

UNSHELTERED POINT-IN-TIME COUNT: PIT COUNT SAMPLING TOOL FOR COCS

Version 1.0

OVERVIEW

Background

HUD requires Continuums of Care (CoCs) to conduct Point-in-Time (PIT) counts of people in unsheltered locations at least every other year. Many CoCs conduct their counts by sampling parts of their jurisdictions and using those samples to derive a full estimate of people experiencing unsheltered homelessness. The **PIT Count Sampling Tool** is intended to help CoCs to develop their sampling approaches and estimate of the total number of people experiencing unsheltered homelessness based on the sample data they collect.

Purpose of this Tool

This tool helps CoCs 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. This tool is intended to support communities who are in the planning phase of their PIT count to determine sample sites and then to use that framework after the count to calculate a weighted estimate that would represent all people experiencing unsheltered homelessness in the CoC.

CoCs that have already conducted a PIT count using sampling and that are seeking assistance with using data collected to estimate the total unsheltered PIT count should use the **PIT Count Estimating Tool**, available on the [How to Use Sampling within a CoC to Conduct an Unsheltered PIT Count page](#) on HUD Exchange.

This tool will only help a CoC establish an estimate of total people experiencing unsheltered homelessness and will not calculate estimates for household, demographic, or subpopulation data. To estimate more specific data about the sampled population, a community would need to conduct a complete sampling of their community.

Guidance on conducting a complete sample of a community is provided in the [How to Use Sampling within a CoC to Conduct an Unsheltered Point-in-Time \(PIT\) Count](#) guidance document. To extrapolate for demographic data, CoCs can use the **PIT Count Extrapolation Tool**, an Excel file available for download [here](#).

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:

- Takes less time to complete than a full census count.
- Requires fewer staff and volunteers to conduct PIT count activities.

Drawbacks of sampling:

- 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.

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 should 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. This tool is designed to guide CoCs through these steps.

TOOL USER GUIDE

Summary of Tabs

Communities will undertake a three-step process to complete the exercises in the PIT Count Sampling Tool. Tabs 1 through 3 each constitute a separate step in this three-step process. The exercises for each tab are outlined in the detailed instructions below. Communities should complete the tabs in sequential order.

- Tab 1—Subarea Types
- Tab 2—Subarea Names
- Tab 3—Sampled Subarea PIT Count

The Extrapolated PIT report is the final roll-up report of the exercises completed in Tabs 1 through 3. No data entry is required on this tab, which provides a weighted total unsheltered PIT count number in Tab 4 to be used for the unsheltered PIT count data submission. Tab 5 is a reference tab where updates to the tool will be recorded over time.

- Tab 4—Extrapolated PIT Count
- Tab 5—Version Change Log

NOTE: Cells are colored by expected user interaction with the PIT Count Sampling Tool:

White: User entry
Blue: Pre-populated from entries in other tabs
Grey: Calculations

Example Tabs

To show how to use this tool in practice, each Tab's section includes screenshots with an example of how this tool could look when completed. The example uses the same numbers throughout to demonstrate how each Tab connects with data entered in prior Tabs. This example was created for illustrative purposes only. It does not reflect real PIT count data or any CoC's geographic subareas or sampling approach.

Detailed Tool Instructions

Instructions for each of the three primary tabs requiring manual entry can be found below.

Tab 1—Subarea Types

Purpose:

Guides communities in planning their subarea sampling strategy. Communities will determine how subareas should be grouped into categories and what percentage of subareas should be sampled within each category. These decisions will have implications for both the accuracy and feasibility of the count effort. Communities should enter subarea categories and sampling percentages before moving on to Tab 2.

Key Terms:

- 1) **Subareas:** Smaller regions within a CoC's geography. 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 in a given subarea type. 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.
 - Example: Six square block in Chicago, Illinois
 - Example: Six square-miles near Bloomington, Illinois
- 2) **Subarea Types:** Categories of subareas based on predicted levels of homelessness. 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. Recommended subarea types are provided below, but communities can use more-granular categories based on local data and knowledge.
 - Example: "Certainty" subareas that must be counted to obtain an accurate count.
 - Example: "High-density" subareas that include known locations where a relatively larger number of people experiencing homelessness frequently sleep.
 - Example: "Medium-density" subareas that include areas where fewer people experience unsheltered homelessness.
 - Example: "Rare or unlikely" subareas that include areas where very few or no people have been known to experience homelessness.
 - Example: "Uninhabitable" subareas where it would be implausible for someone to experience unsheltered homelessness (e.g., densely forested area).
- 3) **Sampling Percentage:** The percentage of subareas to be sampled for a given subarea type.

Community Inputs:

- 1) **Column A:** Input subarea types. These should be limited to subarea types where it is possible for someone to experience homelessness. CoCs should exclude any uninhabitable and unsamplable subarea types. Column A is prepopulated with recommended subarea types, but communities can input as many subarea types as makes sense for their jurisdiction. Enter only one subarea type per cell.
- 2) **Column B:** Input a sampling percentage for each subarea type. There is no universal rule about what these percentages should be; it depends on what is feasible for each community. In general, communities should keep the population density of the subarea type in mind. For high-density subarea types, the sampling percentage should be as high as possible. For medium-density and rare/unlikely subarea types, the sampling percentage can be lower, but do not sample so few that an unexpected (and possibly unrepresentative) count in one of the sampled subareas will spoil the final extrapolated PIT estimates for all medium-density or rare/unlikely subareas. The task of assigning sampling percentages to each subarea type is a chance for communities to take stock of how many subareas they will be able to canvass or survey and what level of confidence they want to have in the extrapolated PIT estimates. Higher sampling percentages will increase confidence in the extrapolated estimates.
- 3) **Column C:** Input the total number of subareas within each subarea type.
- 4) **Submit:** Click the "Submit" button. This will:
 - a) generate calculated counts in Column D and
 - b) transfer the inputted subarea types to Tab 2.If communities later make changes to the inputs in Column A, they must click "Submit" again.

Calculated Values:

- 1) **Column D:** Displays the calculated counts of subareas to be sampled, based on the sampling percentages communities entered in Column B and the counts of subareas communities entered in Column C. These counts represent the number of subareas in which communities should conduct PIT count activities for each subarea type. Make sure these numbers are neither too low (risking biased extrapolations) nor too high (risking a PIT count effort that is not feasible, given CoC resources). Adjust the sampling percentages if needed. Once satisfied, advance to Tab 2.

Example Tab 1:

In the example Tab 1 below, the user has input four subarea types (Column A), each with a different sampling percentage (Column B) and each with a different total number of subareas (Column C). Column D shows that the CoC should plan to sample or conduct PIT count

activities in 19 total subareas, including three “Certainty Subareas,” four “High-Density Subareas,” nine “Medium-Density Subareas,” and three “Rare Subareas.”

Example Tab 1

	A	B	C	D	E	F	G
1	Subarea Types and Sampling Percentages				Submit		
2	Subarea type	Sampling percentage	Total # of subareas	Total # of subareas to be sampled			
3	Certainty Subareas	100%	3	3			
4	High-Density Subareas	75%	5	4			
5	Medium-Density Subareas	45%	20	9			
6	Rare Subareas	20%	15	3			
7							
8							
9							
10							

Tab 2—Subarea Names

Purpose:

Prepares data for the sampling algorithm. Communities will input names or labels for all subareas within each subarea type. This should be an exhaustive list, with the exception of uninhabitable subareas.

Pre-Populated Fields:

- 1) **Row 7:** Subarea types imported from Tab 1.

Community Inputs:

- 1) **Columns A, B, C, ...:** Starting in Row 8, input labels for specific subareas under each subarea type. The number of columns requiring input will vary according to how many subarea types you specified in Tab 1. There are no limitations on how many subareas may be entered under each subarea type. Keep in mind that the subareas do not need to be the same size geographically, but subareas within a given subarea type should be relatively balanced in their expected rates of unsheltered homelessness. Enter only one subarea per cell.
- 2) **Cell A3:** Input a random seed in Cell A3. This can be any number greater than zero. The random seed specifies a starting point for the random process used to create the sample. The purpose of Cell A3 is to give communities the ability to "lock in" particular sampling results, or to generate a different sample if needed. To ensure that the same sampling results will be produced each time the sampling process is executed, enter a

random seed and leave it unchanged. To obtain different results, enter a different random seed.

- 3) **Sample:** Click the "Sample" button to initiate the sampling process. For each subarea type, X% of the subareas that communities entered will be randomly selected, where "X" equals the sampling percentage that communities entered for that subarea type. The number of subareas sampled will be rounded to the nearest whole number. This process might take several minutes to complete if a large number of subarea types and subareas were input. Once complete, communities will be automatically advanced to Tab 3.

Example Tab 2:

Example Tab 2

	A	B	C	D
1	Random Seed	Sample		
2				
3			9	
4				
5	Subarea Names			
6				
7	Certainty Subareas	High-Density Subareas	Medium-Density Subareas	Rare Subareas
8	Certainty 1	High 1	Medium 1	Rare 1
9	Certainty 2	High 2	Medium 2	Rare 2
10	Certainty 3	High 3	Medium 3	Rare 3
11		High 4	Medium 4	Rare 4
12		High 5	Medium 5	Rare 5
13			Medium 6	Rare 6
14			Medium 7	Rare 7
15			Medium 8	Rare 8
16			Medium 9	Rare 9
17			Medium 10	Rare 10
18			Medium 11	Rare 11
19			Medium 12	Rare 12
20			Medium 13	Rare 13
21			Medium 14	Rare 14
22			Medium 15	Rare 15
23			Medium 16	
24			Medium 17	
25			Medium 18	
26			Medium 19	
27			Medium 20	

In the example Tab 2 above, the user has clicked "Submit" on Tab 1 to generate the subarea types listed in Row 7. The user has named all specific subareas (Column A, C, E, and G) within each subarea type (Row 7). The total number of subareas named within each subarea type equals the total number of subareas by subarea type listed in Column C of Tab 1. The user has entered 9 as a "Random Seed" in Cell A3.

Tab 3—Sampled Subareas & PIT

Purpose:

Provides a worksheet for communities to enter PIT counts for the subareas randomly sampled by the algorithm. Communities will conduct PIT count activities in the subareas listed, then input the unsheltered PIT counts in the adjacent cells.

Pre-Populated Fields:

- 1) **Row 5:** Subarea types imported from Tab 1.
- 2) **Columns A, C, E, ...:** Subareas randomly selected to be sampled as a result of communities' inputs in Tab 2.

Community Inputs:

- 1) **Columns B, D, F, ...:** After conducting PIT count activities, enter the final unsheltered PIT count for each subarea listed. Then advance to Tab 4.

Example Tab 3:

In the example Tab 3 below, the user has clicked “Sample” on Tab 2 to generate the sample subareas – that is, the list of subareas in which the CoC will conduct PIT count activities (listed in Columns A, C, E, and G). The CoC has conducted PIT count activities, and the user has input the total number of people identified in each of the sampled subareas.

Example Tab 3

	A	B	C	D	E	F	G	H
1	Sampled Subareas & Unsheltered PIT Counts							
2	Certainty Subareas		High-Density Subareas		Medium-Density Subareas		Rare Subareas	
3	Subarea Name	PIT Count	Subarea Name	PIT Count	Subarea Name	PIT Count	Subarea Name	PIT Count
4	Certainty 3	145	High 4	25	Medium 17	17	Rare 2	2
5	Certainty 2	72	High 2	57	Medium 13	9	Rare 14	16
6	Certainty 1	52	High 1	34	Medium 12	25	Rare 13	9
7			High 5	45	Medium 8	27		
8					Medium 14	32		
9					Medium 7	16		
10					Medium 11	14		
11					Medium 2	12		
12					Medium 5	21		
13								

FINAL REPORT AND REFERENCE TAB

Information about the uses of the final Extrapolated PIT and reference tab contained in this tool can be found below.

Tab 4 – Extrapolated PIT Count

Purpose:

Compiles the PIT counts entered for the sampled subareas and extrapolates these counts to obtain an overall weighted unsheltered PIT count estimate for the community.

Note: There is no manual data entry required in the Extrapolated PIT Count tab. The Extrapolated PIT Count tab is the final roll-up report of the community's completed exercises in Tabs 1 through 3.

How to read the Extrapolated PIT Count tab:

Cell B2	The total number of people experiencing unsheltered homelessness in all sampled subareas on the PIT count night.
Cell B3	The estimated total number of people experiencing unsheltered homelessness in the community on the PIT count night (across all sampled and unsampled subareas). This is the final estimate, extrapolated from the sampled subarea PIT counts.
"PIT Subtotals" table, Column A:	Subarea types prepopulated from Tab 1.
"PIT Subtotals" table, Column B:	Summed totals of subarea PIT counts by subarea type, calculated from inputs on Tab 3.
"PIT Subtotals" table, Column C:	Within a given subarea type, the probability that a given subarea was randomly sampled. These are the sampling percentages that were entered on Tab 1.
"PIT Subtotals" table, Column D:	Weight, calculated as the inverse of the probability of selection (i.e., $1 / \text{probability of selection}$).
"PIT Subtotals" table, Column E:	The extrapolated estimate of people experiencing unsheltered homelessness in the subarea type, calculated as the summed total of subarea PIT counts (Column B) multiplied by the weight (Column D).

Example Tab 4:

The example Tab 4 below has pulled information from the example Tabs 1 through 3 to calculate a weighted PIT count sample. The example user’s CoC counted 630 people across all their sampled subareas (the “Unweighted Sample” in Cell B2), and to account for all of the subareas not sampled they will include 1003 people in their PIT count (the “Weighted Sample” in Cell B3).

To break down how it arrived at these sample totals, Tab 4 shows the total number of people counted by subarea type (Column B, Rows 8–11), which it pulls from Tab 3. The sum of these totals equals the “Unweighted Sample” (Cell B2). It then documents both the probability of each sample subarea’s selection (Column C, Rows 8–11, pulled from Tab 1) and the inverse or weight of that probability (Column D, Rows 8–11) to calculate the total number of people to include in the PIT count by subarea type (Column E, Rows 8–11). The sum of these Column E totals equals the “Weighted Sample” in Cell B3.

Example Tab 4

	A	B	C	D	E
1	Unsheltered PIT Count Totals		Instructions: All values on this tab are auto-calculated. (Row 8 in Column B) looks right. Cell B2 should equal the sum of the number of people counted in the PIT count for this subpopulations when using a sampling approach.		
2	Unweighted Sample	630			
3	Weighted Sample	1003			
4					
5					
6	Unsheltered PIT Count Subtotals				
7	Subarea Type	Summed Total # People Counted	Probability of Selection	Inverse of Probability (i.e., Weight)	Total # People to Include in PIT Count
8	Certainty Subarea	269	100%	1.0	269
9	High-Density Subarea	161	75%	1.3	215
10	Medium-Density Subarea	173	45%	2.2	384
11	Rare Subareas	27	20%	5.0	135
12					

Tab 5 – Version Change Log

Purpose:

This reference tab includes three columns that provide information on updates made to this tool over time.

- **Column A** describes the version of the tool when the update was applied.
- **Column B** is the date the update was made.
- **Column C** is a description of the change made. As a user of the PIT Count Sampling Tool, **you cannot edit this tab**. This tab includes only reference information about updates made to the tool over time.

ADDITIONAL RESOURCES

To view all of HUD’s PIT count and HIC resources visit the [PIT and HIC Guides, Tools, and Webinars Page](#). This is a one-stop site where communities can access all of HUD’s guidance on how to plan and implement their PIT counts. The resources include:

- [How to Use Sampling within a CoC to Conduct an Unsheltered PIT Count](#) guidance
- The PIT Count Data Extrapolation tool (available for download [here](#))
- The [PIT Count Methodology Guide](#)
- PIT count [survey tool resources](#)
- [PIT count implementation resources](#).

If you have technical questions about this PIT Count Sampling Tool, please submit your questions to the [Ask A Question \(AAQ\) portal](#). In Step 2 of the question submission process, select “HDX Homelessness Data Exchange (including PIT, HIC, AHAR, Sys PM, LSA, and PDX)” from the “My question is related to” drop down list.