

# Inclusion of Census 2011 Geographic Variables in NIDS (household level)

---

February 2014

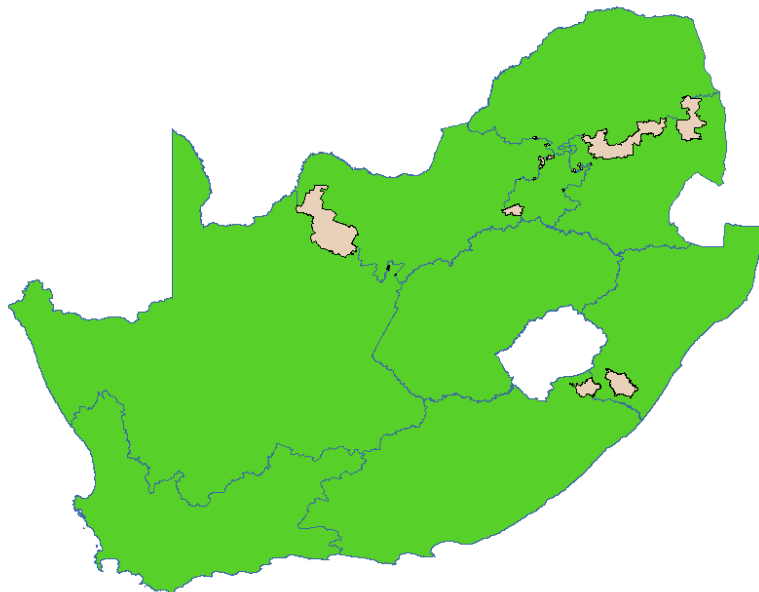
StatsSA made Census 2011 data publicly available late 2013. This creates the opportunity to update the geographical variables in the NIDS Datasets. The household level geographic variables presented in NIDS are Province, District Council and Geo-type. The Secure dataset also includes the Main Place and EA number. Prior to the public release of Census 2011, NIDS had calculated these variables based on the 2001 Census boundaries. This document outlines the differences and has important cautionary notes about the differences between 2001 and 2011 geography.

To assist users all **previously released geography variables** are still included in all waves, they **have just been renamed to include the suffix '2001'**. The new geography variables have the suffix '2011'. See the detail of the changes in the respective Change documents from version 1.1 to version 1.2 (Wave 3), version 2.1 to version 2.2 (Wave 2) and version 5.1 to version 5.2 (Wave 1).

## Provincial Boundary Changes

The Provincial Boundaries changed between the 2001 Census and the 2011 Census. Below is a map showing the provincial boundary changes.

Figure 1: Province changes between 2001 and 2011



Across the NIDS sample there are 375 households that have different provinces in the 2011 boundaries compared to the 2001 boundaries.

Below is a table outlining the changes:

	<i>w`i'`x'prov2011</i>						
<i>w`i'`x'prov2001</i>	Eastern Cape	Northern Cape	KwaZulu-Natal	Gauteng	Mpumalanga	Limpopo	Total
Eastern Cape	0	0	40	0	0	0	40
KwaZulu-Natal	2	0	0	0	0	0	2
North West	0	116	0	75	0	0	191
Mpumalanga	0	0	0	2	0	39	41
Limpopo	0	0	0	0	101	0	101
Total	2	116	40	77	101	39	375

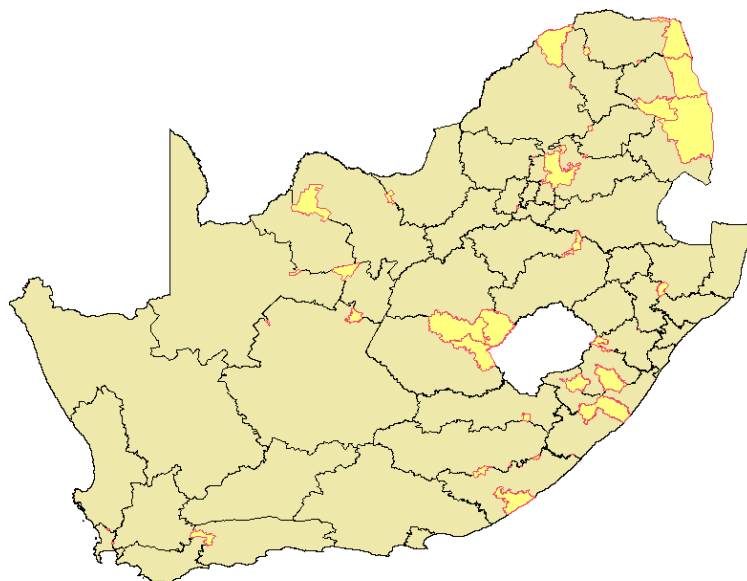
All the provincial codes have stayed the same.

## District Council Changes

There were relatively significant changes to the District Municipal Boundaries between 2001 and 2011. The number of metropolitan municipalities increased from 6 to 8, while the the number of district municipalities remained 52 (44 plus 8 metros). District Council codes have also been changed. For example, City of Cape Town used to be code 171, now it is code 199. In order to assist users we release the original District Council Variable renamed to *w`i'`x'dc2001* as well as the new District Council codes (*w`i'`x'dc2011*). We also include the Municipal Demarcation Board code (*w`i'`x'mdbdc2011*). The Municipal Demarcation Board variable is only available for 2011. Also note that this is a string/text variable, not a numeric.

**It is very important to note that the 2001 and 2011 District Council codes are not comparable at all.** Given the change in numbering and the change in boundaries comparisons cannot be made. Below is a map outlining where the District Councils changed shape.

Figure 2: District Council changes between 2001 and 2011



## Geographical Type Variables

In the 2011 Census data there are two geographical type variables EA\_GTYPE and EA\_TYPE. Previously there was only one variable that classified EAs as Rural Formal, Traditional Authority Area, Urban Informal or Urban Formal. The new geographical type classifications are Urban, Traditional and Farms. Below is a table that gives a description of each new category.

EA_GTYPE_C	EA_GTYPE	Description
1	Urban	A continuously built-up area that is established through township establishment such as cities, towns, 'townships', small towns, and hamlets. The areas are identified by "erf/erven/cadastre" from the Surveyor General or Municipal planning units.
2	Traditional	Communally-owned land under the jurisdiction of traditional leaders. Settlements within these areas are villages.
3	Farms	Land allocated for and used for commercial farming including the structures and infrastructure on it. The areas are identified by farm- and farm portion cadastre from the Surveyor General.

Table 1: EA\_GTYPE Source StatsSA

Given that the categorisation changed we present a change matrix showing the percentage of households in each category using the 2001 definitions and where they are now in the 2011 definitions, across the NIDS sample.

	<i>w'i' _x'geo2011</i>			
<i>w'i' _x'geo2001</i>	Traditional	Urban	Farms	Total
RURAL FORMAL	5.3	17.57	77.13	100
TRIBAL AUTHORITY AREA	95.01	2.48	2.51	100
URBAN FORMAL	1.32	98.11	0.57	100
URBAN INFORMAL	6.24	93.76	0	100
Total	33.32	56.46	10.22	100

EA\_TYPE classifies the EA by land use and human settlement in the area. Below is a table giving the descriptions of the land use and settlement types.

EA_TYPE_Code	EA_TYPE	Example
1	Formal residential	Single houses, town houses, high rise flats, scheme housing, estates
2	Informal residential	Illegal informal structures
3	Traditional residential	Villages in tribal areas
4	Farms	Farms
5	Parks and recreation	State forests, military training ground, holiday resorts, nature reserves, national parks
6	Collective living quarters	School hostels, tertiary education hostels, workers' hostels, military barracks, prisons, hospitals, hotels, old age homes, orphanages, monasteries

7	Industrial	Factories, large warehouses, mining areas, saw mill, railway stations and shunting areas, airports
8	Small holdings	Small holdings, agricultural holdings
9	Vacant	Open areas within urban and traditional areas
10	Commercial	Mixed CBD, office parks, shopping malls

Table 2: EA\_TYPE Source StatsSA

This variable is only available in the Secure datasets and is named `w`i'_`x'eatype2011`

**Very important:** Do not merge across census periods, i.e. 2001 to 2011 variables. Matches might be false and not represent the same space or values.

## Impact of geography variable changes on other variables

### Weights

Weights are calculated using the appropriate mid-year population estimates from [Stats SA](#). The mid-year population estimates have used the latest provincial boundaries since 2007. However, as described above, NIDS has been reporting provincial boundaries as they appeared in the sample originally provided by StatsSA, which reflected the 2001 boundaries. All the weights calculations in all waves have been updated to use the 2011 Census boundaries. The result is that almost all weights have changed slightly. Although individual cases might have shifted by seemingly significant proportions, the overall changes are insignificant. We also believe that this revision reflects the most accurate data available.

### Household quintile

Wave 1 includes a weighted household quintile variable. Due to the changes in weights there are 43 households that have changed quintile.

### Migration variables

There are a range of migration questions asked in all waves of NIDS. Where appropriate the data has been geocoded to the District Council level. These variables are still based on the 2001 boundaries. Care should be taken when comparing to household level variables to only use the 2001 equivalents.

### Imputed Income and Expenditure variables

All derived files for Waves 1, 2 and 3 use the 2001 geographic variables. No changes have been made to the derived variables. The system files are available from our [Program Library](#) for users who want to recreate the variables using the 2011 geographic variables.