



# IPUMS – Integrated Public Use Microdata Country Report: Brazil

**Eduardo Pereira Nunes**

(President of the Brazilian Institute of Geography and Statistics - IBGE)

**Zélia Bianchini**

(Deputy Director of Surveys - IBGE)

Workshop “Integrating Global Microdata: the Challenge for the 2010 Census Round”  
Columbia University, New York  
February 23, 2008

## Outline

- IBGE products
- IBGE cooperation with IPUMS
- The 2007 Brazilian Censuses
- The Brazilian 2010 Population and Housing Census
- Joint Population Censuses Project
- Future Plans - IPUMS

## IBGE products

- A policy of free dissemination of all products through Internet was adopted in IBGE, since 2001
- Besides the electronic publications the IBGE web page has two important databases to stand out:
  - Aggregated Statistical Tables (SIDRA) - Database with information grouped in territorial levels that allows the users to design tables according to selected information –
  - Multidimensional Statistical Database (BME) – Database with microdata information that allows users to design tables according to selected information and confidentiality constraints (requires Internet subscription)
- IBGE has been releasing public use microdata files for household statistics since the early 1990s
  - Measures taken to protect the confidentiality of these microdata include data suppression of geographical detail

## **IBGE microdata products (household sample surveys)**

- Population and Housing Census (long-form)
- National Household Sample Survey (annual)
- Monthly Employment Survey (6 metropolitan areas)
- National Household Expenditure Survey
- Urban Informal Survey



Endereço <http://mapas.ibge.gov.br/divisao/viewer.htm>

Ir Links

Ministério do Planejamento, Orçamento e Gestão



<< Página Inicial

Ver mapa:

▶ Mapa de Divisões Territoriais



Zoom para UF:

Zoom para Escala:

1:  OK

CAMADAS

- 2005
  - UF 2005
  - Mesorregião
  - Microrregião
  - Município 2
- 2001
- 1997
- 1994
- 1991
- Regiões Metro

LEGENDA

UF 2005



IBGE - 2005

0 1459km

Mapa: -76.45 , -31.39

Intranet local

## **IBGE cooperation with IPUMS**

- IBGE is a founding member of the IPUMS initiative
- Brazilian researchers constitute one of the largest groups of all the nationalities using IPUMS data
- Samples for Brazil were added in 2004
- IPUMS-Latin America project, funded by the National Institutes of Health, which began in mid-2003

# IPUMS Census Microdata Inventory

## Brazilian Samples Availability

Census	Sample Fraction (%)	Household Records	Person Records
1960	5	613,273	3,001,439
1970	5	1,022,207	4,953,759
1980	5	1,343,377	5,870,467
1991	5,8	2,012,276	8,522,740
2000	6	2,652,356	10,136,022

# Brazilian Variables Availability

- **HOUSEHOLD VARIABLES**

- Technical
- Group Quarters
- Geography
- Economic
- Utilities
- Appliances, Mechanicals, Other Amenities
- Dwelling Characteristics
- Constructed Household

- **PERSON VARIABLES**

- Technical
- Constructed Family Interrelationship
- Core Demographic
- Fertility and Mortality
- Nativity and Birthplace
- Ethnicity
- Education
- Work
- Income
- Migration
- Disability



## Example: Utilities Variables - Household

Variable	Variable Label	Type	Codes	1960 Census	1970 Census	1980 Census	1991 Census	2000 Census
<b>ELECTRC</b>	Electricity	H	codes	X	X	X	X	X
<b>WATSUP</b>	Water supply	H	codes	X	X	X	X	X
<b>SEWAGE</b>	Sewage	H	codes	X	X	X	X	X
<b>FUELCK</b>	Cooking fuel	H	codes	X	X	X	X	.
<b>PHONE</b>	Telephone availability	H	codes	.	.	X	X	X
<b>TRASH</b>	Trash collection	H	codes	.	.	.	X	X

## The 2007 Censuses: An integrated operation

- Integration among the
  - **Agricultural Census**
    - All 5,564 Brazilian municipalities
  - **Population Count**
    - 5,435 municipalities with up to 170,000 inhabitants plus 21 above this range located in states in which only one or two municipalities exceeded this population estimation
  - **National Address List for Statistical Purposes**

## The 2007 Censuses Innovations

- National Address List for Statistical Purposes – CNEFE
  - Prepared from records of units surveyed in 2000 Census
  - This file aims to improve the survey, treatment and dissemination of statistical information
  - In the rural area, the geographical coordinates for geocoding agricultural holdings was collected, as well as for health and education establishments
- Replacement of the traditional questionnaire on paper by a Personal Digital Assistant – PDA
- GPS for address location and enumerators orientation integrated with Google Earth® to accomplish a better image display of the collected points

## The 2007 Censuses: Dimensions

- 68,000 enumerators
- 18,000 supervisors
- 162,770 enumeration areas
- 82,000 PDAs with GPS
- 3,500 microcomputers connected to national network
- More than 5,000 data collection stations  
(1,200 microcomputer-equipped and 4,624 access by modem and phone line)
- Broadband access extended to all 534 IBGE offices
- 700 radio antennas in remote areas
- Units surveyed: 109 million people (60% of whole pop.), 30 million households and 5.2 million agricultural holdings
- Total budget app. US\$ 270 million

## The 2007 Censuses: Challenges, lessons learned and actions

- The great technological innovation implemented in the operation of the 2007 Censuses, was the utilization of handheld computers – **PDA**s - in the field work, before and during the data collection of households and agricultural establishments
- This innovation significantly affected the **organization of the processes** involved in the gigantic operation of the 2007 Censuses
- Besides this, the construction of the **National Address List for Statistical Purposes** - CNEFE and its **integration** with cartographic information and with the territorial base which give support to collection and supervision operations and lead to the generation of new digital products, represented a strong effect on the activities of the statistical and geographical Brazilian information systems

## The 2007 Censuses: Challenges, lessons learned and actions

- Considering the technological advances and the new available inputs, the **experience obtained by IBGE with the 2007 Censuses** is extremely important
  - for the adequate **planning of the 2010 Population Census**
  - for the **Project of Alternative Methodologies for Population Censuses**
  - for the current surveys, as it is the case of the use of PDAs in the 2007 National Household Sample Survey

# The Brazilian 2010 Population and Housing Census

- Methodological and technological aspects in planning
  - Traditional Brazilian Population and Housing Census
  - Long and short enumeration forms
    - the short form contains general information about the characteristics of the dwelling and each person in the dwelling
    - the long form is applied to a sample of the population and will contain general and more specific questions about dwelling characteristics, families, and persons

# Joint Population Censuses Project

- 2000 Round Members: Argentina, Brazil, Paraguay, Uruguay (MERCOSUR), Bolivia, Chile
- 2010 Round Members: Argentina, Brazil, Paraguay, Uruguay, Bolivia, Chile, Mexico, Ecuador and Venezuela

## General Objectives

- To obtain comparable statistical data
- Optimization of the Methodological, Human and Technological Resources of the countries for Census operation
- Make easier Horizontal Cooperation
- Potencialize the Technical International Support



# Joint Population Censuses

## Products

- Almost all the Mercosur Censuses carried between 2000 and 2002
- Data Capture with scanners and ICR: Joint Cooperation Network (5 countries)
- Common Classification for Industry (CAES Mercosur)
- Convergent Classifications for Occupation
- Automated Coding: principles homogenized
- Multidimensional Data Base
- Aggregated Data Base

# Joint Population Censuses

## Final Remarks

- Differences with other Projects: unusual way of organization
- Integration of the teams
- Progress in the solution of common problems
- Socialization of Technological Advances
- Project continuation: Pilot tests for new items to be Harmonized: disability, migration, etc.
- Lessons for 2010

## Future Plans - IPUMS

- **IPUMS-Latin America II (2009-2013)**
  - A long-form sample for 2010 census (traditional Population and Housing Census)
  - High precision option for sampling ranging from 10-20% for 1991 and 2000
  - Lower level geography (currently region, state and for 1991 and 2000, metropolitan area)
  - Integrated boundary files at the level of geography available in the microdata beginning with the 1991 census
  - Detailed plan may be discussed on the Workshop on the Latin American Region