## Contents

Introduction to Census Data & SASPAC ................................................................. 1

1  Introduction ...................................................................................................... 3

2  The Census Dataset ....................................................................................... 3
   2.1 The Census Process ................................................................................... 3
   2.2 2011 Census Questionnaire ................................................................. 4
   2.3 Communal Establishments ........................................................................ 5
   2.4 2011 Census Release Schedule .............................................................. 5
   2.5 Further releases ....................................................................................... 6
   2.6 Samples of Anonymised Records (SARs) .................................................... 8
   2.7 Commissioned Tables ............................................................................. 8
   2.8 Disclosure Control .................................................................................. 9
   2.9 Data Licensing ....................................................................................... 9
   2.10 The 2001 Census ................................................................................... 10
   2.11 The 1991 Census ................................................................................... 11
   2.12 100% and 10% datasets in 1991 ............................................................. 11

3  Census Geography ......................................................................................... 12
   3.1 Boundary Files ....................................................................................... 12
   3.2 2011 Geographies ................................................................................... 13
   3.2.1 Best Fit .............................................................................................. 15
   3.2.2 Merged wards ................................................................................... 16
   3.2.3 Main Census Geographies ................................................................. 16
   3.2.4 2011 Naming Convention ................................................................. 17
   3.3 2001 Geographies ................................................................................... 18
   3.3.1 2001 Geographic Naming Convention .............................................. 18
   3.4 1991 Geographies ................................................................................... 18
   3.4.1 Special Areas ..................................................................................... 19
   3.4.2 Shipping Areas .................................................................................. 19
   3.4.3 1991 Geographic Naming Convention .............................................. 20
   3.4.4 1991 Geography in Scotland .............................................................. 20

4  Table and Variable Naming .......................................................................... 21
   4.1 2011 Table Naming ................................................................................ 21
   4.2 2011 Variable Naming .......................................................................... 22
   4.3 2001 Table Naming .............................................................................. 22
   4.4 2001 Variable Naming .......................................................................... 22
   4.5 1991 Table Naming .............................................................................. 22
   4.6 1991 Variable Naming .......................................................................... 23

5  Introduction to SASPAC ............................................................................... 24
   5.1 Configuration ......................................................................................... 24
   5.2 System Files ........................................................................................... 24
   5.3 Command Files ..................................................................................... 25
   5.4 Other File Types ................................................................................... 26
   5.5 Framework Files .................................................................................. 27
   5.6 Multiple System Files .......................................................................... 28
   5.6.1 Read in Series .................................................................................. 28
   5.6.2 Read in Parallel ................................................................................ 29
   5.7 Creating Tasks ...................................................................................... 29

6  Further Support ............................................................................................. 32
Module 1  Selection of Variables ................................................................. 42
Module 2  Selection of Tables ................................................................. 55
Module 3a  Selection of Areas - by Text Listing ...................................... 63
Module 3b  Selection of Areas - Using Mapping ..................................... 71
Module 4a  Creating New Zones - Individual Area Selection .................. 84
Module 4b  Creating New Zones - Gazetteer Files .................................. 92
Module 5  Output of Information from SASPAC ................................. 100
Module 6  Creating New Variables ......................................................... 110
Module 7  Selection of Areas - SELECT IF ............................................ 118
Module 8  Use of Double Geography datasets within SASPAC .......... 124
Module 9  Use of Google Maps within SASPAC ................................. 140
Module 10 Creating User-Defined Tables ............................................ 152
Appendix 1 Worked example answers .................................................. 164
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. Household 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 household 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 ONS 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 whose 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 Scottish census questionnaire is largely the same as the questionnaire for England & Wales but there are some differences. For instance, the Scottish census includes (among others):
- A question on respondents ability to speak/read/write Scottish Gaelic
- A more detailed religion question
- Different categories in the ethnicity question

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\textsuperscript{th} 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) also released population estimates for Northern Ireland on 16\textsuperscript{th} July while National Records Scotland (NRS) released the first data for Scotland on 17\textsuperscript{th} December 2012.

\textbf{ONS release schedule}
(correct at October 2013)
(\textit{Geography refers to lowest level available})

<table>
<thead>
<tr>
<th>Date</th>
<th>Release Number</th>
<th>Tables</th>
<th>Geography</th>
</tr>
</thead>
<tbody>
<tr>
<td>16 July</td>
<td>1.1</td>
<td>Population estimates (5-year age bands, sex, rounded)</td>
<td>Local Authority</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Household estimates</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Short-term migrant estimates</td>
<td></td>
</tr>
<tr>
<td>24 September</td>
<td>1.2</td>
<td>Population estimates (single year, sex, unrounded)</td>
<td>Local Authority</td>
</tr>
<tr>
<td>23 October</td>
<td>1.3</td>
<td>Second address information</td>
<td>Local Authority</td>
</tr>
<tr>
<td>30 October</td>
<td>-</td>
<td>Statistical geography boundaries and lookups</td>
<td>Output Area, LSOA, MSOA</td>
</tr>
<tr>
<td>23 November</td>
<td>-</td>
<td>Population estimates (5-year age bands, sex, unrounded)</td>
<td>Output Area, LSOA, MSOA, Ward</td>
</tr>
<tr>
<td>11 December</td>
<td>2.1</td>
<td>36 Key Stats 7 Quick Stats</td>
<td>Local Authority</td>
</tr>
<tr>
<td>Date</td>
<td>Page</td>
<td>Release Type</td>
<td>Geographies</td>
</tr>
<tr>
<td>--------------------</td>
<td>------</td>
<td>--------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>30 January</td>
<td>2.2</td>
<td>36 Key Stats</td>
<td>Output Area, LSOA, MSOA, Ward</td>
</tr>
<tr>
<td>30 January</td>
<td>69</td>
<td>Quick Stats</td>
<td>All other geographies</td>
</tr>
<tr>
<td>19 February</td>
<td>2.3</td>
<td>36 Key Stats</td>
<td>Local Authority</td>
</tr>
<tr>
<td>26 March</td>
<td>2.4</td>
<td>4 'AP' tables</td>
<td>All other geographies</td>
</tr>
<tr>
<td>16 May</td>
<td>3.1</td>
<td>42 Detailed Characteristics</td>
<td>Local Authority, Regional</td>
</tr>
<tr>
<td>30 June</td>
<td>3.2a</td>
<td>Detailed Characteristics</td>
<td>Local Authority, Regional</td>
</tr>
<tr>
<td>30 June</td>
<td>3.2b</td>
<td>Detailed Characteristics</td>
<td>MSOA, Ward</td>
</tr>
<tr>
<td>16 May</td>
<td>3.1</td>
<td>42 Detailed Characteristics</td>
<td>Local Authority, Regional</td>
</tr>
<tr>
<td>31 July</td>
<td>4.1</td>
<td>Local Characteristics</td>
<td>Output Area, LSOA, MSOA, Wards</td>
</tr>
<tr>
<td>30 August</td>
<td>4.2</td>
<td>Local Characteristics</td>
<td>Output Area, LSOA, MSOA, Wards</td>
</tr>
<tr>
<td>November 2013</td>
<td>3.5</td>
<td>Detailed Characteristics</td>
<td>MSOA, Ward</td>
</tr>
<tr>
<td>February 2014</td>
<td>3.6</td>
<td>Detailed Characteristics</td>
<td>MSOA, Ward</td>
</tr>
<tr>
<td>September 2013</td>
<td>4.3</td>
<td>Local Characteristics</td>
<td>Output Area, LSOA, MSOA, Wards</td>
</tr>
<tr>
<td>December 2013</td>
<td>4.4</td>
<td>Local Characteristics</td>
<td>Output Area, LSOA, MSOA, Wards</td>
</tr>
<tr>
<td>January 2014</td>
<td>4.5</td>
<td>Local Characteristics</td>
<td>Output Area, LSOA, MSOA, Wards</td>
</tr>
</tbody>
</table>

**NRS release schedule**
The Nation Records Scotland have a different release schedule for data from the 2011 Census. The first release of Scottish data came on 17 December 2012, six months after the first England & Wales data. Since then the following data has been release/has been scheduled for release:

<table>
<thead>
<tr>
<th>Date</th>
<th>Release Number</th>
<th>Tables</th>
<th>Geography</th>
</tr>
</thead>
<tbody>
<tr>
<td>17 December 2012</td>
<td>1A</td>
<td>Population estimates (5-year age bands, sex, rounded 000s)</td>
<td>Local Authority</td>
</tr>
<tr>
<td>21 March 2013</td>
<td>1B</td>
<td>Population estimates (5-year age bands, sex, rounded 000s)</td>
<td>Local Authority</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Household estimates (rounded)</td>
<td></td>
</tr>
<tr>
<td>23 July</td>
<td>1C(i)</td>
<td>Population estimates (sya, sex, unrounded)</td>
<td>Local Authority</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Household estimates (unrounded)</td>
<td></td>
</tr>
<tr>
<td>15 August</td>
<td>1C(ii)</td>
<td>Population estimates for households &amp; communal establishments (totals, unrounded)</td>
<td>Output Area, postcode</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Household estimates (totals, unrounded)</td>
<td></td>
</tr>
<tr>
<td>26 September</td>
<td>2A</td>
<td>Various Key &amp; Quick Statistics tables</td>
<td>Local Authority</td>
</tr>
<tr>
<td>November</td>
<td>2B</td>
<td>Various Key &amp; Quick Statistics tables</td>
<td>Output Area</td>
</tr>
<tr>
<td>December</td>
<td>2C</td>
<td>Remaining Key &amp; Quick Statistics tables</td>
<td>Output Area</td>
</tr>
<tr>
<td>Unknown</td>
<td>3</td>
<td>Local Characteristics</td>
<td>Output Area</td>
</tr>
<tr>
<td>Unknown</td>
<td>4</td>
<td>Detailed Characteristics</td>
<td>Ward</td>
</tr>
</tbody>
</table>

**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.

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

2.5 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 License 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.


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.

2.6 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.


Note: Local Authorities in London should be aware 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.

2.7 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.


2.8 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:
Introduction

“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.

2.9 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.

2.10 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

2.11 100% 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 is a lookup available 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 are 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 have been produced. There are also 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. Lookup files can be downloaded from the ONS Geography Portal: https://geoportal.statistics.gov.uk/geoportal/catalog/main/home.page

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

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. 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.

Merged Wards  Where ward populations are small and data releases would be disclosive some administrative wards must be merged in order to create larger, non-disclosive, populations. This is the case for some Detailed Characteristics tables in 2011. (See 3.2.2 for more details).

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.

Merged LAs  The populations of the local authorities of City of London and Isles of Scilly are comparatively small and as a result some census estimates in 2011 were judged to be disclosive. In order to mitigate this issue ONS merged City of London with Westminster and Isles of Scilly with Cornwall.
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):

- 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.


### 3.2.2 Merged wards

In some instances (particularly in the more comprehensive Detailed Characteristics tables) wards may require wards to be amalgamated to meet population thresholds. This is 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. These wards will constitute a geographical hierarchy in their own right with a new set of unique GSS (Government Statistical Service) codes assigned (code E36). 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.

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

### 3.2.3 Main Census Geographies

<table>
<thead>
<tr>
<th>Geography</th>
<th>Description</th>
<th>Min – Max population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output Area</td>
<td>Smallest area for which data are available</td>
<td>100 – 625</td>
</tr>
<tr>
<td>LSOA</td>
<td>4 or 5 output areas</td>
<td>1,000 – 3,000</td>
</tr>
<tr>
<td>MSOA</td>
<td>4 or 5 LSOAs</td>
<td>2,000 – 5,000</td>
</tr>
<tr>
<td>Wards (admin &amp; merged)</td>
<td>Comprised of those OAs which best fit the ward</td>
<td>2,000 – 5,000</td>
</tr>
<tr>
<td>LA/UA</td>
<td>Total of 348 district and unitary authorities in E&amp;W</td>
<td>n/a</td>
</tr>
<tr>
<td>Region</td>
<td>9 regions in E&amp;W</td>
<td>n/a</td>
</tr>
</tbody>
</table>
3.2.4 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.

<table>
<thead>
<tr>
<th>Region</th>
<th>1 alphabetic character</th>
</tr>
</thead>
<tbody>
<tr>
<td>County or UA</td>
<td>2 numeric characters</td>
</tr>
<tr>
<td>Local Authority</td>
<td>2 alphabetic characters</td>
</tr>
<tr>
<td>Ward</td>
<td>2 alphabetic characters</td>
</tr>
<tr>
<td>Output Area</td>
<td>4 numeric characters</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th></th>
<th>Two digit code</th>
<th>01-55</th>
</tr>
</thead>
<tbody>
<tr>
<td>County</td>
<td>Two digit code</td>
<td>01-55</td>
</tr>
<tr>
<td>District</td>
<td>Four digit code</td>
<td>AA - TT</td>
</tr>
<tr>
<td>Ward</td>
<td>Six digit code</td>
<td>FA, etc.</td>
</tr>
<tr>
<td>ED</td>
<td>Eight digit code</td>
<td>01, etc.</td>
</tr>
</tbody>
</table>

**Example:**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>County</td>
<td>03</td>
</tr>
<tr>
<td>District</td>
<td>03BN</td>
</tr>
<tr>
<td>Ward</td>
<td>03BNFA</td>
</tr>
<tr>
<td>Output Area</td>
<td>03BNFA03</td>
</tr>
<tr>
<td></td>
<td>Greater Manchester</td>
</tr>
<tr>
<td></td>
<td>Manchester</td>
</tr>
<tr>
<td></td>
<td>Ardwick</td>
</tr>
<tr>
<td></td>
<td>Output area in Ardwick ward in Manchester, Gtr Manchester</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th></th>
<th>2 numeric characters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region</td>
<td>2 alphabetic characters</td>
</tr>
<tr>
<td>District</td>
<td>2 alphabetic characters</td>
</tr>
<tr>
<td>Postcode sector</td>
<td>2 alphabetic characters</td>
</tr>
<tr>
<td>OA</td>
<td>3 characters (2 numeric followed by 1 alphabetic)</td>
</tr>
</tbody>
</table>

**Example:**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Region</td>
<td>57</td>
</tr>
<tr>
<td>District</td>
<td>5705</td>
</tr>
<tr>
<td>PC Sector</td>
<td>5705AC</td>
</tr>
<tr>
<td>Output Area</td>
<td>5705AC10A</td>
</tr>
<tr>
<td></td>
<td>Central</td>
</tr>
<tr>
<td></td>
<td>Clackmannan</td>
</tr>
<tr>
<td></td>
<td>Postcode sector “FK1 01”</td>
</tr>
<tr>
<td></td>
<td>Output area in sector FK1 01 in Clackmannan, Central Region</td>
</tr>
</tbody>
</table>
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
- AF Armed Forces

Univariate tables, the third, fourth and fifth characters refer to the specific table:

- 101 indicates table 1 (etc., up to 999)

Multivariate tables, the third, fourth, fifth and sixth characters refer to the specific table:

- 1104

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

- EW England & Wales
- WA Wales only (Welsh language)
- SC Scottish

**Examples:**

- KS201EW Key Statistics Ethnic Group
- QS204EW Quick Statistics Main Language (detailed)
- QS206WA Quick Statistics Welsh language Skills
- DC4403EW Accommodation type by household spaces tab
4.2 2011 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.

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 a 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:**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>L06</td>
<td>Ethnic Group of Residents</td>
</tr>
<tr>
<td>S12</td>
<td>Long-term illness in households</td>
</tr>
</tbody>
</table>

**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:**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>L060066</td>
<td>Number of ethnically Indian residents aged 15 (taken from the Ethnic Group Table 06)</td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>File type</th>
<th>Extension</th>
<th>Purpose, content, or use of file</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log files</td>
<td>.LOG</td>
<td>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.</td>
</tr>
<tr>
<td>Report files</td>
<td>.PRN</td>
<td>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.</td>
</tr>
<tr>
<td>Interface files</td>
<td>(various)</td>
<td>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.</td>
</tr>
<tr>
<td>Framework Data files</td>
<td>.FWD</td>
<td>These are ASCII representations of the table layouts necessary to produce the Census Offices table images.</td>
</tr>
<tr>
<td>Framework files</td>
<td>.FWK</td>
<td>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.</td>
</tr>
<tr>
<td>Initialisation file</td>
<td>.INI</td>
<td>Windows initialisation file. This file may be modified and customised by users.</td>
</tr>
<tr>
<td>Gazetteer files</td>
<td>.GAZ</td>
<td>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.</td>
</tr>
<tr>
<td>Data Definition Language file</td>
<td>DDL</td>
<td>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.</td>
</tr>
<tr>
<td>OPCS Text files</td>
<td>.TXT</td>
<td>Raw data files as supplied by the Census Offices, or from other sources.</td>
</tr>
</tbody>
</table>
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.
- TSW5 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](image)

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](image)

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.
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):

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"
Introduction

2001TVS=origin:"SASPAC",fwk:"tv01.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"
1991WS(C)=origin:"SASPAC",fwk:"swsc91.fwk",fdesc:"Special Workplace Statistics (C)",year:"1991",type:"D",lookup:"SWSC91DA"

[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 = &
Introduction Page 35

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
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

<table>
<thead>
<tr>
<th>Entity code</th>
<th>Entity name</th>
<th>Number of live instances</th>
<th>Current code (first in range)</th>
<th>Current code (last in range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E92</td>
<td>Country</td>
<td></td>
<td>E92000001</td>
<td>E92000001</td>
</tr>
<tr>
<td>E00</td>
<td>Output Areas</td>
<td>165,665</td>
<td>E00000001</td>
<td>E00165665</td>
</tr>
<tr>
<td>E01</td>
<td>Super Output Areas, Lower Layer</td>
<td>32,482</td>
<td>E01000001</td>
<td>E01032482</td>
</tr>
<tr>
<td>E02</td>
<td>Super Output Areas, Middle Layer</td>
<td>6,781</td>
<td>E02000001</td>
<td>E0206781</td>
</tr>
<tr>
<td>E04</td>
<td>Civil Parishes</td>
<td>10,493</td>
<td>E04000001</td>
<td>E04012250</td>
</tr>
<tr>
<td>E05</td>
<td>Electoral Wards/Divisions</td>
<td>7,683</td>
<td>E05000001</td>
<td>E05009029</td>
</tr>
<tr>
<td>E06</td>
<td>Unitary Authorities</td>
<td>56</td>
<td>E06000001</td>
<td>E06000056</td>
</tr>
<tr>
<td>E07</td>
<td>Non-metropolitan Districts</td>
<td>201</td>
<td>E07000001</td>
<td>E07002241</td>
</tr>
<tr>
<td>E08</td>
<td>Metropolitan Districts</td>
<td>36</td>
<td>E08000001</td>
<td>E0800036</td>
</tr>
<tr>
<td>E09</td>
<td>London Boroughs</td>
<td>33</td>
<td>E09000001</td>
<td>E0900033</td>
</tr>
<tr>
<td>E10</td>
<td>Counties</td>
<td>27</td>
<td>E10000002</td>
<td>E10000034</td>
</tr>
<tr>
<td>E11</td>
<td>Metropolitan Counties</td>
<td>6</td>
<td>E11000001</td>
<td>E11000006</td>
</tr>
<tr>
<td>E12</td>
<td>Regions (from 1 April 2011 - formerly Government Office Regions)</td>
<td>9</td>
<td>E12000001</td>
<td>E12000009</td>
</tr>
<tr>
<td>E13</td>
<td>Inner and Outer London</td>
<td>2</td>
<td>E13000001</td>
<td>E13000002</td>
</tr>
<tr>
<td>E14</td>
<td>Westminster Parliamentary Constituencies</td>
<td>533</td>
<td>E14000530</td>
<td>E14001062</td>
</tr>
<tr>
<td>E15</td>
<td>European Electoral Regions</td>
<td>9</td>
<td>E15000001</td>
<td>E15000009</td>
</tr>
<tr>
<td>E16</td>
<td>Primary Care Trusts</td>
<td>146</td>
<td>E16000001</td>
<td>E16000151</td>
</tr>
<tr>
<td>E17</td>
<td>Care Trusts</td>
<td>5</td>
<td>E17000001</td>
<td>E17000006</td>
</tr>
<tr>
<td>E18</td>
<td>Strategic Health Authorities</td>
<td>10</td>
<td>E18000001</td>
<td>E18000010</td>
</tr>
<tr>
<td>E19</td>
<td>Pan Strategic Health Authorities</td>
<td>3</td>
<td>E19000001</td>
<td>E19000003</td>
</tr>
<tr>
<td>E20</td>
<td>Cancer Registries</td>
<td>8</td>
<td>E20000001</td>
<td>E20000008</td>
</tr>
<tr>
<td>E21</td>
<td>Cancer Networks</td>
<td>28</td>
<td>E21000001</td>
<td>E21000030</td>
</tr>
<tr>
<td>E22</td>
<td>Community Safety Partnerships</td>
<td>309</td>
<td>E22000001</td>
<td>E22000367</td>
</tr>
<tr>
<td>E23</td>
<td>Police Force Areas</td>
<td>39</td>
<td>E23000001</td>
<td>E23000039</td>
</tr>
<tr>
<td>E24</td>
<td>Local Learning and Skills Council areas</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>E25</td>
<td>Primary Urban Areas</td>
<td>56</td>
<td>E25000001</td>
<td>E25000056</td>
</tr>
<tr>
<td>E26</td>
<td>National Parks</td>
<td>10</td>
<td>E26000001</td>
<td>E26000010</td>
</tr>
<tr>
<td>E27</td>
<td>New Deal for Communities</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>E28</td>
<td>Registration Districts</td>
<td>159</td>
<td>E28000001</td>
<td>E28000217</td>
</tr>
<tr>
<td>E29</td>
<td>Registration Sub-district</td>
<td>160</td>
<td>E29000001</td>
<td>E29000229</td>
</tr>
<tr>
<td>E30</td>
<td>Travel to Work Areas 2007</td>
<td>158</td>
<td>E30000001</td>
<td>E30000158</td>
</tr>
<tr>
<td>E31</td>
<td>Fire and Rescue Authorities</td>
<td>46</td>
<td>E31000001</td>
<td>E31000046</td>
</tr>
<tr>
<td>E32</td>
<td>London Assembly Constituencies</td>
<td>14</td>
<td>E32000001</td>
<td>E32000014</td>
</tr>
<tr>
<td>E36</td>
<td>Merged 2011 Census Wards</td>
<td>7,678</td>
<td>E36000001</td>
<td>E36007678</td>
</tr>
<tr>
<td>E37</td>
<td>Local Enterprise Partnerships</td>
<td>39</td>
<td>E37000001</td>
<td>E37000039</td>
</tr>
<tr>
<td>E41</td>
<td>Merged 2011 Census Unitary Authorities</td>
<td>324</td>
<td>E41000001</td>
<td>E41000324</td>
</tr>
</tbody>
</table>

### Wales

<table>
<thead>
<tr>
<th>Entity code</th>
<th>Entity name</th>
<th>Number of live instances</th>
<th>Current code (first in range)</th>
<th>Current code (last in range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>W92</td>
<td>Country</td>
<td></td>
<td>W92000004</td>
<td>W92000004</td>
</tr>
<tr>
<td>W00</td>
<td>Output Areas</td>
<td>9,769</td>
<td>W00000001</td>
<td>W00009769</td>
</tr>
<tr>
<td>W01</td>
<td>Super Output Areas, Lower Layer</td>
<td>1,896</td>
<td>W01000001</td>
<td>W01001896</td>
</tr>
<tr>
<td>W02</td>
<td>Super Output Areas, Middle Layer</td>
<td>413</td>
<td>W02000001</td>
<td>W0200413</td>
</tr>
<tr>
<td>W03</td>
<td>Super Output Areas, Upper Layer</td>
<td>94</td>
<td>W03000001</td>
<td>W03000994</td>
</tr>
<tr>
<td>W04</td>
<td>Communities</td>
<td>870</td>
<td>W04000001</td>
<td>W04000979</td>
</tr>
<tr>
<td>W05</td>
<td>Electoral Divisions</td>
<td>881</td>
<td>W05000001</td>
<td>W05000980</td>
</tr>
<tr>
<td>Introduction</td>
<td>SASPAC Training</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W06 Unitary Authorities</td>
<td>22</td>
<td>W06000001</td>
<td>W06000024</td>
<td></td>
</tr>
<tr>
<td>W07 Westminster Parliamentary Constituencies</td>
<td>40</td>
<td>W07000041</td>
<td>W07000080</td>
<td></td>
</tr>
<tr>
<td>W08 European Electoral Regions</td>
<td>1</td>
<td>W08000001</td>
<td>W08000001</td>
<td></td>
</tr>
<tr>
<td>W09 National Assembly for Wales Constituencies</td>
<td>40</td>
<td>W09000001</td>
<td>W09000047</td>
<td></td>
</tr>
<tr>
<td>W10 National Assembly for Wales Electoral Regions</td>
<td>5</td>
<td>W10000001</td>
<td>W10000009</td>
<td></td>
</tr>
<tr>
<td>W11 Local Health Boards</td>
<td>7</td>
<td>W11000023</td>
<td>W11000029</td>
<td></td>
</tr>
<tr>
<td>W12 Cancer Registries</td>
<td>1</td>
<td>W12000001</td>
<td>W12000001</td>
<td></td>
</tr>
<tr>
<td>W13 Cancer Networks</td>
<td>2</td>
<td>W13000001</td>
<td>W13000004</td>
<td></td>
</tr>
<tr>
<td>W14 Community Safety Partnerships</td>
<td>22</td>
<td>W14000001</td>
<td>W14000022</td>
<td></td>
</tr>
<tr>
<td>W15 Police Force Areas</td>
<td>4</td>
<td>W15000001</td>
<td>W15000004</td>
<td></td>
</tr>
<tr>
<td>W16 Department of Children, Education, Lifelong Learning and Skills, WG</td>
<td>4</td>
<td>W16000001</td>
<td>W16000004</td>
<td></td>
</tr>
<tr>
<td>W18 National Parks</td>
<td>3</td>
<td>W18000001</td>
<td>W18000003</td>
<td></td>
</tr>
<tr>
<td>W19 National Assembly Economic Regions</td>
<td>4</td>
<td>W19000001</td>
<td>W19000004</td>
<td></td>
</tr>
<tr>
<td>W20 Registration Districts</td>
<td>25</td>
<td>W20000001</td>
<td>W20000039</td>
<td></td>
</tr>
<tr>
<td>W21 Registration Sub-district</td>
<td>26</td>
<td>W21000001</td>
<td>W21000043</td>
<td></td>
</tr>
<tr>
<td>W22 Travel to Work Areas 2007</td>
<td>20</td>
<td>W22000001</td>
<td>W22000020</td>
<td></td>
</tr>
<tr>
<td>W23 Spatial Plan Areas</td>
<td>6</td>
<td>W23000001</td>
<td>W23000006</td>
<td></td>
</tr>
<tr>
<td>W24 Spatial Plan Sub-areas</td>
<td>3</td>
<td>W24000001</td>
<td>W24000003</td>
<td></td>
</tr>
<tr>
<td>W25 Fire and Rescue Authorities</td>
<td>3</td>
<td>W25000001</td>
<td>W25000003</td>
<td></td>
</tr>
<tr>
<td>W26 Strategic Regeneration Areas</td>
<td>7</td>
<td>W26000001</td>
<td>W26000007</td>
<td></td>
</tr>
<tr>
<td>W27 Strategic Regeneration Sub Areas</td>
<td>2</td>
<td>W27000001</td>
<td>W27000002</td>
<td></td>
</tr>
<tr>
<td>W28 Transport Consortia Areas</td>
<td>4</td>
<td>W28000001</td>
<td>W28000004</td>
<td></td>
</tr>
<tr>
<td>W29 Agricultural Regions</td>
<td>7</td>
<td>W29000001</td>
<td>W29000007</td>
<td></td>
</tr>
<tr>
<td>W30 Agricultural Small Areas</td>
<td>235</td>
<td>W30000001</td>
<td>W30000235</td>
<td></td>
</tr>
<tr>
<td>W31 Non-National Park Area</td>
<td>1</td>
<td>W31000001</td>
<td>W31000001</td>
<td></td>
</tr>
<tr>
<td>W32 Non-Strategic Regeneration Area</td>
<td>1</td>
<td>W32000001</td>
<td>W32000001</td>
<td></td>
</tr>
<tr>
<td>W33 Communities First Areas</td>
<td>153</td>
<td>W33000001</td>
<td>W33000153</td>
<td></td>
</tr>
<tr>
<td>W34 Non-Communities First Areas</td>
<td>1</td>
<td>W34000001</td>
<td>W34000001</td>
<td></td>
</tr>
<tr>
<td>W35 Footprint Regions for Public Service Collaboration</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W39 Merged 2011 Census Wards</td>
<td>868</td>
<td>W39000001</td>
<td>W39000686</td>
<td></td>
</tr>
<tr>
<td>W40 Merged 2011 Census Unitary Authorities</td>
<td>22</td>
<td>W40000001</td>
<td>W40000022</td>
<td></td>
</tr>
</tbody>
</table>

**Scotland**

<table>
<thead>
<tr>
<th>Scotland</th>
<th>SASPAC Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>S92 Country</td>
<td>1</td>
</tr>
<tr>
<td>S00 Output Areas</td>
<td>42,604</td>
</tr>
<tr>
<td>S01 Data Zones</td>
<td>6,505</td>
</tr>
<tr>
<td>S02 Intermediate Geography</td>
<td>1,235</td>
</tr>
<tr>
<td>S03 Community Health Partnerships</td>
<td>34</td>
</tr>
<tr>
<td>S04 Regeneration Outcome Agreement Areas – Scotland</td>
<td>1</td>
</tr>
<tr>
<td>S05 Regeneration Outcome Agreement Areas – Community Planning Partnerships</td>
<td>28</td>
</tr>
<tr>
<td>S06 Regeneration Outcome Agreement Areas – Local Areas</td>
<td>180</td>
</tr>
<tr>
<td>S07 Regional Transport Partnerships</td>
<td>7</td>
</tr>
<tr>
<td>S08 Health Board areas</td>
<td>14</td>
</tr>
<tr>
<td>S09 Enterprise Regions</td>
<td>6</td>
</tr>
<tr>
<td>S10 Urban Regeneration Companies</td>
<td>6</td>
</tr>
<tr>
<td>S11 Strategic Development Plan Areas</td>
<td>4</td>
</tr>
<tr>
<td>S12 Council Areas</td>
<td>32</td>
</tr>
</tbody>
</table>
### SASPAC Training

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>S13</td>
<td>Electoral Wards</td>
<td>353</td>
</tr>
<tr>
<td></td>
<td>Westminster Parliamentary</td>
<td>59</td>
</tr>
<tr>
<td>S15</td>
<td>European Electoral Regions</td>
<td>1</td>
</tr>
<tr>
<td>S16</td>
<td>Scottish Parliamentary Constituencies</td>
<td>73</td>
</tr>
<tr>
<td>S17</td>
<td>Scottish Parliamentary Regions</td>
<td>8</td>
</tr>
<tr>
<td>S18</td>
<td>Registration Districts</td>
<td>n/a</td>
</tr>
<tr>
<td>S19</td>
<td>Locality</td>
<td>n/a</td>
</tr>
<tr>
<td>S20</td>
<td>Settlement</td>
<td>n/a</td>
</tr>
<tr>
<td>S21</td>
<td>National Parks</td>
<td>2</td>
</tr>
<tr>
<td>S22</td>
<td>Travel to Work Areas 2007</td>
<td>46</td>
</tr>
<tr>
<td>S23</td>
<td>Police and Fire Areas</td>
<td>8</td>
</tr>
<tr>
<td>S24</td>
<td>Highlands and Islands Enterprise</td>
<td>8</td>
</tr>
<tr>
<td>S25</td>
<td>Community Justice Authorities</td>
<td>8</td>
</tr>
<tr>
<td>S26</td>
<td>Immigration</td>
<td>n/a</td>
</tr>
<tr>
<td>S27</td>
<td>Land Use</td>
<td>n/a</td>
</tr>
<tr>
<td>S28</td>
<td>Transport</td>
<td>n/a</td>
</tr>
<tr>
<td>S29</td>
<td>Environmental</td>
<td>n/a</td>
</tr>
<tr>
<td>S30</td>
<td>Other</td>
<td>n/a</td>
</tr>
</tbody>
</table>

### Northern Ireland

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>N92</td>
<td>Country</td>
<td>1</td>
</tr>
<tr>
<td>N06</td>
<td>Westminster Parliamentary Constituencies</td>
<td>18</td>
</tr>
<tr>
<td>N07</td>
<td>European Electoral Regions</td>
<td>1</td>
</tr>
</tbody>
</table>

### Cross Border

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>K01</td>
<td>Travel to Work Areas 2007</td>
<td>8</td>
</tr>
<tr>
<td>K02</td>
<td>United Kingdom</td>
<td>1</td>
</tr>
<tr>
<td>K03</td>
<td>Great Britain</td>
<td>1</td>
</tr>
<tr>
<td>K04</td>
<td>England and Wales</td>
<td>1</td>
</tr>
</tbody>
</table>

### Channel Islands

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>L93</td>
<td>British Crown Dependencies</td>
<td>1</td>
</tr>
<tr>
<td>L00</td>
<td>Strategic Health Authorities</td>
<td>3</td>
</tr>
</tbody>
</table>

### Isle of Man

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>M83</td>
<td>British Crown Dependency</td>
<td>1</td>
</tr>
<tr>
<td>M00</td>
<td>Strategic Health Authorities</td>
<td>1</td>
</tr>
<tr>
<td>M01</td>
<td>Primary Healthcare Directorate</td>
<td>1</td>
</tr>
</tbody>
</table>
## Annex 5  Output geographies 2011 Census

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

Selection of Variables
In the main task window - whatever the task - there will be a Windows Command button that allows access to various searching mechanisms for identification of variables required by users. This also enables creation of the variable identities through on-screen identification. Use of this button leads to a window with a drop down menu and four buttons leading to different Property Sheets as shown in Figure 8.

![Figure 8: Selection of dataset to be searched](image)

Since the datasets that can be accessed through SASPAC contain different data and have different cell identifiers, the drop down menu is used to tell SASPAC which dataset is being accessed. This is done by the user clicking on the required one.

The four buttons on the 'Select Variables' window are:

**Keywords:** SASPAC contains a pre-defined dictionary of the more common keywords or topics found in the Census datasets. This dictionary may be searched when this Tab (which is the default) is selected. One or more items may be selected in the left-hand 'Keywords' window, and transferred to the right-hand 'Keywords' window through use of the 'Add>>' button. Once the keywords are transferred, the bottom right-hand text box will contain a list of the tables which, according to the dictionary, refer to that keyword. Unnecessary keywords may be removed through use of the '<<Remove' button, following highlighting. Clicking on one of these table numbers will activate the 'Table n' Tab and display the table on screen.

**Text:** If a keyword is not found in the dictionary, or the dictionary references are incomplete, this Tab may be used to initiate a full text search of the table layouts to find occurrences of any text string. For example,
the word "mining" may be entered in the appropriate box to discover that 2001 Key Statistics table KS011 contains the word "mining".

Name/Number: If the approximate title of the table is known, this tab may be used to show a list of all tables in the selected dataset, from which the required table is selected. The selected table is then shown on screen.

Header Variables: In addition to the count variables that are available through the tables, there is a set of ‘Header variables’ as defined in CE-05 of the User Manual: accessed through the SASPAC Help file menu. If any of these are required for use in a task, this Tab can be used to access them.

Once the required table is identified, the layout on screen may then be used to identify the required variable(s). When the variables are selected, their full identities are written into the text box at the top of the window. Variable selection may be undertaken, either by a point and click method or by the ‘rubber-band’ method. In the latter, the top left hand corner of a bounding rectangle is chosen by clicking on it, and then by dragging to the bottom right hand corner. The following example shows the use of this.

Example

From the 2011 Census, for all local authorities in London print the total of usual residents born in Scotland, the total number of Urdu speaking residents aged 65 to 74 and aged 75 to 84, the total number of male employees who work 15 hours or less, and the number of Gypsy & Irish Travellers with very good health. Include the authority name in the output.

1. The first question to ask is - "What sort of task is this?" and the answer is that it is to produce variables as a printed report. The first stage is therefore to select the appropriate menu items for defining such a task. This is done as shown in the following image.

![Example Image]
Alternatively, instead of using the menu route, the user can make use of the button bar to select certain tasks. The ‘Print Variables’ button is shown in this image.

2. Following either of these selections, the main ‘Print Variables’ task window is opened. At this point, we can choose the input System Files, and use the ‘Search Tables’ button to select the required variables. There is no mandatory order in which these operations are done. Since we may not be certain of whether or not all the data are in the same System File, we will initially search for the variables by using the ‘Search Tables’ button. For the first variable - residents born in Scotland - we will use the ‘Keywords…’ button and as shown in the next screen capture, use the keyword ‘Country of Birth’.

We must first of all remember that as we are looking for ward level data, it is advisable, but not mandatory to look in the Detailed Characteristics Tables and must therefore select that dataset from the drop-down list. The dictionary now indicates that tables 79 tables/pages include this keyword. Note that tables that span more than one page have multiple entries in this list.

3. Clicking on the first table number in the right hand box (DC2103EW) and closing the ‘Search Tables Using Keywords’ box, will open the table as a layout. Use of the ‘Zoom in’ button twice will bring up the window shown here.
This shows that the variable identifying the number of persons born in Scotland is the sixth in the table and clicking on this produces a red tick, and the variable identity is written to the text box at the top of the window. Note that clicking again on the cell, will remove both the red tick and the entry in the text box. In addition, the entry in the text box may be deleted manually, which also removes the tick.

4. The next two variables required are found by a similar method using ‘Age’ and ‘Language’ as the keywords (remembering to remove ‘Country of Birth’ from the right hand ‘Keywords’ box) which will bring up Table DC2104EW which has 9 pages.

Firstly, to locate the desired variable, the ‘Zoom in’ button may again be used. Using the scroll bars to move to the part of the table containing the appropriate counts will produce a window that looks similar to that shown below.

By scrolling and zooming, we have lost all the useful information surrounding the table, and selection of the correct counts or variables is made very difficult.
5. To overcome this problem, we right click (i.e. click the right mouse button) anywhere in the body of the table shown above. This has the effect of replicating (on a black background), the text that appears directly above and to the left of the cursor. This is shown in the next screen image.

6. The presence of the row and column headings now allows us to select the appropriate counts. As the two that are required are adjacent, we can use the ‘rubber-banding’ technique that makes use of the right mouse button again. First the left button is clicked and a drag operation is used to draw a rectangle around the required cells. The right mouse button is then used to bring up a menu as shown below.
7. Use of the 'Paste Selections & Clear' will transfer the contents of the 'rubber-band' to the text box at the top of the window. The selected counts will be shown in the window with a red tick. As with individually selected variables, these may be deselected by clicking on them, using the ‘Undo’ button or by editing the text box.

8. The procedure is then repeated to select the next variable, using “Hours worked” as the keyword to identify its code. Remember again that before each search is undertaken, the keyword currently in the right hand box has to be removed, and the other variables then selected. This provides table DC6302EW.

Note that the first page of this table is concerned with all usual residents. We are interested in males only so we need to select the next page (Page 2 of 3). This can be done by clicking on the next entry in the Search Tables selection box or by using the right page select arrow at the top and in the centre of the Select Variables window. Once you have navigated to the correct page you can find the variable you need and select it in the usual way.

An alternative to searching by keyword which is especially useful when the appropriate keyword is not known, is to use the ‘Text…’ button in the ‘Select Variables’ window as is shown in the following image. It can be seen that this window allows for multiple searches on the text contained within the tables, but it is usually best to keep the searching criteria as simple as possible.
After searching the tables for the text ‘gypsy’ and ‘health’ a set of tables containing this text is offered to the user. The appropriate variable is now selected from table DC2301EW.

The five variables which we have identified are DC2103EW0006, DC2104EW0152, DC2104EW0175, DC6302EW0113, DC2301EW0146 and having selected these, we could return to select the appropriate input System Files. To return to the ‘Print Variables’ main task window, we use the ‘Save & Close’ button below the Variable(s) selected window.

9. However, we are also asked to output the area name, and this is obtained through use of the ‘Header Variables…’ button in the ‘Print Variables’ window, as shown below. This window may be accessed again even though we have returned to the main task window, and this shows that there is no pre-defined order in which the windows/facilities may be accessed. On the left of this window can be seen all the header variables that are available with printing variables, and the one that we require is ‘ZLABEL’. This is selected in the left hand box, and then the ‘Paste Selected’ button is used to transfer it to the variable(s) selected box. Note that whenever the header variable is selected, it will appear at the beginning of the list.

![Select Header Variables](image)

10. The ‘Print Variables’ main task window presents the user with up to three Command Buttons linked to directories as defined in the user’s configuration. If the required system file is not in one of these directories, then the standard Windows browse facilities are available through the ‘Other Files…’ Command Button. In this example, the files we want are in the second default directory, and using the appropriate button opens a standard Windows dialogue box as shown in the next screen image.
We can now select and Open the System file containing the dataset of interest (2011 Detailed Characteristics tables) for the geography of interest, which in this case is local authorities in London.

11. The command file is now complete and we are ready to run the task. This is achieved by using the 'OK' Command Button, which generates the 'Task' window. In this window, the user can either run the command file under the default name of SASPAC.CMD, or allocate a new name under which the command file will be saved. If a user-defined name is given to the command file, the output report or print file will default to this name as well, unless the report file is forced to a user-defined name as well.

Providing a user-defined name is good practice as this will obviate the possibility of overwriting an existing command file, and also allows for it to be used later as a template which can be edited.

12. When we are satisfied that all is as required, the 'OK' Command Button is used to run the task. Otherwise the 'Cancel' Command Button may be used to return to the 'Print Variables' window to make further selections or amendments. When the task is run, the 'Running Task' window echoes the command file that has been generated and reports any warning or error messages generated by the task.
14. The command file generated by the above sequence is:

```sas
input system file name = "C:\SASPAC\Training / System Files\2011 DETAILLED CHARACTERISTICS FOR LOCAL AUTHORITIES IN LONDON.SYS"
print variables ZLABEL DC2103EW0006 DC2104EW0152 / DC2104EW0175 DC6302EW0113 DC2301EW0146
output print file = / C:\SASPAC\REPORT\TRAINING_MODULE_1.PRN
end
finish
```

15. When the task is complete, the ‘Print Preview’ Command Button may be used to examine the output. The effect of this is to display the output as shown in the first image below. The ‘Print Preview’ option is also available through the ‘File’ menu option. Also available through this option is the ‘Print Export/Output Text File...’ facility. This would display the output as shown in the second image below. The difference between the two is that the first is a graphic (or picture) image that cannot be edited in any way, while the second is simply text which can be edited or ‘cut and pasted’ into another document.
16. The command file created in this task may, if required, be edited to produce different output. This depends upon the user having sufficient information to be able to provide the correct information without using the facilities of the menu system. The command file is opened for editing either through use of the ‘most recent files’ facility under the ‘File’ menu item, or through the ‘File/Open/Open Command File…’. After editing, the ‘File/Save’ menu option is used, the result of which is to present the user with the following window in which the most appropriate selection is made. If the command file has been edited, but no change made to the name of the output Print file, the user will be asked if they wish to overwrite an existing file. If the answer to this ‘No’, then the command file will be saved, but not run.

![Command File Window]

**Summary of task sequence**

1. **What sort of task is this?**
2. **Print Variables**
3. **Search Tables**
4. **find & select variables**
5. **Run Task**
6. **Input System File**
7. **Print Preview**

*make sure the system file contains the geographical areas and variables you are interested in*
Module 2

Selection of Tables
When selecting tables, the main task window has an area where the user is asked to identify the text file which is to be merged with the data file(s) to produce the appropriate tables. It is important that the framework file chosen matches the data type of the data file. For example, if the input system file contains Standard Table data, then the framework file chosen must be the Standard Tables File.

The standard procedure for selection of tables is similar to that for the selection of variables, with access to the searching facilities being made through a ‘Select Tables’ Command Button rather than a ‘Search Tables’ Command Button. When this button is used, the window offered to the user mirrors that presented while searching for variables, except there are only three buttons - the Header Variables button is not present.

However, there is one major difference between the two procedures. In variable selection, clicking the mouse button on a variable identifier automatically writes that code to the ‘Variable(s) selected’ text box. In table selection, a table identified as required is only written to the text box when the ‘Select Table’ Command Button is clicked in the ‘Table(s) Selected’ window.

**Example**

For all London Boroughs print the table which shows Proficiency in English and Age on arrival in the UK, and also the one which shows numbers of Buddhists by Sex and Age. Convert the tables to HTML format and input a table into Excel.

1. Again the question is “What sort of task is this?” to which the answer is that it is to produce tables as a printed report, and so the first stage is to identify this to SASPAC through the menu options as shown on the following screen.

As with the ‘Print Variables’ task discussed in Module 1, there is now a shortcut to the main ‘Print Tables’ task window, available through the Button Bar, and this is shown below.
2. In the subsequent 'Print Tables' window we can do three things - select the input system file(s), select the appropriate framework file, or select the required table(s). As with the 'Print Variables' task, there is no set order in which these options must be done, and it is sometimes appropriate to leave selection of the System Files until the tables have been identified. As we are looking at London Borough (or district) level data, the dataset that we shall use is the Detailed Characteristics, and so the appropriate framework file can be selected from the drop down 2011 list. If other data, 2001 for example, were required, selection would be made from the drop down 2001 list.

Note that when the framework file is selected, its full path name appears in the text box below the drop down list.

3. The 'Select Tables' Command Button can now be used to enter the searching options. The first action is to select the appropriate dataset to search, which in this case is the 2011 Detailed Characteristics.
As in selecting variables, keywords or text can be searched, and by using the former we can ascertain that table DC2803EW is the required table for proficiency in English and age at arrival. To find the required table for Female Buddhists, we shall use the 'Text …' button and enter the text 'Buddhist' in the 'INCLUDE' text box.

This shows us that there are several tables containing this word in the text, and if the first is selected (table DC2107EW) the table layout is offered to us. This shows that age is also a constituent of this table by that the first page (1 of 3) is concerned with all usual residents. Using the page select arrows at the top of the page to tab through pages 2 and 3 we can see the same layout but for Male and then Females.

When the 'Select Table' Command Button is used, the table identifier is written to the 'Table(s) selected' text box, and the table layout is stamped 'SELECTED'.

4. Once the tables required for printing have been identified, there still remains a requirement to identify the input system files. The file required is 2011 DETAILED CHARACTERISTICS FOR DISTRICTS IN LONDON.SYS, as the task is to print tables from the Detailed Characteristics Tables for all London Boroughs. After this file is selected, the task may be run.
Every time a new task is created through the appropriate menu/window sequence, the last stage is the ‘Task’ window as shown above. This window allows the user to allocate a Command File name other than the default of SASPAC.CMD. To do this, the ‘User Defined Name’ radio button is selected, and an appropriate name is then entered in the text box. There is no obligation to change from the default, but if there is a need to save any command files or output report files, then it is advisable that an appropriate name is allocated.

5. The command file generated by this sequence is:

```plaintext
input system file name = "C:\SASPAC\Training / System Files\2011 DETAILLED CHARACTERISTICS / FOR DISTRICTS IN LONDON.SYS"
input framework file = c:\saspac\frwork\dc11.fwk
print tables DC2803EW DC2107EW
output print file = C:\SASPAC\REPORT\TRAININGMODULE2.PRN
end
finish
```

6. After the task has run the ‘Print Preview’ Command Button may be used to examine the output before it is printed. This produces the output on screen as shown below.

The banner heading on this page indicates that it is page 66, which happens to be the last page. This consists of two tables printed for each of 33 areas. This highlights the potential dangers in using SASPAC without the ability to filter out on areas. The current exercise has produced output for every area on the input system file. If that file had contained 1,000 areas which is entirely possible, then the output produced would have consisted of 2,000 pages. Selecting areas is thus a very useful tool within SASPAC and the next module deals with various ways of doing this.
7. The print file generated by this task can, if required, be converted to a HTML format by use of the ‘File/Print Export/Output HTML’ file menu option as shown below.

Once this route is followed, the user is presented with a set of Windows dialogue boxes which ask first for the name of the print file to be converted, and then the name and location of the index HTML file to be generated by this process. Once these files are located, the software will automatically create the necessary files. These consist of an index file and a set of files presenting each page of the print file as a separate HTML file. The index file contains hyperlinks to each of the page files as shown below.
8. Both the index file and the page files can be opened in either a web browser or in Excel. If a page file is opened in the latter, then the user can manipulate the table layout and data as necessary.

**Summary of task sequence:**

*make sure the system file, framework file and tables selected relate to the same dataset*
Module 3a

Selection of Areas

- by Text Listing
Whatever the task, the user will always have an option within SASPAC to select areas, normally achieved through use of the ‘Select Areas’ Command Button on the right of the main task window. This will lead to a window as shown below, where there are four Tabs allowing selection of areas in different ways. This module covers use of the ‘Explore Zones’ Tab which is the default, while the next module will cover the ‘Use Mapshore’ Tab. Module 7 and 9 covers the two other options on this window.

Note: if a double geography file is input, this window will have slightly different tabs on display.

When the ‘Explore Zones’ Tab is used, SASPAC will examine any input System Files to identify what areas are contained on them. For this functionality to work there must therefore be an input System File. This is one of the few cases where the order in which options are undertaken is important.

To access the list of areas found on file, the ‘View Data’ Command Button is used. This initiates a procedure where the input System Files are opened by SASPAC and their headers examined for a list of area codes. These are then presented to the users in the large text box which covers most of the window. If the areas on the file are hierarchical in structure, that is, they follow the Census Offices coding structure (prior to the 2011 Census), the information will be presented to the user in a form which replicates Windows Explorer’s manner of showing files within directories. Areas can then be selected for inclusion or exclusion.

Note the three special categories of exclusion to the right of the text box. Also notice the ‘Make Command’ Command Button. This button occurs in many places and must be used to convert selections into commands. If this button is not used, in this case, a prompting window is displayed which requests the user to make the command. However, there are instances where this is not the case, and in these circumstances, selections can be lost if the user neglects to click the ‘Make Command’ Command Button.
Example: For the Output Areas of E00006994 to E00006996, E00007302, E00007303 and E00007632 in the London Borough of Enfield, print Table KS101EW.

1. As in the last module, the task is defined as producing tables as printed output, and the same procedure is use for identifying the Framework File, and the input System File. As we are asked to produce Table KS101EW for OAs, we must input a Census Key Statistics file for OAs, and must also use the appropriate KS framework file. The table number has been specifically given, and so there is no need to enter the ‘Select Tables’ facility - the table number can be entered directly in the text box in the ‘Print Tables’ main window. All this is shown in the following screen image.

2. If the ‘Select Areas’ Command Button is used, the window shown on the previous page is displayed to the user. When the ‘View Data’ Command Button is now used, the software will open the input System File and display its contents as shown below.

This initial window shows that the software has found that the highest level is the region level (E12000007 London). Clicking on the ‘+’ sign expands the geography tree to the level below which is sub-regional: Outer and Inner London. Expanding the tree again shows the local authorities within the sub-regions. Within each local authority the output areas are grouped within wards (see picture below) because the Administrative Geography radio button is selected. If the Statistical Geography button was selected the sub-local authority geography would be grouped within MSOAs and LSOASs.
In this case, when the folders, or area levels, are expanded, the display in the window will be as shown in the next screen image. The areas have been expanded down through the hierarchy from county down to output area. Sufficient information is now provided in the window for the required areas to be selected.

3. Using the scroll bar on the right of the text box, and closing unwanted area levels (or folders) allows the user control what is displayed in the box, and allows access to the codes to be selected as above. A zone can be selected by clicking on the tick box next to the zone code.

4. If the box next to a zone is grey this means the zone is not present on the selected System file and so may not be selected for output. Any box coloured green allows for a range of zones below a given hierarchy to be selected with a single click.

5. When the full list of required areas has been selected, the 'Make Command' button must be used to complete creation of the command. If the command is not created, the system will prompt the user to do so. After the 'Make Command' button is used, the 'Select Areas' window remains open to allow further selections to be made if necessary. The window is closed by use of the 'Close' Command Button, at which point the user is returned to the 'Print Tables' main window, and the task may now be run.

6. If at any stage before final execution of the task, the user is unsure of what selections and choices have been made, then the 'Undo Commands' Command Button, which is present on all windows, can be used to investigate the commands which have been generated up to that point. If the user finds that the wrong selection has been made then that command can be deleted. Alternatively, if a required selection has not been made, then the user can return to the option which would generate that selection or command.
6. The command file generated by the previous sequence is:

   input system file name = "C:\SASPAC\Training\2011/CENSUS\2011 CENSUS KEY STATISTICS FOR OAS/IN ENFIELD.LD SYS"
   input framework file = c:\saspac\framework\KS11.fwk
   include E00006994 E00006995 E00006996/E00007302 E00007303 E00007632
   print tables ks101EW
   output print file = /C:\SASPAC\REPORT\TRAINING_MODULE3A.PRN
   end
   finish

7. After the task has run the 'Print Preview' Command Button may be used to examine the output before it is printed. This produces the output on screen as shown below.

   ![Print Preview Image]

This command file only produces six pages of output - one table for each of the six areas. By using area filtering or selection, the user is able to drastically reduce the quantity of output that is produced.
Summary of task sequence:

* make sure the system file, framework file and tables selected relate to the same dataset
** remember to ‘make command’ after your zone selections
Module 3b
Selection of Areas
- Using Mapping
Version 5 of SASPAC added mapping capabilities to the existing functionality of the software following a link-up with Pebbleshore, the developers of the Mapshore mapping package.

This additional capability allows the mapping element to be used in two ways:
- As a front end browser allowing selection of areas from maps;
- As a generator of maps for presentation of results derived from SASPAC.

The ability to produce thematic maps and other graphical presentations of data is discussed in module 5, while this module covers the use of Mapshore in providing an easy-to-use method of identifying and selecting areas from on-screen maps thus obviating the need for large numbers of paper maps which over a period of years have usually become extremely dog-eared and tatty.

The base requirement for use of this facility is a file of digitised boundaries (at the appropriate level) in a valid format, and a raster map of the user’s area as a bitmap file. Putting these two together gives users the full power of on-screen identification of areas.

There is no set location for these files as far as SASPAC is concerned, but the installation process creates a sub-directory of SASPAC named MAPPING which could be used for the process. The SASPAC.INI file does not set any defaults for mapping directories, and the MAPSHORE.INI file holds as default the last used directory.

When selecting areas using the mapping facility, the ‘Use Mapshore’ Tab leads to a window with a ‘Run Mapshore’ Command Button. Use of this button activates Mapshore in a separate window. The user then makes use of the facilities within Mapshore before transferring control - and the necessary information - back to SASPAC.

**Example**

For the area in the vicinity of the junction of the A105 and the A110 in Enfield, obtain the total population (QS601EW0001), and the total number of economically active residents (QS601EW0002). Print Statistics with the output.

1. Again, the task is defined as producing variables as printed output, since the print variables procedure has the added parameter of allowing the production of statistics associated with the areas selected and printed. The task therefore becomes one of selecting areas in the vicinity of this a selected point, and printing the required variables for these areas, using the ‘With Statistics’ parameter (see image).
2. When the 'Select Areas' Command Button is then used, the 'Select Areas' window is presented at the default Tab ('Explore Zones'). To use Mapshore the 'Use Mapshore' Tab must be used, and then the 'Run Mapshore' Control Button is used to activate Mapshore.

3. This action opens the Mapshore application and presents a blank window to the user. To start the mapping process, a boundary file must be selected. This is done through use of the Boundary / Retrieve menu as indicated in the next screen image.

4. On use of this menu, a Windows Dialog box titled 'Retrieve Boundary File' is opened as shown below. The user then makes use of standard Windows dialog boxes to locate the required file, as shown below.
5. In this task, we are looking at small areas and so the boundary file we require is the Output Area file for Enfield (2011_Enfield_OAs.bdy), and opening this file will create a window as shown below. If files covering adjacent areas are required, then this may be done by repeating the above process, and the software will knit the two sets of boundaries together at the appropriate places.

6. This image on its own is not enough, without some spatial information which allows the user to identify precisely where the area of interest is located. This can be provided by a raster map selected through the 'Data / Retrieve Raster Background' menu as shown here.

This menu option allows the selection of raster images as shown in the next screen image. Files of type 'Raster Backgrounds' are shown by the dialog box. Currently the raster types recognised by Mapshore are BMP and TIF. In this case the file selected is SD80.bmp, and is an Ordnance Survey 1:50000 product. This is shown in the screen image on the following page.
7. In order to display the raster image in the correct position with reference to the boundary file, Mapshore must establish the coordinates of the bottom left and top right corners of the tile to be displayed. If the name of the file follows the Ordnance Survey’s convention for naming 5km tiles - as in this case -, the software can deduce these points, as shown in the window bearing the title ‘Raster Background Coordinates’. However, if these cannot be deduced from the file name, then the user must supply them in the appropriate text boxes.

8. When this procedure is followed, the raster map will be placed in the correct position on the screen relative to the boundary file. Other raster files may now be opened, and
these will also be placed in the correct position relative to the boundary file(s) and the existing raster map(s). The user may now zoom into a particular area of interest, or use the image shown to continue with the task.

9. There is more than one way of zooming within Mapshore, and in this case we will use the zoom by rectangle facility as provided by the sixth button at the top of the screen. This allows the user to draw a rectangle covering the precise area of interest as shown on the following screen.

10. Once the necessary image is obtained, the Zoom / Layering menu option may be used to bring up the following window. In this window adjustments may be made to the image to achieve a better display.

If necessary, the user can use this window to set the attributes of the open files at different levels of zoom. For example, the boundary lines could be set to a thickness of 1 at one zoom level, and at another level of zoom to a thickness of 3. Alternatively, the raster image could be invisible at the full size picture, but become visible after zooming in.
In this case, we will change the colour and thickness of the OA boundaries at this level. This is done by using the 'Properties' Command Button which allows access to a dialog box for changing some parameters and through this to a colour palette for changing display colours. This is as shown in the next screen image.

11. The result of this is shown below. This window also shows how the process of selecting areas for use within SASPAC is achieved. The Boundary / Selection menu item is used to access a menu defining various means of selecting areas - in this case we will choose the 'Select Individual Polygons' item.

12. Having selected this option, the user must then point and click in the required areas (in this case, output area). The codes of these areas are then written to the text box on the left of the screen, and the selected areas are shaded on the map as shown on the screen image on the next page. Note that the selections may be toggled on and off. That is, clicking in an already selected area will deselect that area, both on the map, and in the text listing. If the user is satisfied with the selections made, then the 'OK' button will be clicked - otherwise the 'Cancel' button will be used. If the latter action is undertaken, then the selections are cleared and the user may use another selection.
procedure if necessary. However, if the selection is accepted, and the 'OK' button is used, the system will await instructions as to what to do with the selection. In most cases, there will be a need, as in this instance, to transfer the information back to SASPAC.

13. This is achieved by use of the 'Boundary / Selection' menu, and selection of the ‘Export Selected Areas => SASPAC…’ item as shown below.

14. Mapshore's work is now complete, and control is transferred back to SASPAC at the point that it was left as shown below. The selected areas are now shown in the SASPAC window, and the user can use the selected areas as necessary.
In most cases all the selected areas will be required, and so the 'Add All=>' Command Button is used to transfer the area codes from the left box to the right box. If necessary, the 'Sort' Command Button can be used to sort the codes into alphabetic order. The window below shows the result of using the 'Add All=>' Command Button.

15. When the codes appear in the 'Zones Selected' box, all that remains to be done is for the appropriate commands to be created. This is done by using the 'Include Selected Zones' Command Button. Note that the 'Make Command' Command Button will remain greyed out throughout this operation as the use of the 'Include Selected Zones' Command Button automatically creates the required commands.

16. Using the 'Close' Command Button in this window will return to the user to the main task window. The final stage in this task is to ensure that the appropriate variable is output to the print file. If the variable identities are known, there is no need to use the 'Search Tables' Command Button to find the identity, as they may be input directly to the 'Selected Variable Name(s)' Text Box, as shown in this window. Notice ‘Print with Statistics’ is also selected.
Having input the appropriate System file (2011 Census Quick Statistics for OAs in Enfield), the task can now be run following the usual method of clicking on the ‘OK’ Command Button in this window, and then on the ‘OK’ command Button in the ‘Task’ window - having first allocated a user-defined name to the command file.

17. The command file generated by this sequence is as follows:

```
input system file name = "C:\SASPAC\Training\2011/Census\2011 CENSUS QUICK STATISTICS FOR OAs IN /ENFIELD.SYS"
include E00007226
include E00007223
include E00007222
include E00007227
include E00007221
include E00007220
include E00171152
include E00171153
include E00007218
include E00006978
include E00006979
include E00007232
include E00006982
include E00007233
include E00007231
print variables with labels QS601EW0002 QS601EW0001
output print file = c:\saspac\report\mod3b.prn
end
finish
```

The output as viewed through use of the 'Print Preview' option is shown below.
This form of the output where the 'With Statistics' option was checked gives the total for the listed areas for each variable.

**Summary of task sequence:**

1. What sort of task is this?
2. Print Variables
3. Select Areas - Run Mapshore
4. Retrieve boundary file
5. Search Tables
6. Export OAs to SASPAC
7. Select OA polygons
8. Retrieve Raster Background
9. find & select variables
10. Input System File
11. Run Task
12. Print Preview
Module 4a

Creating New Zones - Individual Area Selection
Areas of interest to the user may not be immediately available. However, SASPAC again provides the user with facilities for creating new areas from existing ones. This can be done in one of two ways:

by combining explicitly identified existing zones; or
by reference to a gazetteer containing allocations of existing zones to new areas.

This Module deals with the former case, while the latter is dealt with by Module 4b.

New zones may be created by adding together existing areas (which need not necessarily be at the same geographic level), or by subtracting an area from another area (the second area must be at a higher geographic level than the first). SASPAC also allows for scaling factors to be used to allocate proportions of existing areas to new zones. If a scaling factor is used, then it is applied equally to all the variables for the existing area.

Example

Create one new area as the aggregate of OAs E00006957 to E00006960, plus one third of E00007550, plus all of ward E05000210 (Town) less OA E00007631

1. As the task here is to create new zones, it is by definition one of creating a new System File. When new zones are created, SASPAC must first save all the aggregated information to a System File, which is then available for interrogation in exactly the same way as the files containing standard areas. However, the menu system has a specific entry for creating new zones, and it is this rather than the creating new system file path that is followed. The menu path for starting this is shown in the following image.

Alternatively, the ‘Create New Zones’ button may be used as shown here:
2. Use of either of these options will open the 'Create New Zones' main task window, as shown in the next image. This window is similar to those main task windows already seen in other modules. However, there is a major difference in the central area in that there are two Command Buttons which refer to the manner in which the new zones are to be created. In the current exercise we are concerned with creating new zones by explicitly identifying their codes, and to do this we select the 'Using Areas' Command Button.

3. When this button is used the user is presented with a window (as shown below) which contains three text boxes into which the following information is to be entered:
   - the identity of the new area - up to 12 standard alpha/numeric characters without embedded blanks;
   - the name of the new area - up to 32 alpha/numeric characters including embedded blanks if required;
   - the identities of the existing areas, which together will constitute the new area - the default operator is ‘+’, so if there is a need for subtraction, that must be identified by a ‘-’. 

![Image of Create New Zones window]

![Image of Using Areas window]
Creating New Zones - Individual Area Selection  
SASPAC Training

The first two boxes are completed by placing the cursor in the appropriate box and typing in the required information. The third box may be completed either by entering codes directly into the box, or through using the facilities discussed in Modules 3a and 3b by use of the ‘Select Areas’ Command Button. In this case we will use the text listing method, and so we must have selected input System Files first.

If no System Files have been specified when the ‘View Data’ Command Button is used, then the software offers the user the opportunity to select those files through a modified window as shown here.

![Input System File Selection Window](image)

4. This window is effectively the left hand portion of the standard main task window, and the usual procedure is followed to select the file(s) and if necessary, the method of reading those files. In this case as we are combining wards and OAs we must input two system files, one for wards and one for OAs, as shown below.

![Input System File Window](image)

Note also that as we are combining the two area levels, we must use the lowest common denominator in terms of the datasets, which means that both have to be Quick Statistics. We cannot mix the Quick Statistics for OAs with the Key Statistics for wards but could input Quick and Key Statistics for OAs (i.e. identically geography).
6. As two input files have been selected, we must choose whether the files contain the same data, or the same areas, so that the appropriate ‘READ IN…’ command may be generated. In this case, it is the data that is common to the two files, and so the ‘Same Data, Different Areas’ button is selected. When the input system files are selected, we can view the areas contained on them as shown here. The two files selected are shown in the same hierarchy. Use of the ‘+’ symbols to open the hierarchy will display the OAs and wards in Enfield.

The methods previously employed in Module 3a are now used to identify the required areas. When the full list is compiled, the ‘Copy Selected Zones’ Command Button is used to transfer the highlighted identities to the text box at the top of the window, and finally the ‘Cut & Paste’ Command Button is used to transfer these zone identities back to the ‘Using Areas’ window. This empties the text window and allows the user to identify another set of identities if required. In this case we only want to create the single new zone, and so we use the ‘Close’ Command Button to return to the ‘Using Areas’ window. This window now contains the selected zone identities, and we can edit this text box to create the new zone as defined, and as shown below.
Finally the 'Make Command' Command Button is used, followed by the 'Close' Command Button to return to the main task window. At this stage, as we are creating a new system file, we must allocate it a name and a label, at which stage we can run the task. Note that the 'Task' window only contains an option to rename the command file, as this time there is no output print file.

7. The command file generated by this sequence is:

```
input system file name = "C:\SASPAC\Training\2011/Census\2011 CENSUS QUICK STATISTICS FOR /WARDS IN ENFIELD.SYS"
input system file name = "C:\SASPAC\Training\2011/Census\2011 CENSUS QUICK STATISTICS FOR /OAS IN ENFIELD.SYS"
read in series
new zone id = Zone001 name = New Zone 1
USING AREAS E00006957 E00006958 E00006959
E00006960 E00007550*0.33 E05000210 -E00007631
output system file name = C:\SASPAC\SYSFILES\Training_Module_4a.SYS
label = Training Module 4a
end
finish
```

8. As the task only produces a system file output, there is no output that can be examined directly on screen. However, we can run tasks on this system file and we can use 'Tools / System File Details...' to provide some information on the new files contents. If this is done, then the details of the file will be shown as on the following screen image.

This screen image shows the result of such a check, and it shows that the file created contains the single zone, and 66 tables, as expected.
A summary of the task sequence:

What sort of task is this?  →  Create New zones  →  Using Areas  →  Define zones - Select Areas*

Run Task  ←  Output System name/label  ←  Define zones*  ←  Input System Files

Tools/System file details

*remember to ‘make command’
Module 4b

Creating New Zones

- Gazetteer Files
Sometimes, when new zones are to be created, their definition in terms of existing areas may already be available in a machine-readable format as a constitution file. Such a file may be introduced to SASPAC as a Gazetteer File - subject possibly to some editing.

A Gazetteer File is an ASCII text file in a user-defined format where each line of the file links an existing area (or part of an existing area) to a new area, through their zone identities. A typical Gazetteer File might look like the following:

```
E00007034    AREA01
E00007035    AREA02    0.25
E00007035    AREA01    0.75
E00007036    AREA01    0.33
E00007036    AREA02    0.33
E00007036    AREA03    0.33
```

This file indicates that the whole of E00007034 is to be allocated to new zone AREA01, a quarter of E00007035 goes to AREA02, with its remainder going to AREA01, and that E00007036 is to be equally divided between AREA01, AREA02, and AREA03.

To enable SASPAC to make use of a gazetteer file, the location of the three elements - existing zone ID, new zone ID, and scale factor (if present) - within the file, must be defined. This is done in the INPUT GAZETTEER FILE command, which takes the following form:

```
INPUT GAZETTEER FILE NAME = filename   /
EXISTING ZONE COLS a TO b/
SCALE FACTOR COLS c to d /
NEW ZONE COLS e to f
```

where a and b indicate the start and end column locations of the existing zone code, c and d the start and end column locations of the scale factor, and e and f the same for the new zone code.

Note that there is no fixed order for the elements on the gazetteer file, but the order within the Command File is mandatory, and must be as shown.

The scale factor need not be present in all cases - in fact it may not be present at all - but if it is to be used, then its location must be specified in the command. If it is used, and it is not present on a line, SASPAC assumes that all of the area is to be included.

The existing zone and new zone fields may overlap, which allows users to make use of the hierarchic code structure of the standard areas to aggregate to higher areas. However aggregation of 2001 Census Area Statistics and Standard Tables is not recommended as a means of achieving higher-level statistics due to the Small Cell Adjustments that are present in the data. It is preferable to obtain the higher level data direct from the ONS.

The Gazetteer File described above merely allocates a ZONEID to the new area and does not allocate a label or name to the new zone. There is a variant of the Gazetteer File which allows this. The presence of the words WITH LABELS following FILE in the gazetteer command
instructs SASPAC to expect the file to consist of two parts separated by a blank line. The second part of the file would be the allocation to new zones and would look exactly the same as that already described. The first part would consist of a line for each new zone, where the line would have a new zone identity followed by a new zone label or name. Thus the Gazetteer file described previously would appear as:

    AREA01 This is area 1
    AREA02 This is area 2
    AREA03 This is area 3

    E00007034    AREA01
    E00007035    AREA02    0.25
    E00007035    AREA01    0.75
    E00007036    AREA01    0.33
    E00007036    AREA02    0.33
    E00007036    AREA03    0.33

if it were to be read as a 'WITH LABELS' file.

**Example**

Create new areas using the gazetteer file ENFIELD.GAZ.

1. Again the task is to create new zones or areas, and so the initial part of the task creation is as in Module 4a. This leads us again to the following screen, where this time the 'Gazetteer Files' Command Button is used.

   ![](image)

   On use of this button the user is presented with a window where the various parameters of the gazetteer file have to be entered and these (if not already known) have to be found from the gazetteer file itself. This may be done within SASPAC itself through use of the 'Open / Open Gazetteer File' option, but only before the start of the task creation procedure. Once the procedure is commenced, the 'Open' menu becomes unavailable. In this case, the file may be
opened in Textpad, Notepad, or other such text editor, as available. If we open this gazetteer file in such a text editor, we see that its layout is as follows:

ward01 Bush Hill Park
ward02 Winchmore Hill
ward03 Enfield Town
ward04 Ponders End

E00007208 ward01 0.33
E00007208 ward02 0.67
E00007209 ward01
E00007210 ward03
E00007211 ward02 0.5
E00007211 ward03 0.5
E05000200 ward04

2. The blank line following the assignment of labels to the new ward codes indicates that this is a gazetteer file with labels. It also shows that scaling factors are used and that they appear in columns 21 to 24. The existing zones are to be found in columns 1 to 10, while the new zones are in columns 13 to 18.

Note also that the existing zones consist of both wards and OAs, and so both levels must be present on the input system files. With this information now available, we can proceed to the next stage of the procedure which produces the following window.

3. In this window the parameters discussed above may be entered. Additionally, there are a few check boxes available which should be considered.

Zone Echo On: If this is checked, SASPAC will provide an expanded list of the allocation of existing areas (including scale factors) to new areas. This will appear in the Log file.

Zones Hierarchic On: If the new areas are to nest into a hierarchic structure, then this box should be checked.
With Labels: This box must be checked if the input gazetteer files has zone labels.

Following entry of the parameters in this window, including location of the gazetteer file, as shown below, the ‘Close’ Command Button is use to return to the main task window where the appropriate input system files are identified, as well as the location of the output system file.

4. The standard procedure is then followed for running the task, including allocating a user-defined name to the generated command file, the contents of which are as follows:

```
Input gazetteer file with labels name = /C:\SASPAC\COMMAND\Enfield.gaz existing zone cols 1 to 9 scale factor cols 18 to 21 new zone cols 11 to 16
input system file name = "C:\SASPAC\Training\2011 Census\2011 CENSUS QUICK STATISTICS FOR OAS IN ENFIELD.SYS"
set zone echo on
output system file name = /C:\SASPAC\SYSFILES\TRAINING_MODULE_4B.SYS
label = Training Module 4B Exercise
end finish
```

The effect of using the ‘Zone Echo On’ parameter is shown in the following screen image of the Log File.

As with the previous example, there is no directly visible output from this task, and the only options are to run a task on the system file, or to investigate the system file’ details.
Summary of the task sequence

What sort of task is this? → Create New Zones → Using a Gazetteer → Investigate .gaz file*

Output System File name/label → Input System File → Input column parameters → Input .gaz file

Run Task

*.gaz file can be viewed within SASPAC (Open Gazetteer file) or in a text editor package like notepad, textpad or Word
Module 5

Output of Information from SASPAC
SASPAC’s primary function is as a tool to provide quick and easy access to the very large datasets and allows small sub-sets of areas and variables to be identified for further manipulation and analysis. This further work may be undertaken within SASPAC or via some other item of software. As SASPAC was designed with limited analysis capabilities, the link to other analytic software is important.

In many cases, it will be sufficient for the user to obtain the results of the SASPAC interrogation as printed output, or as a simple text file. In such cases the Report File output will be used. Alternatively, if the requirement is to save a sub-set of variables and/or areas for future access via SASPAC, then the output is saved as a System File. Finally, if the intention is to undertake further analysis in some other item of software, then the output will be in the form of an Interface File.

We have already discussed output of Report and System Files in previous modules. In this module we will discuss Interface Files and how they may be used to export data into other software for analysis and presentation purposes.

Interface files are SASPAC’s method of allowing export of data to other software for further analysis. Generally, Interface Files will be output for:

- a selection of variables for all areas
- a selection of variables for a selection of areas.

Selection of areas and variables has already been discussed, and users should be aware that in addition to the standard count variables, the Header variables identified earlier may be also be output to Interface Files. Additionally, the ZONEID, which is automatically output to Print files and System files, is available for output, but is not automatically saved to Interface Files. Therefore, if there is a need for the ZONEID on the Interface File, it must be explicitly named on the SAVE VARIABLE command. If the command file is created through the Windows system then SASPAC will automatically prompt for the ZONEID if it is not present.

The interface file types available and their descriptions are:

- **FORMATTED** An ASCII text file, where each output record has exactly the same format, with the format of the records being defined by a WIDTH command.
- **DELIMITED** An ASCII text file, where each variable is separated from its neighbour by a delimiter - normally a comma.
- **CSV** An ASCII text file, where each variable is separated from its neighbour by a comma.
- **WKS** A binary file suitable for export of data to spreadsheet software such as LOTUS 1-2-3, and Microsoft Excel.
- **DBF** A binary file suitable for export of data to database software such as dBase.
- **DIF** Data Interchange File suitable for importing data to a variety of software.
- **POSTCODE** A file, used in conjunction with a postcode directory, which allows output of estimated counts at postcode unit level.
When WKS, DBF, CSV, and DIF files are exported from SASPAC, the variable names are automatically exported along with the values. With Delimited, and Formatted files this must be specifically requested by the user.

Two other types of Interface files are available - TRANSFER and MATRIX - but the use of these is restricted to specific instances, and they are not covered in this Module or in the SASPAC Training Manual.

**Example**

Create a CSV file containing the percentage of the population aged 0-4 for all OAs in Enfield, and map the output. Export to Word as .jpg file and export in a .kmz format to view in Google Earth.

1. The task here is defined as one involving the export of data to an interface file. The exported data can be mapped using SASPAC's own mapping utility - Mapshore. The initial selection route for this is 'File / New Task / Export Data / CSV File'. Following this route will open the task window as shown below.

   ![Export CSV File Window](image)

   This window can also be reached through the Button Bar using the CSV button shown in the image below.

   ![CSV Button](image)

   Note that in this window, the 'Set Countycode On' checkbox is already ticked. This is because in most cases, the boundary files to which the data is to be mapped will contain the county code in the zone identifier. If the county code is not required this box can be unchecked.

2. This main task window is similar to others that have been used, and the facilities available through the Command Buttons are the same. In this case we wish to select variable giving the percentage of the population that is aged 0 to 4.
3. When exporting data in CSV format it is possible to choose the information that is displayed in the column headings. The ‘With metadata’ tick box adds the variable name to the column heading (e.g. All Usual Residents). The ‘With headers’ tick box adds the variable ID to the column heading (e.g. KS101EW0001). Both boxes are ticked as default.

The Command File generated is as follows:

```plaintext
input system file name = "C:\SASPAC\Training\2011 / Census\2011 CENSUS KEY STATISTICS FOR / OAS IN ENFIELD.SYS"
set countycode on
save ZONEID KS102EW0020
output CSV file with headers with metadata name = / C:\SASPAC\INTERFAC\Training_Module_5.CSV
end
finish
```

4. On completion of the task, the software detects that there is CSV file now available for use, and the ‘Go Mapping’ Command Button is now made active.
4. Use of the ‘Go Mapping’ Command Button opens the Mapshore application in the same way as was seen in Module 3b, and in the same way the first stage is then to open an appropriate boundary file - in this case the OAs for the London Borough of Enfield. When this is done, the CSV file may be opened using the ‘Data / Retrieve Area Data’ menu item as shown below.

When this option is used, the user is asked to identify the appropriate data file for mapping. Once identified, if there is only one variable in the file, the thematic map will be drawn. If there is more than one variable in the file, then the user will be asked to identify the variable to be mapped.

The thematic map will be drawn using default shading and number of ranges. These defaults will be the values used the last time that Mapshore was used to draw a thematic map. The result in this instance is shown below.

Once drawn the legend can be shown using the ‘Display / Zonal Key’ menu item, and this can then be used to manipulate the displayed shadings and the number of ranges.

5. To change a colour, it is clicked in the legend and a ‘Set Colour’ window is presented to the user as shown on the next page. In this window, the user can select the type of
shading to apply to the map. If non-solid shading is used, then the ‘Transparent’ checkbox becomes active. If this is chosen, then any underlying image will become visible.

To change the colour, then it is clicked in this window, and a palette of colours is offered to the user, as shown in the next image.

![Set Colour](image1)

![Color](image2)

A basic colour is selected, and if there is a need to refine this, then that can be done through use of the ‘Define Custom Colours>>’. This will then allow the user to create colours based on the mix of Hue, Saturation, and Luminosity, or as a unique combination of Red, Green, and Blue.

![Intervals](image3)

6. Numbers of ranges and type of ranges are set by using the appropriate buttons in the Zonal Key. If the ‘Interval’ button is used, then the window shown here will appear. As can be seen, the Number of intervals, and the Interval Type are two of the parameters that may be set by the user. When the appropriate parameters have been chosen, the map is redrawn through use of the ‘Redisplay’ button in the ‘Zonal Key’ window.

7. Once a suitable result is obtained, the map may be printed, or exported using the ‘File/Print...’ menu. This option allows a low or high resolution bitmap or jpeg image to
be exported as shown in the two windows below. Chose to save a ‘Hi-Res’ image in a jpg format with a suitable name.

This jpg file may then, for example, be pasted into a Word document as the basis of a report or brief note related to the data depicted in the map. In Word this would be done through use of the ‘Insert/Picture/From File…’ menu item.

8. Alternatively, the release of SASPAC v8.5 saw the introduction of a Google Earth export facility. This tool allows the final Mapshore image to be exported in a .KMZ format, which can be viewed using Google’s virtual globe browser ‘Google Earth’ (downloadable from: http://earth.google.com/).

9. Use of the “File|Export” menu will open an export map image dialogue box as shown below:
10. Choose to save the file in a Google Earth Geographic File (.KMZ) format and give the file a suitable name. This opens a copyright statement window (shown below) in which the appropriate copyright statement for both the mapped Census data and the boundary file may be added. Finally, we must choose to save the file in a vector or raster format, in this case a vector format will be used.

![Copyright Statement Window]

[Important: choosing to publish the .KMZ file in a vector format means other users will be able to extract and use the underlying geographic boundary data. It is important that you ensure this does not breach the terms and conditions of the licence agreement for these datasets. If you do not wish to make the geographic data available to 3rd parties, export the image in a raster format. If you are unsure of the licence restrictions, we recommend publishing in a raster format.]

11. Finally, the exported .KMZ file is opened by double clicking the file in Windows Explorer, using the left mouse button. This will launch the application and zoom to the position of the exported file on the Earth’s surface. Alternatively, the "File|Open" menu within the Google Earth application itself can be used to open the .KMZ file directly. Whichever method is chosen, the result should look similar to the image below:

![Google Earth Application]
Here the final image has been placed in the correct geographical position of the LB of Enfield on the Earth’s surface. Furthermore, the key and copyright statement are positioned in the top and bottom left corners of the map window, respectively. The menu to the left of the map window lists all the geographic features (OAs) individually. It is possible to display these, or not, using the tick-box next to the feature. Below this window is a slide bar that allows the transparency of the image to be adjusted from 0 to 100 percent. Finally, using the navigation bar in the top right-hand corner of the map window, it is possible to zoom in/out, rotate and adjust the tilt of the image. (To understand further some of the navigational aids and other datasets in Google Earth we recommend visiting: http://earth.google.com/tour.html).

**Summary of task sequence**

*remember to ‘make command’ after creating your new variable*
Module 6

Creating New Variables
Even though the Census Area Statistics and Standard Tables between them contain some 75,000 variables, it is often the case that the information required is not directly available. The user may require a variable that is the sum of two or more other variables, or more commonly, may need to know a percentage or proportion. SASPAC provides facilities for these to be calculated and subsequently used in the same way as variables which already exist in the dataset.

Calculations within SASPAC are undertaken using the standard mathematical operators such as ‘+’, ‘-’, ‘*’, and ‘/’. The use of these operators is facilitated by a set of on-screen buttons.

Example

For all Output Areas in the London Borough of Enfield, calculate the percentage of the population that is aged 11 to 13, and output, along with the total population, to a print file. Improve the column headings and change the report heading.

1. Again, the task is defined as producing variables as printed output, even though the variables required do not exist as such in the database. There is a need to create the variable which represents the required percentage. This is done through use of the ‘Manipulate Variables’ Command Button which appears on the main task window.

2. Use of this command button presents a window with three Tabs, and for this task, where we need to create a new variable, we need to use the third of these which bears the title ‘Create New’ as shown in the screen image.

3. In this window, the new variable is created by first allocating a descriptive name to it. This name can have a maximum of 12 characters which can only be numeric or
alphabet. No blanks or special characters may be used. We shall call it ‘age15to24’, and this name is typed into the text box. We need to create a command of the form:

\[
\text{age11to13} = \text{<mathematical calculation>}
\]

Assistance in creating this mathematical calculation can be provided through use of the ‘Search Tables’ Command Button on the right of the window, and the symbol buttons above the text box. The Command Button performs the same functions as that to be found elsewhere, such as on the ‘Print Variables’ main task window. Using it allows the user to select the required variables on-screen as shown below.

Note that to create the total aged 11 to 13, we need the single year of age data contained in tabel QS103EW. To calculate the percentage, we then need the total population from the same table. By selecting the variables in this order, they are made available in the correct order for creation of the new variable.

4. When the ‘Save & Close’ Command Button is used, the selected variables are copied into the ‘Manipulate Variables’ text box, where the appropriate mathematical symbols can be added as shown here.
Once the command is complete, the ‘Make Command’ Command Button must be used to create
the command, otherwise the command will be lost. Once this button is pressed, the text box is
cleared, and the user has the option to create another new variable. In this case we have created
the only new variable that we need, and so we could now use the ‘Close’ Command Button to
return to the ‘Print Variables’ main task window. However, one of the other tabs in this window
allows us to describe a variable by adding a label. This is done as shown in the following
window.

5. After making this command and closing the window, we must ensure that the
appropriate variables are listed for output, and that a suitable input system file has been
selected as shown in the image below. As we have described one of the variables, we
can use the ‘Print with Labels’ checkbox to make use of that label. The ‘OK’ Command
Button may be used to proceed to the ‘Task’ window.

6. In the ‘Task’ window, a user-defined name is allocated, but instead of using the ‘Save
and Run’ Radio Button, we will use ‘Save Only’, as there is a need to edit the generated
command file to allow for the non-division by zero clause in the task. This is achieved
by using the ‘File / Open / Open Command File’ menu option to open a dialog box in
which the required file is identified and opened. (Alternatively, the file may be opened
by accessing the ‘Latest Files’ list under the ‘File’ menu.). Whichever method is
employed, the window shown on the next page will be presented to the user.
This is a text box into which any editing may be typed directly. Note the presence of two consecutive forward slashes ‘/’ on the first and second lines. SASPAC has inserted these automatically as line continuation symbols. When SASPAC encounters a ‘/’ symbol at the end of a line it will automatically add the text on the next line to the first line.

In this case we have the need to divide by CS0010001 only if the value of CS0010001 is not zero. This means that the line which reads:

\[
\text{age11to13} = \frac{\text{QS103EW0013 + QS103EW0014 + QS103EW0015}}{\text{QS103EW0001}} \times 100.0
\]

must be edited to read:

\[
\text{If QS103EW0001 EQ 0 THEN age11to13 = 0 ELSE age11to13 = } \frac{\text{QS103EW0013 + QS103EW0014 + QS103EW0015}}{\text{QS103EW0001}} \times 100.0
\]

As a single line as shown in the next image. The user need not have any concerns as to the length of the resulting line. SASPAC will split it with a continuation symbol if necessary.
7. When this command file is run by using the ‘Save’ and ‘Close File and Run Task’ menu option, the output that is obtained is as shown below. If the value of CS0010001 is zero for an area, the value of the new variable is set to 0.00, as defined, rather than the missing value symbol of ‘-’. Also note that the variable that was not described has the default label at the head of the column.

![Output Table]

**User Note:** There may be a need to edit the label by adding in blanks at suitable places to ensure that the label splits appropriately over two lines as 12 by 12 characters.

**Summary of task sequence:**

- What sort of task is this? [Print Variables] [Manipulate Variables] [Create New]
- Save and edit .cmd file [Input System File] [Describe variable**] [find & select variables*]
- Run Task [Print Preview]

*remember to ‘make command’ after creating your new variable
**remember to ‘make command’ after describing your new variable
Module 7

Selection of Areas

- SELECT IF
In Modules 3a and 3b we discussed the selection of areas based on their area codes. In addition to selection based on the ZONEID of the required areas, SASPAC allows areas to be selected on the basis of criteria defined by the user and which apply to the value of selected variables. For example, selection may be restricted to all areas where the percentage of households living in local authority accommodation is greater than 50%.

The criteria for selection are constructed using the standard mathematical operators which are:

- GT (or >) Greater than
- GE Greater than or equal to
- LT (or <) Less than
- LE Less than or equal to
- EQ (or =) Equal to
- NE Not equal to

while complex conditions may be constructed using AND, OR, and NOT.

The criterion or criteria for selection is prefaced by the words SELECT IF, so that the format of the command is:

```
SELECT IF VAR1 GT 50
```

The functionality of radial searching is now available through the mapping environment and can be carried out using Google Maps (see Module 9) or the integrated Mapshore system.

**Example**

Output, to a print file, all OAs in the London Borough of Enfield, where the percentage of persons aged less than or equal to 15 is greater than or equal to 35%.

1. In this example, we are asked to create a new variable, and to output it to a print file only where a given criterion is satisfied. We must therefore define the task as one that outputs a report consisting of variables, and we must create a new variable to calculate the required percentage. The usual procedure is thus followed, where a main task window is defined by the 'File / New Task / Output Report / Print Variables' menu option, or by use of the appropriate button on the Button Bar.

The appropriate file for OAs in the London Borough of Enfield is opened, and the new variable is calculated using the 'Manipulate Variables' Command Button, as described in Module 4.
2. The 'Select Areas' Command Button is then used to access the 'Select Areas' window which appears as the default on the 'Explore Zones' Tab. Clicking on the 'Select If' Tab will open the window shown here.

![Select Areas Window](image)

This window is very similar to the one which allows creation of new variables, and it may be used in a similar manner. However, it must be noted that user created variables cannot be selected from any list, and their identities must be typed into this text box. The arithmetic operators may be selected from the buttons above the text box, and combination criteria can be constructed using the logical operators.

3. When the text 'perl15 >= 35' is entered into the text box, the 'Make Command' Command Button is used to ensure the command is constructed and written to the command file. The user then returns to the main task window (assuming that no further constraints on output are required), where the identity of the required output variables is entered in the appropriate text box, or selected via the 'Search Tables' Command Button. Once this is done the task may be run as before.

4. The command file generated by this sequence is:

```plaintext
input system file name = "C:\SASPAC\Training\System Files\2001 CENSUS AREA STATISTICS FOR OAS\IN ENFIELD.SYS"
perl15=(CS0010006+ CS0010036+ CS0010066+ CS0010101) / CS0010001*100.0
SELECT IF PERL15>=35
print variables perl15
output print file = /C:\SASPAC\REPORT\TRAINING_MODULE_7.PRN
end
finish
```
5.  The output from this task is shown below. By using the ‘Select IF’ clause, of the 800+ OAs in Enfield, only nine have been output as meeting the selection criterion.

Summary of task sequence

*remember to ‘make command’ after creating your new variable

**remember to ‘make command’ after defining the Select IF criteria
Module 8

Use of Double Geography datasets within SASPAC
Introduction

The ‘standard’ datasets supplied by the Census Offices are one-dimensional in a geographic sense in that each item of information is related to a single area. That is why they are occasionally referred to as single geography datasets. The CSV file as supplied consists of a series of records (or lines), each of which is a series of numbers associated with a single area code and (possibly) one area name. A part of the CSV file for table UV002 is shown here:

"00GANY0001", 238, 3, 70.00
"00GANY0002", 294, 5, 65.33
"00GANY0003", 280, 11, 25.66
"00GANY0004", 300, 12, 24.61
"00GANY0005", 271, 6, 45.09

On the other hand, the ‘special’ datasets supplied by the ONS are geographically two-dimensional as each item of data is linked to two areas. In the migration datasets, the two areas are the area of residence one year prior to the census, and the area of residence at census day, while in the workplace (or in Scotland, travel) datasets, the areas are area of residence, and area of workplace (both at census day). These datasets are referred to as double geography datasets. As with the single geography datasets, the CSV file supplied consists of a series of records each of which is a series of numbers associated with two area codes and (possibly) two area names. A part of the CSV file for table SMS301 is shown here:

“20UEGL0006”, “20UEGF0017”, 3, 3, 0, 0, 0, 0, 0, 3, 0, 0, 0, 0, 0, 0, 0
“20UEGL0006”, “20UEGF0019”, 3, 3, 0, 0, 0, 0, 0, 3, 3, 0, 0, 0
“20UEGL0006”, “20UEGG0006”, 3, 3, 0, 0, 0, 0, 3, 3, 0, 0, 0, 0
“20UEGL0006”, “20UEGG0008”, 3, 3, 0, 0, 0, 0, 3, 3, 0, 0, 0, 0
“20UEGL0006”, “20UEGJ0002”, 7, 4, 3, 0, 0, 0, 7, 4, 3, 0, 0, 0
“20UEGL0006”, “20UEGJ0003”, 33, 14, 19, 0, 0, 0, 33, 14, 19, 0, 0, 0

Apart from the fact that the double geography datasets have two geographic references per area, their record structure is similar to that of the single geography datasets, and SASPAC may be used to access them with a few minor modifications in the appearance of the windows and choices that are offered to the user.

For any single variable the dataset may be represented as a two-dimensional matrix – hence the alternative name matrix datasets – with the third dimension then being represented by the other variables.

The migration data is available as a United Kingdom dataset, but there are two datasets for workplace related movements. This is because in Scotland the question relating to place of work was extended to cover place of study as well. The two relevant questions in each of the areas was as follows:

England, Wales, and Northern Ireland:
Question 33. What is the address of the place where you work in your main job?
Question 34. How do you usually travel to work?
Scotland:

Question 10. What address do you travel to for your main job or course of study (including school)?

Question 11. How do you usually travel to your main place of work or study (including school)?

The ONS released a series of tables under three generic titles:

Matrices of journeys to work (Areas of residence in England, Wales, and NI)

Within SASPAC these tables are given the table identifier code SWS. There are 14 tables (plus one Northern Ireland variation). Seven of these – numbered SWS101 to SWS107 – are available at local authority district level; six – numbered SWS201 to SWS206 – are available at ward level; and one – numbered SWS301 – is available at Output Area level.

Matrices of journeys to place of work or place of study (Areas of residence in Scotland)

Within SASPAC these tables are given the table identifier code TVS. There are 14 tables. Seven of these – numbered TVS101 to TVS107 – are available at local authority district level; six – numbered TVS201 to TVS206 – are available at ward and postcode sector level; and one – numbered TVS301 – is available at Output Area level.

Matrices of migration moves

Within SASPAC these tables are given the table identifier code SMS. There are 16 tables (plus one Northern Ireland variation). Ten of these – numbered SMS101 to SMS110 – are available at local authority district level; five – numbered SMS201 to SMS205 – are available at ward level (and postcode sector level in Scotland); and one – numbered SMS301 – is available at Output Area level.

Within the United Kingdom there are some 220,000 Output Areas, and about 10,500 wards. The resulting matrices of area to area moves for each of these levels have the potential to be very large even though not all the combinations of origin zone and destination zone will actually produce a flow. The following table shows the number of origin zones, the number of destination zones, and the actual number of flows for a few of the datasets currently available.

<table>
<thead>
<tr>
<th>Table</th>
<th>Number of origin zones</th>
<th>Number of destination zones</th>
<th>Number of flows</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWS201</td>
<td>9,432</td>
<td>10,223</td>
<td>1,367,170</td>
</tr>
<tr>
<td>SMS201</td>
<td>10,610</td>
<td>10,608</td>
<td>851,317</td>
</tr>
<tr>
<td>SWS301</td>
<td>175,428</td>
<td>177,404</td>
<td>5,951,376</td>
</tr>
<tr>
<td>SMS301</td>
<td>221,468</td>
<td>221,817</td>
<td>1,971,282</td>
</tr>
</tbody>
</table>

As table SWS301 for a single pair of origins/destinations contains 36 cells, the full dataset for this table contains over 200 million cells. A consequence of this is that the data is extremely sparse with most of the cells having a value of zero and many others having very low values. The Small Cell Adjustment Methods (SCAM) applied by ONS and NISRA Census Offices to this data mean that a small count appearing in a table cell is adjusted – information on what constitutes a small cell count cannot be provided as this may compromise confidentiality protection – but
basically any value of 1 or 2 is adjusted to become either a 0, or a 3, in such a way that loss of information is minimized. Totals and subtotals in tables are calculated as the sum of the adjusted data so that all tables are internally additive; within tables, totals and subtotals are the sum of the adjusted constituent counts. Because of this SCAMming, it has been found that the OA level matrices are unreliable, and must be used with great caution. The use of OAs as building blocks is therefore not advised for this dataset.

As indicated, the datasets supplied by the Census Offices, and made available to SASPAC users as SASPAC System Files cover areas across the whole of the UK (or parts as applicable). This means that there will be flows on the files that users will have no interest in, and which will not be relevant to the user.

The full matrix can be divided into four component parts:

- Flows within an area of interest
- Flows into an area of interest from the rest of the country
- Flows from an area of interest to the rest of the country
- Flows totally external to the area of interest

and this is depicted schematically in the following diagram.

Most users will only be interested in the shaded area within this diagram, and therefore the first requirement will be to reduce the dataset by removing the unshaded area. The area of interest in this case is termed the Core Area. The two blocks shown as B and C are referred to as the Tails of the Dataset.
Creating a sub-set of the origin/destination datasets for zones within an area of interest (or Core Area).

For this exercise we shall take the Core Area as being the London Borough of Enfield, and we shall create the data for all Special Workplace Statistics Tables at ward level (2001 SWS2 WARDS TO WARDS.SYS).

The requirement is to create a new System File containing only data related to a core area, and so the New System File task window is accessed through use of the File/New Task/Create New System File option, at which point the appropriate national dataset is selected.

Once an origin/destination dataset is selected, the task window changes slightly in that there is an additional tick-box now available as shown in this image. This allows the user to restrict the core area data to only flows within the core area, and all flows with an external component will be excluded.

In terms of the previous diagram, this would mean that only the moves in block A would be included. Above this tick-box under the ‘Variables’ header it is possible to input selected variables to be saved, leaving this blank will automatically save all variables in the original system file.

The selection of the core area zones may now be initiated in the usual way through use of the ‘Select Areas’ command button on the ‘Create New System File’ main task window. As a double geography file has been selected, the window that is now presented to the user reflects this in that the property sheet has tabs titled ‘Origins’ and ‘Destinations’.

To view the zones that are on the input file, the ‘View Data’ command button is clicked, and a full list of the zones found on the input file will be shown. Note that all the area symbols will have a superimposed red ‘C’. This indicates that the zone is part of the core area for this file. As it is a national data file all zones are within the core area and will therefore be indicated as such by the red ‘C’
If the core area is to be symmetrical, that is the areas of origin within it are identical to the areas of destination, then the process of area selection may be simplified by clicking the ‘Apply to both axes’ tick-box. This ensures that the process of zone selection need only be done once. The first zone in the core area is selected by scrolling down the list and clicking on it followed by clicking on ‘Include’. The list is then scrolled to the last zone in the area of interest, and by shift-clicking on this zone, the full range may be selected. Note however, that this may take some time, depending on the power of your PC and the number of zones to be included, and it may be faster to simply click on this last zone for inclusion and to edit the command file created as indicated below.

[Note: An alternative to scrolling is to use the ‘Find’ button to access a search window into which the user enters a part of the code or name of the area that is to be used. In this case, as the code for Enfield is 00AK, we enter this in the text box, after which use of the ‘Find next’ button will move the selection to the first ward in Enfield – 00AKGL Bowes]

Once the zones within the core area (or the start and end zones of the range) have been identified, then the ‘Make Command’ command button is clicked to generate the appropriate command(s). The final stage is to allocate an appropriate name and label in the required locations in the task window which appear when the ‘Close’ command button is clicked. Once this is done, the ‘Close’ command button on the task window is used to enter the task submission window, at which point the user can allocate an appropriate command file name. Also, if only the end zones of the range were selected, rather than the range itself, then the ‘Save only’ radio button on this window must be checked to allow the created command file to be edited as follows:

Change the lines which are of the form:

\[ \text{incl ude } \text{<startzone_code>} \]
\[ \text{incl ude } \text{<endzone_code>} \]

To

\[ \text{incl ude } \text{<startzone_code> to <endzone_code>} \]

This command file can then be saved and run to produce a system file containing only the flows relevant to the core area.

If the Tools/System File Details menu item is now used on this new file, the user will see that only the wards in their area of interest are shown as being part of the Core Area. Other areas will appear as origins and/or destinations, but without the core area indicator, they are automatically parts of the tails of the dataset.
The System File Details display also shows that there are now 20,194 flows within the dataset arising from 2,581 origins and 1,665 destinations. The file is therefore much more manageable as is also shown by the fact that the original file of 611Mb has been reduced to less than 5.9Mb.

Re-zoning areas outside the core area

The file that has just been created will contain ward-to-ward flows related to the Enfield core area. As such it will contain, for example, flows with origins in Scottish wards and destinations in Enfield wards. It is however highly probable that the user will not be interested in origins or destinations at such a fine level of detail where the distance involved is very large. In this case, the user will want to create new zones for areas outside the core area.

The easiest form of external rezoning would see all the external wards aggregated to the local authority district level, although it would be possible to aggregate the more local wards to local authority level, those a little further away to county level, and those further still to region level.

To undertake a rezoning of all external wards to local authority districts, we shall use the gazetteer file ‘Ward_to_District.gaz’. This file allocates wards to local authority districts and also provides labels for these districts. It has the existing zone (wards) in columns 1 to 6, and the new zone (districts) in columns 15 to 18, with no scaling factor. To adapt this for use in Enfield, we need to remove all wards in Enfield from the assignment (or second) part of the gazetteer file, and also the Enfield line from the labelling (or first) element of the file. Once this is done we will save the gazetteer file as ‘Ward_to_District_Enfield.gaz’.

The first stages of the process are the same as for single geography rezoning, and the “File/New Task/Create New Zones...” route is followed, at which stage the system file created in the previous exercise is selected. The resultant window now differs from the window seen in Module 7 in that this is double geography specific. There is a block under ‘Method’ which specifies what is to be done with areas on the input system file that are not used in the new zone creation commands – either through Using Areas or Gazetteer Files. As we are not considering the Enfield wards in the gazetteer file and we want those to appear as they are in the new system file, the ‘Include unused origins and destinations’ button is left checked.

As we are going to use a gazetteer file, the ‘Gazetteer Files’ button is used to take us to the next window, which is again different from that seen in Module 7, as it again caters for double geography. The main difference is that this window has two tabs which allow for gazetteer files to be chosen either to apply to both axes together or individually. This is shown in the next
image. A minor difference is that as hierarchies are not allowed in double geographies, the ‘Zones Hierarchic On’ checkbox has disappeared.

The appropriate input gazetteer file is selected in the usual manner by use of the ‘Input from…’ button.

As noted earlier, the existing zone (wards) code is in columns 1 to 6, and the new zone (districts) code is in columns 15 to 18, with no scaling factor. These values are entered into the appropriate boxes. In addition, the gazetteer file has a label element, and so the ‘With Labels’ box must be checked. For the sake of obtaining a full check on the allocation of zones (wards) in the new zone creation, the ‘Zone Echo On’ box is also checked.

The final stage is to allocate a label and name to the output system file and then to run the rezoning task. However, there is currently an issue that must be addressed before this is done and so the file must be saved and closed using the ‘Save Only’ option.

As the gazetteer file has been generated from a listing of all ward files, and not all of these wards will appear as origins or destinations on the input system files. By definition, areas that appear on a gazetteer file are expected to appear on the input system file, and their absence will cause an error. This is overcome by use of a command that is used for testing the software, but in this case prevents the absence of the area from creating an error. This command is ‘Test 71’, and this is placed as the first command line in the command file.

A check on the System File details reveals that we now have a file in which there are 4,991 flows with 366 origin zones and 310 destination zones, and the file size has been further reduced from 5.9Mb to 1.6Mb.
Using the origin/destination datasets

Once the national dataset has been tailored to the user’s needs as detailed in the previous exercises, it may be used in virtually the same way as any other SASPAC System File. The main difference that the user will notice is that, once a double geography system file has been selected, some windows have a slightly different appearance, as has already been noted.

Example

For areas of origin and destination in London Borough of Enfield, extract the total numbers of persons, and males and females moving to work wholly within Enfield, and map the total persons within SASPAC.

This is a standard mapping task within SASPAC, with the difference that the dataset to be accessed is an origin/destination one. As there is a need to produce a CSV file for the mapping, the start of the task follows the usual File/New Task/Export data/CSV file route. This produces the standard main task window where the user must select the required input file first. Note that whereas with single geography analyses, the input system file may be selected at any time, with double geography, the input system file must be selected at the outset. This is because the functionality available changes slightly when double geography is used.

When the double geography file is selected, the amended window shown here will appear. Note that as soon as a double geography System File is selected, the ‘Input System Files’ buttons are greyed out. This is because only one input file is allowed with double geography analyses. Also note the addition of a check box that allows restriction of the analysis to flows wholly within the areas of interest. Checking this box leads to the insertion of the command ‘SET IN&OUT OFF’ in the Command File.

As with single geography, the ‘Select Areas’ command button may be used to select the required areas. However, as the input system file is double geography, radial searching is not an option, and in place of the ‘Explore Zones’ tab, there is a tab each for origin and destination. Additionally, instead of the check boxes for the various
exclusions associated with single geography, there is now a single check box labelled ‘Apply to both axes’. If this box is checked, selections made in the usual way through either the ‘View Data’ or ‘Use Mapshore’ options will apply to both origin and destination. Otherwise, selections will have to be made twice – once for each of the origin and destination tabs.

With the ‘Apply to both axes’ check-box checked, the areas within the core area may now be selected in the usual manner (remembering to use the ‘Make Command’ button). It is worth remembering that although the use of ‘Shift/click’ for selecting ranges works, sometimes in practice, depending on the size of the file, it can be a very slow procedure, and the most efficient way is to select the first area in the range, and then select the last area in the range. The commands created by this option may then be edited once the complete command file has been created, through use of the ‘Save only’ radio button in the final Task window.

The searching facilities for the double geography datasets can be used to search and select the required variables (NB. Remember to use the ward level SWS dataset to select these). Both the origin (‘ORIGID’) and destination (‘DESTID’) header variable codes will also need to be included at this point although, SASPAC will prompt the user to save these before running the task. Having done this, the output CSV file may be named and the task is run.

Using the ‘Go Mapping’ button in SASPAC will open Mapshore, and we then first open the appropriate boundary file – in this case Enfield wards – followed by the CSV file just created. As the CSV file is a double geography related file, the route to this is through the Data/Desire Line Double Geography option, as shown.

Use of this option will open the usual Windows dialog box where the user navigates to the appropriate drive/directory to find the required file. As there are multiple data items on this file that was created, the system asks which data item is to be mapped. If the variable representing
the total (SWS2010001) is selected, the map is populated with a large number of arrows which effectively obliterate the whole map.

Use of Zoom/Layering

Allows the user to control the presentation of the map. This is done by highlighting the Double Geography file, and then using the ‘Properties’ command button. This brings up the window shown here, in which the user can set arrow widths, arrow colours, and range values. Setting an arrow width to zero will prevent that range from displaying on a map.

Using the settings as shown in the adjacent window, produces a map as shown below. This map can then be zoomed to examine flow patterns in particular areas. Note that because boundaries are not available for areas outside Enfield, cross boundary flows would not be mapped even if they were present in the CSV file. The mapping depends on the availability of boundary files at the appropriate level.

As with any mapping within Mapshore, this can be annotated with text, and all the other mapping functionality is also available as discussed in Module 5.
Double Geography datasets

Additional output facilities available only with Double Geography datasets

In addition to the usual Print Tables and Print Variables options, there are three other Print options that are only available when a double geography dataset has been input. These are:

- Print Summary
- Print Netflow
- Print Matrix

Additionally, it should be noted that Print Profile and Print VarTables are not available with double geography datasets.

Each of the three additional options is initiated in exactly the same way as the other Print commands.

Print Summary
A Print Summary task will print the following for each of the variables selected for output:

- **Row Total**: The total number for the area as an area of origin, i.e. it is the sum of flows from that area of origin to all destinations.
- **Column Total**: The total number for the area as an area of destination, i.e. it is the sum of flows to that area of destination from all origins
- **Intra**: The total number of flows within the area, i.e. the area is the area of origin and destination
- **Row Max**: The value of the largest flow from the area as an area of origin
- **Row Count**: The total number of destinations with an origin in the area, and the number of these with a non-zero value for the selected variable
- **Column Max**: The value of the largest flow to the area as an area of destination
- **Column Count**: The total number of origins with a destination in the area, and the number of these with a non-zero value for the selected variable

Sample output from a Print Summary task is shown below.
Print Netflow

A Print Netflow task will print the following for each of the variables selected:

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Out flow</td>
<td>The value for the selected variable of the flow from the first area to the second area i.e. it is the value when the first area listed is considered as the origin, and the second area is considered as the destination</td>
</tr>
<tr>
<td>In flow</td>
<td>The value for the selected variable of the flow to the first area from the second area i.e. it is the value when the first area listed is considered as the destination, and the second area is considered as the origin</td>
</tr>
<tr>
<td>Ratio</td>
<td>The ratio of the Inflow to the Outflow</td>
</tr>
<tr>
<td>Net flow</td>
<td>The difference between the Inflow and the Outflow</td>
</tr>
</tbody>
</table>

Sample output from a Print Netflow task is shown below:

```
<table>
<thead>
<tr>
<th>Variable: SWS2010001</th>
<th>O to D</th>
<th>D to O</th>
<th>Outflow</th>
<th>Inflow</th>
<th>Ratio</th>
<th>Netflow</th>
</tr>
</thead>
<tbody>
<tr>
<td>00AKGL</td>
<td>00AKGM</td>
<td>12</td>
<td>23</td>
<td>0.52</td>
<td>-11</td>
<td></td>
</tr>
<tr>
<td>00AKGL</td>
<td>00AKGN</td>
<td>12</td>
<td>9</td>
<td>1.0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>00AKGL</td>
<td>00AKGP</td>
<td>41</td>
<td>24</td>
<td>1.0</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>00AKGL</td>
<td>00AKGQ</td>
<td>58</td>
<td>18</td>
<td>1.0</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>00AKGL</td>
<td>00AKGR</td>
<td>42</td>
<td>9</td>
<td>1.0</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>00AKGL</td>
<td>00AKGS</td>
<td>9</td>
<td>9</td>
<td>1.0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
```

Print Matrix

A Print Matrix task will print the matrix of moves between selected zones for the selected variables, with the areas of origin being the rows, and the areas of destination the columns.

Sample output from a Print Matrix task is shown below:

```
<table>
<thead>
<tr>
<th>Variable: SWS2010001</th>
<th>O0AKGL</th>
<th>O0AKGM</th>
<th>O0AKGN</th>
<th>O0AKGP</th>
<th>O0AKGQ</th>
<th>O0AKGR</th>
</tr>
</thead>
<tbody>
<tr>
<td>00AKGL</td>
<td>812</td>
<td>12</td>
<td>12</td>
<td>41</td>
<td>55</td>
<td>42</td>
</tr>
<tr>
<td>00AKGM</td>
<td>23</td>
<td>1207</td>
<td>49</td>
<td>72</td>
<td>106</td>
<td>121</td>
</tr>
<tr>
<td>00AKGN</td>
<td>9</td>
<td>53</td>
<td>1237</td>
<td>49</td>
<td>79</td>
<td>164</td>
</tr>
<tr>
<td>00AKGP</td>
<td>24</td>
<td>1</td>
<td>11</td>
<td>110</td>
<td>63</td>
<td>48</td>
</tr>
<tr>
<td>00AKGQ</td>
<td>18</td>
<td>36</td>
<td>35</td>
<td>27</td>
<td>984</td>
<td>126</td>
</tr>
<tr>
<td>00AKGR</td>
<td>9</td>
<td>43</td>
<td>89</td>
<td>35</td>
<td>152</td>
<td>1439</td>
</tr>
<tr>
<td>00AKGS</td>
<td>9</td>
<td>32</td>
<td>68</td>
<td>35</td>
<td>111</td>
<td>337</td>
</tr>
</tbody>
</table>
```

Note that the largest values are usually along the diagonal, as they represent the ‘within zone’ flows.
**Step 1** – Create a subset of the origin/destination datasets:

- What sort of task is this?
- Create a new System File
- Input System File
- Select Areas*
- System File details
- Run Task
- Output System File name/label

* remember to ‘Apply to both Axes’ as the selection is symmetrical and ‘make command’

**Step 2** – Re-zoning areas outside the core area:

- What sort of task is this?
- Create a New Zone
- Investigate & edit .gaz
- Input System File
- Output System File name/label
- Input .gaz file
- Input column parameters
- Using a Gazetteer
- Save and edit .cmd file*
- Run Task
- System File details

* insert double quotes around the .gaz file if there are any blanks in the directory or file name and the command ‘Test 71’

**Step 3** – Output from rezoned origin/destination file:
What sort of task is this?

Output CSV file

Input System File

Search & select variables

‘Go Mapping’

Run Task

Output file name

Select Areas*

Retrieve boundary file

Retrieve Desireline Area file

* remember to check the ‘Apply to both axes’ tick box to save making origin/destination selections separately
Module 9

Use of Google Maps within SASPAC
The advent of Web Mapping has made mapping technologies and geographic data freely available to users on the World Wide Web. Previously, these facilities would be restricted to companies and mapping agencies with access to expensive hard and software, large quantities of geographic information and skilled cartographers. However, the release of Web Mapping applications like Google and Yahoo Maps have made these technologies accessible to everyone and greatly increased the use of these datasets. Currently, mapping applications typically include maps of the road network, aerial photography and routing tools (see screen shot below).

A screen shot of Google Maps data for Central London depicting a route between two locations as a blue line, with directions describing the route in the panel to the left of the map. It is possible to zoom in/out, scroll across the map, display aerial photography and terrain datasets.

In addition to providing access to these applications via a web address, many mapping providers make their services available for reuse in other web pages by publishing their API (Application Programming Interface). Since the release of SASPAC v8.5 users can take advantage of these technologies and seamlessly use Google Maps within a SASPAC task. The Google Maps facility can be launched through the ‘Select Areas’ and ‘Radial Search’ function in SASPAC. This allows a user to search for a place name or postcode and visualise this on a map via www.google.co.uk. There are also a number of other SASPAC/Google Map utilities that are demonstrated through the following worked example.
Example

Output to a Print file the population resident within 1 mile of London Bridge rail station, the population aged 16 to 74 in employment and the number of these that travel to work on foot. Use the ‘Explore’ tool to find the appropriate System File and Google Maps to locate and visualise the search area. Choose to print the file with statistics.

1. In this example, we are asked to define a specific search area, and output a number of variables to a print file. We must therefore define the task as one that outputs a report consisting of variables, and select areas using a radial search command. The usual procedure is thus followed, where the main task window is defined as:

   ‘File / New Task / Output Report / Print Variable’.

2. In the subsequent ‘Print variables’ window there are a number of different options available to us. This example asks us to use the ‘Explore’ tool to find the appropriate System File, it is advisable to define the variables required and select the areas that are within 1 mile of the search location first.

3. In this case we are asked to find the population resident within 1 mile of London Bridge station and to do so we need to select the Census Output Areas that fall within this zone. This is done through ‘Select Areas’ and the ‘Radial Search’ tab within the Print Variables task window.
4. The ‘Radial Search’ tab allows users to select areas selected based upon a defined distance from a fixed point or ‘search centroid’. This point can be defined by directly inputting, if known, an Ordnance Survey grid reference, or Easting and Northing, in the appropriate format. Alternatively, if this information is not known, a place name or UK postcode can be typed into the text box and Google Maps used to define a location.

5. This example asks us to select areas that are within 1 mile of London Bridge rail station and as we do not know the coordinates of this position, the first stage is to define the location of this point. To do this we type the text “london bridge station” into the text box and select ‘Go Google’

6. This opens an internet browser window, using Internet Explorer, with a red placemark over the approximate location of London Bridge station. In this case, the placemark is slightly southeast of the station itself, so to correct this we centre the map on the station by positioning the cursor over the station symbol and right-click the mouse button. Of the five options that are presented, choosing to ‘centre map here’ will shift the centre of the map to the cursor position.
7. Now that we are happy with the map position we can send this location back to SASPAC using the ‘Link to this page’ tool above the top right-hand corner of the map. This opens a ‘Paste link in email or IM’ window and by right-clicking the mouse on the highlighted text (or ‘CTRL+C’ on your keyboard), the location will be copied to your computer’s clipboard.

Now we can return to SASPAC and select the ‘Extract Grid Reference from Google Data’ button to paste the geographical position into SASPAC. Notice that the latitude and longitude have been converted to the equivalent Ordnance Survey grid and easting/northing references.

8. The next step is to define the radial search diameter; in this case we set the units to ‘miles’ and type ‘1’ into the ‘Search inside circle’ text box. Having done this the ‘View/Edit in Google Map’ button becomes available. It is recommended that you always check the location of the radial search on map before continuing. Selecting this button will open the following window:
NB. Depending on the PCs internet security settings you may have to allow any blocked content to be displayed using the information toolbar above the web page.

This window correctly shows the radial search area as a circle with a 1 mile radius and a centre point positioned at London Bridge rail station.

9. If necessary, it is possible to alter the position of the placemark and the radius of the circle following the instructions above the map window. We can also zoom in/out and choose to display aerial photography if desired, as shown below.

10. Now that we are happy with our search parameters we select the ‘Save & Close’ button, to the right of the map, closing this window and returning to SASPAC. This returns us to the ‘Select Areas’ window where we can choose to ‘Make Command’ and save these parameters.

11. This returns us to the main task window where we can use the ‘Explore’ button to search for and select the variables required. To do this we use the ‘Explore’ and produce the following window:
12. The purpose of this window is to help us find the System File(s) that contain the data we are interested in. This may be defined as a series of variables with or without radial search parameters.

13. The variables of interest can be searched for in the usual manner under the ‘WHAT to explore for’ heading. If the variable codes are known, these can be typed directly into the text box. Alternatively, the ‘Search Tables’ button can be used to launch the searching ‘Select variables’ window and select the desired variables.

14. Now that we have defined the variables of interest, we specify where we would like SASPAC to search for this data under the ‘WHERE to explore’ menu.
The default locations in this case are the standard System File directories (SYSFILES, SYS2 and SYS3) but it is also possible to ‘Browse’ to a folder of choice or search a CD/DVD. In this case, the default search location of ‘System Files #1’ is suitable. Before we run the search, note that because we have already defined radial search settings, the ‘WITHIN Geographical Area’ box is checked and our parameters are displayed. To run the search we select the ‘Explore System Files’ button at the bottom of this window.

15. After a short wait, the results will appear as a list of System files that contain some or all of the variables requested. To display the details of the files shown left-click the tick box next to the path name and file description is displayed in the window below.

NB. The only datasets to contain Ordnance Survey grid references from the 2001 Census are those at Output Area level. Therefore, when carrying out a radial search task, Output Area data must be used.

16. In this case only one System file should be found to contain all the variables requested and selecting the ‘View results on Map’ button will visualise these radial search results on a Google Map:
The above map shows a red disc representing a 1 mile circle around London Bridge station. Red placemarks appear at the (population) centres of each Output Area that fall within the search radius. The OA codes are also listed to the right of the map and are dynamically linked to the placemarks, so when the zone ids are selected the associated placemark is identified.

NB. A radial search query includes an Output Area if the population’s centre (centroid) is contained within the specified distance. For example, if only 25% of an Output Area is within a search radius but that happens to include the population centre, then that OA, and 100% of its population, is considered to be within that search area. If you would like to carry out more complex zone selections, i.e. zones ‘completely’ or ‘partly within’ search areas, we recommend using SASPAC’s integrated mapping software - Mapshore.

17. Closing this window and returning to SASPAC, allows us to select the System File using the ‘Select Files and Close Explore’ button. This identifies the selected file in the ‘Current Selection(s)’ window and after choosing to ‘Print with Statistics’, the task can be run using the ‘OK’ button.

18. Navigating to the final page in the print file presents the following summary statistics for the data requested:

```
      008GGJ00023  008GGJ00026  008GGJ00029  008GGJ00032  008GGJ00035  008GGJ00038  008GGJ00041  008GGJ00044  008GGJ00047  008GGJ00050
     172       227       282       255       286       265       206       223       190       225
      008GGJ00053  008GGJ00056  008GGJ00059  008GGJ00062  008GGJ00065  008GGJ00068  008GGJ00071  008GGJ00074  008GGJ00077  008GGJ00080
     187       147       119       181       144       170       139       197       193       152
      008GGJ00083  008GGJ00086  008GGJ00089  008GGJ00092  008GGJ00095  008GGJ00098  008GGJ00101  008GGJ00104  008GGJ00107  008GGJ00110
     226       301       301       197       301       197       301       197       301       197
  TOTAL       57707       26690       4279
  MEAN        276.11       127.70       31.00
  STAND. DEV   113.40       44.24       15.49
  MAXIMUM      1229       491        120
  MINIMUM      106        57         6
  MISS. VAL    0          0         0
```
This demonstrates that of the 57,707 people resident within 1 mile of London bridge rail station, 26,690 (or 46%) are aged between 16 and 74 and in employment and 9,939 (24%) of these travel to work on foot.

19. The command file generated by this sequence is as follows:

```
input system file name = "C:\SASPAC\SYSFILES\KEY / STATISTICS - OAS - CENTRAL LONDON (JULY 08).SYS"
select if distmile(TQ32868024) le 1
* <<radial=1, 532865, 180245, l, mi, 1,,>>
print variables with statistics KS0010004 KS0150001 /    KS0150011
output print file = c:\saspac\report\Module9.prn
end
finish
```

Summary of task sequence

- What sort of task is this?
- Print Variable
- Select Areas – Radial Search
- Search and extract grid reference from Google Maps*
- Explore for a .SYS
- Search Tables / Select variables
- 'Explore' for data
- Define radial search parameters
- View on a Map
- Select .SYS
- Print Preview
- Run Task

- remember to 'Make Command'.
Module 10

Creating User-Defined Tables:

Ward Profile Training
SASPAC allows users to create their own table layouts, and in this module, we shall create a small table showing summary statistic for an area. The task will also require the creation of a bespoke System File containing new variables which we will define.

**Section 1: Create a new table layout**

The first part of this task requires us to design and create the layout for our ward profile. This is where we define all of the metadata necessary for SASPAC to publish the profile and add all of the non-statistical information we require in the final published table.

Our example will include an age breakdown of the population, an ethnicity breakdown and details of the housing tenure in the area. For each variable we will include the total population and the percentage of the total found in that group.

We will have 13 variables as follows:

1. Total Population
2. Age 0-15
3. Age 16-64
4. Age 65+
5. White ethnicity
6. Mixed ethnicity
7. Asian ethnicity
8. Black ethnicity
9. Other ethnicity
10. Total households
11. Owner occupied households
12. Social rented households
13. Private rented households

The table layout is created in a text editor. A simple text editor programme such as Wordpad or Notepad is included as standard on most computers and will be perfectly sufficient for the purposes of creating a table layout. However, more complex text editor programmes, such as Textpad, do offer more features and greater flexibility and so can make the process easier and quicker.

Open your text editor.

The first lines of the table layout are instruction to SASPAC about how to read the file. You will need to type the following:

```
Table WP001   This is the table name
Default format = *,***,*** This is how numbers will be presented
Zoneid @5001   The  Zone Identifier will be placed at position 5001
Zlabel @5002   The  Zone label will be placed at position 5002
End of list
```

Page 154 Module 10
We can now proceed to creating the table layout itself. As well as entering titles, row and column headings and the locations of the variables we also need to build the table structure itself: the borders around and within the table.

1. Begin with a long line which represented the top border of the table. Use the underscore for this. The underscore is access by pressing shift and the key to the right of the 0 on the top line of the keyboard.

2. The next line begins with a vertical bar accessed by pressing the shift key and the key to the left of the Z key on the keyboard.

3. Use the space bar or tab key to move across the page until the cursor on the second line sits under the end of the line you created on the first line and add in a vertical bar.

The vertical bar character begins and ends every line of the table except the first. From now on it is assumed that this will be done on each line although it is not stated.

Line 3 type: “2011 Census Ward profile”
Line 4 is blank
Line 5 type: “Ward: @5002”
Line 6 type: “GSS Code: @5001”
Line 7 is blank
Line 8 is a horizontal line using underscore

The @5001 and @5002 references that we just placed in the layout refer to the list we typed at the beginning of the file. When we run the profile in SASPAC the software will replace these references with data from the System File.
4. Now we are ready to place the variables in the layout.

Line 9 is blank
Line 10 type “Total population”, 10 spaces, “@1”
Line 11 is blank
Line 12 type “Age Structure”
Line 13 type “0-15 years old”, 16 spaces, “@2”, 5 spaces, “@3”
Line 14 type “16-64 years old”, 15 spaces, “@4”, 5 spaces, “@5”
Line 15 type “65 years and over”, 8 spaces, “@6”, 5 spaces, “@7”
Line 16 is blank
Line 17 type “Ethnicity”
Lines 18-22: As age structure but for the 5 ethnic groups listed at the start of the task
Line 23 is blank
Line 24 type “Households”
Line 25 type “Total households”, 9 spaces, “@18”
Line 26-28: As age structure but for the 3 tenure groups listed at the start of the task
Line 29 is blank
Line 30 is a horizontal line
Line 31 is blank
Line 32 type “End of file”

Save your file as “Ward Profile.FWD” in the folder C:\SASPAC\FRWDATA
Section 2: Creating a bespoke System File

Some of the data that we require for our ward profile is not available directly from the census in the way we want to publish it. This is often the case when we want to output percentages from SASPAC as only the limited Key Statistics tables contain percentage information.

We will therefore need to create a new System File that contains all of the variables for our Ward Profile.

1. Open SASPAC

2. From the File menu choose ‘New Task’ and then ‘Create New System File’. This opens the following window:

   ![Create New System File Window]

   Some of the variables needed for our ward profile already exist while some will need to be created. The process for creating new variables is the same as that which was followed in Module 6 of this manual.

3. Click on the ‘Manipulate variables’ button in the centre of window and then open the ‘Create New’ tab.

4. The first variable we need to create is the total population aged 0 to 4. In the white box type ‘Age0to15= ’

5. Now click the ‘Search tables’ button on the right hand side of the window. Now selected the existing variables that will be needed to create our new variables, putting a plus (+) sign in between each so that they are added together.

The resulting equation should look like this:

\[
\text{Age0to4} = \text{KS102EW0002} + \text{KS102EW0003} + \text{KS102EW0004} + \text{KS102EW0005} + \text{KS102EW0006}
\]
6. Now click ‘Make command’ to add this expression to the command file. Make a note of what this variable was called (Age0to4) as you will need this information later on.

7. Now create the following variables using these names:

a. Perc0to15 percentage aged 0 to 15
b. Age16to64 total aged 16 to 64
c. Perc16to64 percentage aged 16 to 64
d. Age65plus total aged 65 and over
e. Perc65plus percentage aged 65 and over
f. White total white ethnicity
g. PercWhite percentage white ethnicity
h. Mixed total mixed ethnicity
i. PercMixed percentage mixed ethnicity
j. Asian total asian ethnicity
k. PercAsian percentage asian ethnicity
l. Black total black ethnicity
m. PercBlack percentage black ethnicity
n. Other total other ethnicity
o. PercOther percentage other ethnicity
p. Owned total owner occupied households
q. PercOwned percentage owner occupied households
r. Social total social rented households
s. PercSocial percentage social rented households
t. Private total private rented households
u. PercPrivate percentage private rented households

8. Once you have created all of the variables you need you can close the Manipulate Variables window. This returns you to the Create New System File window.

9. In the Selected Variable Name(s) box type the following:

a. KS101EW0001
b. KS106EW0001
c. The names of all 21 variables that you just created

10. In the System File Label box type “2011 ward profile”

11. Input the Key Statistics system file that contains ward data for your local area. Using the ‘Select Areas’ button select the wards in your local authority.

12. Click on the ‘Output to…’ button and save the new system file in C:\SASPAC as “System File for Ward Profile.SYS”
13. Click OK and run the command.

We now have a system file for use specifically with our ward profile.

Section 3: Add the variable references to the layout

1. Open Ward Profile.FWD in your text editor.

We are going to add in to the list at the beginning of the file references that link the cell references in our layout to the data contained in the System File. The list at the beginning of the file should currently look like this:

Table WP001
Default format = *,***,***
Zoneid @5001
Zlabel @5002
End of list

2. After Zlabel add a new line as follows:
   KS101EW0001 @1

   This is telling SASPAC that where the ‘@1’ symbol appears in the layout the value for Key Statistics variable 0001 from table KS101EW is needed.

3. We must now do this for each variable in our layout/ System File. Where we have created percentages we need to tell SASPAC not to use the default number layout. Instead we want percentage values to display as ***.* (eg. 100.0). The list should now look like this:
Table WP001
Default format = *, ***, ***
Zoneid @5001
Zlabel @5002
KS101EW0001 @1
Age0to15 @2
Perc0to15 @3 ***.
Age16to64 @4
Perc16to64 @5 ***.
Age65plus @6
Perc65plus @7 ***.
White @8
PercWhite@9 ***.
Mixed @10
PercMixed @11 ***.
Asian @12
PercAsian @13 ***.
Black @14
PercBlack @15 ***.
Other @16
PercOther @17 ***.
KS106EW0001 @18
Owned @19
PercOwned @20 ***.
Social @21
Percsocial @22 ***.
Private @23
Percprivate @24 ***.
End of list

Our layout is now almost complete. The final process is to add some formatting to the file. We are going to make the section heading bold.

4. In the list after PercPrivate add the following:
   Age Structure @5003 [B]
   Ethnicity @5004 [B]

5. Now go to the table layout and replace:
   Age Structure with @5003
   Ethnicity with @5004

6. We are going to underline the name of the ward.
   After Zlabel @5002, add “[U]”

7. Save the file.

The final table layout should look like this:
### Table WP001

**Default format:** 
```
*,***,***
```

<table>
<thead>
<tr>
<th>Zoned</th>
<th>$5001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>$5002 [U]</td>
</tr>
<tr>
<td>PS102EW0001</td>
<td>$1</td>
</tr>
<tr>
<td>Age60-64</td>
<td>$2</td>
</tr>
<tr>
<td>Per60-64</td>
<td>82 *<strong>,</strong></td>
</tr>
<tr>
<td>Age65-66</td>
<td>$4</td>
</tr>
<tr>
<td>Per65-66</td>
<td>85 *<strong>,</strong></td>
</tr>
<tr>
<td>Age65plus</td>
<td>$6</td>
</tr>
<tr>
<td>Per65plus</td>
<td>87 *<strong>,</strong></td>
</tr>
<tr>
<td>White</td>
<td>$8</td>
</tr>
<tr>
<td>PerWhite</td>
<td>89 *<strong>,</strong></td>
</tr>
<tr>
<td>Mixed</td>
<td>$10</td>
</tr>
<tr>
<td>PerMixed</td>
<td>11 *<strong>,</strong></td>
</tr>
<tr>
<td>Asian</td>
<td>$12</td>
</tr>
<tr>
<td>PerAsian</td>
<td>13 *<strong>,</strong></td>
</tr>
<tr>
<td>Black</td>
<td>$14</td>
</tr>
<tr>
<td>PerBlack</td>
<td>15 *<strong>,</strong></td>
</tr>
<tr>
<td>Other</td>
<td>$16</td>
</tr>
<tr>
<td>PerOther</td>
<td>17 *<strong>,</strong></td>
</tr>
<tr>
<td>PS102EW0001</td>
<td>$18</td>
</tr>
<tr>
<td>Owner</td>
<td>$19</td>
</tr>
<tr>
<td>PerOwner</td>
<td>20 *<strong>,</strong></td>
</tr>
<tr>
<td>Social</td>
<td>$21</td>
</tr>
<tr>
<td>PerSocial</td>
<td>22 *<strong>,</strong></td>
</tr>
<tr>
<td>Private</td>
<td>$23</td>
</tr>
<tr>
<td>PerPrivate</td>
<td>24 *<strong>,</strong></td>
</tr>
<tr>
<td>Age structure</td>
<td>$5003 [B]</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>$5004 [B]</td>
</tr>
</tbody>
</table>

---

**2011 Census Ward Profile**

**Wards:** $5002

**GSS Code:** $5001

<table>
<thead>
<tr>
<th>Total population</th>
<th>$1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$5003</td>
<td></td>
</tr>
<tr>
<td>0-15 years</td>
<td>$2</td>
</tr>
<tr>
<td>16-64 years</td>
<td>$4</td>
</tr>
<tr>
<td>65 years and older</td>
<td>$6</td>
</tr>
<tr>
<td>$5004</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>$8</td>
</tr>
<tr>
<td>Mixed</td>
<td>$10</td>
</tr>
<tr>
<td>Asian</td>
<td>$12</td>
</tr>
<tr>
<td>Black</td>
<td>$14</td>
</tr>
<tr>
<td>Other</td>
<td>$16</td>
</tr>
<tr>
<td>Total households</td>
<td>$19</td>
</tr>
<tr>
<td>Owner occupied</td>
<td>$19</td>
</tr>
<tr>
<td>Social rent</td>
<td>$21</td>
</tr>
<tr>
<td>Private rent</td>
<td>$23</td>
</tr>
</tbody>
</table>

---

End of file
Section 4: Convert the FWD layout to a SASPAC readable Framework File (FWK)
The file we have created is essentially a text file with an FWD file extension. In order for SASPAC to output data in our new layout we must convert the file to a Framework File (FWK). This is a simple process in SASPAC:

1. In the Tools menu choose Import data and then Framework File. This opens the Create Framework File window.

2. Click on Input from and find the FWD file that we created: C:\SASPAC\FRWDATA\Ward Profile.FWD

3. Click on Output, name the file Ward Profile.FWK and save it in the folder C:\SASPAC\FRWORK\.

Section 5: Run the ward profile
The ward profile is now ready to view.

1. Open a Print tables task
2. Input System file: C:\SASPAC\System File for Ward Profile.SYS
3. Browse all Framework Files: C:\SASPAC\FRWORK\Ward Profile.FWK
4. In the Table Selected box type “WP001”
5. Click OK and give the task a user defined name
6. View the ward profile that you have created, it can now be printed, published to HTML or exported to Excel.
Summary of task sequence

What sort of task is this? → Create a new table → Design table in text editor → Save .FWD file

Browse & select new .FWK file ← Input System File ← New Task - Print Table ← Convert .FWD file to a .FWK

Select new Table → Run Task
Module 1: Selecting variables

Command file:
input system file name = "C:\SASPAC\SYSFILES\2001 STANDARD TABLES FOR WARDS IN ENFIELD.SYS"
print variables ZLABEL ST0150049 ST0020255 / ST0020270 ST0590031 ST0360229
output print file = C:\SASPAC\REPORT\MODULE_1.PRN
end
finish

Print file:
Module 2: Selecting Tables

Command file:

input system file name = "C:\SASPAC\SYSFILES\2001 STANDARD TABLES FOR DISTRICTS IN LONDON.SYS"
input framework file = c:\saspac\frwork\st01.fwk
print tables ST052 ST103
output print file = C:\SASPAC\REPORT\MODULE_2.PRN
end
finish

Print file: (Page 1 of 66)
Module 3a: Selecting Areas
Command file:
input system file name = "C:\SASPAC\SYSFILES/  
2001 CENSUS AREA STATISTICS FOR OAS IN ENFIELD.SYS"
input framework file = c:\saspac\frwork\C501.fwk
include 00AKGP0004
include 00AKGY0010 to 00AKGY0013
include 00AKHH0029
print tables C5001
output print file = C:\SASPAC\REPORT\MODULE_3a.PRN
end
finish

Print file: (Page 1 of 12)
Module 3b: Selecting Areas

Command file:

input system file name = "C:\SASPAC\SYSFILES/"
2001 CENSUS AREA STATISTICS FOR OAS IN ENFIELD.SYS

include 00AKHE0044
include 00AKGT0007
include 00AKHB0031
include 00AKHE0041
include 00AKHE0040
include 00AKHE0028
include 00AKHE0036
include 00AKHE0043
include 00AKHE0037
include 00AKHE0039
include 00AKHB0030
print variables with statistics CS0010001 CS0210010
output print file = C:\SASPAC\REPORT\MODULE_3b.PRN
end
finish

Print file:

```
SASPAC CENSUS ANALYSIS PACKAGE

<table>
<thead>
<tr>
<th>ZONE ID</th>
<th>CS0010001</th>
<th>CS0210010</th>
</tr>
</thead>
<tbody>
<tr>
<td>00AKHE0044</td>
<td>296</td>
<td>175</td>
</tr>
<tr>
<td>00AKGT0007</td>
<td>260</td>
<td>167</td>
</tr>
<tr>
<td>00AKHB0031</td>
<td>207</td>
<td>108</td>
</tr>
<tr>
<td>00AKHE0041</td>
<td>309</td>
<td>148</td>
</tr>
<tr>
<td>00AKHE0040</td>
<td>214</td>
<td>100</td>
</tr>
<tr>
<td>00AKHE0028</td>
<td>374</td>
<td>166</td>
</tr>
<tr>
<td>00AKHE0036</td>
<td>333</td>
<td>169</td>
</tr>
<tr>
<td>00AKHE0043</td>
<td>309</td>
<td>148</td>
</tr>
<tr>
<td>00AKHE0037</td>
<td>317</td>
<td>155</td>
</tr>
<tr>
<td>00AKHE0039</td>
<td>300</td>
<td>150</td>
</tr>
<tr>
<td>00AKHB0030</td>
<td>254</td>
<td>101</td>
</tr>
</tbody>
</table>
```
Module 4a: Creating New zones – area selection

Command file:
input system file name = "C:\SASPAC\SYSFILES\2001 CENSUS AREA/ STATISTICS FOR OAS IN ENFIELD.SYS"
input system file name = "C:\SASPAC\SYSFILES\2001 CENSUS AREA/ STATISTICS FOR WARDS IN ENFIELD.SYS"
read in series
new zone id = Zone001 name = New Zone 1
USING AREAS - oalevel 00AKGT0007 +(oalevel 00AKGW0001 to/ 00AKGW0003) + oalevel 00AKGW0008*0.33 + wardlevel 00AKGT/
output system file name = C:\SASPAC\SYSFILES\Module_4a.SYS / label=Training Module Exercise 4a
end
finish

Module 4b Creating New zones – gazetteer files

Command file:
Input gazetteer file with labels name = C:\SASPAC\COMMAND\Enfield.gaz existing zone cols 1 to 10 scale factor cols 21 to 24 new zone cols 13 to 18
input system file name = "C:\SASPAC\SYSFILES\2001 CENSUS AREA/ STATISTICS FOR WARDS IN ENFIELD.SYS"
input system file name = "C:\SASPAC\SYSFILES\2001 CENSUS AREA/ STATISTICS FOR OAS IN ENFIELD.SYS"
read in series
set zone echo on
output system file name = C:\SASPAC\SYSFILES\Module_4B.SYS / label=Training Exercise Module 4B
end
finish
Module 5: Exporting variables from SASPAC

Command file:

```sas
input system file name = "C:\SASPAC\Training / System Files\2001 KEY STATISTICS FOR OAS / IN ENFIELD.SYS"
set countycode on
save ZONEID KS0029002
output CSV file with headers name = / C:\SASPAC\INTERFAC\Training_Module_5.CSV
end
finish
```

CSV file:

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZONEID</td>
<td>PLESSTHAN15</td>
<td></td>
</tr>
<tr>
<td>00AKGL0001</td>
<td>17.22</td>
<td></td>
</tr>
<tr>
<td>00AKGL0002</td>
<td>21.61</td>
<td></td>
</tr>
<tr>
<td>00AKGL0003</td>
<td>24.01</td>
<td></td>
</tr>
<tr>
<td>00AKGL0004</td>
<td>9.32</td>
<td></td>
</tr>
<tr>
<td>00AKGL0005</td>
<td>12.82</td>
<td></td>
</tr>
<tr>
<td>00AKGL0006</td>
<td>22.18</td>
<td></td>
</tr>
<tr>
<td>00AKGL0007</td>
<td>25.</td>
<td></td>
</tr>
<tr>
<td>00AKGL0008</td>
<td>10.96</td>
<td></td>
</tr>
<tr>
<td>00AKGL0009</td>
<td>20.6</td>
<td></td>
</tr>
<tr>
<td>00AKGL0100</td>
<td>15.49</td>
<td></td>
</tr>
<tr>
<td>00AKGL0101</td>
<td>16.81</td>
<td></td>
</tr>
<tr>
<td>00AKGL0102</td>
<td>16.81</td>
<td></td>
</tr>
<tr>
<td>00AKGL0103</td>
<td>16.45</td>
<td></td>
</tr>
<tr>
<td>00AKGL0104</td>
<td>15.24</td>
<td></td>
</tr>
<tr>
<td>00AKGL0105</td>
<td>19.66</td>
<td></td>
</tr>
<tr>
<td>00AKGL0106</td>
<td>16.54</td>
<td></td>
</tr>
<tr>
<td>00AKGL0107</td>
<td>17.25</td>
<td></td>
</tr>
<tr>
<td>00AKGL0108</td>
<td>16.62</td>
<td></td>
</tr>
<tr>
<td>00AKGL0109</td>
<td>26.3</td>
<td></td>
</tr>
<tr>
<td>00AKGL0200</td>
<td>24.49</td>
<td></td>
</tr>
</tbody>
</table>

Exported as a *.jpg to MS Word:
Exported as a *.kmz to Google Earth:
Module 6: Creating new variables

Command file:

```
input system file name = "C:\SASPAC\SYSFILES\2001 CENSUS AREA STATISTICS FOR OAS IN ENFIELD.SYS"

IF CS0010001 = 0 then percent1524 = 0  else percent1524= (CS0010096+CS0010126) / CS0010001 *100.0

describe variable percent1524 dp= 2 label=percentage aged 15-24

header all Percentage Population 15-24 OAs in Enfield

print variables with labels cs0010001 percent1524

output print file = c:\saspac\report\MODULE_6.prn

end

finish
```

Print file: (Page 1 of 17)
Module 7: Creating new variable using Select IF

Command file:

input system file name = "C:\SASPAC\SYSFILES\2001 CENSUS AREA STATISTICS FOR OAS IN ENFIELD.SYS"

perl15 = (CS0010006+ CS0010036+ CS0010066+)/ CS0010101 ) / CS0010001*100.0

SELECT IF perl15>=35

print variables perl15

output print file = /C:\SASPAC\REPORT\Module_7.PRN

end

finish

Print file: (Page 1 of 1)
Module 8: Using Travel to Work datasets

Command file:
Step 1 – create a new system file containing just the flows within the area of interest (Core Area)
Command file:
input system file name = C:\SASPAC\SYSFILES\SWS2XX_WITH_LABELS.SYS
include 00AKGL to 00AKHH
output system file name =/ C:\SASPAC\SYSFILES\SWS2XX_Enfield_core.SYS label=SWS2XX /London core
end
finish

Step 2 – rezone the areas outside the core area
Command file:
TEST 71
Input gazetteer file with labels name = /C:\SASPAC\COMMAND\Ward_to_District_Enfield.gaz /
existing zone cols 1 to 6 new zone cols 15 to 18
input system file name = C:\SASPAC\SYSFILES\SWS2XX_Enfield_core.SYS
set zone echo on
output system file name = C:\SASPAC\SYSFILES\SWS2XX_Enfield_Rezone.SYS label=SWS2XX /
London rezone
end
finish

Step 3 – Extract and map the number of persons working within the area
Command file:
input system file name = C:\SASPAC\SYSFILES\ SWS2XX_Enfield_Rezone.SYS
include 00AKGL to 00AKHH
set in&out off
set countycode off
save ORIGID DESTID SWS2010001 SWS2010002 SWS2010003
output CSV file with headers name = C:\SASPAC\INTERFAC\SWS2XX_Enfield.CSV
end
finish

CSV file:
Data mapped in Mapshore:
Module 9: Use of Google Maps and radial searching

Command file:
input system file name = "C:\SASPAC\SYSFILES\KEY STATISTICS - OAS – CENTRAL LONDON.SYS"
select if distmile(TQ32868024) le 1
* <<radial=1,532865,180245,1,mile,1,,>>
print variables with statistics KS0010004 KS0150001 / KS0150011
output print file = c:\saspac\report\Module9.prn
end
finish

Print file:

<table>
<thead>
<tr>
<th></th>
<th>KS0010004</th>
<th>KS0150001</th>
<th>KS0150011</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL</td>
<td>90310</td>
<td>40360</td>
<td>9939</td>
</tr>
<tr>
<td>MEAN</td>
<td>282.22</td>
<td>126.13</td>
<td>31.06</td>
</tr>
<tr>
<td>STAND. DEV</td>
<td>107.04</td>
<td>40.71</td>
<td>19.97</td>
</tr>
<tr>
<td>MAXIMUM</td>
<td>1229</td>
<td>431</td>
<td>130</td>
</tr>
<tr>
<td>MINIMUM</td>
<td>100</td>
<td>63</td>
<td>5</td>
</tr>
<tr>
<td>MISS. VAL</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Module 10: Creating User defined tables

Command file (Create FWK file):
input framework data file name = C:\SASPAC\FRWDATA\MYTABLE.fwd
output framework file name = C:\SASPAC\FRWORK\MYTABLE.FWK
end
finish

Command file (Print new table):
input system file name = "C:\SASPAC\SYSFILES\2001 STANDARD/ TABLES FOR DISTRICTS IN LONDON.SYS"
input framework file = C:\SASPAC\FRWORK\MYTABLE.FWK
include ualevel 00AC
print tables my01
output print file = C:\SASPAC\REPORT\Module10.PRN
end
finish

Print file:

Age Profile of Resident Population of:
Area Name: Barnet
Area ID: 00AC

<table>
<thead>
<tr>
<th>Total Population:</th>
<th>314,559</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population aged</td>
<td></td>
</tr>
<tr>
<td>0 - 4</td>
<td>20,215</td>
</tr>
<tr>
<td>5 - 9</td>
<td>20,537</td>
</tr>
</tbody>
</table>