How to Use Sampling within a CoC to Conduct an Unsheltered Point-in-Time (PIT) Count

Background on 2021 Unsheltered PIT Count

As Continuums of Care (CoCs) respond to the COVID-19 crisis nationwide, the U.S. Department of Housing and Urban Development (HUD) has allowed some flexibilities in CoCs' 2021 unsheltered Point-in-Time (PIT) count requirements.

In *Conducting the 2021 Unsheltered PIT Count*, HUD states that it "will allow CoCs to request an exception to some or all of the unsheltered PIT count requirements. CoCs can request an exception to only conduct a head count of people experiencing unsheltered homelessness (i.e., not data collection on demographics, subpopulation, or household characteristics) or to collect some but not all of the data elements HUD has required in the past count." HUD also notes that CoCs may adjust their PIT count approaches to minimize the need for volunteers. These flexibilities illustrate HUD's commitment both to prioritizing the safety of people experiencing unsheltered homelessness and those who serve them, as well as to collecting sufficient data to understand basic trends among people experiencing homelessness, with the ultimate goal of improving our collective understanding of and responses to this population's needs in an unprecedented time.

This document describes how CoCs can use a sampling approach to estimate a total count of the number of people within their communities who are experiencing unsheltered homelessness on the night of the PIT count. While this sampling guidance is applicable for any year, CoCs that typically canvass and count their full geographic area may want to use a sampling strategy for the 2021 count in light of the COVID-19 pandemic. Sampling reduces the number of places that need to be visited on the night of the count, which may reduce the risk of COVID-19 transmission from interactions associated with PIT count activities and reduce the number of volunteers needed to conduct the PIT count.

What is Sampling?

Sampling is a way to account for information about a whole group of people by only collecting data about a portion of that full group. In the context of an unsheltered PIT count, this means that a CoC could estimate the number and characteristics of all people experiencing unsheltered homelessness by counting, and potentially surveying, only some sub-regions of its full geography and then extrapolating for uncounted and un-surveyed areas.

Benefits of Sampling	Drawbacks of Sampling
• Takes less time to complete than a full census count • Requires fewer staff and volunteers to conduct PIT count activities • Increases safety and decreases risk of COVID-19 transmission for people experiencing unsheltered homelessness, staff, and volunteers	• Relies on assumptions about where people experience unsheltered homelessness, which are limited by a CoC's initial or known data • Not as accurate as a full census count • Unlike full counts, which can be divided by subarea, extrapolated information about demographic characteristics may not be valid for subarea reporting

The contents of this document, except when based on statutory or regulatory authority or law, do not have the force and effect of law and are not meant to bind the public in any way. This document is intended only to provide clarity to the public regarding existing requirements under the law or agency policies.

Determining PIT Count Data Collection Activities and Timeline to Support a Sampling Approach

A CoC's data collection activities or methods and the timeline during which it conducts those activities will help to determine its sampling approach.

There are two main methods for conducting unsheltered PIT counts: observation-only counts and survey-based counts[1]. **Observation-only counts** produce a head count without verbal or physical interaction between people experiencing unsheltered homelessness and those conducting the PIT count. Thus, observation-based data collection activities must be conducted late at night, when it is reasonable to assume someone is experiencing unsheltered homelessness by virtue of the time and place they are located or sleeping. Typically, observation-only counts are conducted only on the night designated for the PIT count (i.e., between the hours of 10pm on the night of the count and 6am the next morning), though HUD is providing some flexibility to count people in distinct geographies on consecutive nights, if needed to reduce the number of volunteers needed to conduct the count.

Survey-based counts can be conducted over a broader period of time, because respondents can be asked questions to validate whether they experienced homelessness on the night designated for the count and where. Surveys can also include questions to collect demographic information and any other information the CoC deems important for local planning purposes. While survey efforts are typically concentrated on the night designated for the count, they may be completed over a span of a few days and in various locations (such as day-time service provision locations or places in which people generally sleep). For 2021, HUD is allowing CoCs to conduct survey-based counts over a longer period, up to 14 days; most years, survey-based counts are limited to no more than a 7-day period. Surveys conducted after the night designated for the count must ask where the person slept on the actual night of the count to determine if the person was experiencing unsheltered homeless on the night of the PIT count.

Table 1. Crosswalk of PIT Count Activities and Timing

		Timing of PIT Count Activities	
		On the Night of the Count (e.g., 10pm-6am, starting on the night designated for the count)	**After the Night of the Count** (e.g., starting at 6am the day after the night designated for the count)
PIT Count Activities	**Observations**	Yes	No
	Surveys	Yes	Yes

When deciding which data collection activities and timeline to use, CoCs should take a number of factors into consideration:

- **What information does the CoC want to collect during the PIT count?** If the CoC wants to collect data on demographics, other subpopulation data, or information on COVID-19 risks and impacts, the CoC will need to conduct a survey. HUD advises that CoCs only consider brief surveys to limit the length of interactions between people conducting the survey and those experiencing unsheltered homelessness.

---

[1] HUD publishes example observation and survey tools that CoCs can adapt for local use: https://www.hudexchange.info/resource/3322/point-in-time-survey-tools/

- **What is the current COVID-19 environment in the community?** Observations involve less risk of COVID-19 exposure for people experiencing homelessness, staff, and volunteers. Observations allow for greater physical distancing and take less time to complete than surveys, which also means that fewer staff and volunteers would likely be needed to count the same geographic area or number of people. Since observation counts can primarily provide head counts, CoCs interested in collecting more information should consider working with local public health officials to determine whether a survey approach is feasible.

- **How much staff capacity is available to support PIT count activities, and over what time period? Does the CoC intend to recruit, train, and support volunteers in conducting any PIT count activities?** The human capacity available to undertake PIT count activities within a set time period (either a single night count or service-based counts over several days) may help determine which approach is more feasible—observation-based counts or surveys. Surveys may require additional staff or volunteers, and will require additional training.

- **For how many days does the CoC intend to conduct PIT count activities?** As noted above, observation-based PIT counts must occur between sunset and sunrise on the night designated for the PIT count, preferably between 10pm and 6am. PIT counts conducted after the night designated for the count, whether over a single day or on multiple days, must have a survey that, at a minimum, verifies where the person interviewed slept on the night designated for the count, unless the CoC is counting distinct areas on consecutive nights due to the CoC's large geographic region. If a CoC is conducting an *observation-only* count over multiple days in distinct geographies, the count should be limited to two to four days, depending on the number of areas in which the CoC is counting, to reduce the likelihood of duplication. HUD is allowing *survey-based* counts to stretch up to 14 days past the night of the count to reduce the number of staff and volunteers needed at one time.
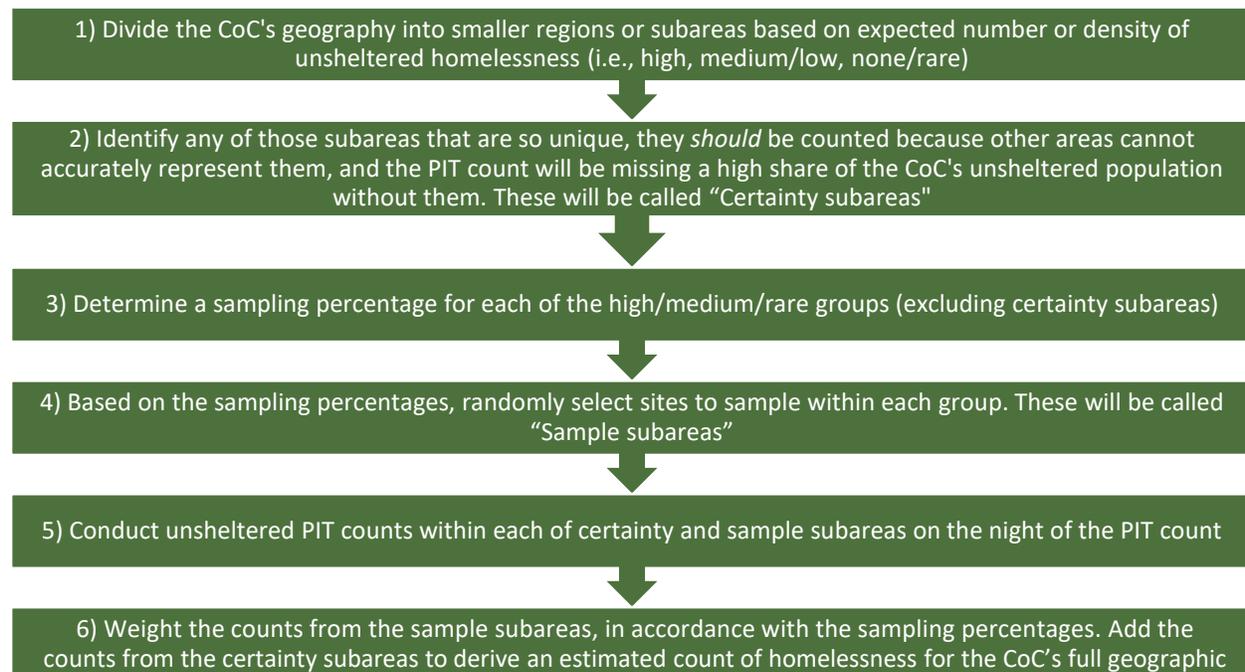
Developing a Sampling Approach

To implement a sampling approach for the unsheltered PIT count, the CoC will need to divide its geography into smaller subareas and categorize them based on the likely number of people experiencing unsheltered homelessness within each subarea. The CoC will then need to randomly select the subareas in which it will conduct PIT count observation or survey activities. Finally, the results from these sample subareas will be weighted to represent the subareas that are not counted to estimate an overall count for the full area.

It is not feasible to provide explicit guidance that would apply to every CoC on the best way to construct a sample or determine a sampling percentage. CoCs vary considerably by geographic type, geographic size, population type, population size, and several other features that make each CoC unique. Therefore, HUD always recommends that CoCs work with local research partners (e.g., university faculty or students with both local knowledge and expertise in statistical methods) to help refine PIT count data collection approaches, support implementation, and complete data cleaning and analysis.

However, CoCs can use the information in this paper to understand the general step-by-step process for developing a sampling approach.

Exhibit 1. General Step-by-Step Sampling Process

> **1) Divide the CoC's geography into smaller regions or subareas based on expected number or density of unsheltered homelessness (i.e., high, medium/low, none/rare)**

> **2) Identify any of those subareas that are so unique, they *should* be counted because other areas cannot accurately represent them, and the PIT count will be missing a high share of the CoC's unsheltered population without them. These will be called "Certainty subareas"**

> **3) Determine a sampling percentage for each of the high/medium/rare groups (excluding certainty subareas)**

> **4) Based on the sampling percentages, randomly select sites to sample within each group. These will be called "Sample subareas"**

> **5) Conduct unsheltered PIT counts within each of certainty and sample subareas on the night of the PIT count**

> **6) Weight the counts from the sample subareas, in accordance with the sampling percentages. Add the counts from the certainty subareas to derive an estimated count of homelessness for the CoC's full geographic**

Each of the steps above is described in more detail below, as well as in the *Point-in-Time Count Methodology Guide*.

1. **Divide the CoC's geography into smaller regions or subareas categorized based on predicted levels of homelessness.** Sampling relies on being able to divide the full geography into smaller regions (i.e., subareas) that can be the basis of sampling. The subareas in the same category should be roughly comparable to one another in terms of the number of people experiencing unsheltered homelessness that the CoC would expect to count. Since people experiencing unsheltered homelessness are not usually evenly dispersed within a CoC's geography, the CoC should consider identifying subareas with higher expected counts (i.e., high-density), subareas with medium or lower expected counts (i.e., medium-density), and subareas where unsheltered homelessness is unexpected or expected to be rare. Then sampling should be done separately within the high count subareas, within the medium count subareas, and within the unexpected or rare areas.

   The subareas do not need to be the same size geographically, if different sized subareas help to define more balanced rates of unsheltered homelessness among subareas. Consider defining subareas by census tracts, zip codes, county or city lines, neighborhoods, specific streets or highways, or some other locally specific markers that make the subarea easy to identify or natural to divide. CoCs should use local data and knowledge to categorize the subareas. Last year's unsheltered PIT count results, HMIS outreach data, actively managed by-name or prioritization lists, and the knowledge of community stakeholders can be helpful in defining comparable subareas. CoCs may also want to consider locations of bus stations, freeways with protected overpasses, 24-hour businesses, and parks with public restrooms when defining subareas that might be expected to have a higher presence of unsheltered homelessness.

After subareas are defined, assign labels based on at least three categories (or more, if the CoC wants to use more granular categories):

   a. **High-density subareas**: Subareas that include known locations where a relatively larger number of people experiencing homelessness frequently stay
   b. **Medium-density subareas**: Subareas that include areas where fewer people experience unsheltered homelessness
   c. **Rare or unlikely subareas**: Subareas that include areas where very few or no people have been known to experience homelessness.
   d. **Uninhabitable subareas**: Subareas where it would be impossible for someone to experience unsheltered homelessness (these include areas like large bodies of water.)

2. **Identify any subareas that are so unique, they must be counted to obtain an accurate count. These will be called "Certainty" subareas.** If there is a small number of subareas that are expected to have a much higher share of persons experiencing unsheltered homelessness than the other subareas, they need to be included in the CoC's PIT count activities to obtain an accurate count. These are the certainty subareas that would not be well-represented in the count by other subareas nor could they represent other subareas. For both of these reasons, subareas with uniquely high patterns of unsheltered homelessness should be selected as certainty subareas. These subareas will only represent themselves in the PIT count results.

3. **Determine the sampling percentage for each of the high/medium/low/rare groups.** There is no universal rule about how many or what percentage of subareas must be sampled for a sampling approach to be valid, but there are a couple of general principles CoCs should follow:

   a. Sample as many of the "high-density" subareas as possible. Since these are the subareas where one would expect to find the highest number of people experiencing homelessness, the CoC's efforts are best invested in these places. The more of these subareas that are sampled, the more accurate the PIT count will be.
   b. Sample fewer of the medium-density and rare/unlikely subareas, but do not sample so few that an unexpected (and possibly unrepresentative) count in one of these subareas will spoil the sampling estimate for all medium-density or rare/unlikely subareas.

In the mock CoC illustrated in Exhibit 2 below, the CoC has 5 high-density subareas, one of which is identified to be counted as a "certainty" subarea; 24 medium-density subareas; 5 subareas where unsheltered homelessness is expected to be rare; and 1 uninhabitable subarea. The CoC might decide to count the certainty site plus 50 percent of (or 2 of the remaining 4) high-density subareas; 25 percent of (or 6 of the 24) medium-density subareas; and none of the subareas where unsheltered is expected to be rare. By not sampling any of the rare/unlikely subareas, no people experiencing unsheltered homelessness in these subareas on the night of the PIT count will be counted. This approach assumes that there are truly no people experiencing homelessness in any of these subareas on the night of the PIT count, or, if there are any, the number is so small it would not meaningfully affect the CoC's PIT count. Conversely, the CoC could choose to sample a small percentage of the rare/unexpected subareas to add more precision to their estimates of the total population experiencing unsheltered homelessness.
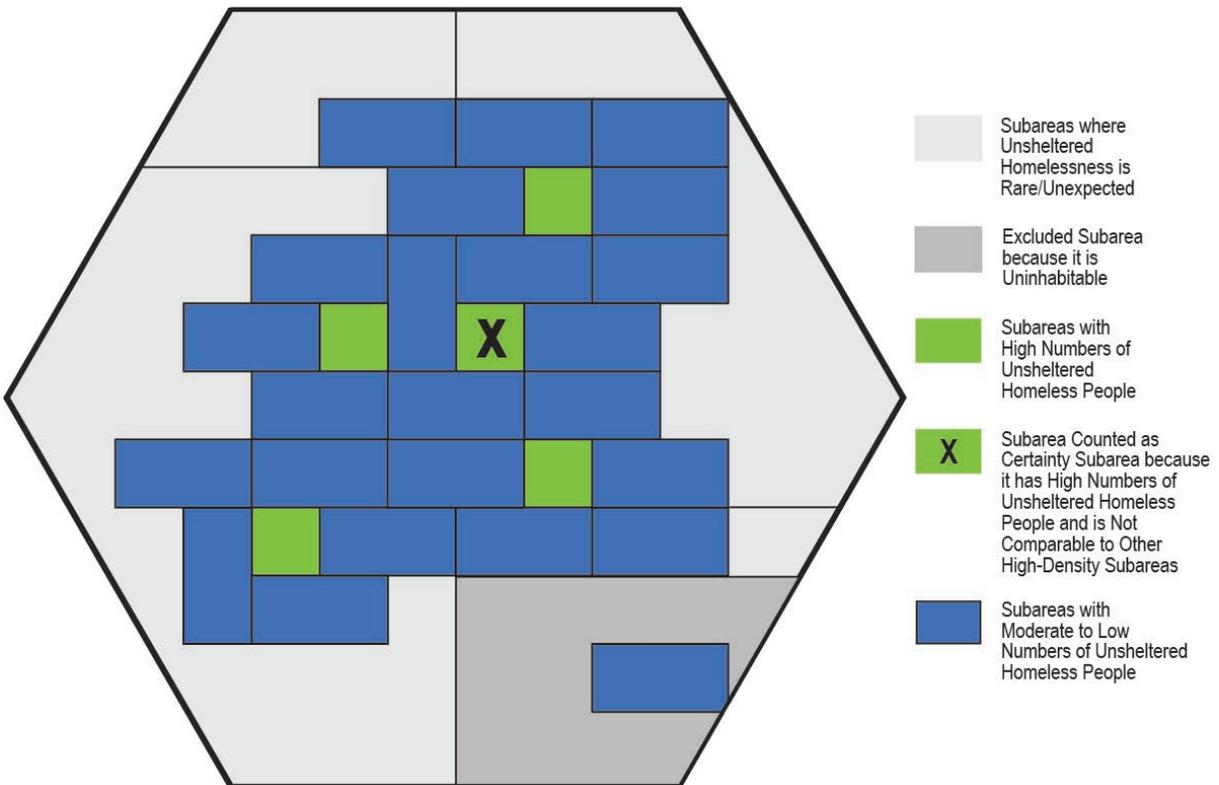
4. **Based on the sampling percentages, randomly select the subareas to be sampled.** All subareas within the CoC will ultimately be assigned to one of the following groups:

   a. **Certainty subareas:** These should include any subareas with high rates of unsheltered homelessness or unique circumstances that were identified in the previous step. These subareas will be included in the PIT count.

   b. **Selected sample subareas[2]:** Within each group (e.g., high-density, medium-density groups), sites should be *randomly selected* consistent with the percentage identified for each group. (Random selection means that no site should be prioritized for selection over another, as this would introduce bias into the results.) The selected subareas should then be included in the unsheltered PIT count. After the count, the results from these subareas will be weighted to represent the subareas that are not selected for inclusion.

   c. **Unselected subareas:** All remaining subareas should be excluded from the unsheltered PIT count activities; instead, the sample subareas will be used to account for homelessness in these subareas.

   d. **Excluded subareas:** Subareas with uninhabitable land (e.g., lakes, swampland) should be excluded from the PIT count.

---

[2] This document presents a simplified method of random sampling. There are more sophisticated methods that can be implemented in collaboration with a sampling expert.

5. **Conduct unsheltered PIT counts within each of certainty and sample subareas on the night of the PIT count.** Using the PIT count methods, the counts should be conducted within the certainty and sample subareas. Sampling works best for counts that are conducted on one night (either observation or survey-based counts), since people are unlikely to move between subareas within the CoC over that time. Conducting a service-based count with a sampling methodology is more challenging, since the survey would need to determine whether the individual slept within a sample subarea.

6. **Weight the counts from the sample sites, in accordance with the sampling percentages. Add the counts from the "certainty" sites to derive an estimated count of homelessness for the CoC's full geographic subarea.** To obtain the overall estimate of people experiencing unsheltered homelessness, the CoC should multiply the number of unsheltered homeless persons counted in the sample subareas by the weight associated with the sample subarea. The weight for the sampled subareas is the inverse of the probability of being selected for the sample, calculated as the total number of subareas in the sampling category divided by the number selected for the sample. For example, if the CoC sampled 25 of 50 subareas, the weight would be equal to two (50 divided by 25). Weights should be calculated and applied separately for high-density versus medium-density subareas. For example, if a CoC samples 8 out of 10 subareas where it expects to identify a high number of people experiencing homelessness and 5 out of 10 subareas where it anticipates finding a lower number of people experiencing homelessness, the weight for the count in the high-density subareas is 1.25 (10 divided by 8) and the weight for the count in the medium-density subareas is 2 (10 divided by 5).

Additional References and Resources

- HUD Exchange landing page for PIT Count and HIC Guides, Tools, and Webinars
  - Any additional guidance related to the 2021 PIT count will be posted to the 2021 PIT Count and HIC Guidance and Training tab on this page.
- *Conducting the 2021 Unsheltered PIT Count*
- *Point-in-Time Count Methodology Guide*