

Census
Census
Sensus
Sensus
Sensus
Sensus
Sensus
Sensus

Saspac

&

geowise™
Underlying Geography

Web Publisher

Training Manual

Website: www.saspac.org
Email: info@saspac.org
Helpdesk: 020 7983 4348

Contents

Introduction and background information	7
Module 1 Publishing to the ‘Area Profile’ Template.....	17
Module 2 Publishing to the ‘Area Profile Comparison’ Template... <i>(not currently available)</i>	
Module 3 Publishing to the ‘Indicator Map’ Template..... <i>(not currently available)</i>	
Module 4 Publishing to the ‘Double Indicator Map’ Template..... <i>(not currently available)</i>	
Module 5 Publishing to the ‘MapShore’ Template..... <i>(not currently available)</i>	
Module 6 Publishing to the ‘Table and Charts’ Template..... <i>(not currently available)</i>	

Introduction and general overview of the GeoWise Publisher for use with SASPAC

Introduction and background information

Since its inception in the early 1980s, SASPAC has evolved continually to meet the needs of changing technology. In the early days, output of information from the Census as stand-alone paper copies of the individual tables, or as printed lists of variables was sufficient for most uses. In some (rare) cases, data was exported in electronic file format for use with other applications. As that decade developed, so did SASPAC with a migration from mainframe computers to the desktop personal computer (PC), but the functionality remained basically the same. With the 1991 version of the software came a greater ability to tailor the output to a user's specific needs, and later enhancements saw the addition of a mapping/GIS module (in the form of MapShore) in addition to the strengthening of the interactive facilities made possible by the move to PCs.

The 2001 version of SASPAC built on the strengths of the previous software whilst still offering the wide range of facilities available to users. Tabular or variable output in paper copy form as well as in electronic form is still widely used by the many users of SASPAC. However, the advent of the World Wide Web (WWW) has meant that end users of Census data have come to expect delivery of their information in a web-browser compatible form. These end users – for example, senior managers in local and health authorities, members of the general public, elected councillors in local authorities – do not have the resources themselves to extract this information directly through SASPAC. Increasingly, the professional user of SASPAC has to obtain the information, and, crucially, deliver it in a form that is easily digestible by that end user.

In early 2004, the Greater London Authority, on behalf of the SASPAC consortium, contracted with GeoWise Ltd for the development of a module (the Publisher module for SASPAC) which would enable users to create Hypertext Markup Language (more commonly known as **HTM** or **HTML**) files, developed from templates supplied with the software. These files may then be placed in a networked location for access by any authorized users. By adopting this approach, it was hoped that the professional SASPAC user could follow the process through from start to finish without requiring support or intervention from an IT department. The only time when such IT assistance may be required (depending on individual site set-up) is when the final **HTM** files (or pages) are made available to other users. As **HTM** files can become infected the IT department will require they be scanned for viruses or other infection before they are made available on any network.



In order to develop the Publisher module for SASPAC, GeoWise have made use of Microsoft's **.Net Framework**. This is Microsoft's strategy for connecting systems, information, and devices through Web services. GeoWise is an accredited Microsoft Certified Partner, and as such is licensed to distribute the **.NET** application. Users will therefore not have to download this from the Microsoft website as it will be bundled with

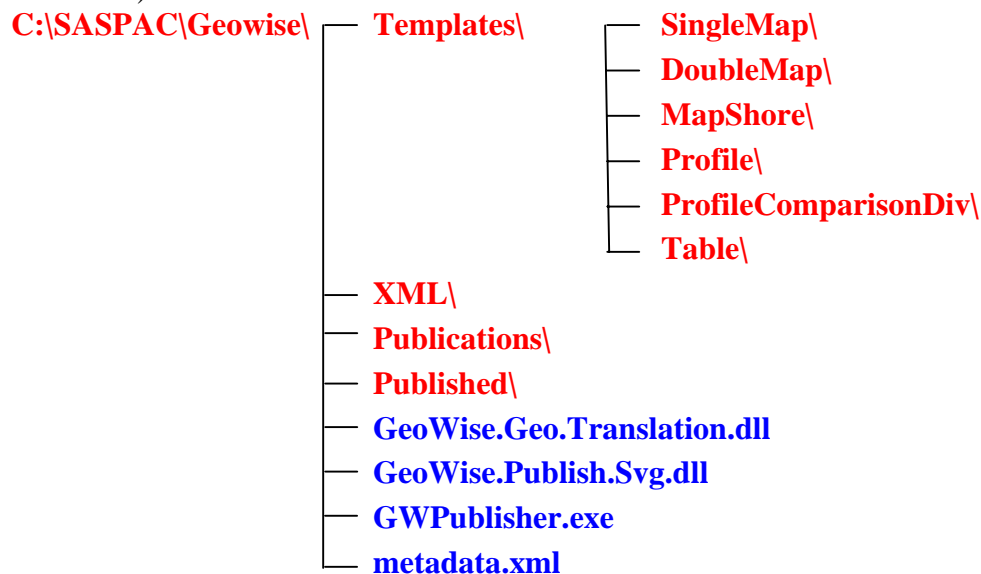
SASPAC and will be installed along with the Publisher. The **.NET** application is only required on the SASPAC user's machine and is not required for viewing the **HTM** files by the end-user.

The **HTM** files make use of Scalable Vector Graphics (**SVG**) which is a language for describing two-dimensional graphics and graphical applications in **XML**. Adobe Systems Incorporated (known for their Adobe Reader software) has developed a **SVG Viewer** which is downloadable from their website. This **SVG Viewer** is essential for all users wishing to view the final **HTM** files. As this viewer is commonly required for viewing other websites – for example, the Office for National Statistics (ONS) requires it for viewing pages on its Neighbourhood Statistics website – most users may find that it is already available on their system.

Installation

Version 8.00 of SASPAC contains the GeoWise Publisher as a part of the installation package. By default, the Publisher and all associated folders/files will be installed in the SASPAC directory on the C: drive of the users PC. Customised installation is possible within the installation procedure. Further guidance relating to the installation procedure can be found in the “**SASPAC v8 Installation & Configuration**” user guidance note.

When installed, the user will see the following folder/file structure (assuming a default installation):



The **Templates** folder itself contains a series of folders, each of which contains the files necessary to define a template to which the final output, or **HTM** file is produced. Each template will be created as you work through this training manual, in addition, further information may be found in UGN GP03 to UGN GP08. The **XML** directory contains the **XSD** (XML Schema Definition) files necessary for the Publisher and SASPAC to co-exist and ensure that the final output or published files are correct. Neither of these directories, with one small exception, will require access or editing by users outside of SASPAC or the Publisher.

*[The exception mentioned above is the **toplogo.gif** file that will be found in each of the folders within the **Templates** folder. This file, as supplied, contains an image depicting the GeoWise logo. Users may substitute their own logo by creating a small version of their*

organisation's logo as a GIF file, rename it to toplogo.gif and copying it into each of the folders within the **Templates** folder, thus overwriting the supplied image.]

On installation, the **Publications** and **Published** directories will be empty, but as the Publisher is used, they will become populated with files created by the Publisher, with the **Published** directory being the default location for the final published file, and its associated files. As this use of the **Published** directory as a default location for published files will lead to overwriting and deletion of files published previously, users will be encouraged to create a new folder for each publication.

In addition to the folders and files that are installed in the **SASPAC\GeoWise** folder, a series of files is also added to the **SASPAC** folder. These are the **XVM** (XML Variable Metadata) and **XFF** (XML Framework Files) files necessary to enable SASPAC to export **XML** files containing all the information that is required to meet the specifications of the **XSD** files mentioned above.

Eventually there will be a pair of **XVM/XFF** files for each of the main 2001 Census datasets, but initially there are no **XFF** files available for the Univariate Tables, or the Key Statistics. These will be supplied at a later date. In addition, the current **XVM** and **XFF** files do not cover all tables in Scotland and Northern Ireland.

Within the SASPAC.INI file, in the [Framework, Table & Cell Metadata] block, a new parameter is added to lines relevant to datasets covered by the XVM and XFF files. This parameter is of the form:

```
xml:"KS01"
```

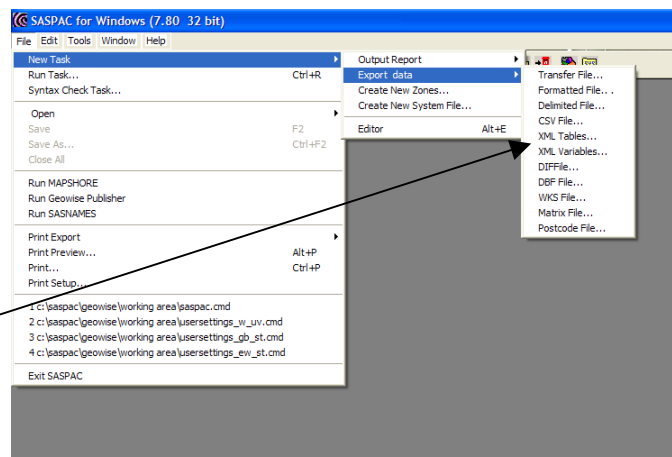
and its presence indicates the existence of an associated **XVM** and/or **XFF** file.

Publication

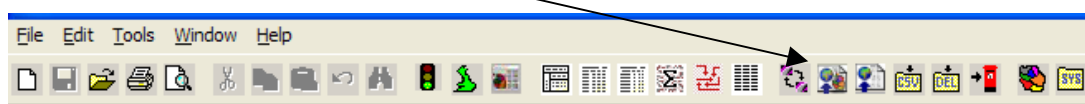
There are three steps the user must follow to create their own customised **HTM** file (or page). Firstly, SASPAC is used to create an **XML** file containing the variables and zone information to be published. Secondly, the Geowise Publishing application is used to convert the **XML** file into a publishable form, i.e. a **HTM** file. Finally, the Publisher is used to edit a template's settings and customise the appearance of the final **HTM** file.

a) Create XML variable/table file

The first step in publication is to run a SASPAC task to output either XML variables or an XML table. The process of achieving this is identical to the process for any other type of output. Two new options have been added to the **File / New Task / Export Data** menu item, as shown here.



Alternatively, the same options may be accessed from two new buttons found on the button bar as shown below:

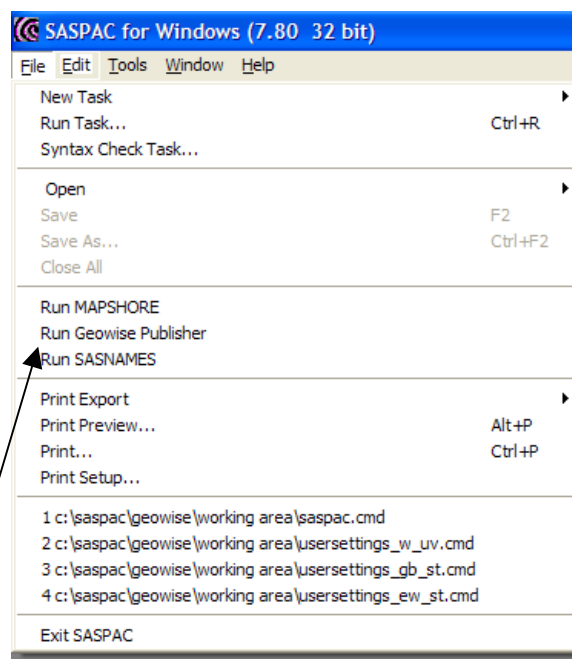


Use of either of these options will lead the user into a series of windows similar to those commonly seen when outputting other types of export data. In fact, with one minor exception, the main task window is identical to that seen in an Export **CSV** Task. The main point is, that as far as the user is concerned, the SASPAC component is changed only in that there is now the option to output the **XML** files, but the means of doing that are identical to previously available options.

It is during the SASPAC processing stage that the software makes use of the **XVM** or **XFF** files to add information to the **XML** file output by SASPAC. The **XVM** file is basically a text file that allocates a label to each of the variables in the Census datasets, while the **XFF** file is the XML equivalent of the **FWK** files. They add metadata to the data that is held within the SASPAC System Files, and enable the final data to be published in a manner that is intelligible to non-professional users of the Census data.

b) Convert XML file to publishable form

Once the **XML** file is created by SASPAC, it may be converted into a publishable form (as **HTM**) by running the GeoWise Publisher, either from within SASPAC, or as a stand-alone application. This mirrors the operation of Mapshore from within SASPAC which may either be called from the File menu option, or from the “Go Mapping” button which becomes operational at the successful end of an Export **CSV** File Task. The GeoWise publisher may also be called from the File menu option, shown in this screen shot, or from the “Go Publish” button available at the conclusion of a successful **XML** File Export Task.



c) Select data output format/settings to be used with chosen Publisher template

The process of publishing to each template is described through each module of this training manual and further details can also be found in UGN GP-03 to UGN GP-08.

Common to each template, however, is the manner in which the data output by SASPAC, is read into the supplied template by the Publisher.

When the Publisher is called, it looks for a file named **geopubsettings.xml** which will be located, by default, in the same folder as the **SASPAC.INI** file. This file contains references to the data, and crucially, the order on the data file, as output from SASPAC. It also groups the variables into **Topics**, so that the final output can be structured in a logical fashion.

The default **geopubsettings.xml** supplied with the software will contain no definitions of **Topics**, but to assist users, an enhanced template user settings file is supplied to facilitate the creation of **Topics** relevant to the Standard Tables dataset.

When a Publisher run is applied to the default [geopubsettings.xml](#) the user must create **Topics** relevant to that data. Within the [geopubsettings.xml](#) file, there will be a block which begins with the characters:

```
<TopicsList/>
```

and concludes with the characters

```
</TopicsList>
```

In the supplied default file, shown here:

```
- <UserSettings xmlns="http://www.geowise.co.uk/xml/namespaces/publisher/user">
  <TopicsList />
  - <PublisherSettings>
    <Setting name="ShowWelcome" value="True" />
    <Setting name="GeoIDColumn" value="CODE" />
    <Setting name="PublicationFileDirectory" value="C:\SASPAC\GeoWise\PUBLICATIONS" />
    <Setting name="PublishToDirectory" value="C:\SASPAC\GeoWise\PUBLISHED" />
    <Setting name="TemplatesLocation" value="C:\SASPAC\GeoWise\TEMPLATES" />
    <Setting name="UserHeading" value="" />
    <Setting name="ONSLicense" value="" />
    <Setting name="Copyright" value="" />
    <Setting name="MetadataFilePath" value="C:\SASPAC\GeoWise\METADATA.XML" />
  </PublisherSettings>
</UserSettings>
```

there is only one line that reads:

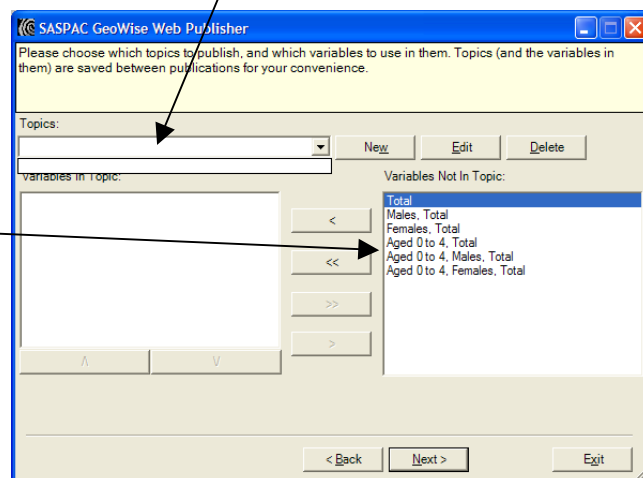
```
<TopicsList/>
```

This means that when a SASPAC output **XML** file created by the following Command file:

```
INPUT SYSTEM FILE NAME = "C:\SASPAC\SYFILES\STANDARD TABLES /
- NATIONAL TO DISTRICT- ENGLAND (FEB 05).SYS"
SAVE ST0010001 ST0010004 ST0010007 ST0010010 ST0010013 /
ST0010016
OUTPUT XMLVARIABLES FILE NAME = C:\SASPAC\INTERFAC\JOB_1.XML
END
FINISH
```

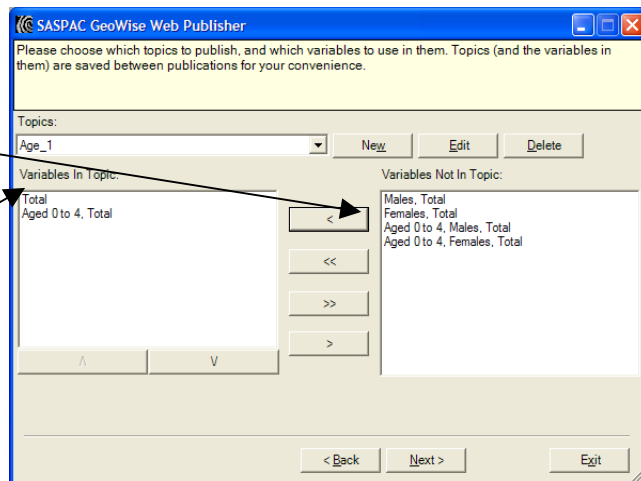
is published against this [geopubsettings.xml](#) file, no **Topics** are pre-defined, and the the user has to define them within the Publisher as shown in this image.

Here it can be seen that the variables output by SASPAC have had the metadata obtained from the file [ST01.XVM](#) associated with them. Thus, variable ST0010013 appears as "Aged 0 to 4, Males, Total". Note that the order in which the data are output from SASPAC, is the order by which the Publisher references them.



It is therefore essential that the user maintains a relationship between the [geopubsettings.xml](#) file, and the data to be published.

Topics may now be created by the process of highlighting them in the “Variables Not in Topic” block and moving them (using the ‘<’ button) into the “Variables in Topic” block. Before this window is seen, the **Topic** is given a (default) ID, and a user supplied Name. If it is to be used in the publication, it must also be marked as for “Use in publication”.



Other **Topics** may then be created using any of the variables in the SASPAC output **XML** file that is being published. Variables may be used in different **Topics**, that is, there is no limit on the number of times an individual variable may be used in a publication. When creation of **Topics** is complete, the user progresses to publication of the final **HTM** file.


Following the first publication, the [geopubsettings.xml](#) file contains the following lines:

```
- <TopicsList>
  - <Topic source="" id="__Topic_1" name="Age_1" use="yes">
    <Indicator id="0"/>
    <Indicator id="3"/>
  </Topic>
  - <Topic source="" id="__Topic_2" name="Age_2" use="yes">
    <Indicator id="1"/>
    <Indicator id="4"/>
  </Topic>
  - <Topic source="" id="__Topic_3" name="Age_3" use="yes">
    <Indicator id="2"/>
    <Indicator id="5"/>
  </Topic>
</TopicsList>
```

It can be seen that the contents of these lines have been generated by the entries made during the process described above. The “**Indicator id=**” entry indicates the absolute position within the SASPAC output **XML** file of the variable being used within the **Topic**, with the first variable on the **XML** file having the indicator id 0, and the n^{th} having the indicator id (n-1).

If a **Topic** is used in a subsequent publication, a “**Position**” parameter is added to each “**Indicator**” line. The [geopubsettings.xml](#) file will thus look slightly different after the second and further publications.

The end result of this process (using the Area Profile template) when viewed within a browser is the following:



Profile for English Districts
Unspecified
Population: Unspecified
Level: Zone

Profile Area: Profile Layout:

Age_1		Age_2		Age_3	
Total	121023	Males, Total	60448	Females, Total	60575
Aged 0 to 4, Total	7949	Aged 0 to 4, Males, Total	4036	Aged 0 to 4, Females, Total	3913

Click use Licence Number ONS12345 ©Crown Copyright

The **Topic** name supplied within the Publisher (or the [geopubsettings.xml](#) file) appears in the final page that is accessed by the end-user. The variable names supplied by the **XVM** file also appear in that page. It is therefore very important for the clarity of the end result that proper consideration is given to the definition of these elements before the publication process is started.

To assist users of the Publisher in obtaining meaningful output, a few variations of the **XVM** and [geopubsettings.xml](#) files have been distributed with the software, and others will be made available through the SASPAC website.

For the **XVM** files, comparison of a few lines of the ST01 files will serve to indicate the difference:

ST01.XVM	ST01_enhanced.XVM
"ST0010013", "Aged 0 to 4, Males, Total"	"ST0010013", "Aged 0 to 4"
"ST0010014", "Aged 0 to 4, Males, Household residents"	"ST0010014", "Aged 0 to 4, Males, Household residents"
"ST0010015", "Aged 0 to 4, Males, Communal Establishment residents"	"ST0010015", "Aged 0 to 4, Males, Communal Establishment residents"
"ST0010016", "Aged 0 to 4, Females, Total"	"ST0010016", "Aged 0 to 4"
"ST0010017", "Aged 0 to 4, Females, Household residents"	"ST0010017", "Aged 0 to 4, Females, Household residents"
"ST0010018", "Aged 0 to 4, Females, Communal Establishment residents"	"ST0010018", "Aged 0 to 4, Females, Communal Establishment residents"

ST01.XVM gives a full description of ST0010013 and ST0010016, but in **ST01_Enhanced.XVM**, these variables lose the description of their gender. This is because the gender of the variable will appear in the **Topic** name as supplied by the variant **geopubsettings.xml** file.

The variant **geopubsettings.xml** file (which is named **(ew_st)geopubsettings.xml**), contains a large number of **Topic** definitions in terms of associated variables.

An example of part of the **(ew_st)geopubsettings.xml** file is shown here:

```
<Topic source="" id="TotEthnic(ST)" name="All Residents by Ethnic Group" use="yes">
  <Indicator id="301" position="0" />
  <Indicator id="302" position="1" />
  <Indicator id="303" position="2" />
  <Indicator id="304" position="3" />
  <Indicator id="305" position="4" />
  <Indicator id="306" position="5" />
  <Indicator id="307" position="6" />
  <Indicator id="308" position="7" />
  <Indicator id="309" position="8" />
  <Indicator id="310" position="9" />
  <Indicator id="311" position="10" />
  <Indicator id="312" position="11" />
  <Indicator id="313" position="12" />
  <Indicator id="314" position="13" />
  <Indicator id="315" position="14" />
  <Indicator id="316" position="15" />
  <Indicator id="317" position="16" />
</Topic>
```

This shows that there are 17 variables associated with a **Topic** that is named “All Residents by Ethnic Group”, and these variables occupy positions 302 to 318 on the SASPAC output **XML** file that is being used by the GeoWise Publisher.

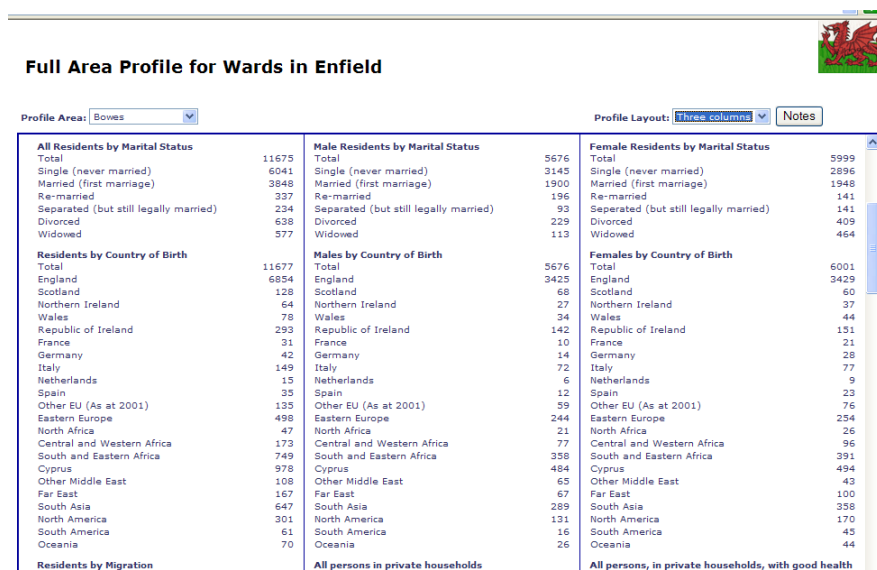
There is no reference within the **geopubsettings.xml** file to the identity of the variable to be used – *only an absolute position within the SASPAC output XML file. It is therefore*

essential that the user ensures that there is compatibility between the SASPAC command file being used, and the [geopubsettings.xml](#) file that is to be used for the publication.

To assist with this, a series of command files are also available with the software, and these have file names that mirror those of the associated [geopubsettings.xml](#) file. These will need to be edited to ensure that appropriate System Files are input, and the required areas are selected for output.

To use all these additional files, the user will need to follow this procedure (taking Standard Tables in England as an example).

1. Rename the **ST01.XVM** file found in the **C:\SASPAC** folder to (for example) **ST01_Master.XVM**.
2. Copy the **ST01_Enhanced.XVM** file found in the **Added_Files** folder on the SASPAC Version 8 distribution CD to the **C:\SASPAC** folder, and rename it **ST01.XVM**.
3. Copy the **(ew_st)geopubsettings.xml** file found in the **Added_Files** folder on the SASPAC Version 8 distribution CD to the folder that contains the **SASPAC.INI** file. Rename the existing **geopubsettings.xml** to (for example) **master_geopubsettings.xml**, and rename the copied xml file to **geopubsettings.xml**.
4. Edit the **(ew_st)geopubsettings.cmd** file found in the **Added_Files** folder on the SASPAC Version 8.00 distribution CD by inputting an appropriate System File, and ensuring that the necessary areas are output.
5. Run the edited SASPAC Command File.
6. Run the GeoWise Publisher taking the appropriate actions at the necessary stages, (as described in the relevant training manual module or User Guidance Notes).
7. View the resulting **HTM** page.



Full Area Profile for Wards in Enfield

Profile Area: Profile Layout:

All Residents by Marital Status		Male Residents by Marital Status		Female Residents by Marital Status	
Total	11675	Total	5676	Total	5999
Single (never married)	6041	Single (never married)	3145	Single (never married)	2896
Married (first marriage)	3848	Married (first marriage)	1900	Married (first marriage)	1948
Re-married	337	Re-married	196	Re-married	141
Separated (but still legally married)	234	Separated (but still legally married)	93	Separated (but still legally married)	141
Divorced	638	Divorced	229	Divorced	409
Widowed	577	Widowed	113	Widowed	464
Residents by Country of Birth		Males by Country of Birth		Females by Country of Birth	
Total	11677	Total	5676	Total	6001
England	6854	England	3425	England	3429
Scotland	128	Scotland	68	Scotland	60
Northern Ireland	64	Northern Ireland	27	Northern Ireland	37
Wales	78	Wales	34	Wales	44
Republic of Ireland	293	Republic of Ireland	142	Republic of Ireland	151
France	31	France	10	France	21
Germany	42	Germany	14	Germany	28
Italy	149	Italy	72	Italy	77
Netherlands	15	Netherlands	6	Netherlands	9
Spain	35	Spain	12	Spain	23
Other EU (As at 2001)	135	Other EU (As at 2001)	59	Other EU (As at 2001)	76
Eastern Europe	490	Eastern Europe	244	Eastern Europe	254
North Africa	47	North Africa	21	North Africa	26
Central and Western Africa	173	Central and Western Africa	77	Central and Western Africa	96
South and Eastern Africa	749	South and Eastern Africa	358	South and Eastern Africa	391
Cyprus	978	Cyprus	484	Cyprus	494
Other Middle East	108	Other Middle East	65	Other Middle East	43
Far East	167	Far East	67	Far East	100
South Asia	647	South Asia	289	South Asia	358
North America	301	North America	131	North America	170
South America	61	South America	15	South America	45
Oceania	70	Oceania	26	Oceania	44
Residents by Migration		All persons in private households		All persons, in private households, with good health	