Corrections for the printed version of SAS Press Book 60502 – Data Preparation for Analytics © 2006

Version 1.3 per June 12th , 2008
Version 1.4 per October 10th, 2008
Version 1.5 per December 16th, 2008
Version 2.1 per July 20th, 2010
Version 2.2 per July 25th, 2011

For the most actual version of this document and the most recent SAS programs, please check out: http://www.sascommunity.org/wiki/Data_Preparation_for_Analytics

Note that all code changes, have already been made to the newest version of the macros and programs that can be downloaded.

The following corrections and enhancements have been made after printing

Page	Navigation	Correction
xxi	Acknowledgements last paragraph	Julie Platt, Donna Faircloth, **Mary Beth Steinbach**, and the team from SAS Press supported me.
6	First Bullet Point at top of the page	"… NOTE that are dealing mostly …." Instead of NOT that we are dealing
8	Second half of the page	"What is the" comes already with the first bullet, can be removed from the second level bullet points
17	Middle of the page	Refer to chapter 23 Scoring and Automation, instead of Chapter 14
34	Chapter 4.6.1	Exchange SIMON to DANIELE at the beginning of the first 2 paragraphs
35	Chapter 4.6.6	Exchange SIMON to ELIAS in the last sentence of the 2nd paragraph
58	Table 7.6	Header should be: Multiple-Row-per-Subject data mart for PATIENT and LABORATORY_VALUES data
72	Chapter 9.3.2	Reference should be to chapter 14, instead of chapter 12
91	Chapter 11.2.3 First Bullet Point	Note that only on Windows and UNIX machines numeric variables have a length of 3 bytes at least, it is different on Mainframes.
118	Macro Code	Note that the macro has been extended, and the new downloadable version supports more than on VAR variable.
119 123	Middle of the page	Note that there is a new set of macros %MAKELONG and %MAKEWIDE, that can transpose more than one variable in one go. See also SAS Tipps and Tricks under http://support.sas.com/kb/32/122.html and http://support.sas.com/kb/32/121.html

121	Code segment in the 2nd half of the page	The $ sign in the line
		
		`time = INPUT(TRANWRD(_measure,"weight",''),$8.);`
		
		Shall be removed. The new version shall look like this:
		`time = INPUT(TRANWRD(_measure,"weight",''),8.);`

122	Macro Code	Note that the macro has been extended, and the new downloadable version supports

122	Macro Definition of %MAKELONG	more than on VAR variable.

122	Macro Definition of %MAKELONG	The $ sign in the line
		
		`&Measurement = INPUT(TRANWRD(_measure,upcase("&root"),''),$8.);`
		
		Shall be removed. The new version shall look like this:
		
		`&Measurement = INPUT(TRANWRD(_measure,upcase("&root"),''),8.);`

124	Table 14.8 and code on top of next page	Table 14.8 should contain an additional variable:
		
		Note that the Code has to contain a VAR statement:
		
		`PROC TRANSPOSE DATA = assocs(obs=21)`
		`                OUT = assoc_tp (DROP = _name_);`
		`  BY customer;`
		`  ID Product;`
		`  VAR Count;`
		`RUN;`

127	Middle of the page	It should be `IF prod{i} = . THEN prod{i}=0;` instead of `IF prod{i}  . THEN prod{i}=0;` The equal sign is missing
127	Bottom	Macro parameter is RPLC (not RPCL)
142	Upper half of the page	At the 2 macro calls of %INTERACT, the parameter QUADR should be omitted. `%INTERACT(age weight height, quadr=1);` and `%INTERACT(age weight height, quadr=0);` is correct. Note that the code in the SAS program is correct.

142	Section 16.2.3	The variable-list in the SUM and MEAN function has to be separated by commas;
		`SUM_USAGE = SUM(USAGE1, USAGE2, USAGE3);`
		`MEAN_USAGE = MEAN(USAGE1, USAGE2, USAGE3);`

173	Bottomo	Obs 5 in the table: actual_tariff should be ADVANCED (not ADVANED)

149	Syntax of the calculation of age	In the book I reference a variable BIRTHDATE instead of '16MAY1970'd. The correct syntax is:
		`AGE = FLOOR((INTCK('MONTH','16MAY1970'd,&current_date)-`
		`(DAY(&current_date) < DAY('16MAY1970'd)) ) / 12);`

224	Lower half of the page	In the definition of WoE is should read: WoE(category) = log(p(event,category)/p(non-event,category)) instead of WoE(category) = log(p(event,category)/p(non-event/category))
228	Macro Target Chart	6th line of the macro definition should be %IF **&class** NE %THEN … instead of %IF &interval NE %THEN 8th line of the macro definition should be %IF **&interval** NE %THEN … instead of %IF &class NE %THEN In the electronic downloadable macro file, the code has already been changed.
237	Middle of the page	The code should be `FEATURE = CATX('=',`**`key`**`,value);` instead of `FEATURE = CATX('=',`**`_key_`**`,value);`
250	Middle of the page	The reference should be to Appendix B instead of Appendix A.

260	First Sentence	It should be 6.000 Observations instead of 10.000 Observations

285
286
288
		In the macro definition
		`%RememberCategories`
		`%CheckCategories`
		`%RememberDistribution`
		`%CheckDistribution`
		the default value of the parameter LIB shall be changed from WORK to SASUSER

329	Figure 27.1	In the ORDERS Table, the PK,FK3 should be OrderTypeID (instead of OrderType)