MICS/NICS 2016-17 Cluster GPS Coordinates -Displaced for Public Use

Dale Rhoda, Biostat Global Consulting

07 May 2019

Contents

1	Background	1
2	Dataset contents	2
3	Original data collection - identifying a representative undisplaced coor- dinate for each cluster	2
4	Calculating displaced coordinates for each cluster	3
5	A Note Concerning Boundaries	4

1 Background

Key results from Nigeria's 2016-17 MICS-NICS survey are available in the final reports from UNICEF and the National Primary Health Care Development Agency (NPHCDA) and the National Bureau of Statistics (NBS). Those reports and microdata are available at https://www.nigerianstat.gov.ng/nada/index.php/home.

This document is meant to accompany an auxiliary dataset of cluster location coordinates (latitude and longitude) that may be helpful to persons who wish to conduct spatial analysis of MICS-NICS microdata. The coordinates appear in a spreadsheet named MICS-NICS_2016-2017_displaced_cluster_gps_coordinates.xls.

The coordinates there are intended to fall near, but not precisely on the locations where the data were collected. In order to protect the privacy and safety of survey respondents, the locations of individual households are not released, but a coordinate is released for 2,701 of the 2,702 clusters in the survey. To further protect respondent privacy, each cluster

coordinate has been displaced in a random direction and by a random distance from the original cluster location. The purpose of this document is to give a short description of the dataset and how it was compiled.

2 Dataset contents

The dataset holds six variables:

- hh1: cluster ID number (integer)
- area: either Urban or Rural (text string)
- **gps_source**: indication of how the original (undisplaced) coordinate was collected (text string)
- **state name**: state in which cluster should appear (text string)
- **latitude** (decimal number)
- longitude (decimal number)

When the cluster locations are plotted on a map, some of the coordinates do not fall within the state listed in the *state_name*. This reflects the fact that some of the undisplaced locations must have been measured with some error or imprecision and even the undisplaced coordinates fall outside the state boundaries. Spatial analysts should check each cluster's location using boundaries of its respective state before deciding whether and how to include that cluster in their analyses. (Note that there may also be differences in different GIS shapefile definitions of Nigerian state borders, so a point that falls inside a particular state, but near the border using one shapefile, may fall in a neighboring state if another shapefile is used in the analysis. Be sure to use a shapefile that is approved for your project).

3 Original data collection - identifying a representative undisplaced coordinate for each cluster

There is no single comprehensive source of cluster coordinates for this survey. The undisplaced coordinates that were used to construct this dataset were compiled from three sources:

1. When MICS interview teams visited a cluster, they usually recorded three GPS locations around its periphery. In some cases the three coordinates for a single cluster fall very near each other and are probably error-free. In other cases one or two of the coordinates are very far away from the others, and so some of the coordinates are probably wrong.

- 2. A single coordinate per cluster was furnished for most NICS clusters.
- 3. After combining coordinates from the MICS and NICS field teams, a total of 88 clusters were missing GPS coordinates, so in 2018, a small team at the NBS conducted a desktop exercise using survey microplans and GIS software to pinpoint those cluster locations in retrospect. The team succeeded in finding coordinates for 87 of those clusters.

Each cluster had from one to three undisplaced coordinates. A single location was selected to characterize each cluster, using the following steps:

- 1. For clusters with 2+ periphery coordinates collected during household interviews:
 - (a) Calculate the average latitude and average longitude if all three points fall near each other. Use the average as the undisplaced cluster coordinate.
 - (b) Calculate the average latitude and longitude of two if only two of the three points fall near each other. Use the average as the undisplaced cluster coordinate.
 - (c) If none of the three points fall near any other, arbitrarily select the first of the three to use as the undisplaced cluster coordinate.
- 2. For clusters with only one periphery coordinate, use that coordinate as the undisplaced cluster coordinate.
- 3. For clusters with no GPS coordinates, the NBS desktop team identified a single undisplaced cluster coordinate using GIS software and the archived survey microplans.

Undisplaced coordinates were found or calculated for 2,701 of the 2,702 clusters that appear in the MICS/NICS dataset, but as noted above, some of those coordinates do not appear to fall inside the state that they should, so some of the coordinates are wrong or imprecise.

4 Calculating displaced coordinates for each cluster

A new displaced coordinate was calculated for each of the 2,701 clusters, using the following logic from the Demographic and Health Systems as described in the document found at https://dhsprogram.com/pubs/pdf/SAR7/SAR7.pdf.

- For each urban cluster, select a random direction, and displace the cluster coordinate by a random distance between 0 and 2 km.
- For a randomly selected 99% of rural clusters, select a random direction and displace the coordinate by a random distance between 0 and 5 km.

 \bullet For a randomly selected 1% of rural clusters, select a random direction and displace the coordinate by a random distance between 0 and 10 km.

Random directions and distances were selected independently for each cluster. Wherever possible, the displaced coordinate was only approved if it falls within the state that it should. For some clusters, the original undisplaced coordinate falls so far outside the correct state boundary, that no displaced point could fall in the correct state and stay within the random range limit from the undisplaced location. In those cases, the displaced coordinate falls outside the boundaries of the correct state.

5 A Note Concerning Boundaries

The following pages show the undisplaced and displaced coordinates for each state.

With 2,701 clusters that have coordinates, many clusters are fall near a state boundary. Precisely which coordinates fall inside or outside a state boundary depends on which set of boundary coordinates the analyst uses. Note that when the undisplaced coordinates clearly fall outside the appropriate state boundaries, the displaced coordinates fall outside those boundaries, too, because the displaced coordinates are constrained to be within 2 or 5 or 10 km of the undisplaced coordinate. Each analyst is encouraged to carefully evaluate the coordinates with respect to state boundaries before incorporating them into spatial analyses.

Finally, note that the undisplaced coordinates are not included with the dataset that accompanies this document.









































































