



Census Microdata and the 2030 Agenda: IPUMS for SDG Measurement and Disaggregation

Kristen Jeffers*

University of Minnesota, Minneapolis, United States of America – kjeffers@umn.edu

Sula Sarkar

University of Minnesota, Minneapolis, United States of America – sanb0027@umn.edu

Lara Cleveland

University of Minnesota, Minneapolis, United States of America – cleveland@umn.edu

Rodrigo Lovaton

University of Minnesota, Minneapolis, United States of America – lovat003@umn.edu

Abstract

The United Nations 2030 Agenda for Sustainable Development proposes 17 goals and 169 targets that carry on the momentum generated by the Millennium Development Goals. The proposed framework for monitoring the SDGs emphasizes the need for disaggregated indicators that measure progress among different demographic and social groups at various levels of sub-national geography. The Sustainable Development Solutions Network recommends spatial disaggregation of indicators and stratification by sex, gender, age, income, disability, ethnicity, indigenous status, economic activity, and migrant status where relevant. While enhanced data collection will be necessary to monitor several SDGs, high-density census microdata samples, like those disseminated by IPUMS-International, represent useful data that are part of the existing statistical infrastructure of most developing countries. These data are highly representative of national populations, are collected at regular intervals, and include measures of the population characteristics required for SDG indicator disaggregation. In this study, we calculate several SDG indicators for targets related to housing conditions, youth inactivity, and gender equality. We present measures calculated using data from 2000- and 2010-round censuses to illustrate the utility of harmonized census data for tracking development across census years and for establishing baseline measures for monitoring SDG progress at a sub-national level.

Keywords: Sustainable Development Goals (SDGs); census microdata; IPUMS-International; data harmonization.

1. Introduction

The United Nations 2030 Agenda for Sustainable Development proposes 17 goals and 169 targets that aim to carry on the momentum generated by the Millennium Development Goals. The Sustainable Development Goals (SDGs) and targets concentrate on the eradication of poverty, hunger and inequality; access to education and healthcare; gender equality; environmental sustainability; economic, social, and technological progress, and the establishment of new partnerships (Sustainable Development Solutions Network 2015). The proposed framework for monitoring the SDGs emphasizes the need for disaggregated indicators that measure progress among different demographic and social groups at various levels of sub-national geography. The Sustainable Development Solutions Network recommends spatial disaggregation and stratification by sex, gender, age, income, disability, ethnicity, indigenous status, economic activity, and migrant status for nearly half of the proposed monitoring indicators. While enhanced data collection will almost certainly be necessary to monitor several SDGs, high-density census microdata samples, like those disseminated by IPUMS-



International, represent useful data that are part of the existing statistical infrastructure of most developing countries. More than twenty SDG indicators are measurable with census microdata included in IPUMS-International. In this study, we calculate indicators for several SDG targets as operationalized by the United Nations using variables widely available from IPUMS-International.

2. Data

The Integrated Public Use Microdata Series-International (IPUMS-International) disseminates high-precision census microdata samples from around the world, spanning five continents and covering three-quarters of the world's population. The data include variables representing a broad range of population characteristics, including fertility, nuptiality, life-course transitions, migration, disability, labor-force participation, occupational structure, education, ethnicity, and household composition (Ruggles et al. 2003; Sobek et al. 2011). Censuses are taken at regular intervals, commonly every 10 years or so, and data in IPUMS are available for multiple census years for most countries in the collection. These pooled data are available to researchers and students world-wide free of charge through an online dissemination system. The principal advantage of IPUMS-International is the reconciliation of sample-specific variable coding schemes to produce datasets that integrate records across time and space, making the data ideal for comparative analysis. IPUMS has also developed a method for creating spatially consistent first and second administrative boundaries (Sarkar et al. 2015).

In this short paper, we calculate and visualize indicators for targets 11.1) By 2030, ensure access for all to adequate, safe and affordable housing and basic services and upgrade slums; 8.6) By 2020, substantially reduce the proportion of youth not in employment, education or training; and 5.b) Enhance the use of enabling technology, in particular information and communications technology, to promote the empowerment of women. We present measures calculated using data from 2000- and 2010-round censuses to illustrate the utility of harmonized census data for tracking development across census years and for establishing baseline measures for monitoring SDG progress at a sub-national level.

3. Analysis and results:

We disaggregate national trends by first- and second-level administrative units, identifying areas in Mexico with poor housing conditions (target 11.1), gendered trends with respect to youth not in employment, education, or training in Mozambique (target 8.6), and gender disparities in cell phone ownership in Ghana (target 5.b)

Urban slum population in Mexico (target 11.1): The United Nations guidelines recommend operationalizing the assessment of this goal as the percentage of urban households living in slums. UN-HABITAT, the agency responsible for global monitoring of the indicator, defines a slum household as a group of individuals living under the same roof lacking one or more of the following conditions: access to improved water, access to improved sanitation, sufficient living area, durability of housing, and security of tenure¹. Data on water access, sewage disposal, and size and quality of dwellings are commonly collected by censuses in the developing world. IPUMS integrates these data to build variables with conceptually consistent categories and harmonized codes across all samples. Using IPUMS integrated variables (Table 1), we computed dichotomous dummy variables to identify households that lacked access to improved water, improved sanitation, sufficient living area, or

¹ Secure tenure is included as a fifth condition in the official definition of slum, but information on secure tenure is not available for most countries, and only the first four conditions are used estimate the proportion of urban population living in slums in practice. For detailed definitions of improved water, improved sanitation, sufficient living area, durable housing, and secure tenure, see: http://www.un.org/esa/sustdev/natlinfo/indicators/methodology_sheets/poverty/urban_slums.pdf



durable housing according to the UN operationalization of each of the four slum conditions. Urban households lacking one or more of these amenities were considered to be living in slum dwellings.

Table 1: IPUMS integrated variables used to identify urban households living in slums

URBAN	Urban-rural status ²
WATSUP	Means by which household receives its water
SEWAGE	Access to sewage system or septic tank
ROOMS	Number of rooms occupied by the household
PERSONS	Number of persons in the household
WALL	Wall or building material
FLOOR	Floor material
ROOF	Roof material

The percentage of urban households living in slums in Mexico decreased from 20.1 percent in 2000 to 12.4 percent in 2010; the trend is visualized in Figures 1 and 2. IPUMS harmonized geographic variables enable us to calculate the same measures at sub-national levels, holding spatial units constant across sample years. The sub-national analyses identify the 133 municipalities (10 percent) where slum populations stagnated or increased between 2000 and 2010. Slum-dwelling remained high or increased in several municipalities in southeastern Mexico, particularly in the state of Tabasco, where poverty rates exceed the national average³. In 2010, 27 percent of households in urban Tabasco experienced at least one slum condition, and 5 percent of households experienced two or more slum conditions.

Percentage of young people in Mozambique not in education, employment, or training (NEET) (target 8.6): This indicator tracks the share of youth (age 15 to 24) who are neither in formal employment nor in full-time education or training. It is a measure of the percentage of youth who are either unemployed, work in the informal sector, or have other forms of precarious jobs.

Table 2: IPUMS integrated variables used to define NEET

SCHOOL	School attendance at, or prior to, time of census
EMPSTAT	Economic activity status (labor force participation) at time of census or during specified time period prior to census

By cross-classifying the IPUMS integrated variables SCHOOL and EMPSTAT we arrive at four categories for youth ages 15-24: attending school, working, both attending school and working, and none of the above. In Mozambique in 2007, 19.6 percent of youth age 15-24 were not in education, employment, or training. NEET in Mozambique was lower than in neighboring countries in the region like Zambia (2010) and Malawi (2008), where NEET rates were 41.7 percent and 29.5 percent, respectively. Disaggregation by sex and administrative unit identifies enclaves of high NEET, particularly among females. Figure 3 shows the percentage of NEET youth in Mozambique in 2007 by sex at the first and second levels of geography. Gender disparities exist throughout the country, but are particularly pronounced in the capital region of Maputo (southern tip of Mozambique) and in the urban center of Tete in the northwest. NEET among young men is concentrated in the province of Gaza, where male employment levels are the lowest in the country. Among young women, high levels of NEET are visible throughout the country. Preliminary results indicate parity in labor force participation at the national level, but much lower levels of school attendance for young women (27.6 percent) compared to young men (45.3 percent).

² In Mexico, any locality with a population of 2,500 or more is considered urban

³ Banco de Información INEGI, <http://www3.inegi.org.mx/sistemas/biinegi/?ind=6300000269>



Gender disparities in cell phone ownership in Ghana (target 5.b): The suggested indicator for target 5.b--enhance the use of enabling information and communications technology to promote the empowerment of women--is the proportion of individuals who own a mobile phone, by sex. We use the latest census from Ghana, which asked respondents whether they owned a mobile phone, and disaggregate by sex. Such technology questions are becoming increasingly common in the developing world. We calculate and visualize phone ownership among adults age 12 and older by sex at the second administrative level of geography (district) for Ghana (Figure 4). Cell phone ownership is more common in the urbanized south among both sexes. Gender disparities exist throughout the country. Mobile phone ownership rates are higher among men in every district in Ghana, though gaps are less pronounced in the urban centers of the country.

5. Conclusions

Our paper demonstrates the use of IPUMS data for monitoring SDG indicators at the regional and local level. Integrated microdata, measures of several demographic and social characteristics, and spatially consistent geographic units provide monitoring agencies and national governments with reliable data for the establishment of disaggregated baseline measures for a number SDGs. As new census data are collected and integrated in IPUMS, SDG progress can be measured against these baselines according to methods presented in this paper.

Along with a comprehensive assessment of targets 11.1, 8.6, and 5.b, our final paper will include calculation goals 4) ensure inclusive and equitable quality education and promote lifelong learning opportunities for all; 6) ensure availability and sustainable management of water and sanitation for all, and 7) ensure universal access to affordable, reliable, and modern energy services. Indicators will be measured for countries of differing levels of development from several regions of the world to demonstrate the diverse applicability of data from the IPUMS collection.

Figure 1: Percentage of the urban population living in slums, Mexico 2000

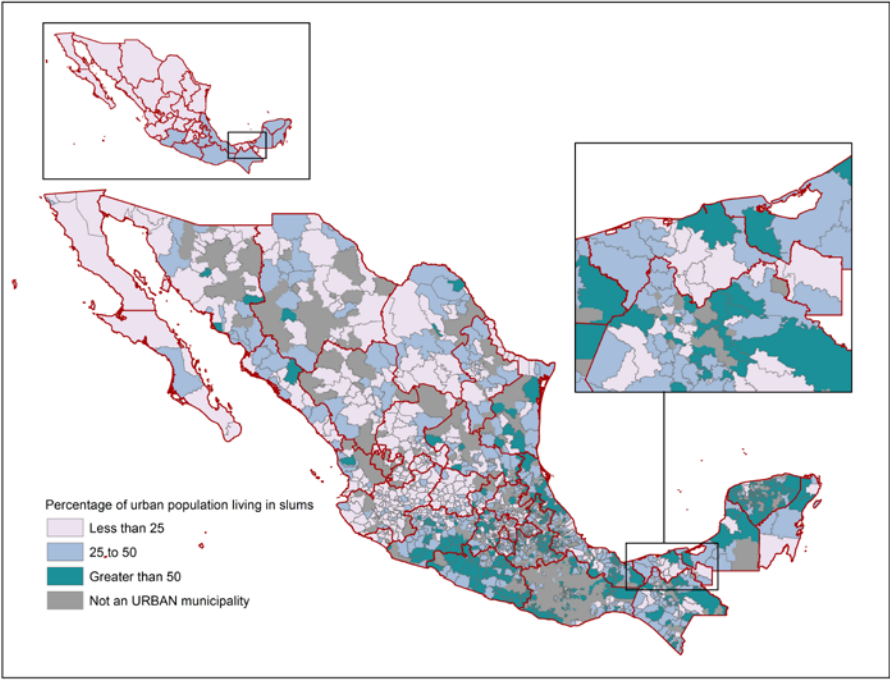




Figure 2: Percentage of the urban population living in slums, Mexico 2010

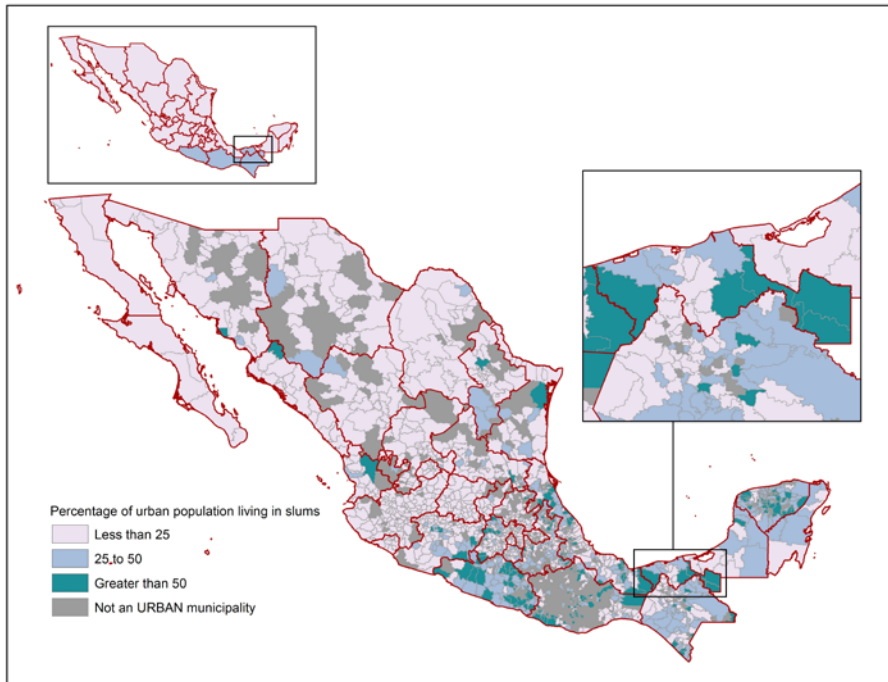


Figure 3: Percentage of youth age 15-24 not in education, employment, or training, Mozambique 2007, by sex

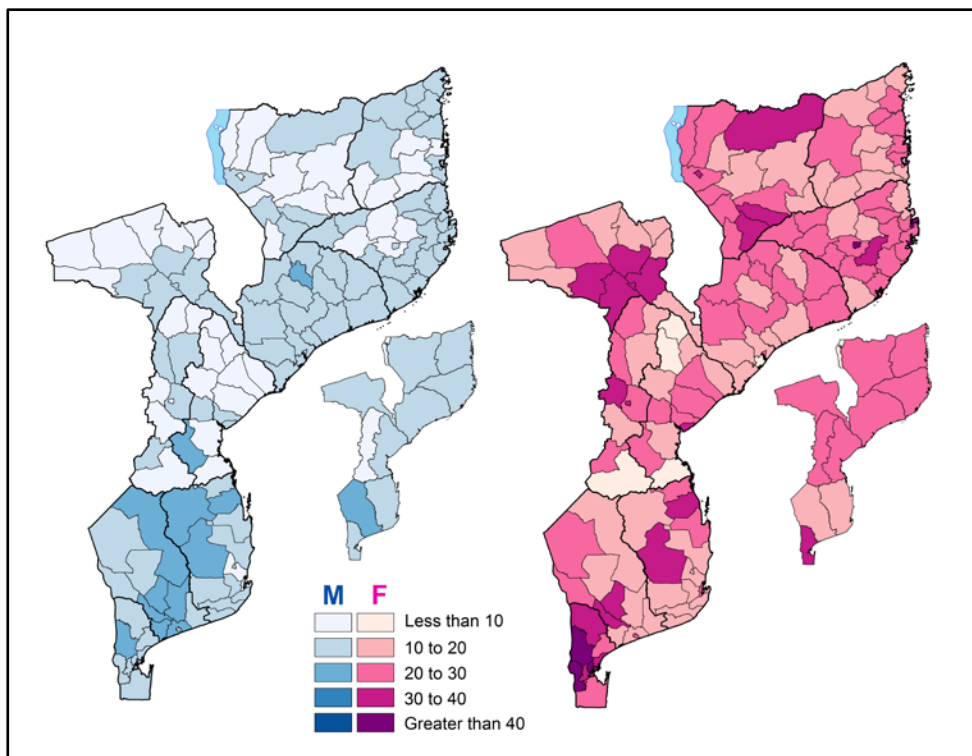
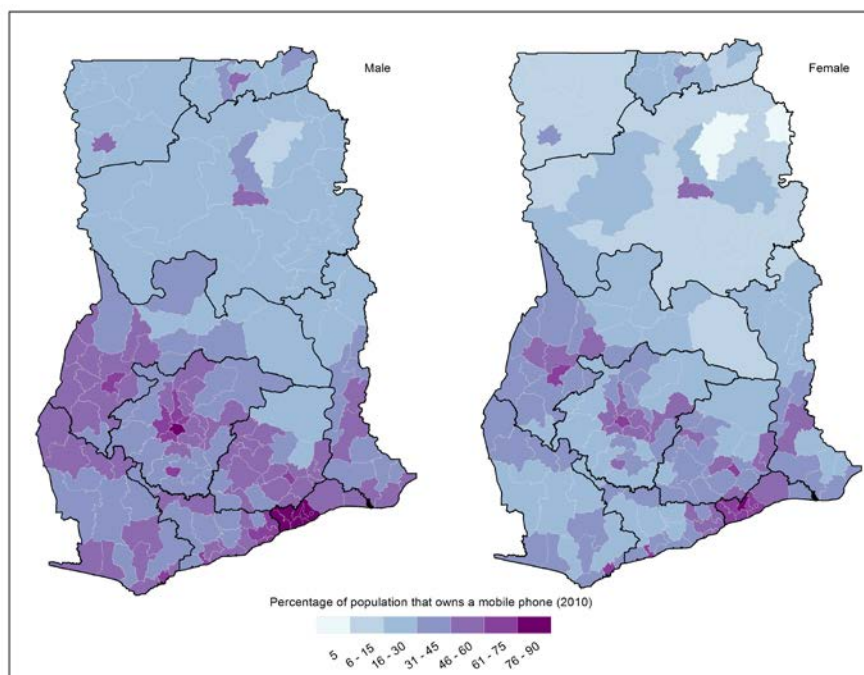




Figure 4: Percentage of the population that owns of mobile phone by sex, Ghana 2010



References

Minnesota Population Center. *Integrated Public Use Microdata Series, International: Version 6.4* [dataset]. Minneapolis: University of Minnesota, 2015. <http://doi.org/10.18128/D020.V6.4>.

Ruggles, S., King, M.L., Levison, D., McCaa, R. & Sobek, M. (2003). IPUMS-International. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 36(2), 60–65, DOI 10.1080/01615440309601215.

Sarkar, S., Cleveland, L., Silisyene, M. & Sobek, M. (2015). Harmonized census geography and spatio-temporal analysis: gender equality and empowerment of women in Africa. Paper presented at the Annual Meeting of the Population Association of America, San Diego, CA in April.

Sobek, M., Cleveland, L., Flood, S., Kelly Hall, P., King M.L., Ruggles, S., & Schroeder, M. (2011). Big Data: Large-Scale Historical Infrastructure from the Minnesota Population Center. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 44(2), 61–68, DOI 10.1080/01615440.2011.564572.

Sustainable Development Solutions Network. (2015). Indicators and a Monitoring Framework for the Sustainable Development Goals: Launching a Data Revolution for the SDGs. A report to the Secretary-General of the United Nations by the Leadership Council of the Sustainable Development Solutions Network. Available for download from: <http://unsdsn.org/wp-content/uploads/2015/05/150612-FINAL-SDSN-Indicator-Report1.pdf>