# The use of handheld computers in the 2010 Brazilian Population Census

## Américo Vicente S. de Miranda Júnior
### Brazilian Institute of Geography and Statistics

**Uso de las Tecnologias en los Censos y Operaciones Estadísticas**

*Seminario*
*Caracas, Venezuela,* 6 de mayo de 2013

# 2010 Population Census

- **Population Census produces information on the main characteristics of people and households for each 5,565 Brazilian municipalities**
- **Population Census is the main challenge for a Statistical Office, specially in a country like Brazil, with 8,515,692 km2 distributed in an heterogeneous and sometimes inaccessible territory, made of 27 Federation Units and 5,565 Municipalities, encompassing approximately 67 million housing units.**

## Territorial Database
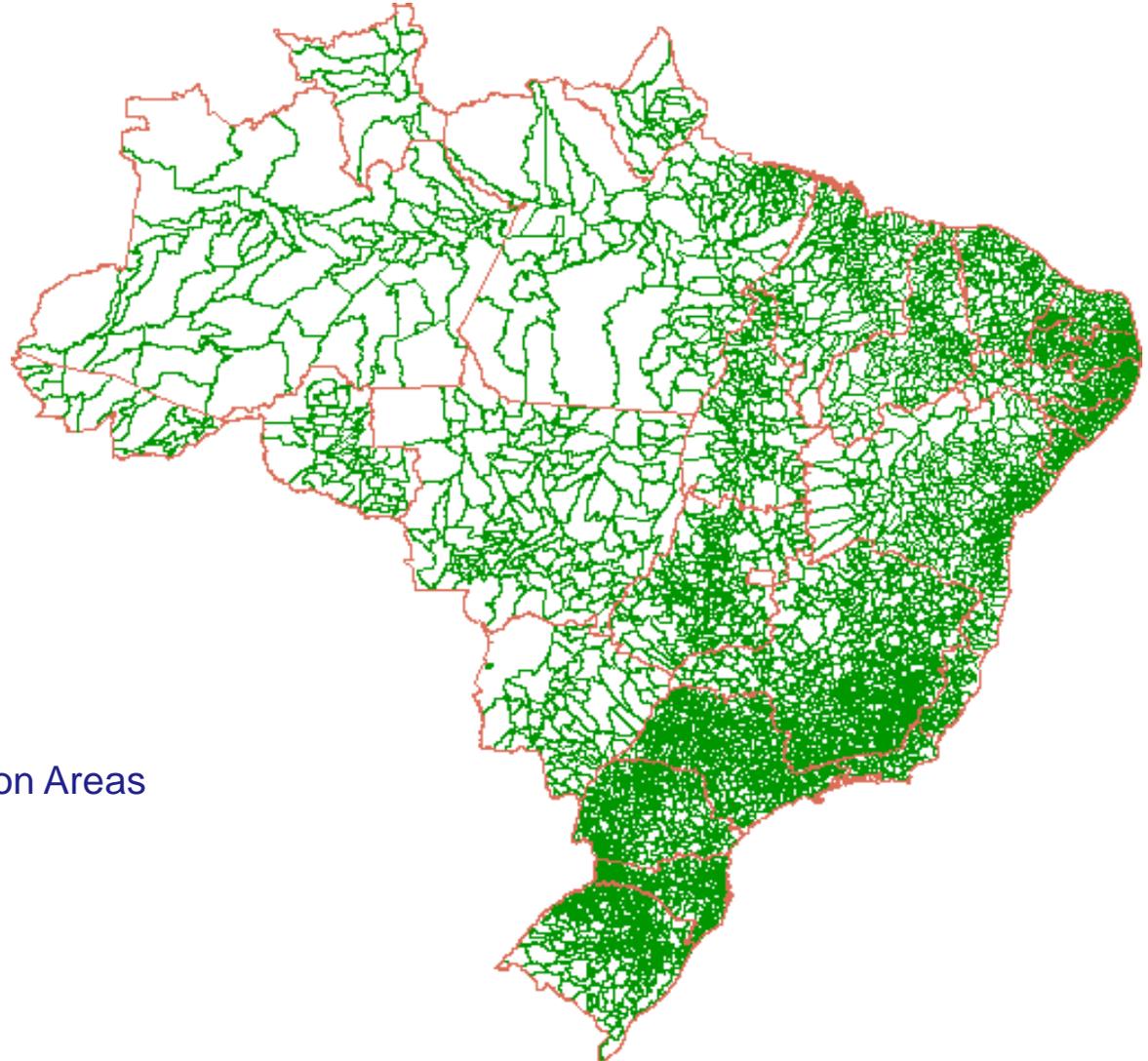
27 Federation Units
(States)

5,565 Municipalities

10,283 Districts

662 Subdistricts/R.A.

316,574 Enumeration Areas
240,382 urban EAs
 76,192 rural EAs

## Pre-Census (from March to June 2010)

**Review of urban Enumeration Areas performed by 32,000 supervisors hired for the Census:**

- Loading enumeration area maps and the list of streets and block-faces associated with the map
- Review / updating digital maps with PDA
- Review / updating association of addresses to digital maps
- Work unit:  block-face

**Duties of Supervisors:**

- For each street: confirm, delete, include, change the name and/or change name's orthography
- For each block-face: include, exclude , confirm or classify as "NAR" (no addresses in the block-face) and collect surrounding characteristics (electricity, sanitation, security, etc.)
- Tracking with GPS the shape of block-faces not represented on the map

**·150,000 LG-750Q, smartphones locked (Phone, 3G) to be used solely as data collection equipment and programs certified by the IBGE**
**·70,000 PDA MIO-P550B (already used in 2007 Censuses)**

- All 57 million occupied housing units were surveyed to collect the housing characteristics and people living there.

- IBGE used two questionnaires:
  - **short form**, with 23 questions for housing unit and 14 for people;
  - **long form** (sample) with 37 questions for housing unit and 71 for people (including same questions of the short form).

- The choice of questionnaire was made by random selection in the handheld computer during the data collection operation, according to the sampling fraction defined for each municipality.
  - The average sampling fraction is 11%. The fractions depend on the number of inhabitants of the municipality and varies from 5% (more then 500,000) to 50% (less then 2,500 inhabitants).

•Data collection was carried out using handheld computers in a face to face mode interview (CAPI) and concurrently the National Address File for Statistical Purposes (CNEFE) was updated

•IBGE staff was responsible for the development the Information and Communication Technology (ICT) for 2010 Census.

•There was no paper questionnaire.

  ◦The GPS receiver allowed the georeferencing of units in the rural areas as well the monitoring of the field operation.

•Each handheld computer was equiped with:

  ◦Enumeration Area Map

  ◦list of addresses developed during pre-census operation (for the EA that this operation was done)

  ◦list of collective living quarters

  ◦short and long form questionnaires

  ◦GPS

  ◦summary information collected to facilitate monitoring enumerator's work.

## Data Collection

Table 1 - Figures of census operations by year of research.

| Description | Population Census 2000 | Population Census 2010 |
|---|---|---|
| Enumerators | 189,937 | 191,972 |
| Households | 44,795,101 | 57,324,167 |
| *Household/Enumerator* | *208* | *302* |
| Population | 169,799,170 | 190,755,799 |
| *Population/Enumerator* | *863* | *1,004* |
| Enumerations areas | 215,811 | 316,574 |
| *Enumerations areas/Enumerator* | *1.14* | *1.65* |
| Questions in short form | 16* | 37 |
| Questions in long form | 90 | 107 |

Source: IBGE, Survey Directorate.
* In the 2000 Census the people responsible for the household answered 25 questions.

## Checks

Table 2- Types of checks laid on the handheld by type of questionnaire.

| Type of Checks | Type of Questionnaire | |
| --- | --- | --- |
| | Short Form | Long Form |
| Validation response | 26 | 57 |
| Warnings | 7 | 10 |
| Functions | 5 | 44 |
| Conditioned Skips | 34 | 111 |

Source: IBGE, Survey Directorate.

## Coding

Table 3-Proportion of people coded during phases of data collection and editing of the 2010 Census, compared to the total number of people reporting each variable.

| Theme | Percentage of people coded in the collection (handheld) | Percentage of people coded in editing step |
|---|---|---|
| **Universe data** | - | - |
| Indian ethnicity | 88.79 | 11.21 |
| Indigenous language | 90.02 | 9.98 |
| | | |
| **Long form data** | - | - |
| Indian ethnicity | 89.73 | 10.27 |
| Indigenous language | 89.54 | 10.46 |
| Religion | 94.30 | 5.7 |
| Undergraduate degree | 84.95 | 15.05 |
| Master degree | 72.44 | 27.56 |
| PhD degree | 76.16 | 23.84 |
| Last municipality people lived | 92.39 | 7.61 |
| Municipality where people lived 5 years ago | 93.45 | 6.55 |
| Municipality where people study | 93.12 | 6.88 |
| Municipality where people work | 90.48 | 9.52 |

Source: IBGE, Survey Directorate.

**LOCAL**

.SIGPC – Management System for Data Collection Station

- Tool for the administrative and operational organization of data collection
  - registration of temporary enumerators
  - to associate enumerator's workload
  - to load the program and Enumeration Area data into handheld computer
  - to download collected data from the handheld computer,
  - to transfer data to Central Processing
  - to summarize the quantitative of data collected.

**CENTRAL**

.SIGC – Data Collection Management Indicators System

□Tool for monitoring the data collection

□Agility to access the evolution of fieldwork: status of EA, quantities of housing units and population enumerated

□Information could be aggregated at different  management levels (EA, Data Collection Station, municipal, state and national)

□Comparative analysis at small area levels with previous Census

□Real time analysis enabled decision for related adjustments on operational aspects of collection process.

□Web application running in the main Datacenter that allowed to create management  reports, summaries, indicators and cartograms.

**IBGE**
Instituto Brasileiro de Geografia e Estatística

• **In process data collection**
  - People counted per day
  - Average of people per housing unit
  - Rate of men on the total counted
  - Occupied private housing units
  - Rate of occupied private housing units in the sample
  - Difference between average of people in housing unit and out of the sample
  - Rate of people who informed the month and year of birth
  - Rate of people under 3 and over 69 years old on the total of counted people

• **Concluded data collection (additional indicators)**
  - Occupied private housing units
  - Number of unities registered (residential or not)
  - Rate of occupied private housing unit without interview and vacant
  - Rate of dwellings of "occasional use"

•A supervision system was developed providing a tool for evaluating the enumerators' work quality

•The system periodically was fed by collection-produced information. The programmed supervision automatic selection of housing units was run on the handheld computer

  ▫The supervisor did up to 3 request per Enumeration Area, depending on the type of EA, the management indicators and the results of previous visits

•The data collected by the supervisor were compared with the data collected by the enumerator; in case of divergence the information was corrected and added on the handheld computer allocated to the enumerator

• **The experience of the integrated operation conducted in 2007 was fundamental for the success achieved in the 2010 Census**

• **The technological innovations implemented in the 2010 Census significantly affected the organization of the processes involved in this gigantic operation.**

• **Unquestionably this electronic tool turned the data collection process much faster allowing IBGE to execute a task larger than the 2000 Census, but with approximately the same number of staff.**

• **The handheld was fundamental for the construction of geospatial database when it permitted the association of addresses to the faces of the digital map.**

• **The technological and methodological innovations introduced in the 2010 Census have a significant influence and constitute a step forward on the institute's future activities**

  □ through the increase of analytical possibilities provided by georeferencing and geospatial data visualisation features, as well as the unlimited possibilities of the graphical interface.