Title of template

**Skill level needed:**Basic/Intermediate/Advanced

**Sample designs supported:**

1, 2, 3, 4, 5, 6

Designing and Implementing Gridded Population Surveys

**gridpopsurvey.com**

**B5. PSU review – Excel to drop & replace**

Last updated: Aug 2022

**PSU substitution in Excel**

## Examples: Uganda & Nigeria

**Motivation:** Use this tutorial while you are reviewing each PSU over satellite imagery (Tutorial B1, B2, or B3). When you discover a PSU that is in an **insecure** or **inaccessible** location, and/or visual inspection of satellite imagery shows that there are **none (or very few) habitable buildings** (e.g., because the PSU is located over a power plant, airport, or graveyard), then use this tutorial to (systematically or randomly) select a suitable replacement PSU.

**Example 1 (Uganda) – when you know the exact order of sample frame units during PPS sampling:** In this first example, survey implementers selected 40 GridEZ units from 12 refugee settlements in Uganda (see Tutorial A6). Because the team generated the sample frame and performed sample selection themselves in R, they knew the exact order of the sample frame when PPS sampling was performed. Thus, they are able to replace any dropped PSUs with the next unit (within the stratum) from the ordered sample frame.

**Steps:**

1. Below is a section of the ordered Uganda first stage (GridEZ) sample frame. Let us say that the team learned that sampled GridEZ unit 2876 was no longer accessible; the survey team would replace it with GridEZ unit 2868 because it is next in the ordered sample frame. The assumption is that the replacement PSU is similar to the dropped PSU.



**Example 2 (Nigeria) – when you do not know the order of sample frame units during PPS sampling:** In this second example, a team selects a sample following the Nigeria MICS design (see Tutorial A1). While the survey team has access to the sample frame as a raster spatial file (as part of GridSample output), they do not have ready access to the sort-order of the frame when PPS sampling was performed. Therefore, this team has selected 10% back-up PSUs within each stratum, and they will now divide the PSUs into “main” and “backup” samples, by strata.

**If using GridSample output, refer to the provided Excel spreadsheet to implement and document the following steps.**

**Steps:**

1. Open the Excel file of PSUs provided with GridSample output; this file is identical to the shapefile attribute table and includes information about strata, clusters (PSUs), and grid cell units. See the GridSample User Manual for a description of each attribute: [www.gridsample.org/tutorial](http://www.gridsample.org/tutorial). In this tutorial we focus on these three attributes:
	1. cl\_id: Cluster (PSU) ID
	2. cl\_type: Cluster (PSU) type
	3. st\_name: Strata name
2. Add a column (attribute) and name it “random\_num”. Use the formula “=rand()” to generate a random number for every PSU in the dataset.
3. Select and copy the random number column, right-click, and paste “values” to permanently store a static set of random numbers with the PSU dataset.



1. Sort the dataset by strata name (or ID) and the random number.



1. If the sample is stratified, you need to repeat this step for each stratum. To reduce the risk of typos, we recommend using the column filter to display one stratum at a time. Under cl\_type, change the cluster type from “main” to “backup” for the first *n* sorted records (where n is the number of backup PSUs in the stratum). In this example, each stratum has 6 backup PSUs except Kano and Lagos which have 12 backup PSUs each. The assumption here is that the randomized allocation of PSUs to “backup” preserves the representivity of the sample.
2. Now, as you review PSUs over satellite imagery, follow these steps to drop and replace inaccessible, insecure, or unusable PSUs.
	1. For the unusable PSU, change cl\_type from “main” to “drop”.
	2. Maintaining the sort order (Strata name, then random number), identify the backup PSU with the smallest random number, and change cl\_type from “backup to “main”.
	3. If a second, third, or further PSU is dropped in the stratum, replace it with the remaining backup PSU with the smallest random number.

This tutorial has provided two approaches to replacing PSUs that are not accessible, secure, or usable. We deliberately provide instructions in Excel to ensure that all users understand the processes and can implement them. If you are an Intermediate or Advanced user, these steps can be performed and documented in Stata, R, or another programming language. For teams with GIS skills, perform a table join based on cluster (PSU) ID to update your sample.

 [This work](http://www.gridpopsurvey.com/)© 2022 is licensed under [CC BY 4.0](http://creativecommons.org/licenses/by/4.0/?ref=chooser-v1)