

Introduction to Census Data & SASPAC

1 Introduction

The 2011 Census was conducted on 31 March 2011 across the whole of the United Kingdom. The census provides a snapshot of the population on Census Day and is used in, amongst other things, resource allocation, policy planning, benchmarking and shaping service delivery. In fact the potential uses of census data are limitless.

The SASPAC software application enables users to interrogate this complex dataset by providing a means to access and manipulate variables within census tables at various geographies.

In this introduction section of the SASPAC Training Manual the user is guided through the basics of census data (including table types, geography and variation between 1991, 2001 and 2011 data) as well as some fundamentals of the program itself. Subsequent sections of the manual are concerned with practical exercises and examples which will enable the user to get the most out of SASPAC.

2 The Census Dataset

2.1 The Census Process

In March 2011 a Census questionnaire was posted out to every residential address in England and Wales. In order to distribute the forms in this way the Office for National Statistics (ONS) compiled an address register using data from local authority Local Land and Property Gazetteers (LLPG), Post Office data and residential development information. Households were required to complete the form on Census Day (27th March 2011) and return the form to ONS in the enclosed envelope. For the first time the census could also be completed online using a unique code printed on each form.

In the weeks immediately following Census Day the ONS conducted an independent sample survey of the population called the Census Coverage Survey (CCS). The survey sampled one per cent of postcodes nationally and then asked households within those postcodes to answer a short questionnaire. The data collected in the CCS are compared to the Census results for the same areas in order to determine how many people and households were missed by the census and what their characteristics are. These missed individuals and households are then added to the census count in a process called imputation which leads to the production of the census estimates. More detail on the estimation process can be found in GLA Intelligence briefing 17-2012 which can be downloaded here:

<http://data.london.gov.uk/datastorefiles/documents/2011-census-estimation-process.pdf>

The estimates produced then went through a quality assurance process which saw them compared to administrative sources of population data such as the NHS patient register and ONS mid-year population estimates. Once the estimates had been validated by the quality assurance panel they were signed off for release.

2.2 2011 Census Questionnaire

The Census has changed significantly over the years as questions have been added and removed to reflect the data requirements at the time of enumeration. An example of this was the introduction of a question in the 1951 census designed to capture how many homes had inside toilets. By 1991 most houses had inside toilets and so the question was no longer providing a useful indication of the quality of housing. As a result in the 1991 census it was replaced by a question to capture the number of households without central heating.

In 2011 the new additions to the census questionnaire largely reflecting the diversity in modern society and the increasing mobility of populations. It is important to be aware of differences between censuses in order to know which data are comparable. The following are the changes in 2011 compared to 2001:

- Passports held
- Date of arrival in UK & intended length of stay
- Main language
- English proficiency
- National Identity
- Type of central heating (as opposed to a yes/no question)
- Number of bedrooms (as opposed to number of rooms)
- additional tick boxes in the Ethnicity question (Arab and Gypsy/Irish Traveller)

The questionnaire is two sections beginning with the 'Household questions'. This section asks questions relating to the property and the relationships of those who live there. Data on type of accommodation, tenure, central heating, bedrooms and the number of cars owned by the household are all captured in this section. The second section, 'Individual questions', asks questions about the people who reside in the households. Data is collected on a wide range and variety of characteristics including age and sex, marital status, qualifications, employment, language, ethnicity, religion, travel to work and caring, to name just a few.

The 2011 questionnaire can be downloaded from the SASPAC website here:

<http://saspac.org/support/2011-census/>.

2.3 Communal Establishments

A communal establishment is an institution which provides managed residential accommodation for individuals. Types of communal establishment include hotels, guest houses, student accommodation, prisons and nursing homes. These institutions are enumerated separately from other dwellings for census purposes. Rather than forms being posted out and posted back; instead a Special Enumerator visited each communal establishment and worked with the management there to ensure every individual was captured on a census form.

2.4 2011 Census Release Schedule

The first release of 2011 Census data came on 16 July 2012. This release constituted an estimate of the national populations on England and Wales by sex and single year of age as well as estimates for local authority areas by sex and five-year age banding. This release also included household estimate data for local authorities in England and Wales. The Northern Ireland Statistics Research Agency (NISRA) released population estimates for Northern Ireland on 16th July while National Records Scotland (NRS) will release the first data for Scotland in December 2012.

The release schedule for the remainder of the 2011 Census data is yet to be finalised but broad details have been published. There are to be three further releases of data. Each data release will contain phases with each phase providing the data at different geographies. The schedule is as follows:

Second release (November 2012 – March 2013)

Key Statistics These tables contain information on a single variable although data will often be banded (for example in age bands). These tables also often include percentages. These are equivalent to the 2001 Key Statistics tables.

Example: the Ethnic Group table (KS201EW) contains totals for 18 ethnic group categories and expresses each as a percentage of the total population.

Quick Statistics These tables contain all of the available data on a single variable (ungrouped and without percentages) and are equivalent to the 2001 Univariate tables.

Example: the Ethnic Group (detailed) table (QS201EW) contains totals for 151 ethnic groups but no percentages.

Headcounts Counts for Males, Females and Households at the unit postcode level.

The Key Statistics and Quick Statistics in the second release will be available for geographies from Output Area up to the national level (including ward level). In addition the Key Statistics data will be provided for Postcode Sectors.

Third release (March 2013 – July 2013)

Local Characteristics These tables are multivariate (i.e. they are cross tabulations between two or more variables). As the name suggests these table are concerned with providing data at the most detailed geography possible (Output Area). In some cases other variables will be grouped in order to meet the demands of the Census Offices' (ONS, NRS, NISRA)

disclosure control policy. These are equivalent to the 2001 Census Area Statistics tables.

Fourth release 4 (July 2013 onwards)

Detailed Characteristics These tables are also multivariate but in this case the variables themselves are the focus of the data. As a result these data will not be available at the very smallest geographies. The minimum geography will be wards/MSOA. These are equivalent to the 2001 Standard tables.

Detailed Themes These are profiles of various main characteristics for a particular population group. These tables will also only available at ward/MSOA and higher. These are equivalent to the 2001 Standard Theme and Census Area Statistics Theme tables.

Armed Forces These tables provide data on the armed forces population and are equivalent to the 2001 Armed Forces tables.

2.5 Further releases

Additional data such as travel to work and other flow data and information on population groups such as students, short-term migrants and workers will be released after the fourth release. There are currently no detailed plans for the scheduling of these releases.

Office for National Statistics release schedule: <http://www.ons.gov.uk/ons/guide-method/census/2011/census-data/2011-census-prospectus/release-plans-for-2011-census-statistics/index.html>

National Records of Scotland: <http://www.gro-scotland.gov.uk/census/censushm2011/producing-delivering-data.html>

Northern Ireland Statistics and Research Agency:
<http://www.nisra.gov.uk/Census/2011CensusProposedOutputs.html>

See Annex 4 for an overview of the release schedule for 2011 Census data.

2.6 Samples of Anonymised Records (SARs)

The Samples of Anonymised Records consist of extracts from Census records which are designed to enable researchers to carry out detailed analyses using Census data for individuals or households.

The SARs are a family of datasets drawn from the 1991 and 2001 UK Census. The SARs contain a separate record for each individual, but identifying information has been removed to protect confidentiality. The SARs datasets are similar to data from a survey, albeit with a much larger sample size thus permitting analysis of small sub-groups and small geographic levels. The SARs cover the full range of Census topics including, housing, education, health, transport, employment, ethnicity and religion.

1991 individual SAR (I-SAR)	2 per cent sample (1.1 million records). Small local authorities are aggregated.
2001 Individual licence (IL-SAR)	3 per cent sample (1.75 million records). Down to regional level.
1991 Household SAR (H-SAR)	1 per cent sample (216,000 households and 500,000 people within household)
2001 Special Licence Household SAR (SLH-SAR)	1 per cent sample (200,000 household and 500,000 people within households)
2001 Small Area Microdata (SAM)	5 per cent sample (3 million people). Available at Local Authority with some areas merged.
2001 Controlled Access Micro Data Samples (CAMS)	More detailed versions of 2001 LI-SAR and SL-HSAR. These data are available at Local Authority level.

Unlike Census data the SARs datasets require a licence for their use. There are three types of licenses corresponding to different levels of security:

End User Licence agreement allows access to the 1991 I-SAR and H-SAR, 2001 IL-SAR and 2001 SAM. The 2001 Special Licence Household SARs provides access to 2001 SLH-SAR. The Controlled Access Microdata Samples (CAMS) require a higher level of data stewardship still.

More detail on the SARDS including information on the application process for the datasets is available on the ONS website: <http://www.ons.gov.uk/ons/guide-method/census/census-2001/data-and-products/data-and-product-catalogue/microdata/samples-of-anonymised-records/samples-of-anonymised-records.html>

2011 SARs data is currently being compiled based on user consultation and requirements. The target date for the release of 2011 SARs is autumn 2013.

2.7 Commissioned Tables

Should your organisation require data that is not available in the published tables the ONS operates a commissioned tables service. Users can request specific tabulations of variables, and

subject to disclosure control, the ONS will produce the data. This is a chargeable service. Once a table has been commissioned it becomes publically available to all users free of charge.

<http://www.ons.gov.uk/ons/about-ons/who-we-are/services/unpublished-data/census-data/census-commissioned-tables/index.html>

Note: Local Authorities in London should note that the commissioned tables service is administered and paid for by the Greater London Authority, on their behalf, as part of the Census Information Scheme and any queries should be directed in the first instance to the CIS rather than the ONS.

2.8 Disclosure Control

As noted above some detailed data cannot be released due to disclosure control. This is to ensure that individuals cannot be identified through census statistics.

In order to protect individuals from identification the Census offices used a form of statistical disclosure control for the 2011 Census data called record swapping. Every individual in a household was assessed for uniqueness or rarity on a small number of variables and then every household was given a risk score. A sample of households was then selected for swapping based on the risk score. The household was swapped with one in another area either within the middle layer Super Output Area or a neighbouring local authority. (the household and its swap are matched on some basic characteristics to preserve data quality, e.g. household size so that overall population totals are not impacted).

By using this method before the tables are produced the Census offices ensure that all tables are additive and different tables showing the same variable will have the same figure.

In addition, where there is potential for individuals to be identified in very detailed table outputs variables are grouped or the geographic level at which the data is available is restricted.

The ONS methodology paper on Statistical Disclosure Control can be downloaded from <http://www.ons.gov.uk/ons/guide-method/census/2011/the-2011-census/processing-the-information/statistical-methodology/statistical-disclosure-control-for-2011-census.pdf>

2.9 Data Licensing

Under the terms of the Open Government Licence (OGL) and UK Government Licensing Framework (launched 30 September 2010), anyone wishing to use or re-use ONS material, whether commercially or privately, may do so freely without a specific application for a licence, subject to the conditions of the OGL and the Framework.

These new arrangements replace the previous Click-Use and Value Added Licences. When reproducing Census data without adaptation the follow statement should be included:

“Source: Office for National Statistics licensed under the Open Government Licence v.1.0”

If reproducing Census adapted content the following statement should be included:

“Adapted from data from the Office for National Statistics licensed under the Open Government Licence v.1.0”

The licence arrangements for other 2011 Census products such as special migration/workplace flows and workplace zone statistics are still being defined. For further information see <http://www.ons.gov.uk/ons/guide-method/census/census-2001/data-and-products/copyright-and-licensing/index.html>.

2.10 The 2001 Census

The 2001 Census was conducted on 29 April 2001 and the data was released between September 2002 and spring 2005. There are some differences between the 2001 and 2011 datasets stemming from changes to the questions asked on the respective census questionnaires. Some examples are ethnicity where two additional tick boxes were added in 2011 or central heating where respondents in 2011 were required to choose a type of central heating rather than just answering whether or not they had central heating.

Also different in 2001 were the names of some of the tables:

Headcounts	Number of people (males/females) and households for Unit Postcodes.
Profiles	A number of standard ‘templates’, designed as indicators to be presented as percentages in a limited number of simple tables. Equivalent to 2011 Key Statistics.
Key Statistics	Cover all main Census topics, designed as indicators to be presented as percentages in a limited number of simple tables. Equivalent to 2011 Key Statistics.
Census Area Statistics	Generally presented as cross-tabulations, but also including the simple ‘univariate’ tables, covering all main Census topics and relationships between them. Equivalent to 2011 Quick Statistics and Local Characteristics.
Standard Tables	Generally presented in cross-tabulations giving more depth than the CS, including additional ‘topic’ extension tables for larger populations. Equivalent to 2011 Detailed Characteristics.
Theme Tables	There are two types of Theme Tables: Census Area Theme Tables (CAS Themes) and Standard Theme Tables (ST Themes). These tables combine data on a specific theme into one table. Equivalent to 2011 Detailed Themes.

Origin/Destination Figures on the people with workplaces in an area and on out-migrants are included in the Standard Tables and CAS, but the flows of workers and migrants between areas are provided in additional matrices. This data will form part of the 'further releases' for 2011 Census.

SASPAC enables access to all these datasets, although the Origin/Destination is a specialist dataset not covered in the basic training modules (see module 10), while the Headcounts and Profiles are probably better handled through software such as Excel.

2.11 The 1991 Census

Small Area Statistics Consist of 83 tables for areas in England, 84 tables for areas in Wales, and 86 tables for areas in Scotland. The SAS dataset is available for Enumeration Districts (ED) in England and Wales, and for Output Areas (OA) in Scotland. It is also available for any higher areas such as wards, local authorities etc.

Local Base Statistics Consist of 95 tables for areas in England, 96 tables for areas in Wales, and 99 tables for areas in Scotland. The LBS dataset is available for wards in England and Wales, and for postcode sectors in Scotland. It is also available for any higher areas such as local authorities, counties etc.

Special Migration Stats Consist of 11 tables for wards in England and Wales, and for postcode sectors in Scotland. They relate areas of residence at census date to areas of residence one year prior to the census.

Special Workplace Stats Consist of 26 tables for wards in England and Wales, and for postcode sectors in Scotland.
Set A: Statistics for areas of residence
Set B: Statistics for area of workplace
Set C: Statistics linking areas of residence to areas of workplace

2.12 100% and 10% datasets in 1991

Some of the counts or variables derived from the census questionnaire in 1991 are considered to be hard to code, and are therefore only presented for 10% of the population. Such a question was number 15, which related to occupation, in comparison with the questions on sex. During the 100% processing of returns, a sample of 10% of households and persons within them, plus a sample of 1 in 10 persons in communal establishments, was drawn from the validated records. The data in this sample relating to relationship within families, occupation, industry, workplace, etc was then processed. Tables 71 to 99 (both SAS and LBS) relate to 10% data only.

All datasets in 2011/2001 Censuses relate to 100% of the population estimates.

3 Census Geography

It is useful for any user of census data to be familiar with the levels of geography at which data is available and the various naming conventions. Due to the ten-year interlude between censuses there are often changes to geographic boundaries, to the naming convention and in some cases new geographies are created. The geographies associated with 2011, 2001 and 1991 data are outlined below.

3.1 Boundary Files

Boundary Files are a digitised representation of the underlying geography of the census. They allow census data to be mapped for analysis and visualisation. The boundaries are available at two coastal extents, these are:

- Clipped to coastline
- Extent of the realm

Digital spatial files are median population weighted centroids which will be created for output areas, super output areas and workplace zones. The centroid is a summary single reference point which represents how the population at census time was spatially distributed and grouped within that OA, lower layer super output area (LSOA) or middle layer super output area (MSOA). The provision of centroids allows users to get consistent and comparable best-fit allocations to a higher geography using a GIS.

Census 2011 geographies will be available in two formats, they are:

- Shape file (a format produced by ESRI)
- KML (Keyhole Markup Language)

Lookup files are files which allow users to identify the comparability of geographies as boundaries and names change over time. There will be a lookup from 2001 output areas/super output areas to the new 2011 output areas/super output areas to allow users to compare between data output from 2001 and 2011 geographies. All 2011 Census estimates for all geographies, including wards, will be best-fitted from output areas in line with the National Statistics Geography Policy. Therefore a number of lookup files from 2011 output areas to other output geographies will be produced. There will also be lookups from workplace zones and enumeration postcodes (those identified during the census) to other census geographies. Lookups will be supplied in comma separated value (csv) and delimited text file formats.

The majority of ONS Great Britain coverage digital boundaries are now freely available under the Ordnance Survey (OS) OpenData and the Open Government Licensing agreement. The boundary files can be obtained by contacting the ONS through their website:

<http://www.ons.gov.uk/ons/guide-method/geography/products/digital-boundaries/index.html>

More information can be found in the ONS 2011 Geography Prospectus:

<http://www.ons.gov.uk/ons/guide-method/census/2011/census-data/2011-census-prospectus/new-developments-for-2011-census-results/2011-census-geography/2011-census-geography-prospectus/index.html>

3.2 2011 Geographies

Output Areas The smallest geography for which data are available is the Output Area (OA). They consist of, on average, 125 households (minimum 40 households or 100 people, maximum 250 households or 625 people). They are broadly socially homogeneous and are constrained to the physical geography of the landscape.

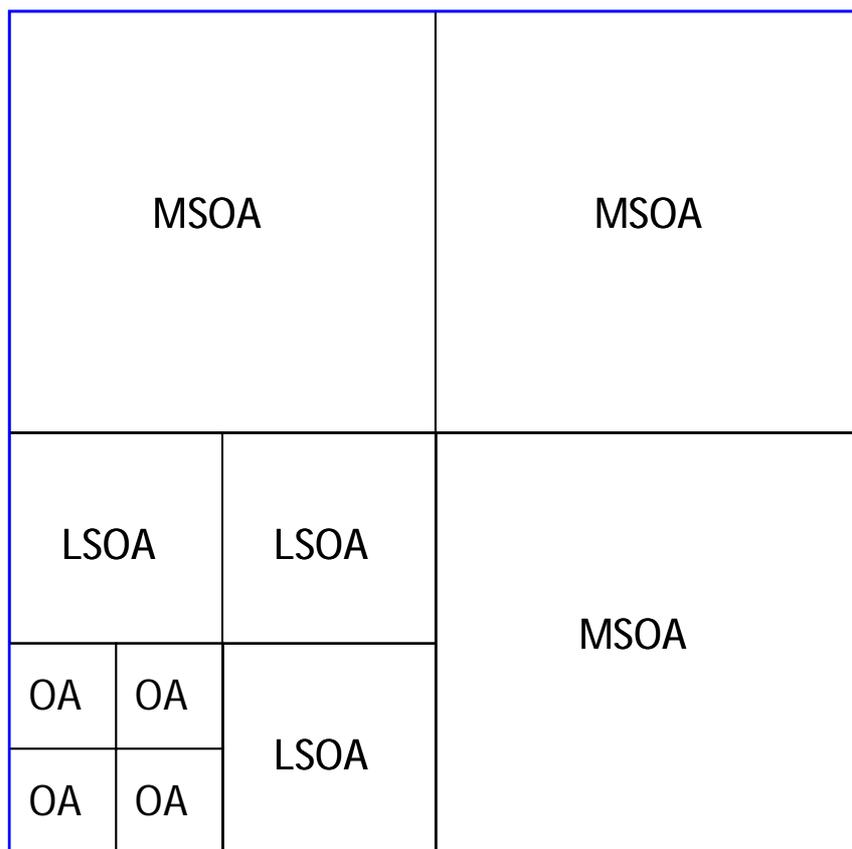
Comparison between 2011 OAs and 2001 OAs is broadly possible, however there will be some changes. Where populations have dramatically increased or decreased since to 2001, to the extent that they now no longer fall within the stated OA thresholds, OAs will either be split or merged to form new boundaries. This is expected to be the case is less than five per cent of OAs in England & Wales.

LSOA Lower Super Output Areas are groups of four or five OAs containing an average of 1,500 people. The minimum size of an LSOA is 1,000 people or 400 households while the maximum size is 3,000 people or 1,200 households. The LSOAs were constrained to 2003 ward boundaries so that, at that time, groups of LSOAs nested within wards. Owing to ward boundary changes, and the potential for changes to OAs themselves, LSOAs may no longer nest within wards.

MSOA Middle-level Super Output Areas are groups of four or five LSOAs containing an average of 7,200 people. The minimum size of an MSOA is 5,000 people or 2,000 households and the maximum size is 15,000 people or 6,000 households. These are constrained to 2003 local/unitary authority boundaries.

Ward A ward is a local authority geography that many users will already be familiar with. In 2011 ward level data will be produced using 'best fit ' OAs (see 3.1.3). This will ensure that ward data is consistent with OA and LSOA data and avoid issues around disclosure control. These are referred to as 'administrative wards' in order to differentiate them from other census geographies (see 3.1.1 & 3.1.2 below).

LA & UA Local Authority and Unitary Authority data will be provided as exact fit data. This means that output areas will nest within LA and UA boundaries allowing data for authorities to be aggregated from their constituent OAs.



blue outline = local authority boundary

Figure 1: Diagrammatic representation of the nesting of census geographies

Higher Geographies	Data for Regions and Counties will be produced on a best fit basis.
Workplace Zones	This is a new geography being produced for 2011 Census geography designed to be more suitable than OAs for disseminating workplace statistics. (OAs being based on residential populations). Workplace zones will be constrained to MSOAs and produced for England & Wales.
Postcode Sector	Key Statistics will be produced at Postcode Sector (everything but the last two characters of the postcode). From sectors it is possible to aggregate up to Postcode Districts and then Postcode Areas.
Other	Data will also be provided at the following geographies (subject to disclosure control): <ul style="list-style-type: none"> • Westminster parliamentary constituencies • National Assembly for Wales constituencies • Parishes • Former counties • Primary care organisations • Local health care boards

- Strategic health authorities
- National parks
- Local administrative units (1 & 2)
- NUTS (1 2 & 3)
- Settlements
- Urban areas

All of these geographies will be produced on a best fit basis from OAs.

See Annex 4 for an overview of the release schedule and geographies of 2011 Census data.

3.2.1 Best Fit

For many in local government the electoral ward is an important geography as it is understood by the electorate and favoured by Councillors. When Output Areas were initially produced, for the 2001 Census outputs, they were constrained to ward boundaries (laid down in statute as at 31/12/02). However changes to ward boundaries since then, and the potential impact of merging OAs that cross ward boundaries, means that this will not be the case for 2011 outputs. 2011 Census ward-level outputs will be produced using OAs in a 'best fit' approach.

The best fit method is a relatively simple way of determining which ward an OA belongs to in cases where the OA is bisected by a ward boundary. First a population centroid is calculated for each OA. This is a geographical point which represents the 'average location' of the population of the OA. The whole OA's population is then assigned to whichever ward its centroid sits within.

As a result, where ward boundaries have changed since 31/12/02, the census data will not be created from OA boundaries that exactly match a ward but the data will still summate to the local/unitary authority total. If no ward boundary changes have occurred since this date: exact fit data will be available.

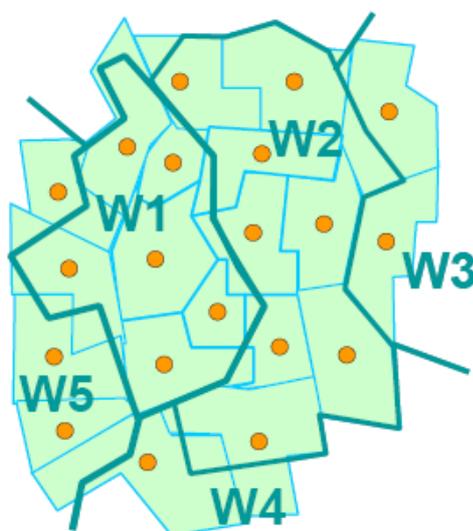


Figure 2: Example of best fit for non-standard geographies using population centroids

Figure 2 shows how the best fitting process works in practice. The blue lines represent output area boundaries while the orange dots represent the population centroids for those output areas. The thick green lines are ward boundaries. The entire population the output area is assigned to the ward within which its centroid lies.

More information on the best fit process can be found in the ONS paper 'Exploring the performance of best fitting to produce ONS data for non-standard geographical areas' (<http://www.ons.gov.uk/ons/guide-method/census/2011/the-2011-census/producing-and-delivering-data/output-geography/best-fit-policy/exploring-the-performance-of-best-fitting-to-produce-ons-data-for-non-standard-geographical-areas.pdf>).

The Census Geography FAQ is also useful resource (<http://www.ons.gov.uk/ons/guide-method/census/2011/census-data/2011-census-prospectus/new-developments-for-2011-census-results/2011-census-geography/2011-census-geography-prospectus/2011-census-geography-faqs.pdf>).

3.2.2 Local Characteristics wards

Some of the more detailed data available from the Local Characteristics tables will not be available at the smallest geographies, in order to avoid the publication of disclosive data. This means that wards whose population fall below the Output Area threshold (40 households/100 people) will be amalgamated to produce populations large enough to prevent disclosure. This was achieved by adding the populations of the relevant wards, determining the population centre, and then assigning the entire population to the administrative ward within which the population centre lies. As a result some administrative wards will not appear in the Local Characteristics ward geography hierarchy, while others will have inconsistent populations across the two ward types (this tends to occur in areas of particularly low residential populations such as the City of London – in 2001 only 25 wards in England and Wales fell below this equivalent threshold). Wards will retain the name and unique GSS code across the two geographies.

3.2.3 Detailed Characteristic wards

As with Local Characteristics wards the data available in the Detailed Characteristic and Themes tables may require wards to be amalgamated to meet population thresholds. Again, this was achieved by adding the populations of the relevant wards, determining the population centre, and then assigning the entire population to the administrative ward within which the population centre lies. However, the Detail Characteristic wards will constitute a geographical hierarchy in their own right with a new set of unique GSS (Government Statistical Service) codes assigned. This means that the not all administrative wards will appear in this hierarchy and there will be inconsistent populations when comparing the two ward types (113 wards fell below this threshold in the 2001 Census).

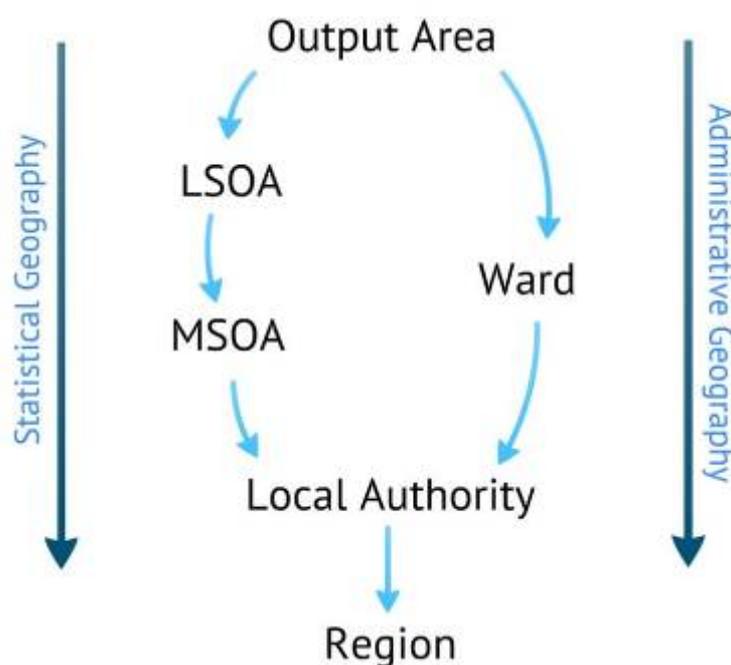
The Detailed Characteristics wards have a minimum population threshold of 5,000 persons or 2,000 households.

Data for higher geographies such as local authority, county, region and country are also available and are produced by aggregating-up data from constituent geographies.

3.2.4 Main Census Geographies

Geography	Description	Min – Max population
Output Area	Smallest area for which data are available	100 – 625
LSOA	4 or 5 output areas	1,000 – 3,000
MSOA	4 or 5 LSOAs	2,000 – 5,000
Wards (LC & DC)	Comprised of those OAs which best fit the ward	2,000 – 5,000
LA/UA	Total of 348 district and unitary authorities in E&W	n/a
Region	9 regions in E&W	n/a

Figure 3: Hierarchy of census geography



3.2.5 2011 Naming Convention

The geographical naming convention for 2011 Census data uses the relatively new system of GSS (Government Statistical Service) codes (implemented on 1st January 2011). These are nine-digit codes where the first three characters indicate the 'entity' and the last six the 'instance'.

Entity	England:	E00 – E32, E37, E92
	Wales:	W00 – W35, W92
	Scotland:	S00 – S25, S92
	N Ireland:	N06 – N07, N92

Others: K01 – K04, L00, L93, M83, M00, M01

Instance 000001 to 999999

See Annex 3 for a full list of GSS Entity Codes.

The Super Output Area geographical hierarchy already uses this format and some important features of this convention are:

unlike 2001 Census codes, these codes are not hierarchic so no relationship to any parent area is incorporated (see 3.2.1)

codes are unique and will not be reused

a new instance of an area will be created with a new code only when a boundary change occurs (i.e. a change in an instance name will not generate a new code)

where available, instances will also include names/labels

For more information see the nomenclature pages of the ONS website

(<http://www.ons.gov.uk/ons/guide-method/geography/products/names--codes-and-look-ups/index.html>)

3.3 2001 Geographies

The geographies used for 2001 Census outputs are broadly consistent in their type and extent to those used for 2011. Data was available at OA, LSOA and MSOA and while there will be some changes to OAs in 2011 which will impact upon LSOAs and MSOAs this will be the case in less than five per cent of OAs.

Ward There are two types of ward data available for 2001 outputs depending on which set of tables are being used. CAS Wards are used for data in the Census Area Statistics tables while ST Wards are used for data in the Standard tables. These wards were created by merging administrative wards in order to ensure that Census table populations were large enough that data would not be disclosive. Where ward populations were large enough merging was not necessary. The minimum threshold for ST Wards was 1,000 persons or 400 households while for CAS Wards it was 100 persons or 40 households.

Other As with 2011 data the 2001 Census outputs were produced for a wide range of geographies such as local/unitary authority, parish, parliamentary constituency, health geographies, counties, regions and more.

3.3.1 2001 Geographic Naming Convention

The naming of areas reflects the hierarchical structure of the geography. For example a four digit code for an Output Area is preceded by first the County code, then the Local Authority code and finally the Ward code.

Region	1 alphabetic character	
County or UA	2 numeric characters	00 for UA, 09-46 for County
Local Authority	2 alphabetic characters	
Ward	2 alphabetic characters	
Output Area	4 numeric characters	0001 - 9999

Example:

Region	J	South West
County	24	Hampshire
District	24UN	Test Valley
Ward	24UNGF	Blackwater
Output Area	24UNGF0010	Output area within Blackwater ward in Test Valley, Hampshire

3.4 1991 Geographies

The geographies for 1991 outputs do differ from 2001 and 2011. This is most notable for the absence of the small level Output Area geographies and those geographies that are aggregated from OAs. In 1991 Enumeration Districts were the smallest geography at which data were produced.

Enumeration District	This is an area used initially in the collection of data and the boundaries are designed for that purpose. The minimum threshold for an ED was 50 people and 25 households. In 1991, and censuses prior to that, data were also output at the ED level as well.
Ward	EDs nested within wards
District	Wards nested within districts
County	Districts nest within counties

3.4.1 Special Areas

In 1991, the Census Offices defined Special Enumeration Districts as communal establishments such as large hotels, hospitals, or defence establishments where 100 or more persons were expected to be present on census night. The characteristics of the populations within SEDs are often markedly different from those in the surrounding area. Statistics for SEDs are not included in the statistics for the OAs or EDs which include them but available separately.

In 2001 and 2011 Special Areas do not exist, as all communal establishments are considered as part of the standard Output Area in which they are located.

3.4.2 Shipping Areas

In 1991, for Census purposes, each local government district included a shipping ward (or postcode sector in Scotland). This shipping area was created for the purposes of enumerating persons on board ships (excluding houseboats) in transit. For each shipping ward or postcode sector, there was at least one shipping ED or OA.

In 2001 and 2011 Shipping Areas do not exist, as all such craft are considered as part of the standard Output Area in which they are located.

3.4.3 1991 Geographic Naming Convention

The naming structure works in the same hierarchical way as in 2001 with smaller geographies having their specific instance code preceded by the codes of the larger units within which they nest.

County	Two digit code	01-55
District	Four digit code	AA - TT
Ward	Six digit code	FA, etc.
ED	Eight digit code	01, etc.

Example:

County	03	Greater Manchester
District	03BN	Manchester
Ward	03BNFA	Ardwick
Output Area	03BNFA03	Output area in Ardwick ward in Manchester, Gtr Manchester

3.4.4 1991 Geography in Scotland

In Scotland the smallest unit of geography in 1991 was not the Enumeration District but the Output Area. The naming structure is as follows:

Region	2 numeric characters
District	2 alphabetic characters
Postcode sector	2 alphabetic characters
OA	3 characters (2 numeric followed by 1 alphabetic)

Example:

Region	57	Central
District	5705	Clackmannan
PC Sector	5705AC	Postcode sector "FK1 01"
Output Area	5705AC10A	Output area in sector FK1 01 in Clackmannan, Central Region

4 Table and Variable Naming

It is also useful for users of census data to understand the naming conventions of census tables. This will help users navigate within SASPAC and understand how census data is structured.

4.1 2011 Table Naming

Details of the naming conventions for 2011 tables are not yet finalised. Outlined here are the recommended table codes for 2011. The first two characters of the name refer to the table type:

KS	Key Statistics
QS	Quick Statistics
LC	Local Statistics
DC	Detailed Characteristics
T	Detailed Themes
AF	Armed Forces

The third fourth and fifth characters refer to the specific table:

001 indicates table 1 (etc., up to 999)

The final two characters refer to the geography of the table:

EW	England & Wales
WA	Wales only (Welsh language)

Examples:

KS201EW	Key Statistics Ethnic Group Table
QS204EW	Quick Statistics Main Language (detailed) Table
QS206WA	Quick Statistics Welsh language Skills Table

There are no details yet about the naming conventions for Scottish or Northern Irish tables.

4.2 2011 Variable Naming

There are no details yet regarding a unique variable identifier for 2011 census results but these are expected to be included with the data.

4.3 2001 Table Naming

The first two characters of the name refer to the table itself:

KS	Key Statistics
UV	Univariate Tables
CT	Census Area Statistics Theme Tables
CS	Census Area Statistics
ST	Standard Tables
TT	Standard Tables Theme Tables
AF	Armed Forces Tables

The fourth, fifth and sixth characters refer to the specific table within the above grouping. In some cases a table may have a national variation whereby a table can have different contents depending on which part of they refer to. These tables are identified by a country character following the table number, for example, ethnic Group table 'KS006' has variations for Wales (KSW06) and Scotland (KSS06). However, generally speaking where a similar table is produced for areas in different nations, different table numbers apply.

001 to 100	tables are applicable to the whole UK.
101 to 200	apply to areas in England and Wales
201 to 300	apply to areas in Scotland
301 upwards	apply to Northern Ireland

Examples:

KS006	Key Statistics Ethnic Group Table
UV046	Univariate Table Household Composition – People

4.4 2001 Variable Naming

A specific variable within a table is identified first by the table name as describe above and then by a four digit number to identify the variable. Therefore a variable identifier has nine characters; the first five refer to the table the final four to the variable.

Example:

KS0060010 – Number of ethnically Pakistani residents (taken from the Ethnic Group Table KS006)

4.5 1991 Table Naming

The first letter of the code indicates which table is being referred to:

L	Local Base Statistics (LBS)
S	Small Area Statistics (SAS)

The second and third characters identify which particular table is being referred to:

01 indicates table 1 (etc., up to 99)

In some instances, tables exist only for areas in Wales or Scotland, or for Great Britain as a whole. In other cases tables for Scotland contain minor differences from those for the rest of Great Britain. In these cases a fourth character is introduced to the Table identifier. This extra character – S, W or G – is placed after the dataset identifier, but before the table number.

Examples:

L06 Ethnic Group of Residents
S12 Long-term illness in households

4.6 1991 Variable Naming

A specific variable within a table is identified first by the table name as describe above and then by a four digit number to identify the variable. Therefore a variable identifier has seven characters, the first three refer to the table the final four to the variable.

Example:

L060066 Number of ethnically Indian residents aged 15 (taken from the Ethnic Group Table 06)

5 Introduction to SASPAC

SASPAC is a software application designed for the storage and interrogation of large datasets. It was initially created for the 1991 Census and since that original release has been enhanced and developed but the addition of many new features. A brand new SASPAC application has been developed for the 2011 Census which incorporates a modern interface and intuitive ease of use. The original version of SASPAC will, as a result, now be known as Legacy SASPAC. Legacy SASPAC will still be useful in interrogating older datasets, and it will be possible to access 2011 data through the older interface. As this manual is concerned with Legacy SASPAC only the name SASPAC will be used to refer to this earlier version of the software.

5.1 Configuration

The operation of SASPAC is controlled by a text file – SASPAC.INI. In general users will not need to amend this file. The file defines the working parameters of the program and specifies the default directories necessary for the operation of SASPAC. These parameters can be amended by use of the Tools > SASPAC Configuration menu.

5.2 System Files

Initially SASPAC converts and compressed raw data into an internal format which is more efficient in terms of data storage, and enables easy access to the statistics. These are called System Files (SYS). Users are provided with the System Files for census data so there is no need to undertake this conversion. If users intend to import other datasets into SASPAC then System Files will need to be created.

A System File is effectively a very large matrix where geographic areas form the rows and the available variable form the columns. The columns are identified by a unique Cell Reference Number determined by the table it comes from and its position within that table. The rows are identified by a unique code assigned by SASPAC called ZONEID. When SASPAC is asked to retrieve a value it accesses the System File, finds the relevant row using the ZONEID and moves along that row until it locates the requested variable using the Cell Reference Number.

Since these System Files are held in a format which is internal to SASPAC, (that is they can only be read by SASPAC), their contents cannot be examined by using a text editor or viewer. There will be occasions when a user is uncertain as to the contents of a system file, and may need some means of examining its contents prior to running a SASPAC task.

Within SASPAC for Windows, there is a facility for doing this, which is accessed through the Tools > System File Details menu.

5.3 Command Files

The Command File (CMD) is at the core of all SASPAC operations. A Command File contains a sequence of instructions which tell the software what to do. Command Files are created using the SASPAC interface and so the user does not need to understand the file's syntax.

In its simplest form the Command File has four elements:

Input - Location of the System File to be used

Selection - which cells within the System File are required and for which geography?

Manipulation – Are any of the variables to be grouped or joined?

Output – Where is the resulting output file to be saved?

The SASPAC interface is used to create System Files. The image below shows how each element of the Command File is selected in the 'Print Variables' wizard.

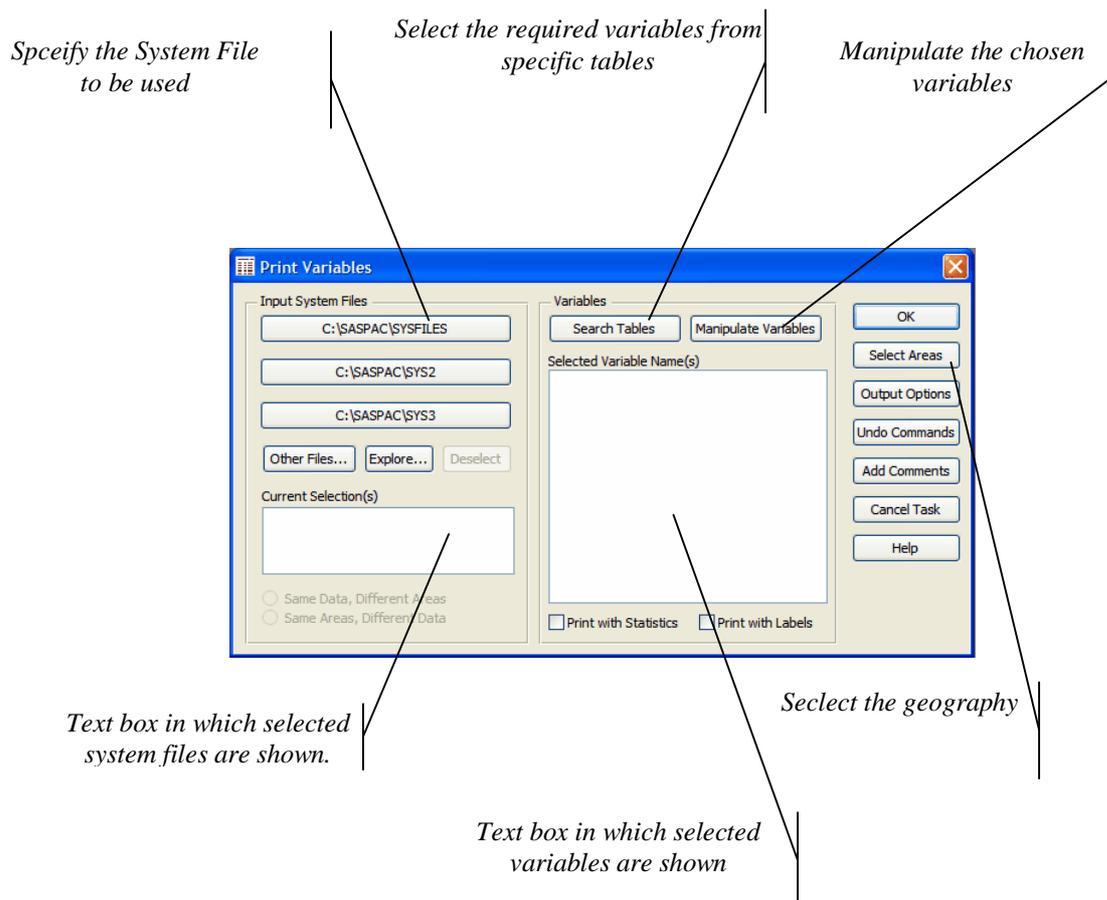


Figure 4: Main Task Window for Print Variables

5.4 Other File Types

In order to function efficiently, SASPAC needs to have access to several different types of file. Each of these file types is identified by a default extension to the name. Users are strongly advised not to change these extensions, since SASPAC will not recognise them if they have non-standard extensions.

File type	Extension	Purpose, content, or use of file
Log files	.LOG	Each time SASPAC acts on a command file, a .LOG file is produced. This file contains information concerning the implementation of the command files, and in the case of problems, will report warning and error messages to alert the user.
Report files	.PRN	If the objective of a command run is to produce printed output, that output will be written to a .PRN file which may be immediately printed or stored for later output.
Interface files	(various)	If the objective of a command run is to produce output for further analysis by other software, or for integration in a published report, such output will be written to an interface file. The particular type of interface file used will depend on the software to which the data is to be exported.
Framework Data files	.FWD	These are ASCII representations of the table layouts necessary to produce the Census Offices table images.
Framework files	.FWK	Internal format versions of the .FWD files. They supply the text necessary to produce tables, and define the location of data within those tables. The relationship between .FWD and .FWK files is similar to that between .TXT and .SYS files.
Initialisation file	.INI	Windows initialisation file. This file may be modified and customised by users.
Gazetteer files	.GAZ	These files contain definitions of new areas or zones in terms of existing areas or zones, and are used in the creation of datasets for new areas as defined by users.
Data Definition Language file	DDL	These files define the content and format of the .TXT files which are to be converted to .SYS files, and are only used at the time of creation of the .SYS files.
OPCS Text files	.TXT	Raw data files as supplied by the Census Offices, or from other sources.

Transfer files	.TRF	Used for transfer of complete (or partial) datasets between SASPAC on two different hardware platforms.
SASPAC4 files	.S81	These files are the 1981 SASPAC equivalents of the .SYS files. They may be converted to .SYS files for use with SASPAC.

5.5 Framework Files

Framework files are needed when a user needs to produce data outputs in table form. The framework file contains the formatting and layout information for the table. Users can either output tables using the layouts provided by the ONS, or they can design their own tables.

These Framework Files are held by SASPAC in a format that is internal to SASPAC, i.e. they cannot be interrogated or edited by other software. In order that users may edit these files, or create their own, ASCII files called Framework Data Files are used. These may be edited, and then loaded into SASPAC for saving as Framework Files. The relationship between Framework Data Files and Framework Files is the same as that between Raw Data Files and SASPAC System files.

The main Framework Files distributed with SASPAC for accessing the 1991 Census datasets were:

TLBS132	To allow printing LBS tables at a maximum of 132 characters per line.
TLBS160	To allow printing LBS tables at a maximum of 160 characters per line.
TSAS132	To allow printing SAS tables at a maximum of 132 characters per line.
TSAS160	To allow printing SAS tables at a maximum of 160 characters per line.
TSWS	To allow printing SWS tables (Sets A & B) at a maximum of 160 characters per line.
SWSC91	To allow printing SWS tables (Set C) at a maximum of 160 characters per line.
SMS91	To allow printing SMS tables at a maximum of 160 characters per line.

Seven Framework Files are currently available for the 2001 Census datasets – one each for the ST, TT, CS, CT, and UV, and two for the KS. The reason that there are two Framework Files for the Key Statistics is that as originally designed by the Census Offices, the KS did not fit into the 'standard' for table layouts. Thus there is a framework which adapts the KS as a 'standard' layout with a single area being output to a page, and there is a new variation which allows for areas to appear as the row variable in a table. These areas are user selectable, and this framework allows the user to replicate the layouts for the KS as they appear in published volumes.

5.6 Multiple System Files

On most occasions, the areas or data required by a user will be found on a single System File. However, there will be instances when the input of more than one system file is demanded. This might occur, for example, when enumeration districts in more than one county are being compared, or when data is required from both the 100% and 10% datasets in 1991.

SASPAC allows up to nine separate System Files to be identified for inclusion in a single task. If more than one file is to be accessed, SASPAC needs to know how the files are to read in relation to each other.

In combining files within the same task, the following options are allowed:

- same areas, different data; or
- same data, different areas.

Note that the 'Different areas, different data' combination is not permitted.

In defining 'same' in this context, the rule is that the sub-set of areas or data selected for manipulation must appear on all files. If no sub-set is selected, then the areas or data found on the first file, must appear on all following files.

As already stated, system files may be considered as matrices, where the rows are the areas, and the columns the data. In simplistic terms, SASPAC reads these files by progressing along a row to the end, before proceeding to the next row and repeating the process through to the end of the file.

If more than one file is input, SASPAC needs to be told whether the areas or the variables are the common factor, in order that it may place the appropriate information as an extension of a row (i.e. the areas are common), or as an extension of a column (i.e. the variables are common).

This is done through the pair of complementary commands:

```
READ IN SERIES  
READ IN PARALLEL
```

If more than one system file is input, the READ IN command must be present in one form or other, and must immediately follow the last INPUT SYSTEM FILE command (except for the possible inclusion of comments).

Files within a single task may be read either in series or in parallel. The two procedures may not be mixed.

5.6.1 Read in Series

If the files to be input contain the same data for different areas, then effectively, SASPAC must add each file onto the end of the previous one to create a matrix which has the same number of columns as the original, but a number of rows equal to the combined total of rows in all files

read. This procedure is termed 'Reading in Series' since SASPAC is placing files in a vertical series.

An example of the use of the 'READ IN SERIES' command would be the input of the CS for OAs in a county, along with the CS for OAs in a different county.

5.6.2 Read in Parallel

If the files to be input contain different data for the same areas, then effectively, SASPAC must place each file alongside the previous one to create a matrix which has the same number of rows as the original, but a number of columns equal to the combined total of columns in all files read. This procedure is termed 'Reading in Parallel' since SASPAC is placing files alongside (or parallel to) the previous one.

An example of the use of the 'READ IN PARALLEL' command would be the input of the 1991 100% SAS for EDs in one county, along with the 1991 10% SAS for EDs in the same county.

The figure below shows how the two methods of reading System Files operate. If the files to be read contain the same data for different areas, then the two (or more) files are effectively put one on top of the other, and read in series. Conversely, if the files to be read contain different data for the same areas, then they are placed one alongside the other, and read in parallel.

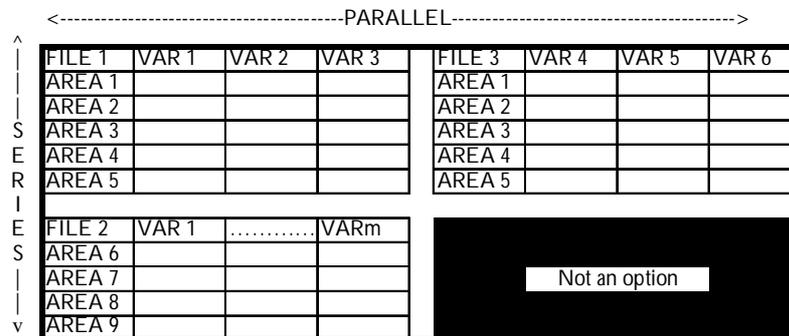


Figure 5: Input of more than one System File

SASPAC allows the appropriate selection to be made, without the user having to consider whether the required keyword is Series or Parallel. All that is required of the user is the need to identify whether it is the areas or the variables that are common.

5.7 Creating Tasks

Each of the modules in this training document relate to the selection and manipulation of the data. With the exception of module 7 - Creating New Zones - each module may lead to any of the types of output available within SASPAC. New Zone creation must lead to the derivation and output of a new system file.

In order for SASPAC to perform a task, it must be provided with certain information by the user. To do this the user must ask and answer questions such as:

What type of output is required and where is it to be written to? The answer to this question will define the type of task that is to be created, and will need to be the first consideration. The answer to this may depend on the answer to the next question?

- What data is required - Key Statistics, Univariate, etc.
- What area level is required - OA, ward, local authority, county, etc.
- What variables (or tables) are required - do they exist or will they need to be created by the user?
- What areas are to be analysed - do they exist or will they need to be created by the user?
- How are the areas to be selected - by name, by geography, or by condition?

Once the answers to these questions are available, the user may start the creation of the command file for the task by acting on the response to the first question and using the File / New Task menu option as shown below.

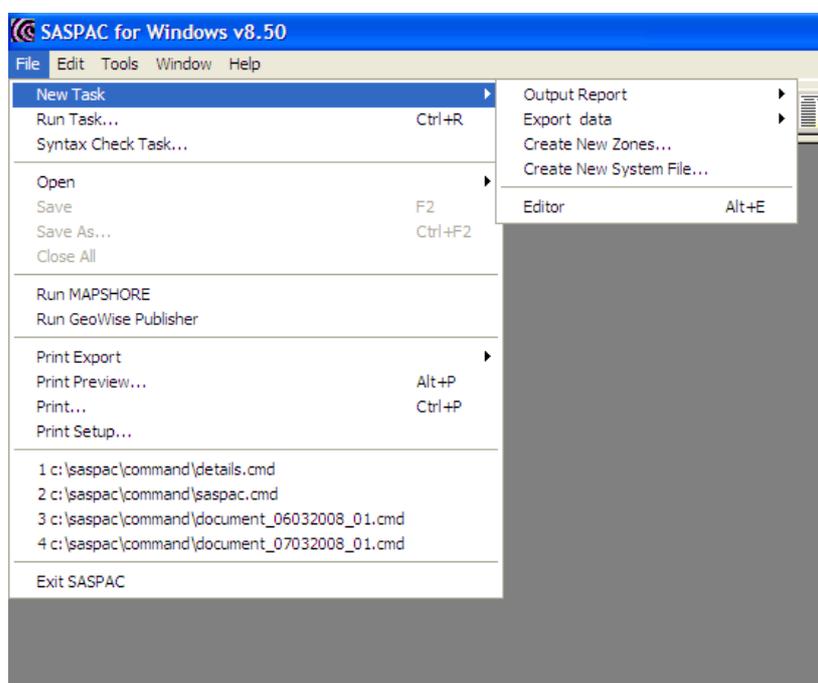


Figure 6: Selection of type of task

Once the type of task has been specified, the answers to the other questions are then conveyed to SASPAC through the various windows and options that are presented to the user. Immediately following the selection of the type of task, the user is presented with the main task window. This window varies with the type of task selected, although there is a certain amount of similarity between them all.

Each main task window contains an area in which the input System Files are defined, and this area (on the left of the window) is the same for all tasks. The other areas of the main task window consist of a central area where, generally, manipulations and selections of the data are undertaken, and an area on the right which consists of buttons linked to other options related to the command file. Figure below shows the main task window for a 'Print Variables' task.

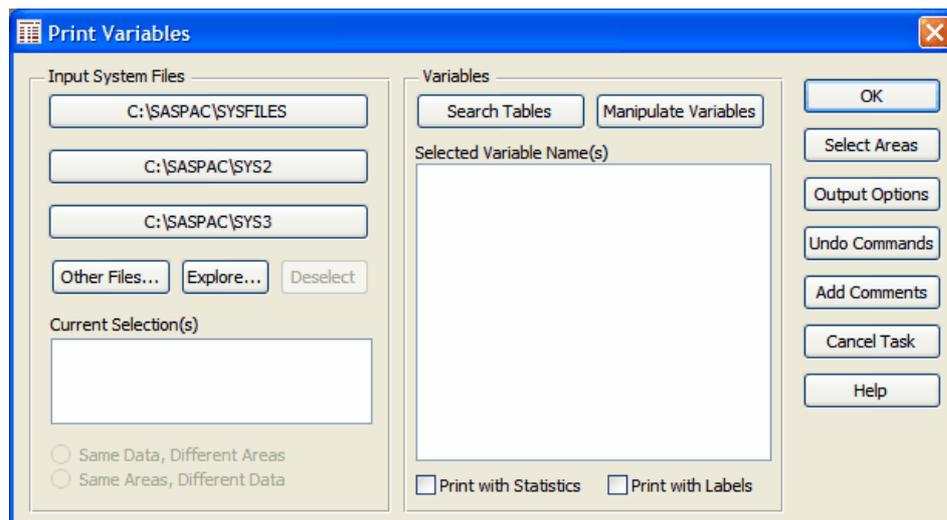


Figure 7: Main Task Window for Print Variables

There is (generally) no set order in which the elements of this window are completed, since the software will automatically order the generated statements into the required logical order. However, there will be instances when an input system file has to be defined before another element can be selected.

However, it should be noted that if the task is to access either the SWSC or SMS datasets, then these must be selected at the outset, as in these cases the windows used by SASPAC differ slightly from those presented to the user as a default, or when one of the other datasets is to be accessed.

What can be output from SASPAC?

The types of output available from SASPAC are:

- "Printed" copy - Variables
- Tables (can also be output as HTML files).
- Profiles

System Files

- Interface Files - Formatted
- Delimited
- CSV
- DIF
- WKS
- DBF

Each of these would be used as appropriate, and in this training document most will be covered. The examples shown in this manual refer to areas in Greater London, while the exercises to be undertaken during the training sessions will refer to areas of interest to the students.

6 Further Support

Visit the Training and Support pages of the SASPAC website (<http://saspac.org>) for more Census and SASPAC information. You can also contact the SASPAC Helpdesk by emailing info@saspac.org.

The websites of the census offices of the constituent countries of the UK have detailed methodology papers for each stage of the census process as well as wealth of census and comparator data resources.

The Office for National Statistics (ONS):

<http://www.ons.gov.uk/ons/guide-method/census/2011/index.html>

The Northern Ireland Statistics and Research Agency (NISRA):

<http://www.nisra.gov.uk/Census.html>

National Records Scotland (NRS):

<http://www.gro-scotland.gov.uk/census/index.html>

The Greater London Authority produces census analysis for London through the Census Information Scheme of which all 33 London Boroughs are members. The London Datastore provides access to reports and interactive tools for the interrogation of census data. The CIS also produces methodological papers which all users may find users.

<http://data.london.gov.uk/census>

Annex 1 Typical SASPAC.INI initialisation file for release 8.5**[DDL Files]**

sas91=C:\SASPAC\METADATA\sas91.ddl
lbs91=C:\SASPAC\METADATA\lbs91.ddl
sas81=C:\SASPAC\METADATA\sas81.ddl
sws81=C:\SASPAC\METADATA\sws81.ddl

[Help File]

help=C:\SASPAC\sashelp.chm

[GIS Information]

GISName= Mapshore
GISProgramPathname= c:\program files\mapshore\mapshore.exe
GISExportToSASPAC= c:\saspac\mapping
GISImportFromSASPAC= c:\saspac\interfac
GISFindWindow= InfomapWClass
GISParameters= /onecopy

[GeoWise]

XMLSVGWarning =
XMLProgramPathname = C:\SASPAC\GeoWise\GWPublisher.exe
XMLParameters =
XMLFindWindow = WindowsForms10.Window.8.app3
XMLImportFromSASPAC = C:\SASPAC\interfac
XMLNAME = SASPAC GeoWise Web Publisher

[Framework, Table & Cell Metadata]

2001KS=origin:"SASPAC",fwk:"ks01.fwk",fdesc:"Key
Statistics",year:"2001",type:"S",lookup:"KS01DA",xml:"ks01"
2001UV=origin:"SASPAC",fwk:"uv01.fwk",fdesc:"Univariate
Tables",year:"2001",type:"S",lookup:"UV01DA",xml:"uv01"
2001ST=origin:"SASPAC",fwk:"st01.fwk",fdesc:"Standard
Tables",year:"2001",type:"S",lookup:"ST01DA",xml:"st01"
2001TT=origin:"SASPAC",fwk:"tt01.fwk",fdesc:"Standard Theme
Tables",year:"2001",type:"S",lookup:"TT01DA"
2001CS=origin:"SASPAC",fwk:"cs01.fwk",fdesc:"Census Area Statistics
Tables",year:"2001",type:"S",lookup:"CS01DA",xml:"cs01"
2001CT=origin:"SASPAC",fwk:"ct01.fwk",fdesc:"Census Area Theme
Tables",year:"2001",type:"S",lookup:"CT01DA"
2001KK=origin:"SASPAC",fwk:"kk01.fwk",fdesc:"Var
Tables",year:"2001",type:"K",lookup:"KS01DA"
2001SWS=origin:"SASPAC",fwk:"sws01.fwk",fdesc:"Workplace-Travel
Tables",year:"2001",type:"D",lookup:"SWS01DA"
2001SMS=origin:"SASPAC",fwk:"sms01.fwk",fdesc:"Migration
Tables",year:"2001",type:"D",lookup:"SMS01DA"

2001TVS=origin:"SASPAC",fwk:"tvs01.fwk",fdesc:"Travel to Work/Study Tables (Scotland)",year:"2001",type:"D",lookup:"TVS01DA"
 2001SCT=origin:"SASPAC",fwk:"sct01.fwk",fdesc:"Specially-Commissioned Tables",year:"2001",type:"H",lookup:"SCT01DA"
 2001AF=origin:"SASPAC",fwk:"af01.fwk",fdesc:"Armed Forces Tables",year:"2001",type:"S",lookup:"AF01DA"
 1991SAS160=origin:"SASPAC",fwk:"tsas160.fwk",fdesc:"Small Area Statistics (SAS)",year:"1991",type:"S",lookup:"SASDA"
 1991LBS160=origin:"SASPAC",fwk:"tlbs160.fwk",fdesc:"Local Base Statistics (LBS)",year:"1991",type:"S",lookup:"LBSDA"
 1991SWS(A&B)=origin:"SASPAC",fwk:"tsws.fwk",fdesc:"Special Workplace Statistics (A&B)",year:"1991",type:"S",lookup:"SWS91DA"
 1991SWS(C)=origin:"SASPAC",fwk:"swsc91.fwk",fdesc:"Special Workplace Statistics (C)",year:"1991",type:"D",lookup:"SWSC91DA"
 1991SMS=origin:"SASPAC",fwk:"sms91.fwk",fdesc:"Special Migration Statistics",year:"1991",type:"D",lookup:"SMS91DA"
 1991LRC=origin:"SASPAC",fwk:"lrc91.fwk",fdesc:"Specially-Commissioned Tables (LRC)",year:"1991",type:"S",lookup:"LRCDA"
 1991NI160=origin:"SASPAC",fwk:"nitab160.fwk",fdesc:"Northern Ireland Tables",year:"1991",type:"S",lookup:"NIDA"
 1981SASTABLE=origin:"SASPAC",fwk:"table81.fwk",fdesc:"Small Area Statistics (Tables)",year:"1981",type:"S",lookup:"81DA"
 1981SASPAGE=origin:"SASPAC",fwk:"page81.fwk",fdesc:"Small Area Statistics (Pages)",year:"1981",type:"S",lookup:"81DA"
 1981SWS=origin:"SASPAC",fwk:"sws81tab.fwk",fdesc:"Special Workplace Statistics (A&B)",year:"1981",type:"S"
 1981JUVOS=origin:"SASPAC",fwk:"juvos.fwk",fdesc:"JUVOS",year:"1981",type:"S"

[Configuration File]

cfg = 8.50

[User configuration]

cfg= C:\DOCUME~1\alewis\LOCALS~1\Temp\saspac.cfg
 system1 file directory = J:\TEAM FOLDERS\SASPAC\SASPAC_DATA\2001 CENSUS RATIONALISED AND CONSOLIDATED SYSTEM FILES (CURRENT)
 system2 file directory = C:\SASPAC\SYS2
 system3 file directory = C:\SASPAC\SYS3
 command file directory name = C:\SASPAC\COMMAND
 report file directory name = C:\SASPAC\REPORT
 log file directory name = C:\SASPAC\LOG
 interface file directory name = C:\SASPAC\INTERFAC
 frame data file directory name = C:\SASPAC\FRWDATA
 opcs data file directory name = C:\SASPAC\OPCS
 framework file directory name = C:\SASPAC\FRWORK
 saspac4 file directory name = C:\SASPAC\SASPAC4
 convert framework border = YES
 key for var in framework = @
 key for title in framework = &

page length = 80
page width = 80
data format of framework = *****
position default of framework = R
maximum page = 200
page symbol for report = \$
text quote for delimited file = '
delimiter for delimited file = ;
missing value symbol = MISS
records file directory name = C:\SASPAC\RECORDS
html file directory = C:\SASPAC\HTML
mapping gazetteer file directory = C:\SASPAC\MAPPING

[Files]

4
c:\saspac\command\details.cmd
c:\saspac\command\saspac.cmd
c:\saspac\command\document_06032008_01.cmd
c:\saspac\command\document_07032008_01.cmd

[Settings]

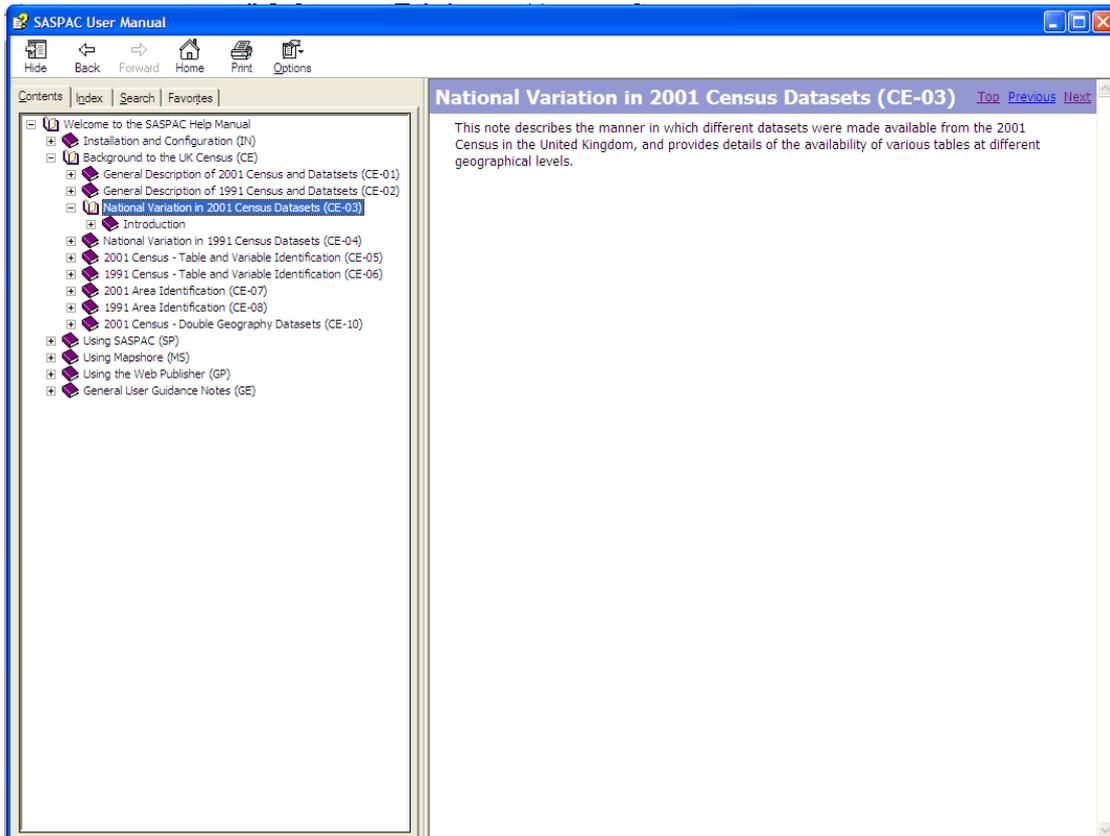
Enhance Include/Exclude readability = 0
Warn if missing XFF or XVM = 0
Large ToolBar Buttons = 1
800x600 warning = 0
RezoneCalcFile = C:\SASPAC\METADATA\KS PER_NUM ID.TXT
CommandLineMax = 140
Default Metadata Folder = C:\SASPAC\METADATA\
XMLHeaderFile = C:\SASPAC\METADATA\XMLHEAD.TXT
Editing font = Courier New,-12
Printing font = Courier New
Internet Browser = "Internet Explorer","C:\Program Files\Internet Explorer\iexplore.exe"
GMsize = 550,550
GMAPIkey =

[Printing]

diags=off
Paper = A4
Width = 210.00
Height = 297.00
LM = 12.00
RM = 12.00
TM = 12.00
BM = 12.00

Annex 2 Valid National Table Identifiers

Valid National Table Identifiers are detailed in the SASPAC User Manual, which can be accessed through the Help menu in the software. Census User Guidance note 'CE03' contains a series of tables that identify which datasets are available for different geographies and countries.



Annex 3 Government Statistical Service (GSS) Codes

Entity code	Entity name	Number of live instances	Current code (first in range)	Current code (last in range)
England				
E92	Country	1	E92000001	E92000001
E00	Output Areas	165,665	E00000001	E00165665
E01	Super Output Areas, Lower Layer	32,482	E01000001	E01032482
E02	Super Output Areas, Middle Layer	6,781	E02000001	E02006781
E04	Civil Parishes	10,493	E04000001	E04012250
E05	Electoral Wards/Divisions	7,683	E05000001	E05009029
E06	Unitary Authorities	56	E06000001	E06000056
E07	Non-metropolitan Districts	201	E07000004	E07000241
E08	Metropolitan Districts	36	E08000001	E08000036
E09	London Boroughs	33	E09000001	E09000033
E10	Counties	27	E10000002	E10000034
E11	Metropolitan Counties	6	E11000001	E11000006
E12	Regions (from 1 April 2011 - formerly Government Office Regions)	9	E12000001	E12000009
E13	Inner and Outer London	2	E13000001	E13000002
E14	Westminster Parliamentary Constituencies	533	E14000530	E14001062
E15	European Electoral Regions	9	E15000001	E15000009
E16	Primary Care Trusts	146	E16000001	E16000151
E17	Care Trusts	5	E17000001	E17000006
E18	Strategic Health Authorities	10	E18000001	E18000010
E19	Pan Strategic Health Authorities	3	E19000001	E19000003
E20	Cancer Registries	8	E20000001	E20000008
E21	Cancer Networks	28	E21000001	E21000030
E22	Community Safety Partnerships	309	E22000001	E22000367
E23	Police Force Areas	39	E23000001	E23000039
E24	Local Learning and Skills Council areas	0	n/a	n/a
E25	Primary Urban Areas	56	E25000001	E25000056
E26	National Parks	10	E26000001	E26000010
E27	New Deal for Communities	0	n/a	n/a
E28	Registration Districts	159	E28000001	E28000217
E29	Registration Sub-district	160	E29000001	E29000229
E30	Travel to Work Areas 2007	158	E30000001	E30000158
E31	Fire and Rescue Authorities	46	E31000001	E31000046
E32	London Assembly Constituencies	14	E32000001	E32000014
E37	Local Enterprise Partnerships	39	E37000001	E37000039
Wales				
W92	Country	1	W92000004	W92000004
W00	Output Areas	9,769	W00000001	W00009769
W01	Super Output Areas, Lower Layer	1,896	W01000001	W01001896
W02	Super Output Areas, Middle Layer	413	W02000001	W02000413
W03	Super Output Areas, Upper Layer	94	W03000001	W03000094
W04	Communities	870	W04000001	W04000979
W05	Electoral Divisions	881	W05000001	W05000980
W06	Unitary Authorities	22	W06000001	W06000024
W07	Westminster Parliamentary Constituencies	40	W07000041	W07000080

W08	European Electoral Regions	1	W08000001	W08000001
W09	National Assembly for Wales Constituencies	40	W09000001	W09000047
W10	National Assembly for Wales Electoral Regions	5	W10000001	W10000009
W11	Local Health Boards	7	W11000023	W11000029
W12	Cancer Registries	1	W12000001	W12000001
W13	Cancer Networks	2	W13000001	W13000004
W14	Community Safety Partnerships	22	W14000001	W14000022
W15	Police Force Areas	4	W15000001	W15000004
W16	Department of Children, Education, Lifelong Learning and Skills, WG	4	W16000001	W16000004
W18	National Parks	3	W18000001	W18000003
W19	National Assembly Economic Regions	4	W19000001	W19000004
W20	Registration Districts	25	W20000001	W20000039
W21	Registration Sub-district	26	W21000001	W21000043
W22	Travel to Work Areas 2007	20	W22000001	W22000020
W23	Spatial Plan Areas	6	W23000001	W23000006
W24	Spatial Plan Sub-areas	3	W24000001	W24000003
W25	Fire and Rescue Authorities	3	W25000001	W25000003
W26	Strategic Regeneration Areas	7	W26000001	W26000007
W27	Strategic Regeneration Sub Areas	2	W27000001	W27000002
W28	Transport Consortia Areas	4	W28000001	W28000004
W29	Agricultural Regions	7	W29000001	W29000007
W30	Agricultural Small Areas	235	W30000001	W30000235
W31	Non-National Park Area	1	W31000001	W31000001
W32	Non-Strategic Regeneration Area	1	W32000001	W32000001
W33	Communities First Areas	153	W33000001	W33000153
W34	Non-Communities First Areas	1	W34000001	W34000001
W35	Footprint Regions for Public Service Collaboration	6	W35000001	W35000006

Scotland

S92	Country	1	S92000003	S92000003
S00	Output Areas	42,604	S00000001	S00042604
S01	Data Zones	6,505	S01000001	S01006505
S02	Intermediate Geography	1,235	S02000001	S02001235
S03	Community Health Partnerships	34	S03000001	S03000044
S04	Regeneration Outcome Agreement Areas - Scotland	1	S04000001	S04000001
S05	Regeneration Outcome Agreement Areas - Community Planning Partnerships	28	S05000001	S05000028
S06	Regeneration Outcome Agreement Areas - Local Areas	180	S06000001	S06000180
S07	Regional Transport Partnerships	7	S07000001	S07000007
S08	Health Board areas	14	S08000001	S08000014
S09	Enterprise Regions	6	S09000001	S09000006
S10	Urban Regeneration Companies	6	S10000001	S10000006
S11	Strategic Development Plan Areas	4	S11000001	S11000004
S12	Council Areas	32	S12000005	S12000046
S13	Electoral Wards	353	S13002476	S13002834
S14	Westminster Parliamentary Constituencies	59	S14000001	S14000059
S15	European Electoral Regions	1	S15000001	S15000001
S16	Scottish Parliamentary Constituencies	73	S16000074	S16000146
S17	Scottish Parliamentary Regions	8	S17000009	S17000016

S18	Registration Districts	n/a	n/a	n/a
S19	Locality	n/a	n/a	n/a
S20	Settlement	n/a	n/a	n/a
S21	National Parks	2	S21000002	S21000003
S22	Travel to Work Areas 2007	46	S22000001	S22000046
S23	Police and Fire Areas	8	S23000001	S23000008
S24	Highlands and Islands Enterprise	8	S24000001	S24000008
S25	Community Justice Authorities	8	S25000001	S25000008

Northern Ireland

N92	Country	1	N92000002	N92000002
N06	Westminster Parliamentary Constituencies	18	N06000001	N06000018
N07	European Electoral Regions	1	N07000001	N07000001

Cross Border

K01	Travel to Work Areas 2007	8	K01000001	K01000008
K02	United Kingdom	1	K02000001	K02000001
K03	Great Britain	1	K03000001	K03000001
K04	England and Wales	1	K04000001	K04000001

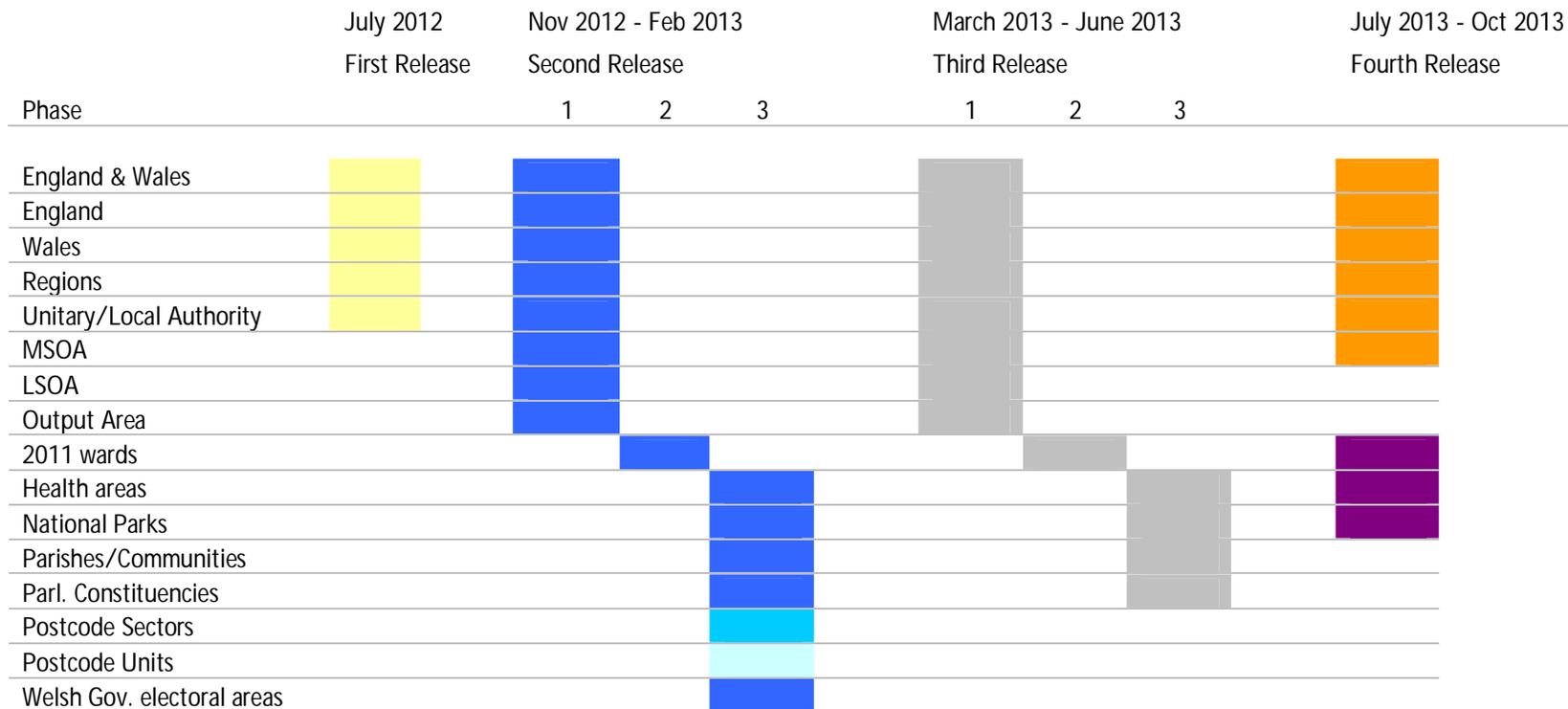
Channel Islands

L93	British Crown Dependencies	1	L93000001	L93000001
L00	Strategic Health Authorities	3	L00000001	L00000003

Isle of Man

M83	British Crown Dependency	1	M83000003	M83000003
M00	Strategic Health Authorities	1	M00000001	M00000001
M01	Primary Healthcare Directorate	1	M01000001	M01000001

Annex 4 2011 Census Release Schedule with Geography



Key

- Population/Household estimates
- Key Statistics & Quick Statistics
- Key Statistics only
- Postcode Unit headcounts
- Local Characteristics
- Detailed Characteristics, Detailed Themes & Armed Forces
- Detailed Characteristics, Detailed Themes

Annex 5 Output geographies 2011 Census

Geography	Currency	Fit	No. of instances
Output areas	December 2011	Exact-fit	To be confirmed
Lower layer super output areas	December 2011	Exact-fit	To be confirmed
Middle layer super output areas	December 2011	Exact-fit	To be confirmed
Workplace zones	December 2011	Exact-fit	To be confirmed
National parks	December 2011	Exact-fit	13
Local authority district	December 2011	Exact-fit	348
Wards	December 2011	Best-fit	8,588
Parishes/communities	December 2011	Best-fit	11,360
Counties	December 2011	Best-fit	27
Former counties	December 2009	Best-fit	34
Regions	December 2011	Best-fit	9
Westminster parliamentary constituencies	December 2011	Best-fit	573
Primary care organisations	December 2011	Best-fit	151
Local health boards	December 2011	Best-fit	7
Strategic Health Authorities	December 2011	Best-fit	10
National Assembly for Wales constituencies	December 2011	Best-fit	40
Local administrative units (LAU) 2	December 2011	Best-fit	8,588
Local administrative units (LAU) 1	December 2011	Best-fit	348
NUTS 3	December 2011	Best-fit	105
NUTS 2	December 2011	Best-fit	32
(NUTS 1	December 2011	Best-fit	10
Settlements	December 2011	Best-fit	To be confirmed
Urban areas	December 2011	Best-fit	To be confirmed
Postcode Unit/Sector	December 2011	Best-fit	To be confirmed