



Constructing the MDM file from raw data

PC users may construct the MDM file directly from different types of input files including SPSS, ASCII, SAS, SYSTAT, and STATA, or indirectly from many additional types of data file formats through the third-party software module included in the HLM program.

Non-PC users may construct the MDM file with one of the following types of input files: ASCII data files, SYSTAT data files, or SAS V5 transport files.

In order for the program(s) to correctly read the data, the IDs need to conform to the following rules:

1. For ASCII data the ID variables must be read in as character (alphanumeric). These IDs are indicated by the A field(s) in the format statement. For all other types of data, the ID may be character or numeric.
2. The level-1 cases must be grouped together by their respective level-2 unit ID. To assure this, sort the level-1 file by the level-2 ID field prior to entering the data into HLM2.
3. If the ID is numeric, it must be in the range $-(10^{13} + 1)$ to $+(10^{13} + 1)$ (*i.e.* 12 digits). Although the ID may be a floating point number, only the integer part is used.
4. If the ID variable is character, the length must not exceed 12 characters. Furthermore, the IDs at a given level must all be the same length. *This is often a cause of problems.* For example, imagine your data has IDs ranging from "1" to "100". You will need to recreate the IDs as "001" to "100". In other words, all spaces (blank characters) should be coded as zeros.
5. For non-ASCII files, the program can only properly deal with numeric variables (with the exception of character ID variables). Other data types, such as a "Date format", will not be processed properly.
6. For non-ASCII files with missing data, one should only use the "standard" missing value code. Some statistical packages (SAS, for example) allow for a number of missing value codes. The HLM modules are incapable of understanding these correctly, thus these additional missing codes need to be recoded to the more common "." (period) code.

SPSS file input

We first illustrate the use of SPSS file input. Data input requires a level-1 file and a level-2 file.

Level-1 file. For our HS&B example data, the level-1 file (HSB1.SAV) has 7,185 cases and four variables (not including the SCHOOL ID). The variables are:

- MINORITY, an indicator for student ethnicity (1 = minority, 0 = other)
- FEMALE, an indicator for student gender (1 = female, 0 = male)
- SES, a standardized scale constructed from variables measuring parental education, occupation, and income
- MATHACH, a measure of mathematics achievement

Data for the first ten cases in HSB1.SAV are shown in Fig. 1.

Note: level-1 cases must be grouped together by their respective level-2 unit ID. To assure this, sort the level-1 file by the level-2 unit ID field prior to entering the data into HLM2.

	id	minority	female	ses	mathach
1	1224	0	1	-1.528	5.876
2	1224	0	1	-.588	19.708
3	1224	0	0	-.528	20.349
4	1224	0	0	-.668	8.781
5	1224	0	0	-.158	17.898
6	1224	0	0	.022	4.583
7	1224	0	1	-.618	-2.832
8	1224	0	0	-.998	.523
9	1224	0	1	-.888	1.527
10	1224	0	0	-.458	21.521

Figure 1 First ten cases in HSB1.SAV

Level-2 file. At level 2, the illustrative data set HSB2.SAV consists of 160 schools with 6 variables per school. The variables are:

- SIZE (school enrollment)
- SECTOR (1 = Catholic, 0 = public)
- PRACAD (proportion of students in the academic track)
- DISCLIM (a scale measuring disciplinary climate)
- HIMNTY (1 = more than 40% minority enrollment, 0 = less than 40%)
- MEANSES (mean of the SES values for the students in this school who are included in the level-1 file)

The data for the first ten schools are displayed in Fig 2.

	id	size	sector	pracad	disclim	himinty	meansas
1	1224	842	0	.350	1.597	0	-.428
2	1288	1855	0	.270	.174	0	.128
3	1296	1719	0	.320	-.137	1	-.420
4	1308	716	1	.960	-.622	0	.534
5	1317	455	1	.950	-1.694	1	.351
6	1358	1430	0	.250	1.535	0	-.014
7	1374	2400	0	.500	2.016	0	-.007
8	1433	899	1	.960	-.321	0	.718
9	1436	185	1	1.000	-1.141	0	.569
10	1461	1672	0	.780	2.096	0	.683

Figure 2 First ten cases in HSB2.SAV

As mentioned earlier, the construction of an MDM file consists of three major steps. This will now be illustrated with the HS&B example.

To inform HLM of the input and MDM file type

1. At the **WHLM** window, open the **File** menu.
2. Choose **Make new MDM file...Stat package input** (see Figure 3). A **Select MDM type** dialog box opens (see Figure 4).
3. Select **HLM2** and click **OK**. A **Make MDM - HLM2** dialog box will open (see Figure 5).

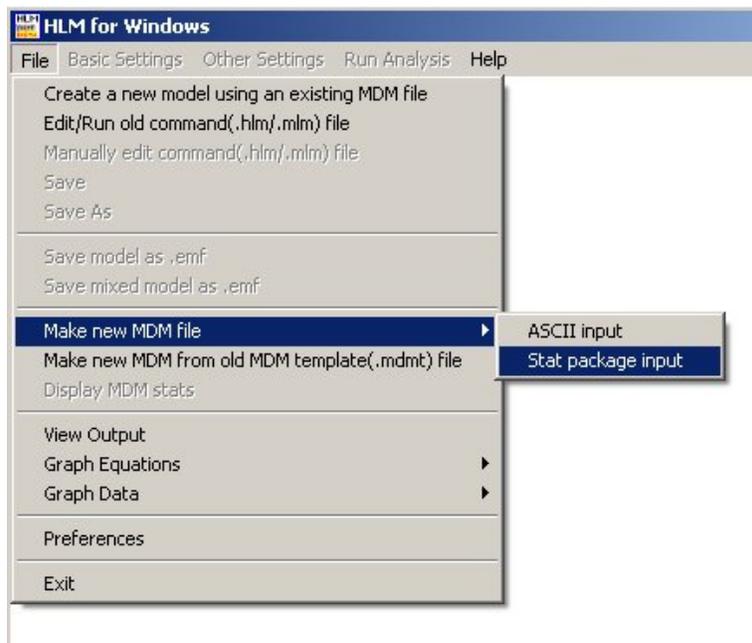


Figure 3 WHLM window

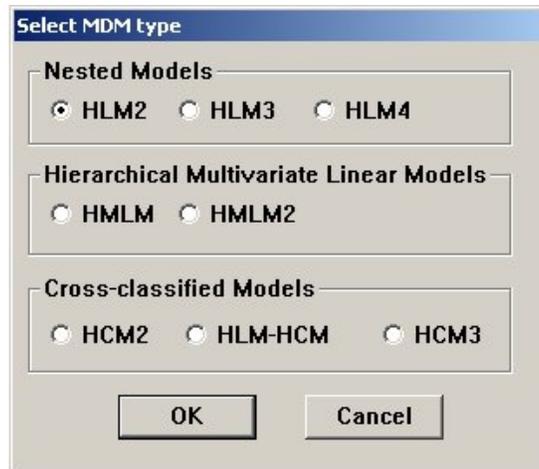


Figure 4 Select MDM type dialog box

To supply HLM with appropriate information for the data, the command, and the MDM files:

1. Select **SPSS/Windows** from the **Input File Type** pull-down menu (see Figure 5).
2. Specify the structure of data. The three choices are cross-sectional, longitudinal, and measures within groups. The data in HSB1.SAV are cross-sectional.
3. Click **Browse** in the **Level-1 Specification** section to open an **Open Data File** dialog box.
4. Open a level-1 SPSS system file in the HLM folder (HSB1.SAV in our example). The **Choose Variables** button will be activated.

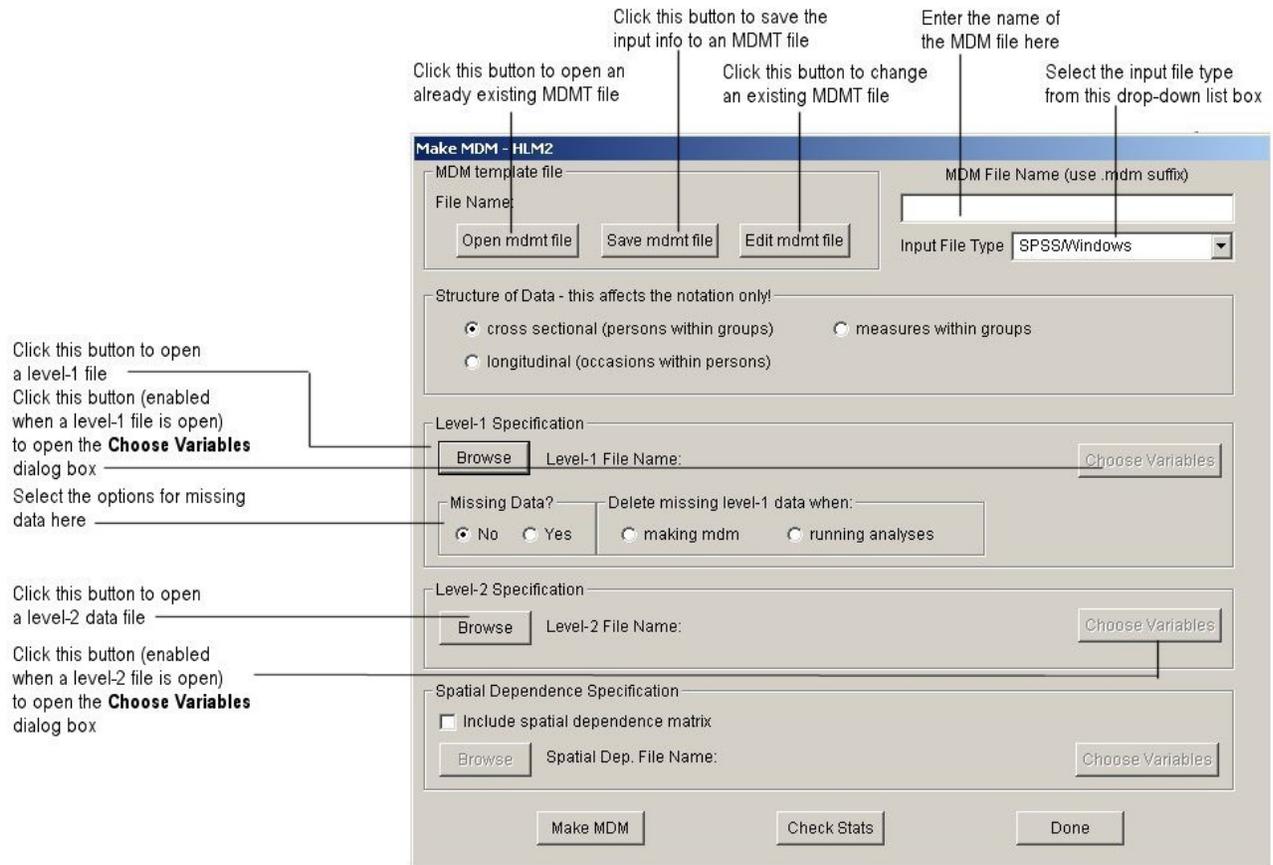


Figure 5 Make MDM - HLM2 dialog box

5. Click **Choose Variables** to open the **Choose Variables - HLM2** dialog box and choose the ID and variables by clicking the appropriate check boxes (See Figure 6). To deselect, click the box again.
6. Select the options for missing data in the level-1 file (there is no missing data in HSB1.SAV).
7. Click the selection button for **measures within persons** for the **type of nesting of input data** if the level-1 data consist of repeated measures or item responses. With this selection, WHLM will use in its displays and output model notations that match those used in *Hierarchical Linear Models* for studies on individual change and latent variables. The default type is **persons within groups**. It is generally used when the level-1 data are comprised of cross-sectional measures. With this option, WHLM will use model notations that correspond to those used for applications in organization research.
8. Click **Browse** in the **Level-2 specification** section to open an **Open Data File** dialog box.
9. Open a level-2 SPSS system file in the HLM folder (HSB2.SAV in our example). The **Choose Variables** button below **Browse** will be activated.
10. Click **Choose Variables** to open the **Choose Variables - HLM2** dialog box and choose the ID and variables by clicking the appropriate check boxes (see Figure 7).

11. Check the box **include spatial dependence matrix** to specify spatial dependence, if applicable (see Section 11.4 for details). The **Spatial Dependence Specification** box should only be used if you have spatial dependence data and wish to run this kind of model.
12. Enter a name for the MDM file in the **MDM file name** box (for example, HSB.MDM).

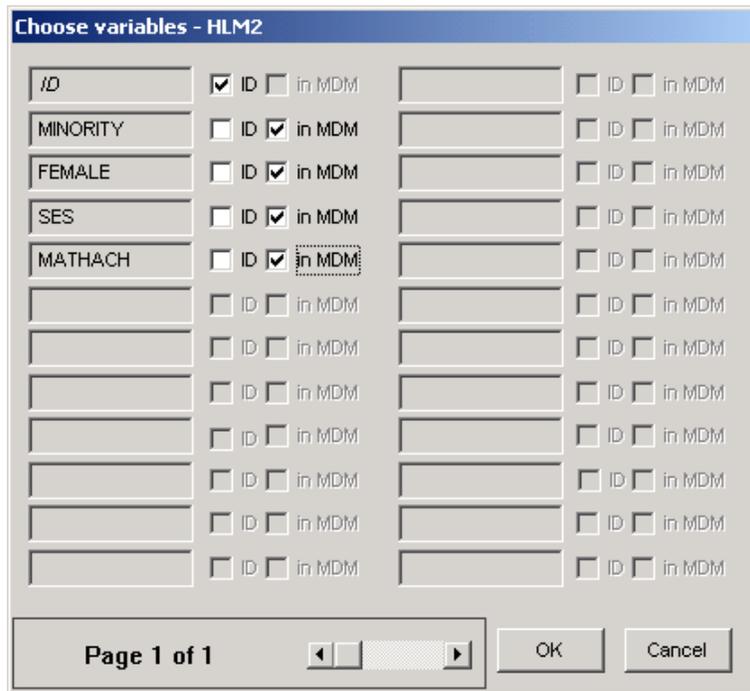


Figure 6 Choose Variables - HLM2 dialog box for the level-1 file, HSB1.SAV

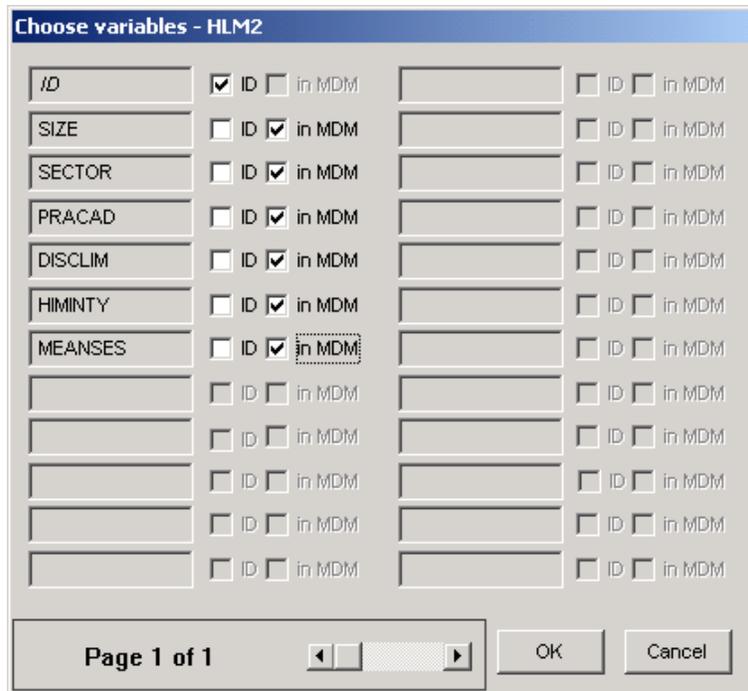


Figure 7 Choose variables - HLM2 dialog box for the level-2 file, HSB2.SAV

13. Click **Save mdmt file** in the **MDM template file** section to open a **Save MDM template file** dialog box. Enter a name for the MDMT file (for example, HSBSPSS.MDMT). Click **Save** to save the file. The command file saves all the input information entered by the user. It can be re-opened by clicking the **Open mdmt file** button (see Figure 5). To make changes to an existing MDMT file, click the **Edit mdmt file** button.
14. Note that HLM will also save the input information into another file called CREATMDM.MDMT when the MDM is created.
15. Click the **Make MDM** button. A screen displaying the prompts and responses for MDM creation will appear.

LEVEL-1 DESCRIPTIVE STATISTICS					
VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
MINORITY	7185	0.27	0.45	0.00	1.00
FEMALE	7185	0.53	0.50	0.00	1.00
SES	7185	0.00	0.78	-3.76	2.69
MATHACH	7185	12.75	6.88	-2.83	24.99

LEVEL-2 DESCRIPTIVE STATISTICS					
VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
SIZE	160	1097.83	629.51	100.00	2713.00
SECTOR	160	0.44	0.50	0.00	1.00
PRACAD	160	0.51	0.26	0.00	1.00
DISCLIM	160	-0.02	0.98	-2.42	2.76
HIMINTY	160	0.28	0.45	0.00	1.00
MEANSES	160	-0.00	0.41	-1.19	0.83

Figure 8 Descriptive Statistics for the MDM file, HSB.MDM

To check whether the data have been properly read into HLM

- 2 When the screen disappears, the level-1 and level-2 descriptive statistics will automatically be displayed (See Figure 8). Pay particular attention to the N column. It is not an uncommon mistake to forget to sort by the ID variable, which can lead to a lot (or most) of the data not being processed. Close the Notepad window when done. Use the **Save As** option to give it a new name if later use of this file is anticipated. The file can also be opened by clicking on the **Display Stats** button.
- 3 Click **Done**. The **WHLM** window displays the type and name on its title bar (**hlm2 & HSB.MDM**) and the level-1 variables on a drop-down menu (See Figure 9).



Figure 9 WHLM: hlm2 MDM File window for HSB.MDM