5. System Modeling Data Guide

August 2022

Contents

Introduction 3

Data Needed for System Modeling with Stella M 4

Households 4

Estimating Households for the Single Year and Multi-Year Options 5

Methodological Approaches for Household Estimates 6

Overall Considerations for Developing Households Estimates 13

Baseline Inventory 14

Known Issues 14

Timeframe Considerations 15

Additional Project Types 15

Subpopulations 15

Geography 15

Data to Inform the Development of Assumptions 15

Demographics and Household Characteristics of People Experiencing Homelessness 16

Current System Performance 19

HMIS Coverage 19

Additional Quantitative Data 20

Qualitative Data 21

Documenting Methodology and Planning for Updates 22

Introduction

The System Modeling Data Guide is the fifth of eight guides that are part of the **System Modeling Toolkit**. It comes after the community has oriented themselves to the process (see 1. System Modeling Introduction) and have identified, engaged with, and invited key partners, creating a workgroup. The workgroup is diverse, inclusive, and demographically representative of the people experiencing homelessness in the community (see 2. Identifying and Engaging with Key Partners for System Modeling Guide and 3. System Modeling Roles and Responsibilities Guide). This guide was developed to assist the person(s) responsible for entering data and assumptions into Stella Modeling (Stella M). That could be the facilitator, or it could be data analyst staff working with the facilitator and the system modeling workgroup.

To engage in system modeling, communities will need to assemble a variety of different types of data. The purpose of this guide is to assist them in understanding what kinds of data they might need, identify and evaluate potential data sources, and understand how to use that data to inform the system modeling process. The types of data to be discussed are:

- Data specific to people experiencing homelessness and the homeless response system, which could be found in the Homeless Management Information System (HMIS), other data systems, or both.
- Additional quantitative data, based on the local context, that will inform the system modeling process and provide a more accurate understanding of what factors intersect with and have an impact on the homeless response system.
- Qualitative data (such as information gained through interviews or focus groups) to understand the story behind the data and make more accurate predictions of how the system will function in the future.

As you work to identify the data most relevant to your community's system modeling process, make sure to consider how equity and data quality will affect the data you will use in modeling. An *equitable analysis* of data requires ensuring that data elements contain information about all the subpopulations and demographics that the system is serving, an analysis plan that regularly disaggregates program and system information by subpopulation and demographics, and an inclusive data collection and analysis team. *Data quality* issues can take various forms – including high levels of missing data, inaccurate data, and gaps in data collection for some projects or subpopulations – all of which can undermine the community's faith in the analysis produced from the data. There are various stages throughout the data life cycle when both equity and data quality should be considered.

The questions below are not meant to be comprehensive. Rather, they serve as examples of questions to ask to understand the data you have available to inform the

system modeling process. They are meant to support the community in thinking through how the homeless response system can be improved over time.

A community does not need to answer yes in order to engage in system modeling – the questions are not prerequisites. Rather, they can help a community begin to understand how their homeless response system is operating; what changes could be most impactful to address equity or data quality concerns; how assumptions might need to be modified based on input from various community members, including people with lived experience of homelessness; and how those changes could help to create the ideal homeless response system being envisioned through the system modeling process.

Questions to Interrogate Data for Use in System Modeling			
	Data Collection	**Data Entry**	**Data Analysis**
Equity	Do the intake workers look like the people from whom they are collecting information? Do they share experiences and identities? Have intake workers been trained in trauma-informed practices?	Is the system set up to comprehensively collect demographic information across projects so that the data can be meaningfully disaggregated by race, ethnicity, gender, household type, age, and disability?	Who is involved in analyzing the data? What perspectives do they bring? What lived experiences do they bring? Who can best tell you why the data shows what it does?
Data Quality	Are the data elements explained and collected in a consistent way across the homeless response system?	Are data entry workflows as simple as possible for end users?	Are there processes in place to determine when a data analysis finding is based on actual system performance and when it is based on data quality issues?

Data Needed for System Modeling with Stella M

The following section provides guidance on data queries and analysis to determine the values to be entered directly into Stella M. See the 7. System Modeling with Stella M User Guide for instructions on that entry. This section reviews two types of data needed in order to use Stella M for system modeling: **Households** and **Baseline Inventory**.

Households

System modeling supported by Stella M uses households (rather than people) to determine inventory needs for units (rather than beds). This is because the homeless response system provides services and housing to people in households.

To determine the amount of inventory needed within a system, Stella M needs to know how many households will be experiencing homelessness in each year of the model – what the model terms the **universe**. Two universes can be modeled for:

- **Inflow** – This is an estimate of the number of households expected to enter homelessness each year. It can be a challenging number to estimate for a single year, let alone to forecast into the future.
- **Long-term homeless** – Stella M also provides a way for users to model separately for households experiencing ongoing homelessness, who might need a different set of interventions or different proportions than households in the inflow universe.

This guide will provide some suggestions for how to estimate both numbers.

Modeling for these two universes can help communities plan for the ongoing inventory needs to serve the annual inflow, as well as surge capacity (or one-time investments) to catch up with unmet needs of people experiencing long-term homelessness. In determining how to model for the two universe options in Stella M, consider the characteristics and service utilization patterns of people currently in the system as well as unmet needs of people experiencing homelessness.

Estimating Households for the Single Year and Multi-Year Options

The simplest model is a **single-year** model, which documents the ideal system that could meet the needs of one or both universes of households without also planning for a transition from the current system to the ideal system. For this model, you need at minimum the estimated number of households for the inflow and long-term homeless universes in the target year. If you want to model for only one of these universes, you can leave the other at 0 households. Users can choose to add baseline year estimates to a single-year model to document a starting point for comparison with the target year without creating a multi-year model.

In the **multi-year** model, baseline year assumptions for inflow and long-term homeless households create estimates for the number of households in the target year. In the multi-year model, inflow assumptions also include an estimate of the percentage of inflow that is first-time homeless and anticipated change in first-time homelessness in future years.

Key Concepts
Household: A single individual or a group of people who want to be assisted together to obtain housing.
Universe: The group of households experiencing homelessness that is the focus of the model. Stella M allows users to model separately for the inflow universe and the long-term homeless universe and can present results for each universe separately or for the two combined.
Inflow: Households expected to enter the homeless system each year. Inflow has two parts – the number of households experiencing homelessness for the first time and the number of households returning to

The following section of this guide addresses data analysis strategies that can be used to develop the inputs for the Households section in Stella M. For technical instructions on using Stella M, see 7. System Modeling with Stella M User Guide.

Methodological Approaches for Household Estimates

There are a few general approaches to developing household estimates, depending on the population you are modeling for and the availability and quality of each data source. This section reviews six approaches: (1) HMIS project enrollment records; (2) Stella Performance (Stella P); (3) annualized Point-In-Time Count; (4) HMIS comparable data and other non-HMIS databases; (5) coordinated entry or by-name list; and (6) other published research.

Some of these approaches will give you all of the household information you need – estimates for both universes, proportions served and unserved, and breakdown by household type. For some of them you might need to use a combination of methods to produce all of the household data needed for system modeling.

1. HMIS Project Enrollment Records. HMIS project enrollment is a starting place for quantifying the number of households experiencing homelessness in a year. It is important to consider HMIS coverage and use additional data sources to account for and deduplicate households not captured in HMIS, including households served by non-participating projects or by victim service providers.

Key Concepts
homelessness after exiting the system in a prior period. • First-time Homeless: Inflow households that were not active in the homeless response system within the 24 months prior to their entry. • Returners: Inflow households that were active in the homeless response system within the 24 months prior to their re-entry.
Long-term homeless: Households that were "unserved" in the prior year, meaning they were homeless in the prior year and were not enrolled in a pathway to housing ("served").
Served: Households that are offered assistance in a pathway to housing within a model. Everyone served in a model year is assumed to exit the homeless system at the conclusion of their pathway, though not all are assumed to exit to permanent housing.*
Unserved: Households that are not expected to be offered or do not enroll in a pathway to housing within the model year. Any households that are unserved within one model year are assumed to NOT exit the homeless system, and therefore will be added to the long-term homeless universe for the following model year.*
Baseline Year: The year representing the current state of the homeless system. Stella M includes optional baseline inputs for the number of households experiencing homelessness (baseline households), the number and type of shelter and housing units currently available (baseline inventory), and the current cost of each project type (baseline cost). In multi-year models, baseline information is an input used to calculate target year estimates.
Starting Year: The calendar year of the first year of a multi-year model.
Target Year: The "end state" envisioned for the homeless system in the year designated for the model, which represents the final year of a multi-year model.

**For the purposes of system modeling, not all households that receive services are counted as "served." Households that receive support services only or shelter only that are not offered housing assistance are considered included in "unserved." To account for households that exit homelessness from shelter or the street with no additional housing support, users can include a pathway for "self-resolvers" in their model.*

Work with your vendor or HMIS administrator to create a report of the deduplicated number of households that were in the system over the course of a year. This includes:

- Households that entered the system over the course of the year (inflow).
- The portion of inflow households that did not have a previous record in the system over the past two years (first-time homeless).
- Households that have been continuously homeless prior to the year (long-term homeless).
- Households in the system that were served only in shelter or street outreach and did not exit (unserved). Once you have an estimate of the unserved population, then you can derive the served population in both universes.

Do not include households in permanent housing projects with a move-in date prior to the start of the year.

Example: District of Columbia

The District of Columbia Interagency Council on Homelessness used an HMIS analysis of homelessness to develop an estimate of "annual demand" including inflow and long-term homelessness. As a right-to-shelter community, with more than 95% of the community's shelter beds participating in HMIS, this was a reasonable starting place to estimate how many households would be in need of services each year. The Council took slightly different approaches for its family and singles population estimates:

***Families**: The Council started with an HMIS analysis to come up with a deduplicated number of households with children experiencing sheltered homelessness in a year. Then it added households that had presented for shelter assistance but had been diverted directly into a rapid rehousing program. Three-year trends showed the annual number of families to be consistent and representative of recent years.*

***Singles**: Because the number of unaccompanied individuals in the system had been quite variable in the years prior to the system modeling, it used a three-year average of households served to generate the baseline. A separate modeling process addressed the needs of transition-aged youth (18-24 years old), so those counts were removed from the total, to estimate for the needs of adults aged 25 and older.*

***Inflow & Long-term Homeless**: Using quantitative analysis and feedback from people with lived experience, the Council examined the service patterns of everyone in the system at a point in time. This analysis resulted in four categories of system utilization, which were used to estimate inflow and long-term homelessness:*

1. *Previously housed with Continuum of Care assistance; returned to shelter (**inflow**).*
2. *Multiple episodes (client exited without housing assistance; services received a year or more apart) (25% of this group included in **long-term homeless**, 75% included in **inflow**).*
3. *First/only episode, has lasted more than one year (**long-term homeless**).*
4. *First/only episode, has lasted less than one year (**inflow**).*

2. Stella Performance (Stella P). Stella P visualizes Longitudinal System Analysis (LSA) data that is uploaded to the Homeless Data Exchange 2.0 (HDX 2.0). The LSA includes data from some participating HMIS projects and can be used as a data source for developing household estimates. Households with days homeless is a count of households that had at least one day in emergency shelter, safe haven, transitional housing, or rapid rehousing or days enrolled in permanent supportive housing prior to move-in during the report period. In the days homeless by pathway analysis in Stella P, users can filter by inflow and long-term homeless to get estimates for those universes.

See the Stella and System Modeling HUD Exchange page for guidance and tutorials about Stella P.

3. Annualized Point-In-Time (PIT) Count. PIT Count data offers another starting place to estimate the annual number of households experiencing homelessness. This could be a good option when HMIS coverage is limited, or where there are large numbers of unsheltered households. Information collected in the PIT Count can be annualized using self-reported length of time homeless data collected during the count. PIT Count can also provide the portion of first-time homeless households and the portion of long-term homeless households. Other data sources might be needed to determine the portion of served and unserved households.

PIT Count methodology can affect how reliable and complete its data is, particularly for the unsheltered portion of the count. Another consideration is how often the Continuum of Care (CoC) conducts an unsheltered count. If modeling for a subpopulation or sub-geography, consider whether PIT Count data is able to be disaggregated appropriately. If PIT Count data does not include household composition data, additional analysis or extrapolation could be needed to estimate the number of households for the household type being modeled for.

Example: Oakland/Berkeley/Alameda County (CA)

In a survey done with a portion of the 5,629 people experiencing homelessness on the 2019 Point-In-Time Count, 2.2% of respondents had been homeless for seven days or less. That number was used to calculate the number of people becoming homeless each week in Alameda County, which was used to estimate the total over the course of the year.

2.2% x 5,629 = 124 people becoming homeless each week

124 people/week x 52 weeks = 6,440 people in becoming homeless annually

6,440 new people + 5,629 people already homeless as of January 2019 = 12,069 total people experiencing homelessness annually

To estimate the number of households experiencing homelessness in the year, local planners reviewed the PIT methodology used to count households in vehicles and found there was a known undercount of households with minor children. Therefore, a ratio of adult-only households to households with minor children from HMIS was used to estimate the portion of households with children and without children. An average household size for each of these household types was used to estimate the number of households based on the number of people.

Example: CoC with Low HMIS Coverage

In CoCs that have low rates of HMIS coverage for different project types, it could be necessary to do additional analysis of data outside of HMIS. For instance, in a CoC that has several providers (including a large emergency shelter) that do not enter data into HMIS, estimating the annual number households experiencing homelessness requires several steps. Using a secure transfer process and site, data from HMIS and the shelter's data system can be combined and deduplicated. To estimate the number people served by victim service providers, who are not allowed to enter data into HMIS for safety reasons, the organizations are asked to provide information on the total number of people and households served over the same time period and to estimate the number that weren't served in other programs in the CoC.

4. HMIS Comparable Data and Other Non-HMIS Databases. Most housing and shelter projects keep client-level data, even if they do not participate in HMIS. To get a more complete estimate of the number of households experiencing homelessness, work with non-participating providers to include data from these projects in your system modeling analysis.

If you are being given access to client-level data, a data use agreement might be necessary. Many households are served by both participating and non-participating projects, so take care to deduplicate household counts as much as possible. For more information about data requirements for HUD-funded victim service providers, see the Comparable Database Manual.

Example: Oakland/Berkeley/Alameda County (CA)

Accounting for Households When Client-Level Data Cannot Be Shared

When client-level data is not able to be shared, such as victim service provider data, the providers still might be able to use their internal program records to provide estimates of the number of unique households that were served only by victim service providers or other non-participating providers and were not also served by HMIS-participating projects.

5. Coordinated Entry or By-Name List. Ideally, all households experiencing homelessness would be accessing the coordinated entry system (CES), which then provides a comprehensive data source to inform the system modeling household assumptions. Even if that is not the case, the number of unduplicated households in CES during the course of a year can still be a starting place for determining need. Additional analysis could identify the portion of those households that were homeless for the first time and the portion that were experiencing long-term homelessness.

When exploring this approach, consider whether there are populations of people experiencing homelessness not accessing CES, and would therefore be left out if you used this approach for system modeling. Explore methods of supplementing this data source with additional data to account for households left out. Also consider whether those populations not included are likely to have different characteristics or different needs than households that are included in your data. For example, are households not accessing CES more likely to be undocumented immigrants, youth, people with limited English proficiency, people with behavioral health conditions, or people living in a certain geography of your CoC?

Also consider the data quality of your CES data. Are all people experiencing homelessness included? Or are people included only when they complete a full assessment? Is the list up to date to reflect current needs, or are there many households on the list from years ago that are no longer in need of assistance? To provide context on this data, review CES policies regarding how long people remain on the list and when they are removed. If households remain on the list indefinitely, CES could show inflated numbers of people currently experiencing homelessness.

Even if there is not a reliable CES dataset for the entire population of people experiencing homelessness, your community might have a by-name list for a subpopulation, such as youth, families, or Veterans, that could provide a reliable data source for that population.

6. Other Published Research. Sometimes existing data or reports from a reliable source address the inputs needed for system modeling. If this data is part of a published report or already is being used in planning and communication, it can be helpful for the system modeling process to align, especially if a better data source is not readily available. Make sure to consider the timeframe of the data, who is included and excluded, and how the data lines up with your target population and modeling goals.

If the published data doesn't include information on first-time homeless, long-term homeless, or the portion of served and unserved, you could use the published data in combination with one of the above methods. For example, using the percentage of first-time homeless from Stella P then applying that percentage to the total number of households expected annually in the published data.

You also might want to consider data sources that would help inform assumptions about inflow, such as eviction data, criminal legal system/discharge data, education data that looks at the number of students and families couch surfing, etc.

Example: Los Angeles Older Adults System Modeling

Using Other Published Research as basis for Household Estimates

Rather than starting from scratch, this model uses forecast data from a reputable published report to estimate the number of older adults expected to experience homelessness over the next five years.

Overall Considerations for Developing Households Estimates

Regardless of the data source, there are some key considerations:

- Is your data at the household or person level? If it is at the person level, it can be converted into household data using average household size.
- Does the data source include information on first-time homeless, returners, inflow, long-term homeless? Are you able to disaggregate the data or calculate percentages for each of these groups?
- Can you tell what portion of households in the system were served in a pathway that included housing assistance? (Including households receiving housing assistance as well as households that are served in shelter or services-only programs that are housed with problem-solving or self-resolve.)
- Who is not included in this data? What is your HMIS coverage rate? Make a plan to account for people who are unsheltered or not connected with services.
- How will you extrapolate for households not captured in the data source? Consider how the needs of households unknown to the system might differ from the needs of known households. For example, are households that are not captured in the data more likely to not speak English as a first language?
- Is this data able to be disaggregated by subpopulation or household type? For example, does PIT Count data have a reliable estimate of adult-only vs. adult and child households?
- Is this data able to be disaggregated by region? For example, if you are modeling for a sub-CoC jurisdiction, such as a city, county, or other service area within a CoC, is HMIS project data easily separated for that geography?
- If the data is for a month or a point in time, determine how the data can be annualized.

- How might external conditions, such as local employment rates, vacancy rates, and housing affordability, affect future changes in first-time homelessness?

The answers to these questions should help you identify limitations in the data source you are considering and guide you to identify supplemental data sources. If you use multiple data sources, be sure to establish a method for deduplication.

Baseline Inventory

Baseline inventory is the number of units of each project type currently available in the homeless response system. In system modeling, baseline inventory is compared to the target year inventory recommendations to quantify gaps and support implementation planning.

Baseline inventory is an optional input in Stella M. If included, the baseline number of units for each project type will be shown in the Results Preview and Results Page with a Net Change calculation to highlight system gaps.

In most cases, the Housing Inventory Count (HIC) is the best starting place for the baseline inventory.

The inventory information should reflect point-in-time capacity rather than utilization and be presented as units (not beds). For adult-only and child-only household types, beds might be equivalent to units, if each unit serves only one person. However, some project types might serve more than one adult in a unit, such as a couple or adult family members. In that case, average household size can be used to estimate the number of units by converting beds to units.

When calculating capacity for rapid rehousing or other voucher-based projects, use the average point-in-time capacity rather than the number of slots available throughout a year.

Be sure to consider the household type or subpopulation that you are modeling for. Though the HIC is a good starting point for the baseline inventory, there are some common situations that could lead you to adjust the baseline inventory in your model, rather than entering exactly what was submitted in the HIC.

Known Issues

If you are aware of issues or discrepancies in the submitted HIC, use the best data you have available at the time of modeling. For example, ensure that project types are correctly assigned, that beds and units are not duplicated, and that units not dedicated to the homeless service system are excluded. This kind of adjustment is often needed for rapid rehousing because of how rapid rehousing is required to be reported. Remember you are estimating capacity at a point in time, not utilization.

Timeframe Considerations

If there have been substantial inventory changes since the most recently submitted HIC, you may want to make adjustments based on the current situation.

Additional Project Types

Your system model might include additional project types that are not included in the submitted HIC, such as street outreach, diversion, prevention, or other permanent housing. If those project types already exist in your community, work with providers to get an estimate of point-in-time capacity for each project to include in your model. If those projects do not already exist as a dedicated part of your homeless inventory, leave the baseline as zero.

Subpopulations

The HIC is helpful in identifying beds for some household types including families, singles, Veterans, youth, and people who are chronically homeless. However, some project inventory might not be specified for a specific household type. If so, an option is to use HMIS analysis to produce a baseline estimate of the number of units for each group, such as households with adults only and adult and child households. However, if those units are not dedicated to those populations, the proportion could vary year to year.

If you are modeling for a different subpopulation, such as older adults, it can be challenging to determine baseline inventory for them. If there are no projects specifically dedicated to them, you might want to consider the baseline as zero. Another option is to use HMIS analysis to determine current system use for that subpopulation as a proxy for baseline inventory.

Geography

If you are modeling for a geographic region that is smaller than a CoC (such as a city within a CoC), you will need to rely on local data sources to determine the baseline inventory specific to that geography.

Remember, the baseline inventory is not required to build a model. However, including the baseline inventory can help with implementation planning by quantifying the gaps between your current system and the ideal system.

Data to Inform the Development of Assumptions

Key Concepts
Average Length of Assistance Expected: The average number of days that households are expected to be in each project in the pathway, including days assisted in the project while experiencing

Understanding the characteristics and needs of people experiencing homelessness will help the workgroup design project types and pathways to effectively and equitably meet the needs of all people experiencing homelessness. This section addresses potential data sources and analysis methods to gather such information. The 6. System Modeling Assumptions Guide addresses how to use this information in the modeling process to design projects and pathways to build a system model.

The data that can be used to develop assumptions discussed in this section are demographic data, current system performance, and additional qualitative data.

homelessness and days assisted in the project once housed.
Cohort: A group of households with similar characteristics, resources, and housing and service needs.
Days Homeless: The sum of the average number of days that households are expected to experience homelessness while enrolled in the projects included in the pathway.
Days Housed: The time that a household is assisted in a project after they have moved into permanent housing.
Exits to Permanent Destinations: Households that exited the homeless response system to permanent destinations.
Pathway: The distinct combination of project types that a household is provided while involved in the homeless response system with the goal of helping the household obtain permanent housing. Also referred to as a "pathway to housing."
Project Type: A way to categorize projects into similar groups, typically based on a project's purpose, service and housing components, and target population (if that group has distinct housing and service needs). Stella M includes default project types from the LSA and user-generated custom project types. Stella M generates inventory projections by project type.
Projected Performance: Expectations about the average number of days homeless, percentage of exits to permanent housing, and percentage of returns to homelessness for households served in a model. Projected performance will be calculated at a number of levels within the model (e.g., pathways, universe, years).

Demographics and Household Characteristics of People Experiencing Homelessness

Purpose of Demographic and Household Characteristics Data Analysis in System Modeling

Demographic and household characteristics information can inform our understanding of people's housing and service needs. This data analysis can help communities understand groups of people with similar needs (**cohorts**) and estimate how many households are in each. For example:

- Households living on a fixed income, including older adults (might need long-term subsidy).
- Households with young children (might need connections to childcare or other family supports).
- Households that report behavioral health or chronic health conditions as a cause of their homelessness (might need support services).
- Households with criminal justice history (might need extra housing navigation support or landlord incentives).

- Households that have earned income but have had a temporary episode of housing instability (might need housing problem solving or other short-term intervention).
- Households that experience barriers to employment (might need longer periods of rental assistance).
- Households that experience housing and other discrimination on the basis of race/ethnicity (might have longer lengths of stay prior to obtaining housing).

Demographic analysis also can identify whether certain groups of people are not being served well by the homeless response system – that is, they are not accessing the system at rates proportionate to their experience of homelessness; or once they are in the system, they obtain housing at lower rates or more slowly than other groups. This type of analysis will not identify the root causes of such disparities and might not provide enough information to design more effective projects and pathways. To do that, other data might be needed. See the **Qualitative Data** section in this guide for more information.

All projects using HMIS must collect and enter HUD's **Universal Data Elements (UDEs)** – including race, ethnicity, gender identity, date of birth, and household composition – for the people they serve. There are limitations to this data collection, including a limited number of response options available for these demographic data points. But if the community's HMIS has high system coverage, this data can be a great place to begin to understand who is accessing your homeless response system. Victim service providers and many other systems used by projects that do not use HMIS collect comparable identifying information about the people they serve.

Victim Service Providers (VSPs) Consideration

VSPs are an integral part of a community's homeless response system. In some communities, VSPs are the only services available. Though VSPs cannot share identifiable, client-level data, they can provide aggregate information about their projects' performance that can help inform the system modeling process.

Beyond these sources, you might need yet additional information about people experiencing homelessness to help inform your system modeling process. This could include specific information about sexual orientation, citizenship status, Tribal Nations status, and the like. Communities should determine the extent to which this type of additional information would help them understand the people served in their homeless response system and work with local partners to obtain it. For example, people who identify as lesbian, gay, bisexual, transgender, queer, intersex, asexual, and more (LGBTQIA+) are disproportionately affected by homelessness; are not easily

identified in HMIS data; and can have safety and privacy concerns with disclosing their identities. For these reasons, you must take special care in gathering information from and about this population. The same is true for people with undocumented citizenship status. Working with trusted partners within these communities likely will give you more useful insight into how they experience and resolve their homelessness given the lack of data that demonstrates their need.

HMIS, comparable databases, and other provider-specific databases will help you understand who is accessing your homeless response system and how those households are being served. But in and of themselves, those systems will not show who is not being served or not even accessing services.

The workgroup should not make assumptions about why certain groups are underrepresented in the homeless system without further exploration. Qualitative information can help communities begin to understand whether underrepresented subpopulations really do not need to access the homeless response system or there are other explanations for the underrepresentation – barriers to accessing services, services being provided in a culturally inappropriate manner, lack of knowledge about services available, etc.

Considerations beyond System Coverage

HMIS system coverage looks at how many projects providing homeless services are entering client-level data into HMIS, but it does not take into account the quality of that data.

Additionally, HMIS coverage takes into account people who are accessing and being served by projects providing homeless services, but it does not provide information about who is not accessing the homeless response system at all.

Communities should explore additional quantitative and qualitative data sources to understand the population that is experiencing homelessness but is not accessing the homeless system, and determine what their needs are.

Coordinated entry or by-name list data can be another source of information about who is experiencing homelessness and what they need. The number of households that are prioritized for various housing interventions can inform system modeling assumptions. Demographics such as income, employment, or disability status can inform how many households could need short- or long-term housing assistance or support services. Identifying groups that have longer than average wait times for referrals can point to resource gaps. Coordinated entry data can also provide information about the characteristics of people returning to homelessness.

Current System Performance

Data analysis about current system performance is critical in informing the system modeling process. By understanding how households are moving through the current system and which pathways result in fewer days homeless, higher rates of exit to permanent destinations, or lower rates of returns, communities can apply what is working well and address gaps. The ability to disaggregate system performance by race, ethnicity, gender identity, household type, age, disability, etc., is critical to ensure an equitable approach to system modeling. A community needs to understand not only how the homeless response system is performing but also who it is serving and serving successfully and who it is not.

The System Performance Measures (SPMs) and Stella P are built off HMIS data and available to all CoCs. These reports look at system performance around metrics such as how long people or households experience homelessness, how many exit to permanent destinations, and how many return to the homeless response system after exiting. Though both the SPMs and Stella P address system performance, there are some key differences in their calculations, described here.

Example: Current System Performance Data

Knowing the number and type of households that return to the homeless response system, a community could include in its system model an after-care program with support that reduces the likelihood of households returning. Additionally, knowing where in the homeless response system households return to, a community could incorporate into its model strengthened housing problem solving and rapid access to prevention supports at those "re-entrance points."

Understanding the system's current performance related to these kinds of metrics will help to inform key focus areas for improvement and also will highlight what might be working well. Additionally, analyzing performance metrics disaggregated by race, ethnicity, gender identity, household type, age, disability, other identities, and the intersectionality between and among these identities will help the community identify where the system might not be serving specific subpopulations well. Then those gaps in the system can be explicitly addressed in the system modeling process.

HMIS Coverage

Communities have varying degrees of system coverage in their HMIS. If a community's HMIS system coverage is high, the ability to use that dataset to understand current system performance is equally high. In that scenario, a community might be able to rely heavily on data in HMIS (with an alternate analysis of the needs of households fleeing

domestic violence). If a community has low coverage, then HMIS cannot provide an accurate picture of system performance.

If a community's HMIS system coverage is low and excludes a large number of service providers or service providers that serve specific subpopulations or groups, the community might need to consider including data not just from HMIS and comparable database data but also from those other service providers and entities.

Additional Quantitative Data

Though the community homeless response data described above is essential, additional data can contextualize that homeless system data and draw a more comprehensive picture of the needs and experiences of people experiencing homelessness. Using these supplementary datasets, communities can make comparisons between households in the homeless response system and the general public, varied subpopulations, and intersectional identities in order to identify and address gaps in their homeless crisis response.

Example: Additional Quantitative Data for Equity Analysis

Using the CoC Analysis Tool: Race and Ethnicity, a community identified that the Native American/Alaska Native population had disproportionately higher rates of unsheltered homelessness. The modeling workgroup, which included Tribal members and shelter providers, explored shelter program models that were more welcoming and accessible to this population, as well as pathways to housing that did not require a shelter stay.

A number of quantitative datasets are publicly available, such as US Census data, local income level data, eviction rates, and educational data. Other datasets are available through partners such as the local public housing authority, hospitals, schools, and courts/jails.

In addition to comparative data analysis, these supplemental datasets can help communities understand other significant systemic influences that could be addressed in system modeling. For example, knowing the overlap of frequent users of hospital emergency departments, police departments, and emergency homeless services, a community could model a permanent housing program with support designed to meet the specific housing needs of that frequent-user population.

Qualitative Data

Qualitative data is non-numerical data, representing the aspects of experiences that are often difficult to measure through tools and assessments or are open-ended in nature. Gathering qualitative data can be as simple as interviewing service providers or as comprehensive as conducting multiple focus groups with people experiencing homelessness. Communities also could use existing qualitative data sources such as client feedback gathered as part of a program or coordinated entry evaluation. The objective is to explore root causes of observed disparities, understand the how and why of the quantitative data, and get feedback on solutions from people most affected by homelessness.

Some topics that might be addressed with qualitative data include:

- Understanding barriers to housing for a group that has lower rates of exits to permanent destinations.
- Understanding risk factors for housing stability for a group that has higher rates of returns.
- Understanding access barriers for a group that has lower rates of CES assessment or referrals compared to their prevalence of homelessness.
- Suggestions for system improvements or new project types to better meet people's needs.

Example: Oakland/Berkeley/Alameda County (CA)

Quantitative data analysis in this community revealed deep racial disparities in who enters homelessness and who returns to homelessness after exiting the system. However, this analysis fell short in understanding the root causes or finding solutions to these inequities. To gather additional information, the community conducted nine 90-minute focus groups with people experiencing homelessness representative of the groups disproportionately affected by homelessness, specifically Black non-Hispanic and Hispanic/Latinx populations. The focus groups explored the root causes of homelessness, housing barriers, and participants' experiences with the homeless response system.

The qualitative data gathered from the focus groups informed the development of project types and pathways in the system model and the proportions of households expected to be served in each pathway. For example, it was determined that rapid rehousing was insufficient to meet the needs of most Black people experiencing homelessness in a community with high rents, limited affordable housing options, and vast gaps in wealth and economic opportunities. Instead, dedicated affordable housing, which provides a non-time-limited housing subsidy, was proposed as a project type more likely to meet the long-term housing needs of this population.

See *COVID-19 Homeless System Response: Qualitative Data 101* for more information and guidance on using qualitative data.

Documenting Methodology and Planning for Updates

The homeless response system is never static – it changes every moment of every day as projects start, end, or shift focus; as households enter, exit, or change in composition; and as community conditions including unemployment and housing costs and availability change. Because of this, the data that is used to inform the system modeling process also will change over time. A community should use the data they have available to them that helps them best understand their homeless response system and the people served and to plan for future updates as shifts occur.

It is critical to document the data sources, how (and when) the data was pulled from the systems (i.e., parameters, filters, data points included), analysis methods, and decisions around what data is used in the system modeling process. This documentation is helpful to understand the context of the system modeling results and

to support communication to key partners, funders, and policy makers. See the 8. System Modeling Results and Implementation Guide for more information about communicating the system modeling process and results. In addition, having clear documentation of the methodological approaches to analysis used will be helpful when the modeling inputs need to be updated over time.

Here are some examples of methodology documentation:

- Centering Racial Equity in Homeless System Design – See Appendix B: Method of Estimating Annual Homeless Population and Geographic Distribution [report]
- Home Together 2026 Community Plan: 5-Year Plan to End Homelessness in Alameda County – See Appendix C: System Modeling Overview and Update [report]
- Los Angeles Older Adults System Modeling: Envisioning a Better System for Those 50+ Experiencing Homelessness – See Methodological Notes for Inputs and Assumptions [report]