

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.

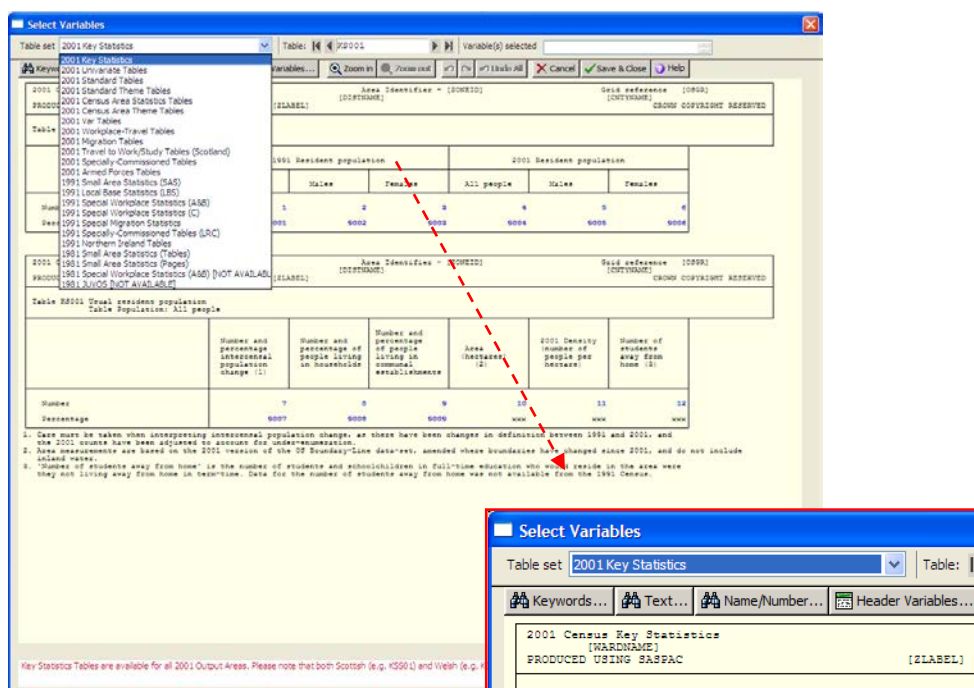


Figure 8: Selection of dataset to be searched

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.

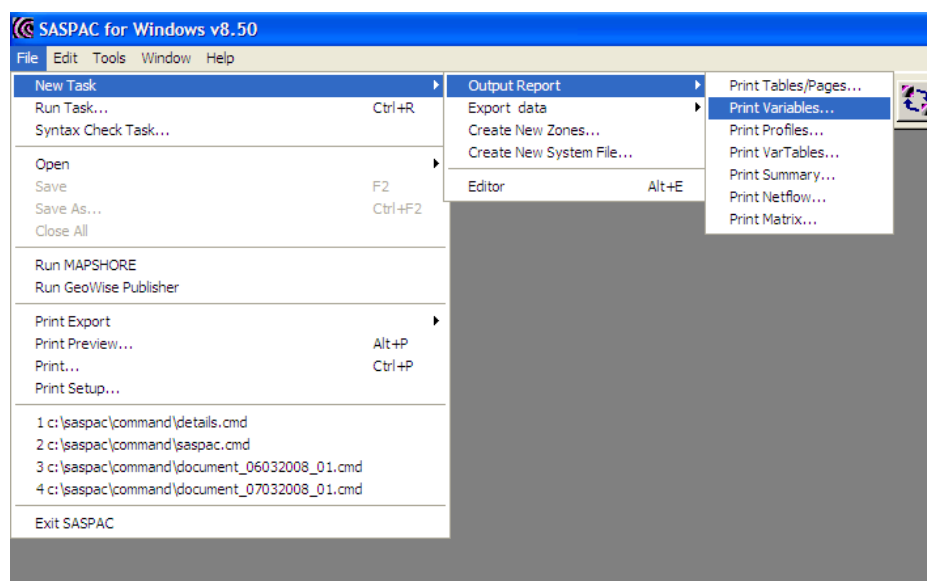
Header Variables: In addition to the count variables that are available through the tables, there is a set of 'Header variables' as defined in CE-05 of the User Manual: accessed through the SASPAC Help file menu. If any of these are required for use in a task, this Tab can be used to access them.

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.

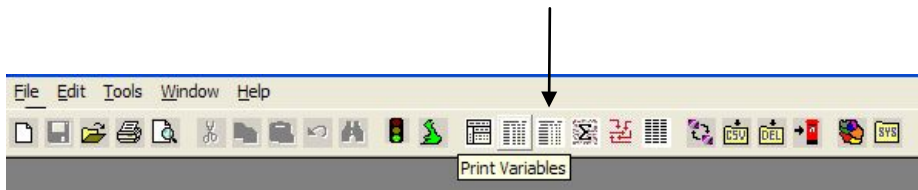
Example

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 males who work 15 hours or less, and the number of Gypsy & Irish Travellers with very good health. Include the ward 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.

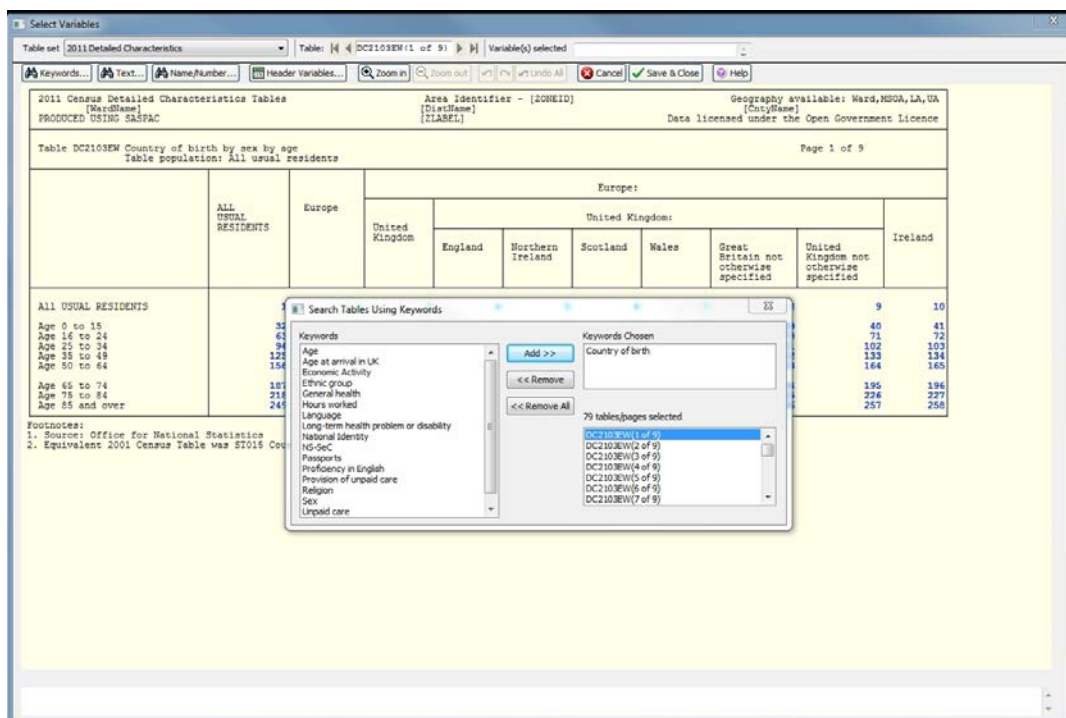


Alternatively, instead of using the menu route, the user can make use of the button bar to select certain tasks. The 'Print Variables' button is shown in this image.



- 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.



- 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.

The screenshot shows the 'Select Variables' dialog box in SAS. The table displayed is 'Table DC2103EW Country of birth by sex by age' with a population of all usual residents. The table has 7 columns representing different regions: ALL USUAL RESIDENTS, Europe, United Kingdom, England, Northern Ireland, Scotland, Wales, and Great Britain/other special. The rows represent age groups from 0 to 15 up to 85 and over. The variable 'Scotland' is selected, indicated by a red tick in the 'Variable(s) selected' box at the top of the dialog.

	Europe:							
	ALL USUAL RESIDENTS	United Kingdom:					Wales	Grea Brit othe spec
		England	Northern Ireland	Scotland	Wales	Grea Brit othe spec		
All USUAL RESIDENTS	1	2	3	4	5	6	7	
Age 0 to 15	32	33	34	35	36	37	38	
Age 16 to 24	63	64	65	66	67	68	69	
Age 25 to 34	94	95	96	97	98	99	100	
Age 35 to 49	125	126	127	128	129	130	131	
Age 50 to 64	156	157	158	159	160	161	162	
Age 65 to 74	187	188	189	190	191	192	193	
Age 75 to 84	218	219	220	221	222	223	224	
Age 85 and over	249	250	251	252	253	254	255	

Footnotes:
 1. Source: Office for National Statistics
 2. Equivalent 2001 Census Table was ST015 Country of Birth

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.

Table set: 2011 Detailed Characteristics | Table: DC2104EW (2 of 9) | Variable(s) selected

Area Identifier - [ZONEID] | Geography available: Ward,MSOA,LA,UA [DistName] [CntyName] [ZLABEL] | Data licensed under the Open Government Licence

Usage by sex by age | Population: All usual residents aged 3 and over | Page 2 of 9

Arabic	West/Central Asian Language	South Asian Language	South Asian Language:					
			Panjabi	Urdu	Bengali (with Sylheti and Chatgaya)	Gujarati	Tamil	Any other South Asian Language
10	11	12	13	14	15	16	17	18
33	34	35	36	37	38	39	40	41
56	57	58	59	60	61	62	63	64
79	80	81	82	83	84	85	86	87
102	103	104	105	106	107	108	109	110
125	126	127	128	129	130	131	132	133
148	149	150	151	152	153	154	155	156
171	172	173	174	175	176	177	178	179
194	195	196	197	198	199	200	201	202

Additional Statistics
Language were first collected in the 2011 Census.

- 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.

Table set: 2011 Detailed Characteristics | Table: DC2104EW (2 of 9) | Variable(s) selected

Area Identifier - [ZONEID] | Geography available: Ward,MSOA,LA,UA [DistName] [CntyName] [ZLABEL] | Data licensed under the Open Government Licence

Usage by sex by age | Population: All usual residents aged 3 and over | Page 2 of 9

Arabic	West/Central Asian Language	South Asian Language	South Asian Language:					
			Panjabi	Urdu	Bengali (with Sylheti and Chatgaya)	Gujarati	Tamil	Any other South Asian Language
AGE 3 AND OVER		12	13	14	15	16	17	18
Age 3 to 15		35	36	37	38	39	40	41
Age 16 to 24		50	59	60	61	62	63	64
Age 25 to 34		81	82	83	84	85	86	87
Age 35 to 49		104	105	106	107	108	109	110
Age 50 to 64		127	128	129	130	131	132	133
Age 65 to 74		150	151	152	153	154	155	156
Age 75 to 84		173	174	175	176	177	178	179
Age 85 and over		196	197	198	199	200	201	202

Additional Statistics
Language were first collected in the 2011 Census.

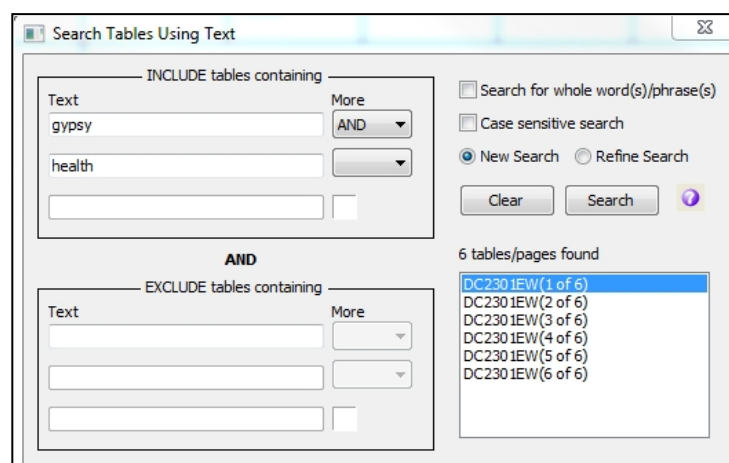
- The presence of the row and column headings now allows us to select the appropriate counts. As the two that are required are adjacent, we can use the 'rubber-banding' technique that makes use of the right mouse button again. First the left button is clicked and a drag operation is used to draw a rectangle around the required cells. The right mouse button is then used to bring up a menu as shown below.

South Asian Language	South Asian Language:			
	Panjabi	Urdu	Bengali (with Sylheti and Chatgaya)	Gur
	12	13	14	15
	35	36	37	38
			60	61
			83	84
			106	107
			129	130
			152	153
			175	176
			198	199

- 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.
- 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.

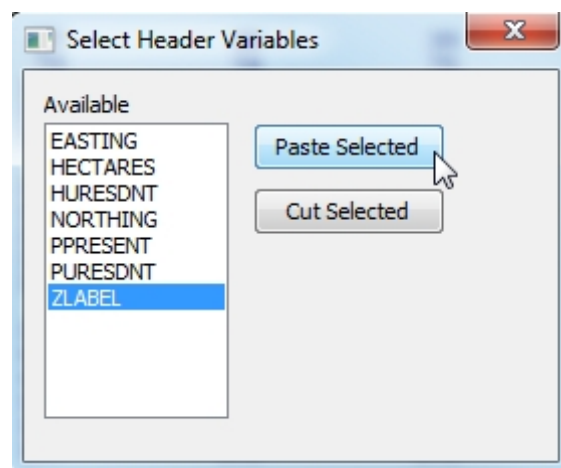
An alternative to searching by keyword which is especially useful when the appropriate keyword is not known, is to use the 'Text...' button in the 'Select Variables' window as is shown in the following image. It can be seen that this window allows for multiple searches on the text contained within the 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.

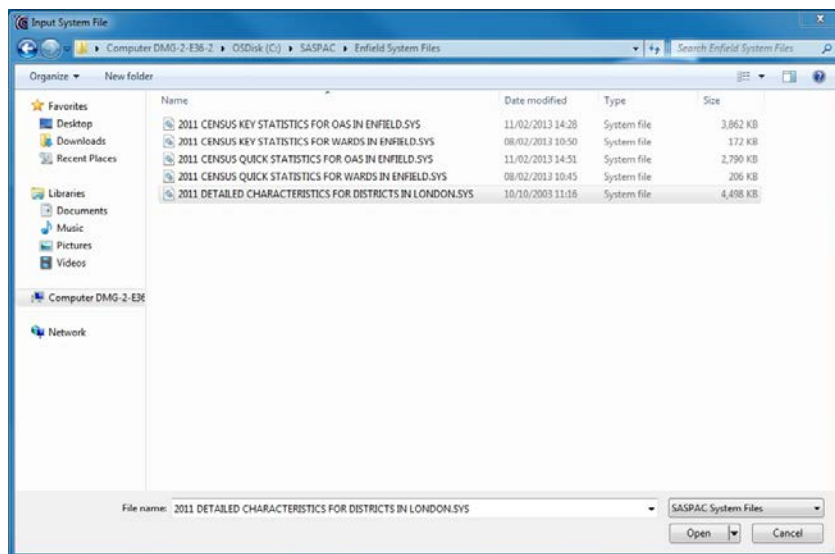
The five variables which we have identified are DC2103EW0006, DC2104EW0152, DC2104EW0175, DC6302EW0113, DC2301EW0146 and having selected these, we could return to select the appropriate input System Files. To return to the 'Print Variables' main task window, we use the 'Save & Close' button below the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



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.

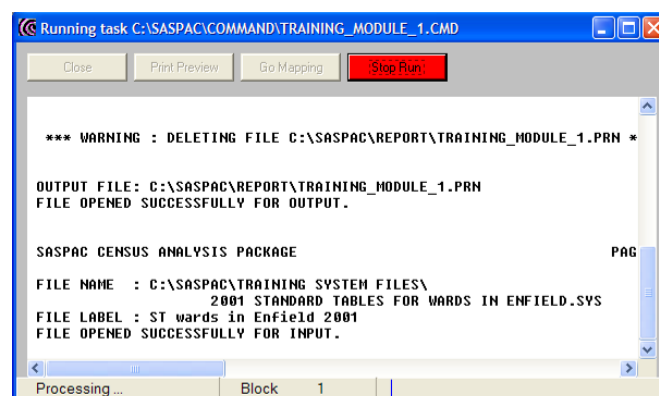
We can now select and Open the System file containing the dataset of interest (2011 Detailed Characteristics tables) for the geography of interest, which in this case is local authorities in London.



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.COMD, 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.

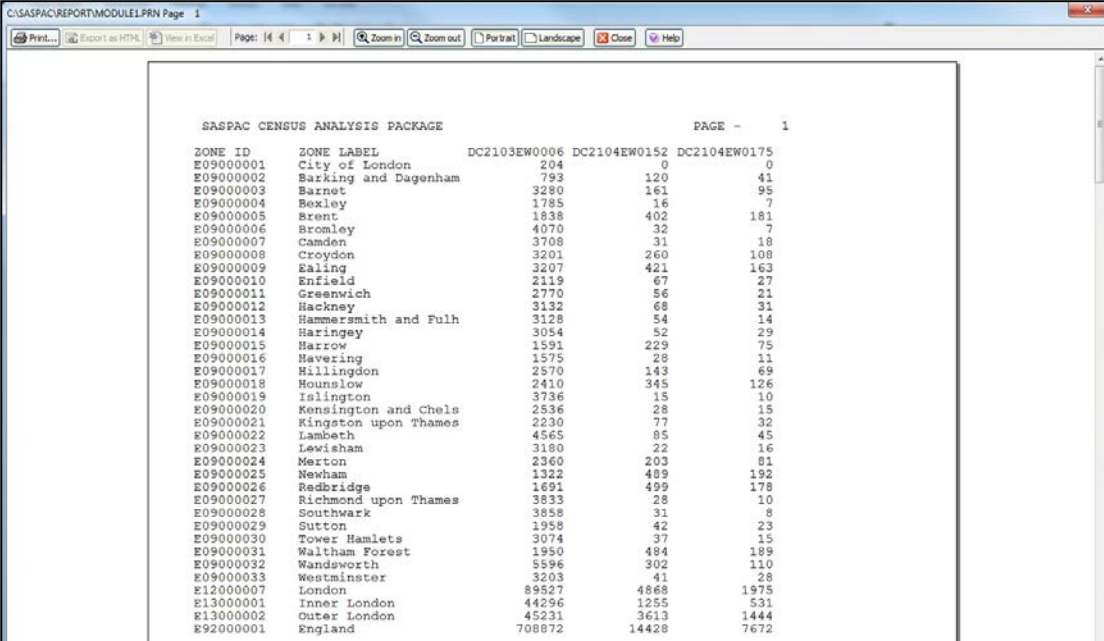
12. When we are satisfied that all is as required, the 'OK' Command Button is used to run the task. Otherwise the 'Cancel' Command Button may be used to return to the 'Print Variables' window to make further selections or amendments. When the task is run, the 'Running Task' window echoes the command 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:

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

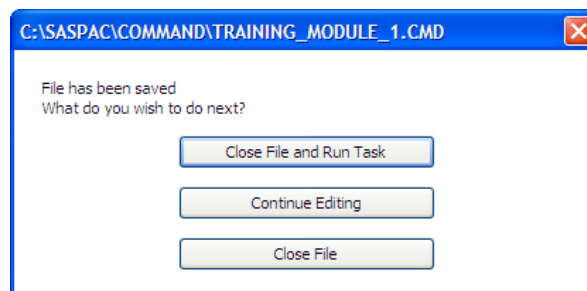
15. When the task is complete, the 'Print Preview' Command Button may be used to examine the output. The effect of this is to display the output as shown in the first image below. The 'Print Preview' option is also available through the 'File' menu option. Also available through this option is the 'Print 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.



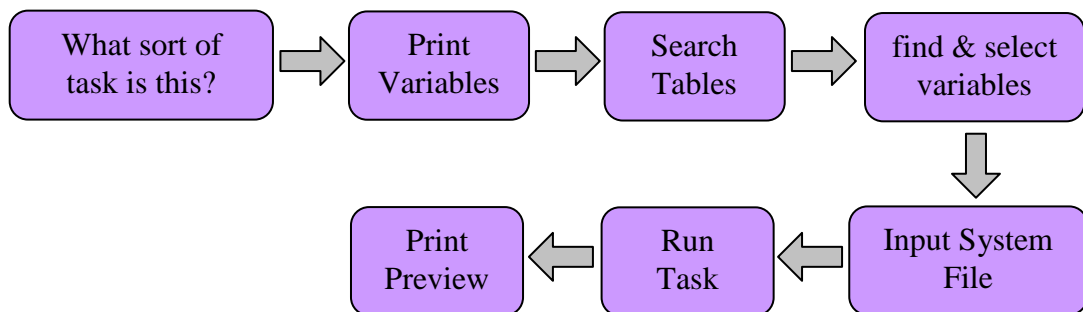
The screenshot shows a window titled 'C:\SASPAC\REPORT\MODULE1.PRN Page 1'. The window contains a table of data with the following columns: ZONE ID, ZONE LABEL, DC2103EW0006, DC2104EW0152, and DC2104EW0175. The data is as follows:

ZONE ID	ZONE LABEL	DC2103EW0006	DC2104EW0152	DC2104EW0175
E09000001	City of London	204	0	0
E09000002	Barking and Dagenham	793	120	41
E09000003	Barnet	3280	161	95
E09000004	Bexley	1785	16	7
E09000005	Brent	1838	402	181
E09000006	Bromley	4070	32	7
E09000007	Camden	3708	31	18
E09000008	Croydon	3201	260	108
E09000009	Ealing	3207	421	163
E09000010	Enfield	2119	67	27
E09000011	Greenwich	2770	56	21
E09000012	Hackney	3132	68	31
E09000013	HammerSmith and Fulh	3128	54	14
E09000014	Haringey	3084	52	29
E09000015	Harrow	1591	229	75
E09000016	Havering	1575	28	11
E09000017	Hillingdon	2570	143	69
E09000018	Hounslow	2410	345	126
E09000019	Islington	3736	15	10
E09000020	Kensington and Chels	2536	28	15
E09000021	Kingston upon Thames	2230	77	32
E09000022	Lambeth	4565	85	45
E09000023	Lewisham	3180	22	16
E09000024	Merton	2360	203	81
E09000025	Newham	1322	489	192
E09000026	Redbridge	1691	499	178
E09000027	Richmond upon Thames	3833	28	10
E09000028	Southwark	3858	31	8
E09000029	Sutton	1958	42	23
E09000030	Tower Hamlets	3074	37	15
E09000031	Waltham Forest	1950	484	189
E09000032	Wandsworth	5596	302	110
E09000033	Westminster	3203	41	28
E12000007	London	89527	4868	1975
E13000001	Inner London	44296	1255	531
E13000002	Outer London	45231	3613	1444
E92000001	England	708872	14428	7672

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.



Summary of task sequence



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