportunities and challenges faced by residents and pose explanations for what the data might show. These discussions can generate ideas for new solutions to create more inclusive communities.

As illustrated in the examples below, having specific numbers, such as the density of Latino residents in a neighborhood and their access to public transit or how many vacant homes are owned by investors, allows advocates to craft more effective and targeted messages about what changes are needed to address fair housing and barriers to opportunity. Having better information helps advocates make their case for policy change and community investment to local government officials and to the public. Secondary data can be used, for example, to improve the implementation of a policy or program by enhancing outreach and targeting efforts. Data can also be used to monitor the effect of the changes and to ensure that the outcomes desired by the community are being generated.

The following examples show how community organizations have used secondary data as part of their explicit advocacy and community engagement strategies by arguing for geographically targeted programs, identifying disparities, and holding government and for-profit entities accountable.

Advocating for Transit Access in the Twin Cities

The Center for Urban and Regional Affairs (CURA) at the University of Minnesota and local community advocacy group Asamblea de Derechos Civiles (Assembly for Civil Rights) teamed up in 2011 to identify

how a proposed bus rapid transit (BRT) project would affect Latinos in the region.[4] The project emerged when the Twin Cities began planning its Corridors of Opportunities Initiative to expand the regional transit system. Access to transit throughout a metropolitan region plays a vital role in expanding community members' access to greater opportunity by connecting residents to employment, city services, and community events. La Asamblea was interested in identifying which transit corridors would have the greatest impact on Latino communities and to prioritize where to get involved in the community planning process. To identify areas with the highest densities of Latinos along seven proposed transit corridors, CURA created a population density map and a list of proposed bus stops serving Latino areas by using the region's municipal transit data as well as US Census data.

CURA's analysis helped La Asamblea focus their efforts on the Cedar Avenue Red Line BRT and hone in on the specific areas that could most benefit from the new bus line. They then involved Latino residents, who were most directly affected, in the planning process. La Asamblea organized a Red Line bus tour with public officials and highlighted the need for the transit authority to include an additional station location in an area with a high population of low-wealth immigrant community members currently excluded from easy and fast access to public transit. Since the community group's organizing efforts in 2013 and 2014, the transit authority has created Spanish-language ridership materials, and in phase two BRT development in 2020 it plans to create a new bus stop in the area identified by CURA and La Asamblea.

Improving Fair Housing Planning in Philadelphia

In Pennsylvania, the Philadelphia Chinatown Development Corporation's (PCDC) participation in the development of Philadelphia's AFH in 2016 is a prime example of how to use data to drive community action and inform local planning processes. As the assessment process began, PCDC saw a gap in who was being represented in the data and community engagement processes, with underrepresentation for immigrants and those with limited English proficiency (LEP) in both the Asian and Latino communities.

Concerned with the low participation rates for Asians, and specifically Chinese-speaking LEP populations, in housing programs run by the Housing Authority and Department of Housing and Community Development, PCDC helped found the Coalition for Fair Housing. This coalition pushed for a greater focus on Asians and Latinos (a population facing similar disparities in participation rates) in the assessment. With data such as housing choice voucher participation by race/ethnicity and LEP status, the coalition developed a concrete, evidence-based foundation to express long-standing community concerns about disparities in access to subsidized housing.

As a result of their advocacy, the city held additional meetings for stakeholders and actively solicited input from marginalized communities. In response to PCDC's and other community

organizations' data-informed input and identification of disparities in access to services, the assessment's focus on immigrant communities increased substantially, with a much deeper discussion of the challenges facing immigrants in the city. The final version of the assessment included multiple goals centered on increasing access to housing programs and public transit for immigrant communities and plans for greater outreach efforts to LEP populations.

Holding Investors Accountable in Oakland

Following the foreclosure crisis, Oakland, CA, faced a serious problem with vacant buildings. Vacant properties, particularly if not well maintained, can affect the safety and quality of life in a neighborhood. The Alliance of Californians for Community Engagement (ACCE) along with the Urban Strategies Council, a local data intermediary, set out to address the numerous blighted and abandoned properties and ensure that banks were paying for the tax burden created by vacant foreclosed properties.[5] To do this they had to first examine administrative property records to untangle the complex network of ownership entities that held controlling interests in these properties. They organized a campaign around the issue that resulted in new City of Oakland legislation to create a vacant property registry. Actually implementing the legislation proved to be a huge lift, but they were able to connect with a champion in the city bureaucracy who helped spearhead the registry's development. Crucially, the final registry was made public, allowing concerned residents and neighborhood advocates access to the data.

The data from the vacant property registry allowed ACCE to hold absentee property owners accountable for property upkeep and maintenance and to ensure the city enforced fines associated with vacant properties. Further, ACCE drew on the registry to develop an evidence base for larger narratives about the wealth and inequality gap and limited enforcement of fines for large investors. Success with the property registry also inspired ACCE to thread data through more of its work demonstrating disparities and holding the government and corporations accountable. For example, ACCE has pushed the city to create a similar database on speculators and routinely uses data in political education workshops.

How Can Data Inform Strategy?

Data alone cannot determine program or policy solutions. But with deliberate planning they can be used to inform a community organization's overall strategy to improve fair housing and access to opportunity. Considering the following questions will assist you to focus your efforts by identifying several key issues that support an effective use of both primary and secondary data.

- What is your organization's primary goal (e.g., provide evidence of disparities, hold institutions accountable, target programs and resources)?

- Who are you trying to influence, that is, who is the intended audience?

- What is the key message to convey?

- What information gaps must be filled so your organization can meet its objectives?

- How can your audience be engaged to share the data?

Once you have answered these basic questions, you can begin thinking about what data exist to fill those needs. Talking to other organizations working on related issues in your community may help clarify your data needs or reveal shared information needs among several organizations. Nobody wants to duplicate data collection and analysis, so plan to share data and insights on your website, blogs, social media, and community forums.

- **Look for partners and existing tools.** Working with data and recognizing what is useful can be challenging, even for experts. Find partners and collaborators. Look for institutions described below in the Resources section for assistance or online tools that offer curated indicators.

- **Prioritize data collection and analysis based on your objectives.** As time and resources are almost always limited, you need to weigh the difficulty of obtaining data against the value the data provide. Are certain data critical or even imperative to making a point to those in leadership positions? Is there only one source for the information? Could other data be used as reasonable proxies? Reassess your data collection and analyses over time to make sure they are still informing your organization's strategies. The section on trade-offs discusses other considerations to weigh in prioritizing working with data.

- **Plan for the long game.** Obtaining access to data may require building relationships and establishing trust with data providers. Trust, not technical or even legal challenges, is the biggest barrier to obtaining data. Though some data take considerable time to obtain, they could provide valuable insights for your community. Find and champion local partners or data intermediaries that could serve as long-term community information resources to help you with data assistance when you need it.

What Types of Data Are Available?

Two major types of data, *administrative data* and *survey data*, are widely used by community organizations to advance access to opportunity and fair housing.

Data created through the operation of government programs or oversight functions (federal, state, or local) are referred to as *administrative data*.[6] For example, administrative data for Temporary Assistance for Needy Families would contain information collected when clients apply for benefits (e.g., income and household type and size) and information needed to pay out benefits (e.g., client address, payment history, and recertification date). Administrative data only cover individuals or households participating in the program or those who may have applied for benefits. They do not cover those who are eligible for a service but not enrolled. Therefore, each data source will reflect a unique population, with some having near universal coverage, like birth certificates, or only a few hundred households, like energy assistance.

Some administrative data are a matter of public record, which means laws or regulations require that the information be made available in some form to the public. For example, in most states transactions involving property are recorded with the local government, and those records can be viewed by anyone (in person or often online). Most administrative data that contain sensitive information about people or businesses are considered confidential. Several laws at the federal, state, and local levels protect the privacy of people and businesses and specify restrictions on access and use of confidential data (e.g., education or income tax records).

In this guide, we mainly discuss administrative data associated with government agencies, but administrative data are also generated through the operation of commercial firms or nonprofit organizations (e.g., sales transactions or client records). For instance, administrative records from mortgage servicing companies are used to understand how many borrowers are delinquent on their mortgages, how many properties are in foreclosure, or how many borrowers have negative equity. Multiple listing services may create reports on real estate listings and indicators of market activity at the jurisdiction or zip code level. Commercial firms usually restrict republishing and sharing their data with other parties.

Survey data are based on questions asked of a subset or sample of a population. Unlike administrative sources, which are limited to the populations served, survey responses are used to generate estimates for the full population of interest. One of the most well-known large-scale surveys in the United States is the American Community Survey (ACS), which is designed to be representative of the entire US population. Surveys can cover a variety of topics, including some that may not have administrative sources, such as monthly rent for a unit. Because surveys are only asked of a sample of

people, a particular estimate will be associated with a range of values above and below the estimate at which it is most likely the true value lies. If the same survey questions were asked of a different group of people, the survey may produce slightly different results. The "margin of error" reflects this variation of estimates. Political polling often displays this error range. For example, a candidate may lead the polls with 48 percent in favor, plus or minus 3 percent, indicating the true value lies between 45 and 51 percent in favor. The size of the margin of error will vary for every estimate. As an example, the estimate for the poverty rate for each city in the United States will have a different size margin of error.

What Are the Trade-Offs to Using Different Sources of Data?

The initial descriptions of survey and administrative data sources already noted two of their differences—the extent to which they reflect the whole population and their precision. Several other important factors, including geographic level and coverage, timeliness, accessibility, ease of use, and cost, influence which data are needed and will provide the most utility.

Geographic Level. Depending on the question you are asking, you may need data that show the differences between neighborhoods in a community or you may decide a county total will be sufficient. Describing black-white segregation in your community, for example, will require data on the share of the population that blacks and whites represent for neighborhoods or other small areas. But to determine whether incomes are generally higher in one jurisdiction in your region compared with another, the median income for each county may be sufficient detail. Some data sources, both administrative and survey based, are only available at higher geographic levels (cities, counties, states), but others are available at the neighborhood or even address level. Part II describes data sources that are available at smaller geographies (subcity or subcounty) and are useful to community organizations interested in advancing opportunity. Small geographies include the following:

- **Address**: Some administrative records about housing units or data on amenities like libraries, recreation centers, museums, and so forth, will have the unit's or building's exact address.
- **Parcels**: A parcel is a piece of land with defined boundaries for the purposes of tracking land ownership and taxation. A parcel may have more than one building and be associated with more than one address (e.g., an apartment complex).
- **Census Bureau–defined geographies**: The Census Bureau defines several levels of small geographies, the smallest of which is the *block*. Blocks are typically bounded by streets or features like railroad tracks and streams.[7] A cluster of blocks forms a *block group*, with a population size ranging from 600 to 3,000 people.[8] One or more block groups form a *census tract*.[9] Census tracts vary in the size of the land area they cover depending on the population

density. There are typically 1,200 to 8,000 people living in each tract, with the goal to have about 4,000 people per tract. Census tracts are defined with local input. They do not cross county lines but may cross municipal boundaries.

- **Zip codes**: Zip codes are defined by the US Postal Service for the purposes of delivering mail efficiently. They are technically collections of addresses rather than geographies with clear boundaries. Zip codes change over time (especially in newly developing areas) and can also cross county or in a few cases state boundaries.[10] For these reasons, data from the Decennial Census or the ACS are not available for zip codes. Instead, the Census Bureau defines special zip code tabulation areas that approximate zip codes.

- **Other subcity and subcounty areas**: Usually other small geographies are defined by local jurisdictions such as city or county council districts, police precincts, voting precincts, or other areas defined for the purposes of planning. Some cities have clearly defined "neighborhood" units, like Chicago's community areas.[11]

Some data sources, like the ACS, provide data at multiple geographic levels; others may supply data at only one level. Understanding the area the geography represents and the implications the defined area may have for the data can be critical. As recently documented, the poor alignment of zip codes and the boundaries of Flint, MI, caused state officials to misunderstand the magnitude of the lead problem in Flint's water system.[12]

Geographic Coverage. The geographic coverage of a data source may also be important. If your organization is involved in a regional planning process, you may need to have comparable data across several counties. National data sources like the ACS or HUD's Picture of Subsidized Housing are helpful because they allow comparisons of the same measure across different levels of geography— neighborhood, city, county, or region.[13] Local administrative data sources are usually more difficult to compare across jurisdictions due to differences in how data are recorded internally and how they get reported. For example, one jurisdiction may collect the exact number of apartment units in a building in their property assessment database, while another may report the number of units in categories, such as 1 to 4, 5 to 10, and more than 10 units. Regions that cross state boundaries may face more of these challenges to using administrative data.

Timeliness. Most nationally available data will not be timely enough to capture the changes in rapidly evolving communities. Many of the data products from national data sources take time to produce and are usually released annually several months, if not years, after the time period the data represent. For example, the data collected on mortgage applications and loans in 2015 as part of the Home Mortgage Disclosure Act became available in September 2016. For this reason, many of these data sources are more appropriate for looking back at what happened than in aiding in developing a

proactive approach. State and local agencies may release data more often or have the ability to release them in a timelier manner than the federal government. Some jurisdictions are even publishing certain data, like crime reports or property sales, online daily. Commercial data such as credit scores, rental listings, or retail sales transactions may be available in real time but are more difficult to access, though some corporations are doing more to increase access and promote the use of their data (Greene and Pettit 2016).

Accessibility. The accessibility of data varies by source. The next section on resources covers many online platforms that make data like the ACS easy to access. For state and local administrative data, accessibility also varies by the jurisdiction. Some data are public records and may be more readily available or even posted on open data portals. Open data portals are websites that collect the datasets, tables, or maps that an agency or set of agencies has released to the public. The portals typically have data tagged by topic so users can search for data and to provide some information about the fields included in the data.[14] Public records that are not posted online may also be obtained by working with agency staff or submitting a request under the jurisdiction's freedom of information process.[15]

Obtaining confidential administrative data is also possible but requires negotiating data use agreements and complying with technical data security guidelines, permissible uses defined by the agencies, and privacy laws. The negotiations for confidential data can take anywhere from several months to several years. Having the data provider's trust is critical, and building that relationship also takes time. In the next section, we describe several types of organizations that may already have built these trusted relationships.

Occasionally, commercial firms may make data more readily accessible, with additional indicators or analysis and without cost to data users. For example, Zillow provides a variety of home sales and rental market indicators for various levels of geography free on their website.

Ease of Use. Much of the data that are easy to access has already been transformed into useful indicators by agencies like the Census Bureau or national organizations and can be easily downloaded from one of the online platforms described below. The next section also describes the various local institutions that may maintain useful data.

Administrative data may also be easy to access, but they usually require staff with substantial technical skills and expertise to transform them into meaningful indicators for advocacy, planning, or policy development. The data are collected for program management, and so the information provided may not be in the form most useful for analysis. Raw administrative data need to be evaluated for quality (e.g., are fields missing a lot values?) and require engaging with data providers to understand the meaning of fields and how the information is collected. As an example, a raw dataset with all property

sales transactions for a county can be used to create the monthly median single-family home sales price by neighborhood. But before that can happen the nonmarket transactions and those for property types that are not single-family homes must be identified and removed, the data quality and outlier observations evaluated, and addresses matched to neighborhoods. After these adjustments are made, the median neighborhood sales price can be calculated.

Costs to Consider. Several types of data-related costs may be incurred. Paying for data may be necessary if no free sources exist for needed information or if obtaining data from free sources would take too long. For example, foreclosure data are hard to pull together because sources may not release the needed data in bulk or users may need to merge information from several datasets. Buying foreclosure data from commercial firms like CoreLogic or RealtyTrac may save considerable time. Similarly, population projections for small areas are hard to come by, but firms like Esri and Geolytics sell these products. Additionally, administrative agencies may levy processing fees to pay for the staff time to fill data requests. Fees for data requested through freedom of information processes may be waived if the disclosure request is in the public interest, depending on the jurisdiction.

Even if administrative data are free, the labor costs involved in cleaning and analyzing data may be significant. Depending on the analysis desired, you might consider whether it is more efficient to augment your organization's internal capacity to manipulate and analyze data or find an entity that can be hired to perform the work.

What Resources Aid in Finding and Analyzing Data?

Wading through data sources and indicators and deciding which ones you need is a daunting task. In part II we describe data sources that might be useful in your work, but here we discuss a few resources that may help your organization find the right data and analyze them: state and local institutions, data platforms, and tools to analyze and display data.

STATE AND LOCAL INSTITUTIONS

Many private and public institutions and individuals use some of the data types discussed above. The amount of work being done will vary by the number and types of institutions in your community, of course, but they may already be doing the hard work of transforming raw administrative data into usable pieces of information and sharing the data and analyses directly in online portals, reports and briefs, or presentations. We recommend building relationships with these institutions to help you navigate administrative data and leverage existing resources.

State, regional, and local agencies. Government agencies often make data and reports available, and more and more state and local governments are publishing data on open data portals. Open Knowledge International, Code for America, and the Sunlight Foundation collaborate on the Open Data Census Project, which documents which cities have open data and what domains they have made open (http://us-city.census.okfn.org/). Cities and counties may also have geographic information system or data departments that produce and publish data (even if they don't have a portal). Other local agencies, such as planning departments, public health departments, police departments, and school districts, may consider it part of their mission to publish data and reports. The data that are readily accessible may not always be in the format or geography that you need, but they are a good starting point and may help you identify which people in the agencies to connect with. Regional agencies such as councils of governments or metropolitan planning organizations need to use data as part of their own planning efforts and are also good sources of information for community organizations.

Local data intermediaries. As the name implies, a local data intermediary acts as a mediator between data and local stakeholders. Local data intermediaries use data to describe their communities and empower communities to use data in their activities such as community building, program planning, policymaking, and advocacy (Hendey et al. 2016). Local data intermediaries may focus on one topical domain or geographic scale. In our experience, local data intermediaries are usually nonprofit organizations or applied research centers at universities, but they occasionally involve local foundations, government agencies, or collaborations among these various institutions. Local data intermediaries such as those mentioned above in the Minneapolis and Oakland examples follow a specific model focused on neighborhood-level information and building the capacity of institutions and residents to use data to improve low-income neighborhoods. This model was developed by the National Neighborhood Indicators Partnership (NNIP, www.NeighborhoodIndicators.org). NNIP has partners operating independently in more than 30 communities across the country. If your community participates in NNIP you could contact the local partner as a first step.

Universities and colleges. If your community does not have a local data intermediary or an organization working in that capacity, local and regional universities and colleges may be a good alternate resource. Some universities have applied research centers focused on serving the community. Two coalitions of universities may be helpful in identifying these centers: the Coalition of Urban Serving Universities (http://usucoalition.org) and the Consortium of University Public Service Organizations (www.cuspso.org). Individual researchers who are interested in the issues of fair housing and advancing opportunity are also possible sources of relevant data and expertise. In addition, university students studying urban affairs, sociology, planning, geography, or public policy may be learning data and mapping skills in experiential learning programs that could be used in projects or internships for

community organizations. For example, the University of Minnesota's Center for Urban and Regional Affairs has several programs that provide space and resources for community-initiated research projects.[16]

State Data Centers. Another potential data resource to identify is the entity or entities participating in the Census Bureau's State Data Center Program.[17] These entities may be nonprofits, government agencies, or universities. The Census Bureau established this program in the late 1970s to make its data available and educate communities on how to use the data. The Census Bureau also has a Data Dissemination Branch that is an outreach and training resource whose mission is to make Census Bureau statistics more accessible, useful, interesting, and relevant. Data dissemination specialists are available to provide free presentations, workshops, consultations, or trainings in person or via webinar. To identify the data dissemination specialist serving your area, contact 1-844-ASK-DATA or census.askdata@census.gov.

Federal Reserve Banks. One of the goals of the Federal Reserve System is to promote consumer protection and community development. The 12 regional Federal Reserve Banks, among their other responsibilities, have community development departments that engage in a variety of activities, including research and data analyses. Some banks provide research reports and data on their websites. The community development staff may also serve as conveners on important issues for the community. For example, the Federal Reserve Bank of Richmond co-convened Charlotte Data Day 2015 with the University of North Carolina–Charlotte Urban Institute to explore demographic trends in the Charlotte region and conduct workshops on accessing and using data.[18]

Community-based organizations. Service providers, advocacy organizations, fair housing groups, community services boards, and community development corporations, among others, can also be good sources of information about protected groups whose size or dispersion makes it difficult to produce reliable estimates in surveys like the ACS or who may not be well served by government agencies and represented in administrative data. Fair housing organizations often conduct audits and test for discrimination. They may be able to share information about local trends and challenges different protected classes face. Advocates for people with disabilities may be good references for information on problems encountered by their constituents. The example above on the Philadelphia Chinatown Development Corporation shows how such organizations have the ability to convene populations that may not be traditionally well represented in forums or focus groups to solicit their input on issues of fair housing and barriers those groups face.

DATA PLATFORMS

Several online platforms make secondary data available. These platforms typically provide already-calculated indicators as well as charts, graphs, and maps to visualize the data. Some platforms may allow you to download the data or visualizations. These platforms can lower the burden on organizations and remove the need to manipulate large data files or calculate new statistics. Here we briefly describe a few platforms that provide data below the city level and cover multiple topics and data sources. We describe platforms to access specific data sources in part II.

American FactFinder (http://factfinder.census.gov). The US Census Bureau makes much of its data available through American FactFinder, including the Decennial Censuses, population estimates, and the ACS. The Community Facts section of the platform can be used to quickly access a set of basic indicators and tables on a state, county, city, town, or zip code. The Guided Search and Advanced Search can also be used to find particular ACS tabulations at specific geographies, including census tracts and legislative districts.

Community Commons (http://www.communitycommons.org/). Community Commons offers data on health, the environment, social and economic issues, and demographics. Some tabulations by race and ethnicity are available. The platform allows users to build reports, view interactive maps at geographies down to the zip code level, and share maps they create on Community Commons. It includes other collaborative features, including the ability to create a "hub," a group of users with whom to share data and information.

HUD's affirmatively furthering fair housing tool (https://egis.hud.gov/affht/). HUD's AFH requires jurisdictions to analyze national publicly available data as well as local data and knowledge to assess local and regional fair housing issues and contributing factors. HUD grantees are required to use the tables and maps provided in HUD's online affirmatively furthering fair housing tool in their AFH. HUD provides data on race and ethnicity, national origin, LEP, publicly supported housing, housing problems, housing tenure, poverty, and disability and indices on environmental health, school proficiency, job proximity, labor market, and transportation (HUD 2017).

HUD CPD Maps (https://egis.hud.gov/cpdmaps/). CPD Maps was designed for HUD grantees and the public to compare data on housing and economic conditions as part of their consolidated planning process. Data are available down to the census-tract level. You can create maps and generate reports with data from a number of national public data sources (CPD 2016).

National Equity Atlas (http://nationalequityatlas.org/). The National Equity Atlas aims to provide a tool for communities that want to create a "more equitable, sustainable, and resilient economy." The data available for the 100 largest cities and 150 largest regions are helpful in identifying disparities by

race and ethnicity, as well as gender, nativity, and ancestry when possible. Visualizations are available in the form of charts and maps, and images can be downloaded. Only limited data are available at geographies below the city level.

PolicyMap (http://www.policymap.com/). PolicyMap's interactive maps let users visualize a large number of data sources with indicators on demographics, real estate, health, and jobs. The data can be mapped down to the census-tract level for most indicators, and some tabulations by race and ethnicity are available. Additional features are available with the subscription service, including the ability to download data and upload your own datasets.

TOOLS TO ANALYZE AND DISPLAY DATA

Data can be complicated to analyze and understand. Data visualizations allow stakeholders to understand information quickly and grasp its meaning. Not every type of visualization—charts, maps, or tables—is appropriate for every type of data. Tables, for example, are a simple way to display information, but they do not usually tell clearstories. Interpreting tables of neighborhood- or census tract–level data can be particularly difficult. As an example, figure 1 shows the beginning of a table that classifies all 179 census tracts in the District of Columbia by their majority racial composition. Figure 2 displays the same information as a thematic map. In contrast to the table, the map vibrantly and decisively illustrates the nature and extent of racial segregation in DC.

FIGURE 1

Table of Racial Composition of Census Tracts in the District of Columbia

Geo2010	tract_comp	majblack	majwhite	mixedngh
Full census tract ID (2010)	Tract Racial Composition 1=White 2=Black 3=Other	Tract Population in 2010-14 is at least 75% Black	Tract Popuation in 2010-14 is at least 75% White	Tract Population is not 75% white or 75% Black
11001000100	1	0	1	0
11001000202	1	0	1	0
11001000300	1	0	1	0
11001000501	1	0	1	0
11001000502	1	0	1	0
11001000600	1	0	1	0
11001000701	1	0	1	0
11001000702	1	0	1	0

Source: NeighborhoodInfo DC analysis of the American Community Survey 2010–14.
Note: Only a portion of the table is displayed.

FIGURE 2

Map of Racial Composition of Census Tracts in the District of Columbia
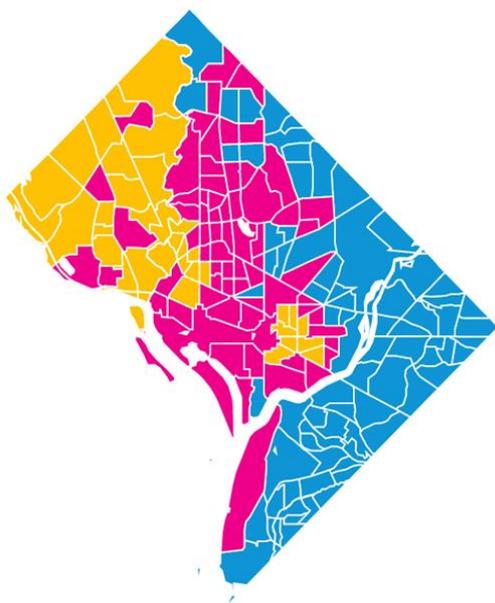
Source: NeighborhoodInfo DC analysis of the American Community Survey 2010–14.

Most of the data platforms described above allow users to visualize the platforms' data in maps or charts. But you may want to do your own visualizations to tell the story of your community. One of the most widely available and popular programs is Microsoft Excel. More recent versions, such as Excel 2016, enable users to create maps as well as charts and graphs. Free software or online tools such as Tableau and Google Charts and mapping software like QGIS or CartoDB are available, but they frequently require a wide range of technical skills and capacity to be used effectively. These tools may also have paid versions that include more features. The website PolicyViz.com documents many resources to help you create data visualizations.

ADDITIONAL RESOURCES

Finally, we highlight a few additional resources on data sources, working with data, and negotiating for access to data. Many of the data sources described in the first two resources are discussed in more detail in part II.

Catalog of Administrative Data Sources for Neighborhood Indicators (Coulton 208). This catalog documents 42 sources of state and local administrative data that can be used to create neighborhood indicators.

List of National Datasets with Small-Area Data.[19] NNIP maintains a list of national, publicly available data sources that have data available at geographies smaller than counties or cities. The list includes links to find the data and brief descriptions of the source.

Measuring Performance: A Guidance Document for Promise Neighborhoods on Collecting Data and Reporting Results (Comey et al. 2013). This document was developed for Promise Neighborhood grantees, but it has good information on data collection strategies (such as neighborhood surveys) and developing data-sharing agreements (particularly with school districts).

Neighborhoods by the Numbers: An Introduction to Finding and Using Small Area Data (Mallach 2017). This guide from the Center for Community Progress focuses on data related to community development and provides additional detail on conducting your own physical conditions surveys.

NNIP's Lessons on Local Data Sharing.[20] NNIP offers brief guidance on negotiating for local administrative data including key elements of a data-sharing agreement and sample agreements from NNIP partners.

Part II: National and Local Data Sources for Fair Housing and Access to Opportunity

This section discusses key topics to include in an analysis of fair housing and access to opportunity. Though there are many ways to organize fair housing and opportunity access issues, the topics included in this guide are modeled on those in HUD's Assessment of Fair Housing (AFH) template from the Affirmatively Furthering Fair Housing (AFFH) rule. Although this guide is designed to be useful for any attempt to understand and contextualize local fair housing and opportunity access conditions, it is particularly useful for organizations seeking to use local data to supplement the data HUD provides through the AFH process

For each topic, we include example indicators, data sources that are readily available nationally or in most jurisdictions, and sources that require more advanced data capacity or are harder to obtain. Readily available national data sources can be used to create a baseline picture of a community or to provide comparative data for communities across the United States. It is often useful to supplement these readily available data with data that may be more technically difficult to compile, require more manipulation, or require building relationships with the agency or organizations that house the data. (For advice on the trade-offs between using easily accessible data and those that may require more time and effort, see part I.) In some cases, such as the AFH process, using local data may be a required step; in other instances, technically complex data may yield more accurate or timely information. Even within the advanced data category there is variation in difficulty; in the following sections we note when advanced data capacity or long-term relationship building would be required.

A few general reminders and cautions are handy before delving deeper into the topics and their corresponding data sources. If possible, disaggregate (i.e., break out) the data by the characteristics of interest. For instance, housing affordability might be a useful indicator, but housing affordability by racial group or gender of the head of household tells a more nuanced story, one that is essential for advancing choice and opportunity. However, keep in mind when disaggregating data that estimates become more unreliable when you narrow your populations of interest. Be careful of reporting survey estimates for small geographies (such as neighborhoods) or for small subpopulations (such as those representing less than 10 percent of the total population). Because estimates like these are less reliable than estimates of larger geographies or populations, considering any associated confidence intervals or margins of error is critical.

Remember that local data intermediaries and other institutions (see part I) may already routinely obtain, clean, and share some of these more complicated data and indicators that are described in the sections that follow as "advanced options." When protected classes of people are involved, it is important to connect with associated organizations and government offices, as they may have data related to the populations they serve or knowledge on how to interpret the data about their community. For instance, organizations representing people with disabilities may have conducted surveys of sidewalk conditions or building accessibility points, or a fair housing group may have audited a developer's compliance with the Fair Housing Act.

The data sources discussed in part II are listed in appendix A (available online). If you are using this guide for the AFH, please see appendix B (available online), which provides a crosswalk of each topic and the relevant sections of the AFH. The following sections draw heavily on Coulton's "Catalog of Administrative Data Sources for Neighborhood Indicators" and the list of national publicly available data sources maintained by the National Neighborhood Indicators Partnership described in part I.[21]

Demographics and Segregation

A snapshot of demographics and patterns of segregation in a community is a common starting point for fair housing and access to opportunity analysis. Demographic characteristics include protected classes under the Fair Housing Act of 1968, such as race, national origin, disability, religion, sex, and household type (such as families with children). In a neighborhood analysis of fair housing and opportunity access, you may be interested in other characteristics such as age or education level. The size of population groups and their location in a region have implications for fair housing and strategies to advance opportunity. In addition to being valuable in their own right, population estimates are used to calculate indicators for many other topic areas, serving as the denominator in percentages and rates (e.g., the number of violent crimes per 1,000 people).

The most widely used sources of demographic data are the Decennial Census and the American Community Survey (ACS). The Decennial Census has the most accurate population counts, especially for small geographies (population counts by race, ethnicity, age, and gender are available down to the census-block level). However, it is only conducted every 10 years. The ACS provides more frequent estimates on population characteristics down to the block-group or census-tract level that include breakdowns by race, ethnicity, age, gender, household type, disability status, and other factors (see box 2 for more detail on the ACS).[22]

The American Community Survey

The ACS is an ongoing, large-scale survey that is sent to a sample of US households each month. It covers a range of topics including housing, education, language, health insurance, disability, employment, income, and poverty. Because the ACS is administered nationwide, it is useful for comparing conditions and populations across neighborhoods, cities, and regions.

Estimates are produced at different geographic levels from the block-group level up to the nation overall. The estimates are available annually for geographies with populations over 65,000; estimates for geographies with smaller populations are released every five years. The most recent releases in the fall of 2016 were the one-year data for 2015 and the five-year data for the 2011–15 period. The five-year estimates are created by pooling the survey sample across time to increase the reliability of the estimates.

As the ACS is a survey, each estimate produced has an associated margin of error. Generally, the margins of error are biggest for estimates representing small geographies or small subpopulations (e.g., Native American children under age 5). The margins of error for these groups can be quite large, even when using the five-year data, indicating there is considerable variability around the estimate. Users should be cautious about using these estimates.

The margins of error should also be taken into account when comparing any two estimates. In general, for example, if a 2015 estimate plus or minus its margin of error overlaps with the corresponding 2014 estimate and its margin of error, then you cannot conclude that those two estimates are different. In other words, the difference is not statistically significant (see the instructions on statistical testing noted below). The Census Bureau has released a tool to help with statistical testing.[a]

Changes are more difficult to detect using the five-year data, as each year only one-fifth of the sample is replaced. For example, 2010–14 data were released in 2015; the 2011–15 data drop the sample from 2010 and add new sample households that were surveyed in 2015. The US Census Bureau generally recommends not comparing data from overlapping time periods. For suggestions on using advanced statistical methods to make such comparisons correctly, see appendix 4 of *A Compass for Understanding and Using American Community Survey Data*.[b]

ACS online resources include the following:

- general handbook on ACS basics[c]
- ACS statistical testing tool[d]
- definitions, codes, and instructions for statistical testing[e]
- ACS data users group (https://acsdatacommunity.prb.org/)

Note: The short-form census questionnaire used by the Decennial Census is not considered a survey. It is meant to count every person living in a household or group quarters in the United States for a complete count of the population. The old long-form questionnaire, in contrast, was a survey that only a sample of households received. For more details, see https://www.census.gov/geo/reference/gtc/gtc_ct.html.

[a] "Statistical Testing Tool," ACS, US Census Bureau, last modified February 15, 2017, https://www.census.gov/programs-surveys/acs/guidance/statistical-testing-tool.html.

[b] US Census Bureau, *A Compass for Understanding and Using American Community Survey Data: What Researchers Need to Know* (Washington, DC: US Department of Commerce, 2009), https://www.census.gov/content/dam/Census/library/publications/2009/acs/ACSResearch.pdf.

[c] See note b.

[d] "Statistical Testing Tool," American Community Survey, last modified February 15, 2017, https://www.census.gov/programs-surveys/acs/guidance/statistical-testing-tool.html.

[e] "Code Lists, Definitions, and Accuracy," American Community Survey, last modified December 2, 2017, https://www.census.gov/programs-surveys/acs/technical-documentation/code-lists.html.

For small subgroups or less populated areas, these estimates may either be less precise or suppressed at the block-group level (see box 3 for more details about the small geographic areas we refer to throughout part II). We recommend using census tract–level estimates for greater reliability, unless you are using block-group data to aggregate up to a custom geography. Even at the census-tract level, you should examine margins of error to ensure your estimates are reliable.

BOX 3

Census Bureau–Defined Geographies

Much of the federal data on small geographies are available at one or more of the Census Bureau–defined geographies: block, block group, or census tract.

Block. Blocks are typically bounded by streets or features like railroad tracks and streams. These are the smallest of the Census Bureau–defined geographies.

Block group. A cluster of blocks forms a block group, with a population ranging from 600 to 3,000 people.

Census tract. One or more block groups form a census tract. Census tracts vary in the size of the land area they cover depending on the population density. Typically, 1,200 to 8,000 people live in each tract, with the goal of having about 4,000 people in each tract. Census tracts are defined with local input; they do not cross county lines but may cross municipal boundaries.

Sources: "Geographic Terms and Concepts, Block," US Census Bureau, last modified January 10, 2017, https://www.census.gov/geo/reference/gtc/gtc_block.html; "Geographic Terms and Concepts, Block Groups," US Census Bureau, last modified December 6, 2016, https://www.census.gov/geo/reference/gtc/gtc_bg.html; "Geographic Terms and Concepts, Census Tract," US Census Bureau, last modified December 6, 2016, https://www.census.gov/geo/reference/gtc/gtc_ct.html.
Note: See part I for an in-depth discussion of other small-area geographies.

Mapping demographics and segregation data can help analyze segregation patterns (see box 4 for more detail on segregation data). Several publicly available interactive sites allow mapping of ACS and Census data. The data and mapping tool developed by HUD for the AFFH rule offers maps by

race/ethnicity, race and ethnicity trends, national origin, limited English proficiency status, and disability type down to the census-tract level.[23] The tool was designed in part to evaluate patterns of segregation among protected classes, and it is a useful resource for visualizing spatial disparities. PolicyMap also allows mapping of demographic data and includes some additional segregation and diversity indices.

The following are example indicators (data sources for these indicators are in parentheses):

- total population (Decennial Census, ACS)
- percentage of population that is Hispanic or Latino (ACS)
- percentage of population age 65 and over that is female (ACS)
- percentage of households headed by single parents with children (ACS)

Segregation and Diversity Measures

Demographic data can be used to develop measures of segregation within a city or region by creating segregation indices that average the demographics of a smaller geographic area (e.g., census tracts) to show the extent of segregation in a larger unit (e.g., a city). A dissimilarity index is a measure of the evenness with which two demographic groups are distributed across a geographic area. A higher score along this index indicates greater segregation between where two populations live. Another common measure of segregation involves isolation of a demographic group. For example, an isolation index might measure how likely a black person is to live in a majority black census tract.

See "Measures of Residential Segregation" in Iceland, Weinberg, and Steinmetz (2002, appendix B) for additional details and mathematical formulas for these types of indices. Nate Silver discusses several indices and how they play out across the country.[a] Some of these indices have been calculated and are available for download:

- Brown University provides dissimilarity indices, isolation indices, and measures of exposure for cities and metropolitan areas.[b]
- DiversityData.org provides dissimilarity, isolation, and exposure indices for full population and child population by metro area (http://www.diversitydata.org).
- The National Equity Atlas offers a diversity index for the nation, states, regions, and cities.[c]
- PolicyMap provides a Theil index (county and metro areas) and diversity index (county and census tracts) (https://www.policymap.com).

[a] Nate Silver, "The Most Diverse Cities Are Often the Most Segregated," *FiveThirtyEight*, May 1, 2015, https://fivethirtyeight.com/features/the-most-diverse-cities-are-often-the-most-segregated/.
[b] "Residential Segregation," Diversity and Disparities, Brown University, accessed July 26, 2017, https://s4.ad.brown.edu/Projects/Diversity/segregation2010/.
[c] "Diversity Index," National Equity Atlas, accessed July 26, 2017, http://nationalequityatlas.org/indicators/Diversity_index.

Generally, demographic data will come from federal sources such as the ACS. Some datasets may include data that allow you to approximate demographic information for population groups. For instance, the Longitudinal Employer-Household Dynamics (LEHD) can be used to approximate racial demographic data by looking at the racial composition of working residents. Using proxies in this way may generate a timelier estimate, but at the cost of not measuring the full population.

In some regions, local planning departments or regional agencies may create their own population estimates at the census-tract level. If your community is changing rapidly, several local sources of data might provide more information about those changes (more detail about these data sources are provided in the sections that follow):

- *Vital statistics* follow the trends in the number of births by race and neighborhood.

- *Public school enrollment* monitor how public school enrollment and student demographic characteristics change over time. Interpretation of these figures would depend on the amount of school choice and private school enrollment in your community.

- *US Postal Service Vacant Address Data* monitor patterns in vacancy. These data are based on the information from mail carriers on vacant housing and houses that receive mail. After Hurricane Katrina in New Orleans, these data (available from HUD and direct mail services) were the best information on where and how fast the city was being reoccupied.

Housing

This guide is aimed at helping local community organizations to develop neighborhood-level indicators of fair housing and opportunity access. Unfortunately, little quantified information is available on small-area fair housing complaints or litigation. Your first step in a fair housing analysis should be to connect with local fair housing agencies to see what data they have available on fair housing complaints issues. Fair housing complaints are handled both by private, nonprofit fair housing agencies and state or local government fair housing agencies. Both can be sources of information about fair housing in your area. HUD's Community Assessment Reporting Tool (CART; https://egis.hud.gov/cart/) has information on the location of these agencies.[24]

Although there is a lack of direct fair housing complaints or litigation data for small areas, we can still construct a rich understanding of the housing conditions, quality, affordability, and choice in a neighborhood. In this section, we consider housing in terms of property characteristics, housing tenure,

lending, affordability, subsidized housing, and blight and abandonment and describe data sources for each component.

Property Characteristics

Property characteristics of interest include number and size of units, building age, and type of property (single-family home versus multiunit housing). These characteristics can help describe the local housing stock and reveal gaps in housing needs. For instance, limited numbers of housing units with three or more bedrooms may indicate gaps in housing needs for large households. Building characteristics can also highlight housing needs; the age of properties may point to the need for renovation or remediation programs in certain areas.

The ACS summarizes additional housing characteristics, such as number of units, vacancy rate, age of properties, number of bedrooms, and monthly housing costs. Household characteristics related to housing, such as number of occupants per room and household size, are available from these sources as well. The comprehensive housing affordability strategy (CHAS) special tabulation of ACS data offers similar information, but with more disaggregation by income level, race, age, household size, and disability status. The CHAS tabulation has a one- to two-year lag relative to the ACS. CHAS also includes information on cost, overcrowding, and housing problems. The format of the CHAS data allows for disaggregation by three or four variables, allowing for much more specific analysis. HUD's AFFH data tool includes data from the ACS and CHAS detailing household size and type. HUD's CART details federal investments in housing within an area, including details on individual properties.

The following are example indicators:

- housing units
  - » number of housing units by type (ACS)
  - » percentage of units with three or more bedrooms (ACS)
  - » percentage of buildings built before 1950 (ACS)
  - » number of federally insured multifamily units (CART)

- household characteristics
  - » percentage of households with incomes below 100 percent of the area median family income by number of children (CHAS)
  - » number of family households with five or more people (AFFH data and mapping tool)

Local administrative data on properties are available from several government departments. Tax assessors or auditors have records containing data on taxes and property characteristics. Tax records contain data on the owner, a parcel identifier, assessed value, property class, and tax class of the property (classifications of properties taxed at different rates).[25] Property records may also contain information on type of structure, year built, size, land use, and other characteristics. The recorder of deeds may also be a useful source of information on property transactions. They are likely to have records on completed foreclosures, sales, and transfers that can be used to create indicators of how many of these events have occurred in a neighborhood over a given period.

Administrative records on property characteristics can be used separately, but they are often most effective if linked to create a full picture of characteristics for each property. Local data intermediaries are good sources for assistance in undertaking more comprehensive cataloging of property characteristics, some may have already completed such cataloging. For example, Case Western Reserve University's Center on Urban Poverty and Community Development makes integrated parcel records available through their NEO CANDO system.[26] Additionally, local tenant organizers and advocacy organizations may have more detailed information on properties they are monitoring, such as the condition of the property and the living conditions of residents. Establishing connections with on-the-ground organizations can both improve the detail and accuracy of administrative property records and create a potential end user for property data.

The following are example indicators:

- property characteristics
  - » percentage of properties that are multifamily residential (tax assessor)
  - » median assessed value in the past year (tax assessor)

- property events
  - » number of sales of single-family homes in the past year (tax assessor or recorder of deeds)
  - » number of completed foreclosures in the past year (recorder of deeds)
  - » percentage of properties that are delinquent on property tax (tax assessor)

Housing Tenure

Housing tenure refers to whether a renter or owner household is occupying a housing unit. A fair housing analysis must look at the mix of owners and renters in a community, where there are

concentrations of owners and renters, how those patterns have changed over time, and how those concentrations might overlap with areas that are racially segregated or low income. Fair housing concerns related to tenure include whether members of protected classes are systematically unable to access mortgages or if members of protected classes are unable to access ownership opportunities in certain neighborhoods.

The ACS and the AFFH data tool provide data on occupied housing units by owner or renter households. These data sources allow users to look at tenure by race and ethnicity, and the AFFH mapping tool allows them to see overall homeownership or rental rates geographically. Viewing these rates on a map can highlight disparities in which neighborhoods residents have access to homeownership or where there may be a more even mix of renters and homeowners.

The CHAS special tabulation of ACS data includes more in-depth data on housing characteristics by tenure (such as housing problems, housing burden, age, disability status, and income). The Home Mortgage Disclosure Act (HMDA) provides annual data on the disposition of home purchase and refinanced mortgage applications, including data by race, gender, and income. These data can be used to develop an approximation of the number of new homeowners and the percentage of home purchases by protected class. Pettit and Droesch (2009) document the benefits and limitations of using HMDA data as a proxy for home purchases.

The following are example indicators:

- homeownership rate (owner-occupied units/occupied units) by race and ethnicity (ACS)
- percentage of owner-occupied home purchase mortgages by race (HMDA)

ADVANCED OPTIONS

In rapidly changing neighborhoods, local data may be more useful than five-year ACS estimates to better understand the current picture of housing tenure. Many jurisdictions have property tax credits for owner-occupants, and this information (often in combination with the address to which the tax bill is sent) can be used to determine where owner-occupants are living.[27]

The following is an example indicator:

- percentage of single-family homes and condominium units with an owner-occupancy tax credit (tax assessor or auditor)

Lending

Homeownership has long been the primary driver of wealth accumulation in the United States, so access to the financial products needed to purchase a home is a primary concern for those interested in advancing opportunity. Lending discrimination continues to affect households' ability to obtain financing for mortgages, home equity, and home repair loans. Thus, the lending practices and outcomes in your area are critical pieces of a fair housing and access to opportunity analysis.

HMDA data are the primary source for information on lending practices. These data allow you to drill down into disparities in homeownership and lending access by race, gender, age, and income. You can construct indicators of loan denial rates and high-priced mortgages and identify whether lenders with high proportions of high-cost loans are concentrated in particular neighborhoods. These indicators will allow you to analyze whether there are patterns of disparities in access to conventional financing across protected classes or neighborhoods.[28] Another potential use of the HMDA data is to develop a picture of the levels of investment occurring in areas through indicators of loan value. These data have certain limitations, particularly that they cannot be used to identify discrimination directly due to the lack of variables related to credit worthiness. In addition, HMDA data do not capture cash sales of properties, so they may not accurately reflect the amount of activity in the housing market in areas with high proportions of cash sales.

The following are example indicators:

- home purchase mortgage denial rates by race (HMDA)
- average dollar amount of home purchase mortgages (HMDA)

Affordability

Housing affordability can be captured by indicators that measure the cost of housing at the household level or by overall market trends. Affordability is a significant but common barrier to opportunity that many households face. Neighborhood variations in affordability can be a substantial contributor to segregation, influencing the ability of protected classes to access certain neighborhoods. Further, affordability can serve as a significant barrier to entry in neighborhoods with high-quality public resources such as schools and public transportation that contribute to opportunity.

The most straightforward aspect of affordability is how much housing costs. The ACS has data on monthly housing costs for renters and owners, as well as median home value. These data can be used to

develop indicators that show the cost of housing in a market. Monitoring changes in housing costs over time can also be useful for understanding how neighborhoods are changing.

Housing costs are commonly compared to income levels to measure the affordability of housing relative to a household's economic situation. One common benchmark of affordability is cost burden, which is generally defined as paying more than 30 percent of household income in monthly housing costs (rent and utilities for renter households and mortgage payments, taxes, insurance, and utilities for owner households). Housing cost burden is considered severe if a household is paying over 50 percent of income in housing costs. HUD's AFFH data tool includes a table with data on households facing severe cost burden by race and household size. ACS has summary tables on housing costs as a percentage of income. CHAS data contain more specific tabulations of cost burden across several household characteristics.

Housing cost burden is an important indicator to disaggregate across different demographic groups. For instance, a fair housing and opportunity analysis might be less concerned with wealthy households that are experiencing cost burden (i.e., paying 30 percent of income in rent) but do not necessarily experience poverty as a result of their housing costs.

There are several options for mapping affordability data. HUD's AFFH mapping tool allows you to see the percentage of housing affordable at 50 percent of area median income down to the census-tract level by race/ethnicity and national origin. The AFFH map is a dot density map; one dot represents a certain number of households, and each dot is placed randomly within the census tract. PolicyMap and Community Commons allow you to map housing costs and rates of cost burden disaggregated by housing tenure. Mapping these measures allows you to analyze not just the total costs associated with housing, but also whether there are neighborhood-level variations in affordability.

The following are example indicators:

- median monthly housing costs (ACS)
- median monthly housing costs as a percentage of median income (ACS)
- percentage of renter households paying 30 percent or more of their income in rent (ACS)
- percentage of owner households paying more than 50 percent of their income in housing costs by race (CHAS)

ADVANCED OPTIONS

Local administrative data on property sales can be used to create more timely and precise indicators on how affordability is changing. Local tax assessors and recorders of deeds will have data on housing sales that can be used to create indicators of the number of sales and the median sales price for an area.

These data will be at the address level (i.e., there will be data on each individual sale), so it may need to be geocoded before use. With these data, you can analyze the housing market and highlight trends, disparities, and accelerations in home sales and prices. Regional multiple listing services also may publish reports on properties listed for sale, the inventory of properties available, and sales prices.

Generally, timely data on rents for small areas are more difficult to come by. Some jurisdictions may survey large apartment complexes annually to track these trends. Zillow offers summarized rent (and sale price) indices, and several additional indicators down to the zip code and neighborhood levels.[29] Redfin offers data on home sales and prices down to the same geographic levels.[30] Data from these sites, as well as similar local sites, can be used for a deeper analysis of the real estate market in your area. However, data compiled from these sites should be reviewed for accuracy and coverage; some listings or types of listing may not be published online and therefore may be excluded from your analysis.

Data on evictions may reveal affordability challenges and the possibility of fair housing problems. Evictions can indicate households are struggling to pay their rent. Changes in how landlords pursue evictions may indicate market pressures, such as an attempt to empty a building to renovate it and release at higher rent levels (see box 5 for more on measures of displacement). Evictions data are found in court records and may be more difficult to obtain in bulk. In many places, sheriff's offices carry out the eviction orders and data may be maintained on those orders carried out.

The following are example indicators:

- median sales price for single-family homes (tax assessor)
- percentage change in sales price (tax assessor or recorder of deeds)
- median rent list price by property type (Zillow)
- median rent listed by number of bedrooms (Zillow)

BOX 5

Hard to Measure: Residential Mobility and Displacement

Moving homes can be a positive or negative event for households. A complex set of motivations, incentives, and barriers affects households' decisions about when and where to move, including financial resources, family structure and supports, and employment locations. However, involuntary displacement, when a household is not moving out of choice, is a cause for serious concern and merits attention through public policies and programs. As one example, when investments occur in low-income areas and result in new amenities, landlords may increase rents or the city may raise property

assessments and taxes for owners, making neighborhoods less affordable to current low-income residents. Some households may be forced to move because of the rising housing costs and not be able to take advantage of the improvements in the neighborhood. This displacement from gentrification is a fiercely contested issue for many communities. The rising costs in the housing market also mean that new low-income households are shut out of these places with increasing opportunities.

Unfortunately, there is little direct information on displacement. There are no data sources that track all people who leave an area, where they move to, and why they move. However, you can better understand overall residential mobility in an area by analyzing multiple data sources. The ACS provides indicators of the percentage of the population that moved into their housing unit within the previous year and if they previously lived in another county, state, or country. The five-year averages for tract-level data also make it difficult to discern short-term trends. The ACS does not collect data on whether these moves were voluntary or involuntary, economic or personal, so they should not be used as a proxy for displacement.

Foreclosures and evictions represent involuntary moves (see discussions above in the Property Characteristics and Affordability sections). These measures are useful for forming a picture of whether residents are being increasingly displaced due to market forces. However, these measures may have only limited information on cause (e.g., affordability, discrimination, just cause), and there are no data on where the household moved or if they stayed in the same neighborhood.

Renters are more vulnerable to displacement. Property transactions and characteristics can be used to understand who is buying or currently owns rental buildings in an area and what their motivations may be for selling or maintaining their property. These data may be useful to preserving naturally affordable buildings or subsidized buildings. This type of analysis can also be used to advocate for cities and land banks to consider affordable housing when transferring publicly owned property and land or investing in new subsidized construction.

Though most data cannot measure who leaves a neighborhood, a few data sources can give you a good picture about who is entering a neighborhood. For example, HMDA data can be used to analyze the characteristics of homebuyers (and could be compared to the characteristics of homebuyers 5 or 10 years ago). Depending on the composition of the neighborhood, you may also examine changes in demographics through data on births or student records.

Several researchers have developed methods for classifying neighborhoods based on their risk of gentrification and displacement to guide public and philanthropic decisions.[a] These methods should be evaluated based on your own local knowledge and context.[b]

Note: See Coulton (2014) for a more in-depth discussion of residential mobility data.
[a] See, for example, Chris Bousquet, "Where Is Gentrification Happening in Your City?" Data-Smart City Solutions, June 5, 2017, http://datasmart.ash.harvard.edu/news/article/where-is-gentrification-happening-in-your-city-1055.
[b] See Lisa K. Bates, "Gentrification and Displacement Study: Implementing an Equitable Inclusive Development Strategy in the Context of Gentrification," Urban Studies and Planning Faculty Publications and Presentations no. 83: 2013, http://pdxscholar.library.pdx.edu/usp_fac/83/; see also, Urban Displacement, http://www.urbandisplacement.org/.

Subsidized Housing

Understanding the availability, location, and characteristics of subsidized housing is an important step in analyzing the availability and affordability of housing in an area. Federal, state, and local governments can all provide some form of subsidy for housing. Federal subsidies include public housing, privately owned project-based subsidies, such as the low-income housing tax credit (LIHTC) or Project-Based Section 8, and tenant-based subsidies like the Housing Choice Voucher program. Many states operate trust funds or tax credits that use funds to subsidize housing, and increasingly, cities and counties are developing similar structures to create affordable housing. Local jurisdictions may also fund additional vouchers to supplement federal funds.

Several nationally available sources (discussed below) provide data on subsidized housing and households that live in subsidized units. The National Low-Income Housing Coalition's National Housing Preservation Database aggregates many of the national sources, providing one place to easily access project-level data (http://www.preservationdatabase.org/). The preservation database represents the most comprehensive national data on subsidized housing. However, the database is only updated three times per year and does not contain data on state and locally subsidized housing.

HUD publishes several datasets for subsidized housing. HUD's AFFH data and mapping tool compiles data on subsidized properties (categorized into public housing, Project-Based Section 8, other multifamily, and Housing Choice Voucher program) summarized to the census-tract level and individual project data. The tool includes tables disaggregated by project category on the number of units by number of bedrooms and children; households by race, income, racially and ethnically concentrated area of poverty status, age, family type, and disability; and project-level data on units by race and family type. The tool also allows you to see publicly subsidized housing projects displayed with race/ethnicity demographic data. These data are helpful in examining concentrations or absences of racial groups in publicly subsidized housing units. A Picture of Subsidized Households has data at the census-tract level on households that receive HUD-subsidized housing by program type, income, race, disability status, and family type.[31] HUD's CART website provides similar project-level data on housing subsidy programs and census tract–level data on housing choice vouchers as well as estimated expenditure data for these programs.

Creating a comprehensive picture of subsidized housing in your community has several benefits for analysis of fair housing and access to opportunity. Understanding the demographics of households in subsidized housing can allow you to identify disparities in access to affordable housing. Knowing the location of subsidized properties within your area allows you to identify whether subsidized units are clustered in certain neighborhoods and how access to high-opportunity neighborhoods varies.

The following are example indicators:

- characteristics of subsidized households
    - » percentage of households in Project-Based Section 8 program, by race (Picture of Subsidized Households, ACS household data for denominator)
    - » percentage of Housing Choice Voucher program participants who are Hispanic (AFFH data and mapping tool)
    - » average number of months on waiting list for housing choice vouchers (Picture of Subsidized Households)

- subsidized units
    - » number of HUD-subsidized units by census tract per 100 housing units (Picture of Subsidized Households, ACS)
    - » Section 8 units with subsidies expiring in a year (National Housing Preservation Database)
    - » number of project-based Section 8 units with three or more bedrooms (AFFH data and mapping tool)

ADVANCED OPTIONS

HUD provides data on individual projects and subsidies via the Multifamily Assistance and Section 8 Contracts Database and the HUD Insured Multifamily Mortgages Database. The Multifamily Assistance and Section 8 Contracts Database contains two tables, one at the project level and one at the subsidy level. Viewing the full property data requires linking the two tables by property ID. The HUD Insured Multifamily Mortgages Database is available as a single Excel file with data on properties with HUD-insured mortgages. HUD also provides data on LIHTC buildings, available through an interactive search engine and downloadable database. Due in part to the nature of the LIHTC program, the LIHTC database has fewer variables and is not as up to date as the other HUD databases. However, as LIHTC is the largest subsidy program in the nation, the LIHTC database is a critical source for subsidized housing data despite these limitations. These databases are culled into the National Housing Preservation Database mentioned above, but accessing the files directly allows for more timely information and the ability to combine them with local subsidy data.

There are no nationally available databases of state or locally subsidized housing. However, some state and local departments of housing and community development and state housing finance agencies have publicly available data on these units. These data may come in a wide variety of formats, from simple published lists of new subsidized housing to geocoded data on the full stock of locally subsided housing. Because many properties contain subsidies from multiple sources and funding streams, these

data are most useful when linked together. It is helpful to work with a skilled data analyst or a data intermediary to catalog subsidy data in this way.

Local public housing authorities also have data on where voucher holders are renting units and may have information on which landlords are more willing to rent to voucher holders (and where). It may be difficult to access confidential data on voucher holders, but the housing authority staff may be good sources of qualitative information relevant to fair housing and increasing access to a greater range of neighborhoods for subsidized households.

The following are example indicators:

- unduplicated number of units receiving federal, state, or local subsidies (local housing departments)
- number of new subsidized units constructed in last 12 months (local housing departments)
- number of units affordable at 50 percent or less of area median income (local housing departments)

Blight and Abandonment

Legal definitions for what constitutes a blighted or abandoned property differ between jurisdictions. Generally, blighted properties are in a state of substantial disrepair, and abandoned properties are unoccupied properties that have lapsed on property taxes for some time. The number and location of blighted or abandoned properties and the neighborhood demographics in areas with disproportionate shares of such properties are all possible indicators related to fair housing and opportunity.

Blighted and abandoned properties can have a twofold effect on fair housing and access to opportunity. A concentration of properties that have fallen into disrepair can have a negative impact on neighborhood quality and opportunity. Additionally, an overrepresentation of members of protected classes in neighborhoods with high levels of blight or large numbers of abandoned properties may indicate discriminatory housing practices.

Federally subsidized housing is an easy place to start evaluating housing conditions as HUD releases the results of physical inspections of these properties online via the Real Estate Assessment Center (REAC) Online Systems. Failing scores on physical inspections can trigger the loss of a property's subsidy. The age of a property can also be used as an indicator that approximates housing quality, particularly in low-income neighborhoods. Older properties may be more likely to need repairs, and the building materials associated with particular time periods may present additional concerns (e.g., lead

paint or pipes). Utility companies or government utility departments may have data on where punitive measures, such as loss of water or electricity, have been taken on properties. Additionally, HUD has an agreement with the US Postal Service to publish data on vacant properties; accessing these data requires registering and verifying that you are a member of a government agency or nonprofit.[32]

The following are example indicators:

- housing occupancy status
    - » percentage of long-term vacant residential addresses (vacant over 12, between 12 and 24, and between 24 and 36 months) (US Postal Service Vacancy data)
    - » change in long-term vacancy rate over time (US Postal Service Vacancy data)

- housing quality
    - » number of properties with failing REAC score (HUD REAC)
    - » percentage of housing units built before 1950 (ACS)

ADVANCED OPTIONS

Local data on blight or abandonment are dependent on local laws and regulations. Regulatory or consumer affairs departments may have data available on properties with reported code violations or from vacant property registries. If such consolidated sources are unavailable, data may need to be compiled from multiple sources. Different tax rates for vacant or blighted properties may affect reporting rates for vacant properties. Tax assessors also have data on properties delinquent on property taxes, which may indicate abandonment or financial distress for the owner. Some jurisdictions may also maintain a vacant or foreclosed property registry.[33]

Several agencies may have data that can be used to create indicators that are indicative of housing quality or blight. The local assessor typically maintains information on the year a property was built, construction type, and whether a major renovation was completed. Health departments collect data on children who are screened for lead exposure. Local agencies with lead abatement programs may also have data on properties that have participated in the programs.

Many local governments operate centralized service request systems, also known as 311 systems, which can be good sources of information about neighborhood distress. For example, residents may use 311 to report code violations with properties or other maintenance issues like trash, dirty streets and alleys, or clogged storm drains. This type of data should be analyzed in conjunction with other indicators; low numbers of reports may not indicate that these issues do not exist within a neighborhood, but rather that residents do not believe their complaints will be addressed.

The recorder of deeds records changes in property ownership, including completed mortgage foreclosures. Obtaining information on properties actively in the foreclosure inventory may be more challenging. Some jurisdictions have a judicial foreclosure process in which a *lis pendens* is filed with the recorder of deeds, but it can also be filed in other circumstances when a lawsuit involves the ownership of a property. Nonjudicial foreclosure states may have notifications of default or intent to foreclose that are recorded publicly. None of these sources are definitive proof of blight or abandonment, but taken together they may indicate areas with higher levels of distressed properties.

The following are example indicators:

- housing quality and blight
    - » percentage of residential properties built before 1970 (local tax assessor)
    - » number of blighted properties (regulatory or consumer affairs office)
    - » percentage of children with elevated blood lead level (health departments)
    - » percentage of residential properties with housing code violations (regulatory or consumer affairs office)
    - » number of properties with multiple 311 complaints related to litter or trash (local 311)

- vacancy and distress
    - » number of completed mortgage foreclosures per 1,000 single-family homes and condominiums (recorder of deeds)
    - » number of tax foreclosures (tax assessor)
    - » percentage of homes with delinquent taxes (tax assessor)

Land Use

Permits

Building permits are necessary for a wide range of activities including construction, rehabilitation, alteration, and demolition for both residential and commercial development. Because permits and the corresponding regulations are administered at the local level, there are few national sources on building permits and no nationally available sources for small areas. The US Census Bureau publishes data on building permits down to the city level via the Building Permits Survey. These data include information on number of units, buildings, and value. Data from HUD's Survey of Construction provide similar information on permits by units and type (e.g., single or multifamily). These data sources can be used to

construct indicators for analyzing the level of development and investment in a community, but they are not available at small enough geographic levels to analyze where development is occurring within a jurisdiction.

ADVANCED OPTIONS

In contrast to national sources, local agencies have rich data on building permits. Local building, housing, or regulatory departments have administrative data on permits that include permit type, such as new construction or rehabilitation, and property location. These data allow for mapping of new construction and development and creating indicators of the number of new permits, a key step in analyzing where new residential and commercial development is occurring in your area. Demolition permits may provide useful information for communities focused on revitalization and removing abandoned and blighted properties. Permit data may be publicly available online in open data portals and/or summarized for small geographic areas. If summarized data are not available, the data on individual permits will need to be transformed into indicators for neighborhoods or other relevant geography.

The following are example indicators:

- number of permits for new commercial buildings (local housing and building departments)
- percentage change in new residential construction permits (local housing and building departments)

Zoning

Zoning is a persistent fair housing and opportunity issue that is reflected in the long history of exclusionary zoning practices in the United States. Zoning regulations can dictate what type of buildings and how many units can be constructed in a given space, who can live in those buildings, and where both positive (parks) and negative (trash processing plants) amenities can be built. Certain types of households can still be excluded from an area through zoning regulation and enforcement (e.g., zoning areas for single-family use only). Zoning can also be used to encourage specific types of development of affordable housing (like inclusionary zoning and density bonuses). Because zoning, like permits, is regulated at the local level, there are no nationally available data on local zoning requirements. Any data collected on zoning will likely need to come directly from local records and zoning offices or commissions.

Zoning regulations are a complex network of requirements, restrictions, and exceptions. This complexity can make it difficult to create useful indicators. Local records on zoning regulations and which areas are zoned for a specific use will likely be available online. Many jurisdictions have publicly available zoning maps, but these alone are unlikely to generate useful indicators for analysis. Property records present another avenue for accessing zoning data. The planning department or zoning office will have records on how an individual property or lot is zoned. These records may be mined to create indicators of how many parcels in an area are zoned by use, though depending on the enforcement of zoning regulations in your area the actual use of a property may not match the zoned use. Even if you are unable to access property- or parcel-based zoning data, zoning information can add further nuance to population and segregation data. For example, wealthier areas of cities or suburban communities may have zoning restrictions on the size or height of new buildings or not allow for multifamily housing, which may functionally exclude lower-income families.

The following are example indicators:

- number of parcels zoned for residential use (planning department or zoning office)
- number of parcels zoned for multifamily use (planning department or zoning office)
- percentage of parcels in census tract zoned for single-family detached housing (planning department or zoning office)

Disability Access

Data on people with disabilities are important to consider in all elements of a fair housing and opportunity analysis, and each of the sections of this guide should be considered in relationship to people with disabilities alongside other protected categories. This section describes how common datasets might define disability and considers accessibility in housing for people with disabilities.

Disability and Services

Data sources discussed earlier in this guide can start to develop a picture of where people with disabilities live and which subsidized housing programs they might be participating in. Basic indicators on the number of people with disabilities can be created using the ACS. As of 2008, the ACS uses six nonmutually exclusive definitions of disability:

- hearing difficulty: deaf or having serious difficulty hearing
- vision difficulty: blind or having serious difficulty seeing, even when wearing glasses
- cognitive difficulty: because of a physical, mental, or emotional problem, having difficulty remembering, concentrating, or making decisions
- ambulatory difficulty: having serious difficulty walking or climbing stairs
- self-care difficulty: having difficulty bathing or dressing
- independent living difficulty: because of a physical, mental, or emotional problem, having difficulty doing errands alone such as visiting a doctor's office or shopping

The ACS allows for disaggregation of disability status by additional characteristics such as race, age, sex, employment status, poverty status, and receipt of food stamps. These data can provide more demographic information on people with disabilities. Additionally, the ACS has data on participation in the Supplemental Security Income program, which has an eligibility requirement that a person have a severe enough disability that they cannot work. Participation in this program can be used as a proxy for low-income people with severe disabilities. HUD's AFFH mapping tool uses the ACS definitions to create dot density maps of disability by type and age group. The HUD data also include tables showing disability by type, age group, and publicly supported housing type. The ACS data and AFFH maps can help develop a picture of the prevalence of disability in an area and whether people with disabilities are concentrated in certain neighborhoods.

There are limitations to the ACS data on disability. The ACS definitions are broad categories of disabilities and include wide variations in the potential severity of a disability. Additionally, because of changes in how the Census defines disability, comparisons are not possible before 2008.

The following is an example indicator:

- percentage of population ages 16–64 who have a cognitive disability (ACS)

ADVANCED OPTIONS

Several sources of local data can provide additional detail about the severity and types of disability people are experiencing, as well as their type of residence. State Medicaid agencies and state mental health agencies can yield data on people with mental disabilities. These data can include a client's residence, severity of mental disability, and the types of services they receive. Record-level data will require advanced data capacity and the ability to adhere to the privacy restrictions of the Health Insurance Portability and Accountability Act of 1996 (commonly referred to as HIPAA), but these agencies may also provide summary data on their clients. Service providers such as nursing homes, care facilities, federally funded centers for independent living, and community service boards may also be

sources of data on people with disabilities. These providers may report data on their clients' type and severity of disability and residence to county or state social service departments, which may then provide summarized data on people with disabilities. Data on residences serving people with disabilities are particularly useful in the fair housing context to distinguish whether people with mental disabilities are concentrated in one type of housing, such as group homes, that may indicate a lack of independent living options.

The following are example indicators:

- number of people with severe mental illness receiving state-funded mental health services (state mental health agency)
- number of units in federally funded centers for independent living or group homes (group homes, disability advocates, state health departments)

Housing Accessibility

People with disabilities face unique challenges in the housing market. In addition to the more common struggle to find decent and affordable housing, the accessibility of housing poses an additional burden on people with disabilities. Housing accessibility, per HUD's Fair Housing Accessibility Guidelines, refers to ensuring that "public or common use areas of the building can be approached, entered, and used" by people with disabilities, but the definition of accessible housing will vary according to local laws and regulations.

Limited data are available on which housing units are accessible to people with disabilities. Property characteristics may offer some indication of disability access. Multifamily dwellings first occupied after 1991 that have four or more units are required to be accessible by the Fair Housing Act. This requirement can be used to construct a proxy of how many units should be accessible according to the act. Using the date of construction as an indicator, however, has a number of issues: exceptions to the act (e.g., townhouses and multifamily buildings with multiple floors of residential units but no elevator only need to have accessible ground floor units), potential noncompliance, the distinction between first construction and first occupancy, and data limitations of the ACS (e.g., the data available through American Factfinder is limited to structures built after 1990).

Another method for establishing estimates of accessible housing is to look at program data. Units with subsidies specifically targeted toward people with disabilities, such as HUD's Section 811 Supportive Housing for Persons with Disabilities, are required to be accessible. These data have the

limitation of only covering publicly funded accessible units. Data on Section 811 housing can be accessed via the National Housing Preservation Database.

The following are example indicators:

- number of housing units built after 1990 (ACS)
- number of Section 811 housing units (National Housing Preservation Database)

ADVANCED OPTIONS

Local data on housing accessibility vary widely. Housing agencies at the state or local level may have data on accessibility through local programs or subsidies for housing people with disabilities. These programs may have additional requirements for the production of accessible affordable housing (e.g., 5 percent of new units must be affordable and accessible). In these cases, subsidized housing data, as discussed earlier in the guide, may provide insight into the amount and characteristics of accessible housing. Additionally, housing agencies have data on the accessibility needs for people on the waiting list for subsidized housing. Accessing this information may suggest some of the unmet need for accessible units. Local housing data on property characteristics may also contain information on year of construction, which might serve as a proxy for units that must comply with the Fair Housing Act (see the earlier sections on Property Characteristics and Subsidized Housing).

In some areas, landlords may be required to register accessible properties (as in Massachusetts through the Mass Access Housing Registry). These data may be available as summarized data or individual property records. In other areas, local disability services organizations or agencies may have information on accessible units, their location, and features that can be used to begin to compile data on accessible housing. For example, in Virginia the Fairfax County Department of Family Services' Disability Services Planning and Development and the Fair Housing Council of Greater San Antonio, Texas, develop lists of accessible housing.[34]

Civil rights lawsuits from the Department of Justice are another potential source for information on accessible housing. Housing agencies that have had complaints regarding access to housing for people with disabilities may have agreed to build additional accessible housing. Data on their progress in constructing these units may be tracked at the agency level.

The following are example indicators:

- number of accessible rental units (local housing registry)
- number of multistory, multifamily buildings with elevator access (local housing registry)

Access to Opportunity

A full analysis of fair housing and opportunity goes beyond the housing in which people live. Quality education, economic stability, health, and safety all help determine whether a neighborhood provides opportunities for success to its residents. This section outlines how to access and use data on these elements of neighborhood quality: early care and education, education, employment, poverty and social services, infrastructure and transportation, environment, health, and public safety.

Early Care and Education

Research has shown that the early years of children's lives have an outsized impact on their future. Access to high-quality child care and early education can improve future educational, social, and health outcomes. Although the ACS includes limited information in this domain, its data could be used to calculate the need for services based on the number of young children and young children in poverty. The ACS also provides information on nursery school and preschool enrollment, but it has no measures of quality.

Obtaining a list of child care providers and their locations for an area is possible with a little leg work, and once obtained, the list can be compared to estimates of the number of young children. The federal Head Start office maintains a "Head Start Locator."[35] States have licensing agencies for child care providers, and most states have a child care resource and referral agency that maintains lists of licensed providers, including both center-based and family home–based providers. These agencies may also maintain information about the number of slots, ages served, and ratings for providers who participate in quality rating systems. The parameters and robustness of quality rating systems vary across states. Some early childhood education centers also seek national accreditation such as with the National Association for the Education of Young Children.[36]

The following are example indicators:

- percent of children in nursery school or preschool (ACS)
- number of slots for licensed early child care and education providers by children's age group (local child care resource and referral agency)

ADVANCED OPTIONS

Although describing what types of child care are available is fairly straightforward, getting information on *who* is enrolled in quality child care is challenging. You would need to get this information from each

provider, unless there are programs administered by school districts or Head Start agencies that might have data across several provider locations. Data from prekindergarten programs may be integrated with kindergarten to grade 12 data, particularly where there is state-funded prekindergarten.

It may also be possible to obtain data on federally funded subsidies for child care that are managed by state and local human services agencies. These data can be used to create indicators of the percentage of eligible families receiving child care assistance. Getting this information at the neighborhood level requires significant experience handling sensitive data.

The following is an example indicator:

- percentage of eligible households receiving child care subsidies (departments of human services and ACS [for determining number of eligible households])

Education

The availability of high-quality education is one of the core elements of opportunity. Access to a high-performing school has a direct impact on a child's future employment, wealth, and health status. The availability of high-performing public schools, student demographics, the experience of students with disabilities, and the geographic distribution of quality schools are essential elements of a fair housing and opportunity analysis.

Documenting which schools are in a neighborhood is a good first step for evaluating education access. The National Center for Education Statistics' Common Core of Data has detailed information on the number, grade level, and student demographics of schools. Their website has an interactive tool to access these data as well as the ability to download data files. You will also need to understand your school district's policies on school choice and attendance zones. Some schools are neighborhood based; other schools draw students from wider areas such as entire cities.

HUD's AFFH mapping and data tool provides a school proficiency index based on fourth-grade performance outcomes at schools within a 1.5-mile radius of a census block group (a higher score indicates greater school proficiency). The mapping tool allows you to view this index alongside race/ethnicity, national origin, and family status at the census-tract level. This index is meant to standardize across areas, but because each state has different tests to measure proficiency, comparisons across state lines may not be appropriate.

Other sources have data for more grade levels and indicators of the percentage of proficient students. National data are available from EDFacts for school- and district-level proficiency rates by

race/ethnicity, sex, disability status, limited English proficiency status, economically disadvantaged status, migrant status, and homeless enrolled status. Both EDFacts and Common Core of Data have about a year lag in reported data (e.g., school year 2014–15 data are reported in 2016). The Civil Rights Data Collection offers data disaggregated by race on student absenteeism and discipline in addition to school demographics, advanced placement courses, and SAT and ACT scores, although the data are only collected biannually.

Often public school districts, boards of education, or state departments of education will publish reports detailing aggregated data on school and student performance. These reports may contain enough information to conduct school-level analysis of education outcomes, keeping in mind that the student population may not reflect the neighborhood the schools are in depending on enrollment policies. However, if the data are provided only at a higher geographic level, or if there are additional characteristics you want the data to be cut by (e.g., race, disability status, gender), you may want to pursue student-level data directly from the source. Remember that information on private schools is not captured in these data sources. The ACS provides data on the share of students in private schools, so you can use that context to interpret the indicators for public school students.

The following are example indicators:

- student demographics by school
    » enrollment by race/ethnicity (Common Core of Data)
    » enrollment by sex (Common Core of Data)
    » student/teacher ratio (Common Core of Data)
    » percentage of students on free or reduced-price lunch (Common Core of Data)

- student proficiency
    » math proficiency by grade level by disability status (EDFacts)
    » Reading and language arts proficiency by grade level by homeless enrolled status (EDFacts)
    » high school graduation rate (EDFacts)

- student attendance and discipline
    » number of students with disabilities who received suspensions (Civil Rights Data Collection)
    » number of absentee days by race (Civil Rights Data Collection)

Understanding how students living in a neighborhood are performing in school, rather than looking at school-level indicators, gives you a better picture of that neighborhood's access to educational opportunities. In cities with significant numbers of public charter schools and choice about which school to attend (compared with cities where attendance is based on the student's residence), school-level measures of student performance (suggested above) may not be useful for evaluating educational opportunity for a neighborhood as the students attending the school may come from all over the city. Most school districts do not regularly report data by where students live. Measuring student performance in this manner requires working with the school district to negotiate access to individual student records. In this situation, we recommend partnering with an entity, such as one of those described in part I, that is more experienced with handling confidential data or may already have an established relationship with the school district. Data on individual outcomes of children are highly confidential and are covered by privacy laws such as the Family Education Rights and Privacy Act. Developing datasets and indicators from these data requires careful consideration of security and deidentification (removing identifying features from the data) processes.[37] Other indicators may be useful measures of school quality and student performance, such as the rate of chronic absences (missing more than 10 percent of school days) or student suspensions by race, can be created with access to student records.[38] Having information like this is key for improvements in educational opportunity over the long term and can help organizations advocate for improved policies in areas like student discipline that would keep more students in school and in the classroom.

The following are example indicators:

- percentage of students meeting reading proficiency standards by neighborhood (school district)
- rates of chronic absenteeism by race (school district)
- percentage of students who have been suspended by race (school district)

Employment

Employment serves as the primary income source for most individuals and families, and the availability of quality jobs is an important factor for advancing opportunity and upward mobility. Analysis of employment data may focus on the availability of jobs, the match between the skills and education levels of populations and those required by jobs, disparities in wages or employment by sector, and neighborhood unemployment and employment rates.

The Bureau of Labor Statistics at the US Department of Labor operates two main data collection programs on employment and wages: Current Employment Statistics provides industry data on employment, hours, and earnings at the county level, and Local Area Unemployment Statistics provides monthly data on employment and unemployment for counties and the 50 largest cities.

The ACS updates its employment data only annually, but they are available at lower geographic levels than the data from the Bureau of Labor Statistics (small areas like census tracts, however, will require you to use the rolling five-year estimates). Data on employment status, commuting time and means, industry, and occupation are available at the census-tract level by multiple demographic characteristics. Although some ACS data are available at the block-group level, there are fewer demographic categories by which to disaggregate, and the estimates are less reliable. The ACS data can be used to construct indicators of rates of employment or unemployment and labor force participation, which can be used to identify areas or populations that lack access to employment opportunities.

The LEHD Origin-Destination Employment Statistics (LODES) created by the Census Bureau can be used to understand the characteristics of workers (race, age, gender, wage level, and industry) by their job location and by their residence down to the census-block level.[39] The jobs data are based on wage records collected in the administration of state unemployment insurance, and residence is based on federal administrative tax and program data. Individuals who are self-employed or not covered by unemployment insurance, such as those working in agriculture, private households, or railroads, would not be represented in the data. Federal workers are represented in the data beginning in 2010. The LODES data include a file that has the origin and destination of workers, which allows for analysis of where workers in an area live and where people who live in an area work. The spatial relationship between where people work and where they live is important for analyzing neighborhood opportunity.

The easiest way to access LODES is to use the OnTheMap mapping system (https://onthemap.ces.census.gov/) to explore and visualize the data and download reports. Updates to LODES are usually annual and typically lag by at least two years. HUD's AFFH mapping and data tool provides a jobs proximity index that can be used to access LODES data. The index is a measure of the distance to all potential job locations scaled by the number of jobs and size of the labor market. A higher score on the job proximity index indicates that an area has greater access to jobs and less competition for jobs. The HUD tool also includes a labor market engagement index, which is constructed from the unemployment rate, labor force participation rate, and the percentage of people with a bachelor's degree or higher. This index can be used to identify which neighborhoods have residents who are more connected to the labor market (a higher score indicates greater engagement in the labor force). The

mapping tool allows you to view both indices alongside race and ethnicity, national origin, and family status at the census-tract level.

Obtaining data on business establishments may also help develop a picture of employment opportunities in your area. The Census Bureau's County Business Patterns data provide information down to the zip code level on business establishments by number of employees and industry. County Business Patterns data can also be used to develop indicators of businesses that typically offer services at a neighborhood level (e.g., drugstores and laundries). For an example of this analysis, see Bajaj, Kingsley, and Pettit (2005).

The following are example indicators:

- percentage of population ages 16–64 who are employed (ACS)
- unemployment rate by factors such as age, gender, and race (ACS)
- percentage of commuters spending more than 45 minutes in transit (ACS)
- share of workers traveling more than 10 miles to their primary job (LODES)
- number of retail businesses (County Business Patterns)

ADVANCED OPTIONS

For community organizations with more advanced data and spatial analytic capacity, the LODES data can be downloaded for detailed analyses or combined with other datasets such as transportation information to create indicators of transit time to and from work using geographic information system software.

Another source for employment data is the Quarterly Census of Employment and Wages (QCEW), which is a collection of unemployment insurance data from every establishment with more than one employee. The QCEW includes information on the establishment, such as name and address, as well as information on employees, such as total number of employees and total quarterly wages. The QCEW data are the foundation of the Longitudinal Employer–Household Dynamics data that underlies LODES; accessing the QCEW data directly may provide more timely information as they are updated quarterly. Because access to these data is regulated by restrictions at the federal and state levels, using this data source will require substantial data privacy and security measures. Your state job center may be able to provide summary data below the county level. If you are able to connect with the state job center and access these data, they can be used to construct various useful indicators about employment and job availability at the neighborhood level.

The following are example indicators:

- percentage of residents of high-poverty neighborhoods earning more than $3,333 per month at their primary job (LODES, ACS)
- number of jobs within 30 minutes by transit (LODES, local transit data)
- change in employment, by industry (QCEW data)

Poverty and Social Services

Experiencing poverty can have a negative impact on health and well-being, and when this experience is concentrated in certain neighborhoods or communities, the effect is multiplied. An analysis of fair housing and opportunity must identify communities that are disproportionately poor. Identifying these areas can add context to many of the other data points you gather: Are affordable housing units disproportionately located in high-poverty areas? Are health outcomes lower in these areas? Similarly, data on social services designed to help alleviate poverty can be used to develop additional insight about whether people experiencing poverty have access to services they need and whether demand for services is increasing.

Several measures are related to poverty and receipt of social assistance. The ACS contains data on households with incomes below the federal poverty level as well as data on income as a ratio of the federal poverty level. These data can be disaggregated by household characteristics, such as race, disability status, and age. Measures of participation in the Supplemental Nutrition Assistance Program (SNAP) and cash assistance programs are available through the ACS as well.

Additional measures of low-income households include a Brookings Institution database of summarized zip code–level tax return data on earned income tax credit claimants. The Internal Revenue Service provides zip code–level summaries of tax filings by income bracket as well. Data on use of the earned income tax credit can be used to estimate usage rates and determine whether eligible households are accessing this credit. Both sources have two- to three-year lags. The Social Security Administration publishes annual zip code–level data on the population receiving Old-Age, Survivors, and Disability Insurance. These data allow you to create indicators on the number and household type of people receiving these benefits.

The following are example indicators:

- percentage of people in poverty by race (ACS)
- percentage of children in poverty by race (ACS)

- percentage of tax returns receiving the earned income tax credit (Brookings Institution)
- percentage of people age 65 and over receiving Old-Age, Survivors, and Disability Insurance (Social Security Administration)

ADVANCED OPTIONS

Local administrative records on social services and assistance can create a more detailed and timely view of the social safety net. Unemployment insurance records can be accessed from state agencies. Two sources, the Unemployment Insurance Claimant File and the Unemployment Insurance Wage Record, are collected by state employment services agencies. The claimant file contains information on the filer, address, weekly benefits, average weekly wage, and number of weeks receiving benefits. These records can be summarized to create indicators of the number of people receiving benefits and the average length of time they receive unemployment insurance. The wage file contains information on wages paid and weeks worked. These records do not contain addresses and so cannot be used alone to create neighborhood indicators. However, they can be matched by identifying information (such as name and Social Security number) to other social services records. Accessing these data will require significant relationship building with the state agency that administers the program and substantial privacy and data security capacity. If there is a data intermediary in your area, you should check to see if they have already worked with these data.

Obtaining data from Temporary Assistance for Needy Families; SNAP; the Special Supplemental Nutrition Program for Women, Infants, and Children; and other social services requires accessing confidential files held at departments of human services, although some jurisdictions may already have summarized records into a usable dataset. Public assistance records can be matched across years or programs, which would allow you to create indicators of usage over time and across programs. However, as this effort would require accessing identifiers of program recipients, this process would need to be managed by a person or organization proficient in data security and confidentiality requirements. Tracking these indicators over time and by neighborhood can show whether need is increasing or decreasing in a community (assuming no change to the programs) and where populations in need live.

The following are example indicators:

- median Temporary Assistance for Needy Families benefits (departments of human services)
- number of people receiving SNAP (departments of human services)
- percentage of eligible households receiving SNAP (departments of human services, ACS)
- average number of weeks receiving unemployment insurance benefits (unemployment insurance claims data)

- percentage of former SNAP recipients who found employment (Unemployment Insurance Wage Record, departments of human services)

Infrastructure and Transportation Access

Transportation access undergirds many elements of opportunity access. A household's ability to have and use reliable transportation can determine whether they truly have access to education, employment, and social services opportunities in their neighborhood. Infrastructure and transportation access are areas of particular concern for people with disabilities, as difficulty accessing transportation may be an insurmountable barrier to accessing opportunity for that population (see box 6 for additional discussion).

HUD's location affordability index models transportation costs by household type at the block level. However, the index is a model of what transportation costs should be for a given family type rather than data on what households are actually spending in transportation costs. A transit index based on the location affordability index is provided in the HUD AFFH data tool. The HUD AFFH mapping tool includes maps based on these data, showing estimated transit trips and transit affordability by race, national origin, and family status. Within these indices, higher numbers indicate a higher access to transit and lower transportation costs.

The following is an example indicator:

- transit index (AFFH data and mapping tool)

BOX 6

Transportation Access for People with Disabilities

Accessible transportation and infrastructure can help ensure that people with disabilities are able to move throughout the community. A lack of transportation options and poor infrastructure can create insurmountable barriers for people with disabilities and limit access to opportunity.

Little of the national data on transit touches on accessibility for people with disabilities. However, paratransit organizations (local transportation services for people with disabilities) may have data on the location of riders and number of trips. Local government may have data on the location of paratransit stops that can be used to create indicators of access. For standard transit options, local transit authorities may have data on the percentage of vehicles in their fleet that are accessible.

The National Transit Map[40] includes data from participating agencies on transit stops, including mapping of quarter-mile and half-mile radii around each stop. These data would allow you to create a picture of transit accessibility by constructing indicators of the percentage of block groups within each of these distances.

Infrastructure features such as sidewalk quality, curb ramps, and accessible pedestrian signals are important contributors to accessibility. Unfortunately, there are no standard local data sources for this information. Local departments responsible for installing or maintaining accessible features may maintain data on where installation has taken place. For example, public works departments may maintain data on curb cuts or sidewalk width or geographic layers where sidewalks are present. Complaints to 311 systems for sidewalk repair may indicate areas of the city where accessibility is a problem. In other cases, local disability advocates; research institutions; or departments of public works, planning, or transit may have conducted surveys of this public infrastructure. The KidsWalk Coalition's (2011) report, for example, which inspected sidewalk quality, curb ramps, signage, and crosswalks within a one-block radius of public schools in New Orleans, LA.

The following are example indicators:

- transit proximity
    - » number of people with disabilities living in a census tract less than one-half mile from transit (National Transit Map, ACS)

- infrastructure quality
    - » number of school areas with curb ramps in poor condition (sidewalk surveys)

Environment

Environmental concerns for neighborhoods include the usual measures of the natural environment, such as air and water quality, as well as measures of the built environment, such as land usage, density, and green space. Indicators of pollutants in the air or water, proximity to hazardous facilities, amount of traffic, and access to parks can help illuminate disparities in access to a clean and safe environment.

The Environmental Protection Agency (EPA) produces several sources of small-area data through compilations of indicators from multiple sources and primary data collection. The EPA's Smart Location Database[41] compiles data to construct indicators at the block-group level regarding road density, walkability, and average travel time to jobs. The Smart Location Database allows for interactive

mapping, report generation, and downloading the data directly. Similarly, EPA's EJSCREEN[42] collects environmental data from several EPA sources into an interactive tool that maps environmental and demographic data at the block-group level. EJSCREEN includes indicators of air quality, proximity to environmental hazards, and traffic volume.

You can also access EPA data on air quality through HUD's data and mapping tool, which includes an environmental health index that aggregates several EPA indicators of air quality (a higher score indicates better environmental health). The mapping tool allows you to map this index alongside race/ethnicity, national origin, and family status at the census-tract level.

The EPA also produces data on hazardous wastes through the Toxics Release Inventory, which contains data on hazardous waste and polluting facilities. Data are available via summarized reporting down to the zip code level or downloadable point-level data on individual facilities.

The following are example indicators:

- environmental conditions
    - » environmental health index (AFFH data and mapping tool)
    - » pounds of released waste (Toxics Release Inventory)
    - » superfund proximity (EJSCREEN)

- social environment
    - » gross residential density (Smart Location Database)
    - » total road network density (Smart Location Database)

ADVANCED OPTIONS

Much of the data on environmental quality are collected at the state or county level and are less useful for neighborhood analysis. State departments of health may have additional small-area data on air and water quality. Additionally, departments of parks and recreation or environmental services may have geocoded information that can be used to develop indicators of the number of parks or amount of green space in an area. For example, the Charlotte, NC, Quality of Life Explorer (https://mcmap.org/qol) includes indicators on the density of the tree canopy and the percentage of impervious surfaces at the neighborhood level.

The following are example indicators:

- percentage of residential units within one-half mile of a park by neighborhood (municipal department of parks and recreation, local assessor's office)

- percentage of residential land area covered by tree canopy by neighborhood (local environmental services agency, zoning office)

Health

The health of a community includes indicators of physical, behavioral, and mental health. Many communities face significant health disparities from one neighborhood to the next and between demographic groups. Health outcomes also reflect a person's experience with personal and community-level poverty, education, housing, and employment (referred to as social determinants of health). Physical, behavioral, and mental health conditions can affect a person's ability to seek educational opportunities, find and maintain employment, and live in stable housing. Analyzing measures of disease and health condition prevalence, preventative health activities, and reported mental and behavioral health can create a picture of access to opportunities for a healthy life.

A new source of nationally available health data for small areas is available from the Centers for Disease Control and Prevention through its 500 Cities Project (https://www.cdc.gov/500cities/), which provides census tract–level health estimates for the 500 largest cities in the United States. The project contains a wide range of indicators on adult health outcomes, such as asthma, arthritis, and cancer; prevention measures, such as health insurance, mammograms, and colorectal cancer screening; and behavioral health, such as excessive drinking, smoking, and sleep. The 500 Cities Project data are available through interactive tables and maps, reports, and download. Many of the measures are reported as both crude rates (percentage of the population) and age-adjusted rates (a similar measure that controls for geographies with differently aged populations). The data are modeled estimates of the prevalence of health outcomes that factor in demographic and socioeconomic data. Because these data are modeled estimates, they cannot detect changes over time or the effect of local interventions. Therefore, they should not be used to evaluate interventions. However, you can use them to identify local disparities in health outcomes and access to health systems, and some communities have used them to track various health disparities, including behavioral health issues like binge drinking, across neighborhoods.[43]

The Dartmouth Health Atlas[44] is another national source of small-area data that provides Medicare data at the hospital level, including chronically ill patient deaths, cancer patient deaths, and postdischarge events.

Other important aspects of health can also be measured. For example, access to healthy food is a driver of health outcomes, and there are sources for local data on food access or the location of food

deserts. The US Department of Agriculture maintains a Food Access Research Atlas that displays census tracts by median income, distance to supermarkets, and vehicle availability.

Data on rates of health insurance coverage (insured or uninsured) are another source of information on the determinants of health outcomes at the neighborhood level. The ACS, which does not include general health measures, does include data on individual health insurance status and type of insurance by demographic characteristics.

The following are example indicators:

- primary care
    - » percentage of adults age 18 and over with coronary heart disease, age-adjusted (500 Cities Project)
    - » percentage of adults age 18 and over with health insurance (ACS)
    - » percentage of adults age 18 and over who have received a routine checkup in the last 12 months (500 Cities Project)

- behavioral health
    - » percentage of adults age 18 and over who are current smokers (500 Cities Project)
    - » percentage of adults age 18 and over who reported binge drinking in the past 30 days (500 Cities Project)

- healthy living
    - » percentage of low-income population living beyond one mile from a supermarket (US Department of Agriculture)

ADVANCED OPTIONS

Although a wealth of health data are collected at the state and local levels, many of these data are confidential. Vital statistics (birth and death records) are collected by state and local health departments. These records usually contain identifying information such as names and dates of birth, but many states publish aggregate vital statistics online. Health departments also collect data to monitor the prevalence of various diseases and conditions. For small areas, the confidentiality of the data may require negotiating with the health department to access the data and demonstrating compliance with privacy and data security standards. Partnering with local data intermediaries or researchers who may have already established a relationship with the health department and have access to the data may be helpful.

The following are example indicators:

- percentage of low-weight births, by race (health departments)
- deaths per population, by cause of death (health departments)
- rate of people living with diagnosed of HIV per 100,000 population (health departments)
- rates of positive blood lead tests for 5-year-olds (health departments)

Public Safety

The most common source of information on safety is data on reported crimes, but safety encompasses much more than just the reported crime rate. Reported crime may not reflect the true amount of crime or residents' perceptions of the level of crime in their neighborhoods and how safe they feel walking about. When possible, an analysis of neighborhood safety should include not only crime data, but also data on disparities in police activity, reported discrimination or harassment, and other measures that can inform public safety conditions from multiple viewpoints.

There are no nationally available crime data to support a neighborhood-level analysis. The FBI's Uniform Crime Reporting system collects data nationwide based on summary reports from local police agencies, but it does not produce data below the city or county level. Some police departments publish reports or annual statistics on crime in easy-to-access formats. Checking your local police department's website may provide useful neighborhood-level data on crime statistics.

ADVANCED OPTIONS

Local police departments keep records on crime reports, arrests, and 911 calls for service. These administrative data are generally available upon request, and many local police departments already provide these data through open data portals. Some law enforcement agencies have data compiled at the Police Data Initiative.[45] This is a good place to start to see if your local police agency has data available on the site or to get an idea of which indicators might be available. If these data are not available through a data portal, you may be able to contact the department directly to get access. If summarized data are not available, you may have to use administrative data from individually reported crimes. Using these data in a neighborhood context involves categorizing crimes by type and creating variables to summarize data from individual reported crimes.[46]

Local civil rights offices may have additional data pertaining to safety, such as records of discrimination complaints. Such data may allow for indicators showing disparities in reported discrimination, by area or protected class.

The following are example indicators:

- Part I property crimes reported per 1,000 population (police departments)
- Part I violent crimes reported per 1,000 population (police departments)
- reported discrimination per 1,000 population (office of civil rights)

Notes

1. The Fair Housing Act is Title VIII of the Civil Rights Act of 1968.

2. "Fair Housing—It's Your Right," US Department of Housing and Urban Development, accessed February 20, 2017,
   https://portal.hud.gov/hudportal/HUD?src=/program_offices/fair_housing_equal_opp/FHLaws/yourrights.

3. Secondary data are also valuable when used within an organization to improve internal program management and outcomes, but this use is not the focus of this guide. For example, data on neighborhood conditions and resident characteristics can help service providers better target their outreach efforts, tweak program design, and understand the neighborhood context in which their clients live. The "Data Playbook" outlines an approach to using data for performance management. See Rella Kaplowitz and Lynn Schusterman, "Data Playbook," accessed July 11, 2017, https://www.schusterman.org/playbooks/data/.

4. CURA is a partner in the National Neighborhood Indicators Partnership, a network of local data intermediaries that have a shared mission to help community stakeholders use neighborhood data for better decisionmaking, with a focus on assisting organizations and residents in underserved communities.

5. Urban Strategies Council is a partner in the National Neighborhood Indicators Partnership.

6. Please see Coulton (2008) for a catalog of administrative data sources.

7. "Geographic Terms and Concepts, Block," US Census Bureau, last modified January 10, 2017,
   https://www.census.gov/geo/reference/gtc/gtc_block.html.

8. "Geographic Terms and Concepts, Block Groups," US Census Bureau, last modified December 6, 2016,
   https://www.census.gov/geo/reference/gtc/gtc_bg.html.

9. "Geographic Terms and Concepts, Census Tract," US Census Bureau, last modified December 6, 2016,
   https://www.census.gov/geo/reference/gtc/gtc_ct.html.

10. For more detail, see "The ZIP Code Resources Page," University of Missouri, last modified June 21, 2017,
    http://mcdc.missouri.edu/webrepts/geography/ZIP.resources.html.

11. "Boundaries: Community Areas," Chicago Data Portal, accessed July 25, 2017,
    https://data.cityofchicago.org/Facilities-Geographic-Boundaries/Boundaries-Community-Areas-current-/cauq-8yn6/data.

12. See Richard Casey Sadler, "How ZIP codes nearly masked the lead problem in Flint," The Conversation, September 19, 2016, http://theconversation.com/how-zip-codes-nearly-masked-the-lead-problem-in-flint-65626.

13. For a listing of national datasets with data on small areas (below the county or city level), see Kathryn L.S. Pettit, "List of National Data Sets with Small-Area Data," National Neighborhood Indicators Partners, Urban Institute, January 11, 2017, http://www.neighborhoodindicators.org/library/catalog/list-national-data-sets-small-area-data.

14. See, for example, "DataSF," accessed July 25, 2017, https://datasf.org/opendata/.

15. For guidance from the US Department of Justice on the federal Freedom of Information Act, see "Department of Justice Guide to the Freedom of Information Act," US Department of Justice, last modified July 11, 2016, https://www.justice.gov/oip/doj-guide-freedom-information-act-0. For federal, state-specific, and open government guides, see also "Federal Open Government Guide," Reporters Committee for Freedom of the Press, accessed July 27, 2017, http://www.rcfp.org/federal-open-government-guide; "Open Government Guide," Reporters Committee for Freedom of the Press, accessed July 27, 2017, http://www.rcfp.org/open-government-guide.

16. "CURA programs for community-initiated projects," University of Minnesota, accessed July 27, 2017, http://www.cura.umn.edu/assistance/community.

17. "State Data Center (SDC) Program," US Census Bureau, accessed July 27, 2017, https://www.census.gov/about/partners/sdc.html.

18. For the locations of the 12 Federal Reserve Banks and 24 branch offices, see " The Twelve Federal Reserve Districts," Board of Governors of the Federal Reserve System," last modified April 24, 2017, https://www.federalreserve.gov/aboutthefed/federal-reserve-system.htm; See also "Charlotte Data Day 2015," University of North Carolina at Charlotte, accessed July 25, 2017, https://ui.uncc.edu/programs/dataday2015.

19. Kathryn L.S. Pettit, "List of National Data Sets with Small-Area Data," National Neighborhood Indicators Partners, Urban Institute, January 11, 2017, http://www.neighborhoodindicators.org/library/catalog/list-national-data-sets-small-area-data.

20. "NNIP Lessons on Local Data Sharing," National Neighborhood Indicators Partnership, last modified February 14, 2016, http://www.neighborhoodindicators.org/library/guides/nnip-lessons-local-data-sharing.

21. See note 19.

22. The Census Bureau produces separate estimates through the Population Estimates Program. Although these estimates are more accurate than the ACS, they are produced only for geographies down to the city level.

23. The HUD maps are dot density maps (i.e., maps that display one dot per person, usually colored to signify a characteristic). The dots on these maps are placed randomly throughout the geography (either census tract or block group) and do not represent the location of a specific household or person.

24. For contact information for nonprofit fair housing agencies, see "Get Local Help," National Fair Housing Alliance, accessed July 26, 2017, http://nationalfairhousing.org/get-local-help/.

25. Identifying the owners of rental properties can prove difficult. Rental properties are often owned through holding companies created specifically for the property in question. One strategy to identify owners is to use a company's tax address to identify properties with different holding companies but the same tax address. Another useful strategy is to categorize properties by the type of owner (e.g., individual, for-profit company, nonprofit, church, government). Although categorization does not identify specific owners, it can help identify how much property is held by certain types of entities and let local community organizations begin to explore what that means for development and turnover in the neighborhood (e.g., a concentration of government-owned property in an area could indicate a lack of private development).

26. "Property Data," NEO CANDO, Case Western Reserve University, accessed July 26, 2017, http://neocando.case.edu/cando/housingReport/interface.jsp.

27. An indicator for a tax credit for owner-occupants cannot be used by itself to create an approximate homeownership rate. Characteristics such as the type of property and the number of units in the property need to be considered. Properties may also be subdivided into multiple units, only one of which is occupied by the homeowner.

28. For examples of how to craft and use these indicators, see Pettit and Droesch 2009.

29. "Data," Zillow Research, accessed July 26, 2017, https://www.zillow.com/research/data.

30. "Data Center," Redfin, accessed July 26, 2017, https://www.redfin.com/blog/data-center.

31. HUD's Picture of Subsidized Households includes the Housing Choice Voucher program, Project-Based Section 8, Loan Management Set Aside, Section 236, Section 221(d)(3), Section 202, Section 811, Rent Supplement Program, and public housing. Housing subsidy programs not covered include those under the US Department of Agriculture's Rural Housing Service, Indian housing, HOME, and community development block grants.

32. "HUD Aggregated USPS Administrative Data on Address Vacancies," US Department of Housing of Urban and Development, accessed July 27, 2017, https://www.huduser.gov/portal/datasets/usps.html.

33. Safeguard Properties maintains a national database on property registry ordinances; see "Property Registration," Safeguard Properties, accessed July 27, 2017, http://www.safeguardproperties.com/Resources/Vacant_Property_Registration.aspx.

34. "Guide to Rental Complexes with Units with Accessible Features," Fairfax County Virginia, accessed July 27, 2017, http://www.fairfaxcounty.gov/dfs/disabilities/housing-access.htm; "Property Search," Directory of Accessible Housing, Fair Housing Council of Greater San Antonio, Texas, accessed July 26, 2017, http://www.accessiblehousing.org/property_search.asp.

35. "Head Start Center Locator," Head Start ECLKC, accessed July 27, 2017, https://eclkc.ohs.acf.hhs.gov/hslc/HeadStartOffices.

36. "Find Child Care and Preschool," NAEYC for Families, accessed July 27, 2017, http://families.naeyc.org/find-quality-child-care.

37. For a discussion of how community-based organizations can build trust with school districts and obtain access to student-level data, see StriveTogether (2015).

38. See Comey et al. (2013) for more information about Family Educational Rights and Privacy Act and negotiating for access to student data.

39. LODES data that link workplace and residence are partially synthetic. For the purposes of neighborhood-level data analysis, what this means is that for any individual worker or firm, the characteristics may have been altered to preserve confidentiality. However, the aggregate flows at the neighborhood level are accurate.

40. "National Transit Map," US Department of Transportation, accessed July 27, 2017, https://www.rita.dot.gov/bts/ntm.

41. "Smart Location Mapping," US Environmental Protection Agency, accessed July 26, 2017, https://www.epa.gov/smartgrowth/smart-location-mapping#SLD.

42. "EJSCREEN: Environmental Justice Screening and Mapping Tool," US Environmental Protection Agency, last modified December 19, 2016, https://www.epa.gov/ejscreen.

43. See, for example, Corianne Scally and Kathryn L.S. Pettit, "New data can help communities address their health challenges," *Urban Wire* (blog), Urban Institute, June 13, 2017, http://www.urban.org/urban-wire/new-data-can-help-communities-address-their-health-challenges.

44. "Data by Region," Dartmouth Atlas of Health Care, accessed July 26, 2017, http://www.dartmouthatlas.org/data/region/.

45. "Datasets," Police Data Initiative, accessed July 27, 2017, https://www.policedatainitiative.org/datasets/.

46. For example, as part of the Uniform Crime Reporting system the FBI classifies major crimes into two categories: Part I violent crimes (homicide, aggravated assault, rape, and robbery) and Part I property crimes (motor vehicle theft, burglary, larceny, and arson).

References

Acs, Gregory, Rolf Pendall, Mark Treskon, and Amy Khare. 2017. *The Cost of Segregation: National Trends and the Case of Chicago, 1990–2010*. Washington, DC: Urban Institute.

Bajaj, Beata G., Thomas Kingsley, and Kathryn L.S. Pettit. 2005. *Business Patterns and Trends: National Summary*. Washington, DC: Urban Institute.

Carr, James H., and Nandinee K. Kutty. 2008. *Segregation: The Rising Costs for America*. New York: Routledge.

Comey, Jennifer, Peter A. Tatian, Lesley Freiman, Mary K. Winkler, Christopher Hayes, Kaitlin Franks, and Reed Jordan. 2013. *Measuring Performance: A Guidance Document for Promise Neighborhoods on Collecting Data and Reporting Results*. Washington, DC: US Department of Education.

Coulton, Claudia. 2008. *Catalog of Administrative Data Sources for Neighborhood Indicators*. Washington, DC: Urban Institute and the National Neighborhood Indicators Partnership.

———. 2014. "Using Data to Understand Residential Mobility and Neighborhood Change." In *What Counts: Harnessing Data for America's Communities*. San Francisco: Federal Reserve Bank of San Francisco, and Washington, DC: Urban Institute.

CPD. 2016. *The eCon Planning Suite: A Desk Guide for Using CPD Maps*. Washington, DC: US Department of Housing and Urban Development.

Deitrick, S., F. Ye, J. Childs, and C. Zhang. 2015. "Connecting People and Place—Improving Communities through Integrated Data Systems: Chronic School Absenteeism in Public Schools in Pittsburgh, PA." Policy research brief. Pittsburgh, PA: University of Pittsburgh, University Center for Social and Urban Research.

Greene, Solomon, and Kathryn L.S. Pettit. 2016. "What if Cities Used Data to Drive Inclusive Neighborhood Change?" Washington, DC: Urban Institute.

Hendey, Leah, Jake Cowan, G. Thomas Kingsley, and Kathryn L.S. Pettit. 2016. "NNIP's Guide to Starting a Local Data Intermediary." Washington, DC: Urban Institute and the National Neighborhood Indicators Partnership.

HUD (US Department of Housing and Urban Development) 2017. AFFH-T User Guide: Affirmatively Furthering Fair Housing Data and Mapping Tool v. 4.0. Washington, DC: HUD.

Iceland, John, Daniel H. Weinberg, and Erika Steinmetz. 2002. *Racial and Ethnic Residential Segregation in the United States: 1980–2000*. Washington, DC: US Census Bureau.

KidsWalk Coalition. 2011. Stepping to School: An Assessment of Neighborhood Walkability and Solutions for a Safer, Healthier New Orleans. New Orleans: Tulane University.

Li, Huiping, Harrison Campbell, and Steven Fernandez. 2013. "Residential Segregation, Spatial Mismatch, and Economic Growth across US Metropolitan Areas." *Urban Studies* 50 (13): 2642–60.

Logan, John R., and Brian Stults. 2011. "The Persistence of Segregation in the Metropolis: New Findings from the 2010 Census." Providence, RI: Brown University.

Mallach, Alan. 2017. *Neighborhoods by Numbers: An Introduction to Finding and Using Small Area Data*. Flint, MI: Center for Community Progress.

McCormick, Kathleen. 2017. "Planning for Social Equity: How Baltimore and Dallas Are Connecting Segregated Neighborhoods to Opportunity." *Land Lines*. Winter.

Pendall, Rolf, Elizabeth Davies, Lesley Freiman, and Rob Pitingolo. 2011. "A Lost Decade: Neighborhood Poverty and the Urban Crisis of the 2000s." Washington, DC: Joint Center for Political and Economic Studies.

Pettit, Kathryn L.S., and Audrey Droesch. 2009. *A Guide to Home Mortgage Disclosure Act Data*, Washington, DC: Urban Institute.

StriveTogether 2015. "Student Data Privacy Best Practices: Five Ways Community Organizations Can Ensure Effective and Responsible Data Use." Cincinatti, OH: StriveTogether.

Turner, Margery Austin, and Lynette A. Rawlings. 2009. Promoting Neighborhood Diversity: Benefits, Barriers, and Strategies. Washington, DC: Urban Institute.

Wachter, Susan M., and Lei Ding, eds. 2016. *Shared Prosperity in America's Communities.* Philadelphia, PA: University of Pennsylvania Press.

About the Authors

Leah Hendey is a senior research associate in the Metropolitan Housing and Communities Policy Center at the Urban Institute, where her research is focused on housing policy, including the foreclosure crisis, and on policy to improve neighborhoods. She is experienced in working with national and local administrative datasets to create neighborhood indicators and study neighborhood conditions. Hendey has been involved in Urban's work with the National Neighborhood Indicators Partnership (NNIP) since 2007. She is a member of its executive committee and was appointed its deputy director in 2013.

Mychal Cohen is a research associate in the Metropolitan Housing and Communities Policy Center at the Urban Institute. His work focuses on affordable housing, neighborhood initiatives, and community development. Before joining Urban, Cohen was a policy and development associate at the National Housing Trust, where his work focused on state and local preservation of affordable housing, especially through the low-income housing tax credit.

Statement of Independence

The Urban Institute strives to meet the highest standards of integrity and quality in its research and analyses and in the evidence-based policy recommendations offered by its researchers and experts. We believe that operating consistent with the values of independence, rigor, and transparency is essential to maintaining those standards. As an organization, the Urban Institute does not take positions on issues, but it does empower and support its experts in sharing their own evidence-based views and policy recommendations that have been shaped by scholarship. Funders do not determine our research findings or the insights and recommendations of our experts. Urban scholars and experts are expected to be objective and follow the evidence wherever it may lead.

URBAN INSTITUTE · ELEVATE · THE · DEBATE

2100 M Street NW
Washington, DC 20037

www.urban.org