

Contents

Module 1	2
1 Welcoming from the Course Author	2
1.1 Véronique Bourcier	2
1.1.1 University	2
1.1.2 Working as a SAS Programmer	3
1.1.3 Blogging on SAS Programming	4
1.1.4 xxformat GmbH	4
1.2 Planning Your Training	7
1.2.1 Focusmate	7
1.2.2 Weekly Meeting	7
1.2.3 Taking Notes	7
2 Identifying the Version of Your SAS® System and Other Useful Information for Your CV	8
2.1 The Story of the SAS® System	8
2.2 Four Technical Questions for Your Resume	10
2.2.1 Introduction	10
2.2.2 Which Version is Available?	11
2.2.3 Which Interface is Used?	12
2.2.4 Which Modules are Available?	12
2.2.5 Summary	13
2.2.6 This Course	14
3 Running your First Program in SAS® Studio and Getting Familiar with Your Programming Interface	15

3.1	Creating and Save a Program	16
3.1.1	Opening a New Program	16
3.1.2	Typing In the Program	16
3.1.3	Saving the Program	18
3.1.4	The <code>CODE</code> Tab	19
3.2	Running a Program and Analyzing the Results	19
3.2.1	Running a Program	19
3.2.2	The <code>LOG</code> Tab	20
3.2.3	A New Table in the Tree Structure	21
3.2.4	The <code>RESULTS</code> Tab	23
3.2.5	The <code>OUTPUT DATA</code> Tab	23
3.3	Discovering the Buttons Around the Program	24
3.4	Customizing the Preferences	26
3.4.1	<code>CODE</code> and <code>LOG</code> Tab Preferences	26
3.4.2	<code>RESULTS</code> and <code>OUTPUT DATA</code> Tabs Preferences	27
3.5	Managing Folders and Files	28
3.5.1	<code>Files (Home)</code> Folder	28
3.5.2	What's the Link between Files on SAS® Studio and on Your Computer?	28
3.5.3	Creating a New Folder	29
3.6	Creating a Folder Shortcut	30
3.7	Which Path Should Be Used in SAS® Studio	32
3.8	Five Sections in the Left Menu	33
3.9	<code>work</code> and <code>sashelp</code> , Two Libraries Available by Default	34
3.9.1	Overview	34
3.9.2	The <code>work</code> Library	34
3.9.3	The <code>sashelp</code> Library	35
3.10	Files and Folders vs. Libraries	35
3.11	How to Switch from English to German in SAS® Studio	36
4	Saving a Dataset Permanently and Understanding the Concepts of Member and Library	37
4.1	Table and SAS Dataset	37

4.2	Differences Between a Dataset and a View	38
4.3	Assigning, Using and Disassigning a Library	38
4.3.1	Assigning a Library	39
4.3.2	Using a Library	39
4.3.3	Disassigning a Library	40
4.4	Working with the <code>work</code> Library	41
4.5	SAS Files vs. External Files	42
4.5.1	Files Specific to the SAS® System	42
4.5.2	Different Types of Members	43
4.5.3	Specifying the Type of a Member	43
4.5.4	In Summary	43
5	Understanding the Concepts of Global Statement, Data Step, Procedure, and so on	44
5.1	Typing the Program in your Editor	44
5.2	Global Statements, Data Steps and Procedures	46
5.2.1	Global Statements	47
5.2.2	Data Steps	48
5.2.3	Procedures	48
5.3	Colors Used in the Editor	49
5.3.1	Words Without Color	49
5.3.2	Keywords in Blue	49
5.3.3	Values in Quotes	50
5.3.4	Numeric Values	50
5.3.5	Comments	50
5.4	Choosing Between Single or Double Quotes	50
5.4.1	Using Single Quotes Unless Double Quotes are Needed	50
5.4.2	When are Double Quotes Needed?	50
5.5	Choosing Between Upper- or Lower-Case	51
5.6	Indentations	52
5.6.1	Indentation Width	52
5.6.2	How Many Blanks For One Tabulation?	52
5.6.3	Converting Tabulations Into Blanks	52

5.7	Adding Comments	53
5.7.1	Two Ways to Use Comments	53
5.7.2	Two Ways to Create a Comment	53
5.7.3	Adding a Header and Numbering Sections	54
5.8	Creating, Calling and Deleting a Macro Variable	57
5.8.1	Introduction	57
5.8.2	Creating a Macro Variable	57
5.8.3	Displaying the Value of a Macro Variable in the Log . .	58
5.8.4	Using a Macro Variable in the Definition of Another Macro Variable	58
5.8.5	Using a Macro Variable in Quotes	59
5.8.6	Deleting a Macro Variable	59
5.9	Creating and Calling a Macro	60
5.9.1	Data	60
5.9.2	Program Before	60
5.9.3	Program After: Macro Definition	60
5.9.4	Enabling Macro Specific Global Options	61
5.9.5	Program After: Macro Call	62
5.10	Displaying and Modifying Global Options	62
5.10.1	Creating a Text File	62
5.10.2	A Few Global Options	63
5.10.3	Displaying Global Options with <code>proc options</code> . . .	63
5.10.4	Changing Some Global Options	64
Module 2		66
6	Creating a Dataset and a View - Methods Overview	66
6.1	Creating a Dataset in a Data Step using the <code>output</code> Statement	66
6.1.1	Example with One Observation	66
6.1.2	Example with Several Observations	67
6.2	Creating a Dataset: review of a Few Methods	68

6.2.1	Creating a Dataset in a Data Step using the <code>input</code> Statement	68
6.2.2	Creating a Dataset using <code>proc sql</code>	69
6.3	Creating a Dataset from Another Dataset	70
6.3.1	Both Read and Created Datasets Have the Same Name	70
6.3.2	Both Read and Created Datasets Have Different Names	70
6.3.3	Creating a Dataset in a Different Library	71
6.4	Renaming, Copying and Moving a Dataset with <code>proc datasets</code>	71
6.4.1	Renaming a Dataset	72
6.4.2	Copying a Dataset to Another Library	72
6.4.3	Moving a Dataset to Another Library	73
6.5	Creating a View	73
6.5.1	General Considerations	73
6.5.2	Creating a View	74
7	Working with Table and Variable Attributes	76
7.1	General Considerations	76
7.1.1	Data Step vs. <code>proc datasets</code>	76
7.1.2	Data	76
7.1.3	Listing Table and Variable Attributes and their Value .	77
7.2	Table Attributes	78
7.3	Variable Attributes List	80
7.4	Variable Name	80
7.4.1	Defining the Variable Name	80
7.4.2	Renaming a Variable	82
7.5	Variable Type	83
7.5.1	Definition	83
7.5.2	The Two Types used in the SAS Language	83
7.5.3	Changing the Type of the Variable	84
7.6	Defining the Variable Length	84
7.6.1	General Considerations	84
7.6.2	Character Variable Length	85
7.6.3	Numeric Variable Length	85

7.6.4	Defining the Length of a Variable in a Data Step	86
7.6.5	What Happens When the Length of a Character Variable is Not Explicitly Defined?	86
7.7	Changing the Variable Length	88
7.7.1	Increasing the Variable Length	88
7.7.2	Reducing the Variable Length	88
7.8	Variable Label	90
7.8.1	Definition	90
7.8.2	Temporarily Assigning a Label to a Variable	90
7.8.3	Permanently Assigning a Label to a Variable	91
7.8.4	Temporarily Disassigning a Permanent Label	92
7.8.5	Permanently Disassigning a Permanent Label	92
7.9	Creating a Format and Displaying Its Content	93
7.9.1	Introduction	93
7.9.2	Creating Formats using a <code>value</code> Statement	94
7.9.3	Where are Formats Saved?	95
7.9.4	Displaying Format Catalog Content using <code>fmtlib</code>	96
7.9.5	Extracting Formats in a Dataset using <code>cntlout=</code>	96
7.9.6	Four Essential Elements in a Format	97
7.9.7	Creating a Format from a Dataset with <code>cntl=</code>	98
7.9.8	General Considerations on Formats Provided by the SAS [®] System	99
7.10	Assigning and Disassigning a Format to a Variable	99
7.10.1	Assigning a Format to a Variable	99
7.10.2	Temporarily Assigning a Format to a Variable	100
7.10.3	Permanently Assigning a Format to a Variable	100
7.10.4	Temporarily Disassigning a Permanent Format	101
7.10.5	Permanently Disassigning a Permanent Format	102
7.11	The Global Option <code>fmtsearch=</code>	103
7.11.1	The Problem	103
7.11.2	Listing the Libraries used by the <code>fmtsearch</code> Option	104
7.11.3	Modifying the Libraries of the <code>fmtsearch=</code> Option	105

7.12 Note on Numeric Missing Values	106
7.12.1 The List of Numeric Missing Values	106
7.12.2 Example	106
7.12.3 The Order of Missing Numeric Values	107
7.13 Creating Formats using a <code>value</code> Statement: Extra Features .	107
7.13.1 Creating Formats using a <code>value</code> Statement	107
7.13.2 Extracting Formats into a Dataset using <code>cntlout=</code> .	111
7.14 Parallel Between Format and Informat	114
7.14.1 Summary Table	114
7.14.2 Saving a Informat or Formats in a Format Catalog . .	115
7.14.3 Indicating the Type of the Informat/Format	115
7.14.4 Creating an Informat and a Format Manually	115
7.14.5 The <code>informat</code> and <code>format</code> Statement	115
7.14.6 Format and Informat Role	115
7.14.7 The <code>input</code> and <code>put</code> Functions	116
7.14.8 The <code>input</code> and <code>put</code> Statements	116
7.15 The <code>attrib</code> Statement	116
7.15.1 Defining the Attributes of the Variable	116
7.15.2 Deleting All Labels	117
7.16 The Variable Position	118
7.16.1 General Considerations	118
7.16.2 Data	118
7.16.3 Defining the Position	118
7.16.4 Changing the Position	120
7.17 Listing Table and Variable Attributes	122
7.17.1 Definition	122
7.17.2 Data	122
7.17.3 Method 1, Using the Views of the library <code>sashelp</code> . .	123
7.17.4 Method 2, Using <code>proc sql</code>	124
7.17.5 Method 3, Using <code>proc datasets/proc contents</code> (1/3)	126

7.17.6	What is the Main Difference Between the Two Procedures?	127
7.17.7	Which Method Should be Used?	128
7.17.8	Method 3, Using <code>proc datasets/proc contents</code> (2/3)	129
7.17.9	Method 3, Using <code>proc datasets/proc contents</code> (3/3)	131
7.17.10	Which Dictionaries are Available?	132
7.18	Naming Conventions	133
7.18.1	Maximum Number of Characters	133
7.18.2	Authorized Characters	134
7.18.3	Global Options	135
7.18.4	Summary Table	136
7.18.5	Windows: Words Not Allowed	136
8	Adding Titles and Footnotes	137
8.1	Data	137
8.2	Titles	137
8.2.1	Example 1: Adding a Title	137
8.2.2	Example 2: Modifying an Existing Title	138
8.2.3	Example 3: Adding Up to 10 Titles	138
8.2.4	Example 4: Deleting All Titles From....	138
8.2.5	Example 5: Deleting All Titles	139
8.3	Footnotes	139
8.3.1	Example 1: Adding Several Footnotes	139
8.3.2	Example 2: Deleting All Footnotes From....	140
8.3.3	Example 3: Deleting All Footnotes	140
9	Getting Familiar with Bit/Bytes, Hexadecimal Values and ASCII	
	Table	141
9.1	The Notions of Bit and Byte	141
9.1.1	Binary Values	141
9.1.2	Several Units of Measurement	141

9.1.3	From Bit to Byte	141
9.1.4	From Byte to Yottabyte	142
9.1.5	International Electrotechnical Commission	143
9.1.6	256 Values Stored in 1 Byte	143
9.2	Hexadecimal Values	143
9.3	The ASCII and EBCDIC Tables	144
9.3.1	General Considerations	144
9.3.2	The ASCII Table	144
9.4	Using the <code>\$hex</code> Format to Display Values in Hexadecimal . .	148

Module 3 **150**

10 Dealing with SAS Date, SAS Time and SAS Datetime **150**

10.1	Basic Concepts	150
10.1.1	SAS Date	150
10.1.2	SAS Time	152
10.1.3	SAS Datetime	153
10.1.4	Comparing the Results with -1, 0 and 1	154
10.2	Creating a SAS Date	154
10.2.1	Creating a SAS Date using the <code>today</code> and <code>date</code> Functions	154
10.2.2	Creating a SAS Date using the <code>mdy</code> Function	155
10.2.3	Creating a SAS Date using <code>'DDMMMYYYY'd</code>	155
10.2.4	Creating a SAS Date using the <code>input</code> Function	155
10.2.5	Displaying a Formatted SAS Date	156
10.3	Creating a SAS Time	158
10.3.1	Creating a SAS Time using the <code>time</code> Function	158
10.3.2	Creating a SAS Time using the <code>hms</code> Function	158
10.3.3	Creating a SAS Time using <code>'HH:MM:SS't</code>	158
10.3.4	Creating a SAS Time using the <code>input</code> Function	159
10.3.5	Displaying a Formatted SAS Time	159
10.4	Creating a SAS Datetime	160

10.4.1	Creating a SAS Datetime using the <code>datetime</code> Function	160
10.4.2	Creating a SAS Datetime using the <code>dhms</code> Function . . .	161
10.4.3	Creating a SAS Datetime using <code>'...'</code> dt	162
10.4.4	Creating a SAS Datetime using the <code>input</code> Function . .	162
10.4.5	Displaying a Formatted SAS Datetime	163
10.5	Extracting a SAS Date and a SAS Time of a SAS Datetime .	164
10.6	Extracting the Day, Month, Year and Quarter of a SAS Date .	165
10.7	Extracting the Hours, Minutes and Seconds of a SAS Time .	166
10.8	Computing the Difference Between Two SAS Dates	166
10.8.1	Example 1	166
10.8.2	Example 2	167
10.9	Managing Two-Digit Years using the Global Option <code>yearcutoff</code> =	168
10.9.1	The Problem	168
10.9.2	Using <code>yearcutoff</code> = Default Value	168
10.9.3	Updating <code>yearcutoff</code> = Value	169
11	Discovering Common Functions, Series 1	170
11.1	Data	170
11.1.1	data_01	170
11.1.2	data_02	171
11.1.3	data_03	171
11.1.4	data_04	171
11.1.5	data_05	171
11.1.6	data_06	172
11.1.7	data_07	172
11.1.8	data_08	172
11.1.9	data_09	172
11.1.10	data_10	173
11.1.11	data_11	173
11.2	<code>sum</code> Function and Alternative	173
11.2.1	Example 1	173
11.2.2	Example 2	174
11.2.3	Example 3	175

11.3	<code>mean</code> Function and Alternative	176
11.4	Finding the Minimum and Maximum: <code>min</code> and <code>max</code> Functions and Alternative	177
11.4.1	Finding the Smallest Numeric Value	177
11.4.2	Finding the Largest Numeric Value	178
11.4.3	Finding the Smallest and Largest Character Values . .	179
11.5	Computing the Number of Missing or Non-Missing Values: <code>n</code> , <code>nmiss</code> and <code>cmiss</code> Functions	180
11.5.1	Computing the Number of Non-Missing Numeric Val- ues: <code>n</code> Function	180
11.5.2	Computing the Number of Missing Numeric Values: <code>nmiss</code> Function	180
11.5.3	Computing the Number of Missing Values: <code>cmiss</code> Func- tion	181
11.6	<code>missing</code> Function and <code>call missing</code> Routine	181
11.6.1	Is it a Missing Value or Not: <code>missing</code> Function	181
11.6.2	Setting All Values to Missing: <code>call missing</code> Routine	182
11.7	Rounding Numeric Values	182
11.7.1	<code>round</code> Function	182
11.7.2	<code>round</code> Option in a Picture Format	183
11.8	Extracting the Integer Value of a Numeric Value: <code>int</code> Function	185
11.9	Extracting the Absolute Value: <code>abs</code> Function	185
11.10	Extracting Lower/Upper Integer: <code>floor</code> and <code>ceil</code> Functions	185
11.11	Extracting the Modulo: <code>mod</code> Function	187
11.11.1	Understanding the <code>mod</code> Function	187
11.11.2	Using the <code>mod</code> Function	187
11.12	Concatenating Strings: <code>cats...</code> Functions and Alternative . .	188
11.12.1	Example 1	188
11.12.2	Example 2	189
11.12.3	Example 3	190
11.12.4	Example 4	191
11.12.5	Example 5	192

11.13	Converting Successive <i>Blanks</i> into One <i>Blank</i> : <code>compbl</code> Function	192
11.14	Removing Specific Characters: <code>compress</code> Function	193
11.14.1	Program	193
11.14.2	Output	193
11.15	Replacing a Sub-String with Another One: <code>transtrn</code> Function	195
11.15.1	Replacing with a Sub-String of Same Length	195
11.15.2	Replacing with a Longer Sub-String	196
11.15.3	Replacing with a Shorter Sub-String	197
11.15.4	Removing a Sub-String	197
11.16	Extracting the xth Sub-String from a String: <code>scan</code> Function	198
11.16.1	Example 1	198
11.16.2	Exercise a	199
11.16.3	Example 2	200
11.16.4	Example 3	200
11.16.5	Example 4	201
11.16.6	Exercise b	202
11.17	<code>substr</code> and <code>substrn</code> Functions on the Right Hand Side of the Equation	203
11.17.1	<code>substr</code> Function	203
11.17.2	<code>substrn</code> Function	205
11.18	<code>substr</code> Function on the Left Hand Side of the Equation	206
11.19	<code>repeat</code> Function	207
11.20	<code>reverse</code> Function	207

Module 4 **208**

12 Selecting Observations and Variables **208**

12.1	Data	209
12.2	Keeping/Dropping Variables using Dataset Options	210
12.2.1	Using <code>keep=/drop=</code> in a Data Step	211
12.2.2	Using <code>keep=/drop=</code> in a Procedure	211

12.2.3 Using <code>keep=</code> / <code>drop=</code> in <code>proc sql</code>	211
12.3 Selecting Observations using Dataset Options	212
12.3.1 <code>where=</code> Dataset Option	212
12.3.2 Juggling Between Dataset Options	213
12.3.3 Using Two Dataset Options: <code>firstobs=</code> et <code>obs=</code>	216
12.4 Working with Operators	217
12.4.1 Discovering the Basis	217
12.4.2 Using the Logical Operators <code>and</code> and <code>or</code>	219
12.4.3 Selecting a Range of Values	221
12.4.4 Using Comparison Operator <code>like</code>	222
12.4.5 Listing All Values Starting With...	224
12.5 Selecting Observations using the <code>if</code> Statement	225
12.5.1 Understanding the Implicit <code>output</code>	225
12.5.2 Creating Several Datasets from a Single One	226
12.5.3 Conducting Several Actions	227
12.6 Doubling Observations of a Dataset	229
12.7 Iterating an Action with a Loop	230
12.7.1 Example 1	230
12.7.2 Example 2	231
12.7.3 Example 3	231
12.7.4 Example 4	232
12.7.5 Example 5	232
12.8 Avoiding the <code>if</code> Statement	233
12.8.1 Using a Logical Variable	233
12.8.2 Using the Functions <code>put</code> and <code>input</code> (Exercise)	234
12.9 Using <code>select</code> as Alternative to <code>if</code>	235
12.9.1 Conducting One Action	235
12.9.2 Conducting Multiple Actions	236
12.9.3 Using an Advanced Condition	237
13 Sorting Data and Managing Duplicates using <code>proc sort</code>	238
13.1 Data	238
13.1.1 Dataset <code>data_01</code>	238

13.1.2 Dataset <code>data_02</code>	239
13.1.3 Dataset <code>data_03</code>	239
13.2 Sorting in Ascending and Descending Order	239
13.2.1 Sorting in Ascending Order	240
13.2.2 Sorting in Descending Order	240
13.3 Saving Sorted Data in a New Dataset	240
13.4 Exercise	241
13.5 Sorting Character Variables As If They Were Numeric	242
13.6 Managing Duplicates	243
13.6.1 General Considerations	243
13.6.2 <code>nodupkey</code> Option	244
13.6.3 <code>nouniquekey</code> Option	245
13.6.4 <code>noduprecs</code> Option	246
14 Creating a Counter using <code>retain</code> and Temporary Variables	247
14.1 Data	247
14.2 Discovering the Concept of <i>retain</i>	248
14.2.1 Example 1	248
14.2.2 Example 2	249
14.3 Discovering Concepts of <code>first</code> and <code>last</code>	250
14.3.1 Discovering Fundamentals	250
14.3.2 Using the <code>descending</code> Option of the <code>by</code> Statement	253
14.3.3 Using the <code>notsorted</code> Option of the <code>by</code> Statement	253
14.3.4 Using the <code>groupformat</code> Option of the <code>by</code> Statement	254
14.4 Using <code>first</code> and <code>last</code> with a <code>retain</code>	255
14.4.1 Using <code>first</code> and <code>last</code> with a <code>retain</code>	255
14.4.2 Exercise	255
14.5 Note on the <code>put</code> and <code>putlog</code> Statements	256
14.5.1 Example 1	256
14.5.2 Example 2	257
14.5.3 Example 3	257
14.5.4 Example 4	258
14.6 Discovering the Temporary Variable <code>_n_</code>	259

14.6.1	Displaying the Values of <code>_n_</code> in a Dataset	259
14.6.2	Displaying a Text According to the Value of <code>_n_</code>	259
14.7	Discovering the Temporary Variable <code>eof</code> Created from the <code>end=</code> Option of the <code>set</code> Statement	260
14.7.1	Displaying the Values of <code>eof</code> in a Dataset	260
14.7.2	Displaying a Text According to the Value of <code>eof</code>	260
14.8	Using the <code>end=</code> Option and the <code>retain</code> Statement Together (Exercise)	261
14.9	Be Careful!	262
15	Discovering Common Functions, Series 2	263
15.1	Data	263
15.1.1	<code>data_01</code>	263
15.1.2	<code>data_02</code>	263
15.1.3	<code>data_03</code>	264
15.1.4	<code>data_04</code>	264
15.1.5	<code>data_05</code>	264
15.1.6	<code>data_06</code>	264
15.1.7	<code>data_07</code>	265
15.2	Overview	265
15.3	Using the <code>put</code> and <code>input</code> Functions	267
15.3.1	Preparing the Formats and Informats	267
15.3.2	Scenario 1 - Creating a Character Variable Based on a Numeric Variable	269
15.3.3	Scenario 2 - Creating a Numeric Variable Based on a Character Variable	270
15.3.4	Scenario 3 - Creating a Character Variable Based on Another Character Variable	271
15.3.5	Scenario 4 - Creating a Numeric Variable Based on Another Numeric Variable	272
15.4	Managing the Width of Format Labels	273
15.4.1	Truncating Unintentionally	273
15.4.2	Modifying the Width when Using a Format	274

15.4.3	Specifying the Default Width when Creating a Format	275
15.4.4	Dealing with Unplanned Change in Meaning	276
15.4.5	Avoiding a Change in Meaning using <code>other</code>	276
15.5	Understanding the <code>w.d</code> Format Values	277
15.5.1	Identifying Format <code>w.d</code> in the SAS® Online Doc	277
15.5.2	Example with the Format <code>DATEw.</code>	278
15.5.3	Example with the Format <code>BESTw.</code>	279
15.5.4	Example with the Format <code>w.d</code>	279
15.5.5	Example with the Format <code>Zw.d</code>	280
15.6	Using the <code>put</code> Function: Additional Examples	280
15.6.1	Using a Format from the SAS® System	280
15.6.2	Using the <code>put</code> Function in a Concatenation	281
15.6.3	Using the <code>put</code> Function with Missing Values	283
15.7	Understanding the Informats <code>w.d</code> Values	284
15.7.1	Identifying Informat <code>w.d</code> in the SAS® Online Doc	284
15.7.2	Avoiding Trap 1	285
15.7.3	Avoiding Trap 2	286
15.7.4	Dealing with Not Fully Compatible Data	287

Module 5 **289**

16 Appending Several Datasets **289**

16.1	<code>set</code> Statement: Basis	289
16.1.1	Data for Examples 1, 2, 12 and 15	289
16.1.2	Example 1	290
16.1.3	Example 2	290
16.2	<code>set</code> Statement: Working with Variables Having Different Lengths	291
16.2.1	Data of Examples 3, 4, 5 and 14	291
16.2.2	Example 3	291
16.2.3	Example 4	292
16.2.4	Example 5	293
16.2.5	Data for Examples 6 and 7	294

16.2.6 Example 6	294
16.2.7 Example 7	295
16.3 <code>set</code> Statement: Working with Datasets Having Different Structures	296
16.3.1 Data for Examples 8, 9, 10 and 11	296
16.3.2 Example 8	296
16.4 <code>in=</code> vs. <code>indsname=</code>	297
16.4.1 Example 9	297
16.4.2 Example 10	298
16.4.3 Example 11	298
16.4.4 Exercise	300
16.5 <code>proc append</code> and <code>proc datasets</code> as Alternative to the <code>set</code> Statement	301
16.5.1 Example 12	301
16.5.2 Data for Example 13	301
16.5.3 Example 13	302
16.5.4 Example 14	303
16.5.5 Example 15	304
17 Merging Several Datasets	305
17.1 Introduction	305
17.2 Examples 1 to 3: Basis	305
17.2.1 Data for Example 1	305
17.2.2 Example 1	306
17.2.3 Data for Examples 2 to 8	307
17.2.4 Example 2	308
17.2.5 Example 3	308
17.3 Examples 4 to 6: <code>in=</code> , the Dataset Option	309
17.3.1 Example 4	309
17.3.2 Example 5	309
17.3.3 Example 6	311
17.4 Examples 7 and 8: Merging Three Datasets	311
17.4.1 Example 7	311

17.4.2 Example 8	312
17.5 Examples 9 and 10: Common Variables Not Listed in the <code>by</code> Statement	313
17.5.1 Data for Examples 9 and 10	313
17.5.2 Example 9	313
17.5.3 Example 10	314
17.6 Example 11: Juxtaposing Datasets	314
17.6.1 Data for Example 11	314
17.6.2 Example 11	315
17.7 Examples 12 to 15: 'One-to-Many', 'Many-to-One' and 'Many- to-Many'	316
17.7.1 Data for Example 12	316
17.7.2 Example 12	317
17.7.3 Data for Example 13	317
17.7.4 Example 13	318
17.7.5 Data for Example 14	318
17.7.6 Example 14	319
17.7.7 Data for Example 15	320
17.7.8 Example 15	321
17.8 Example 16: Avoiding an Implicit <code>retain</code>	322
17.8.1 Data for Example 16	322
17.8.2 Example 16	322
17.9 Example 17: Using a Format as Alternative to a <code>merge</code>	323
17.9.1 Data for Example 17	323
17.9.2 Exercise	324
17.9.3 Example 17	325
17.9.4 Example 18	326

Module 6 **327**

18 `proc means` by Example **327**

18.1 Data	327
---------------------	-----

18.2	Displaying Statistics in the Default Destination	328
18.2.1	Global Statistics	328
18.2.2	Statistics by Treatment	329
18.2.3	Statistics by Age	329
18.2.4	Statistics by Treatment and Age	330
18.3	Saving Statistics in a Dataset	331
18.4	Adding Missing Values of <code>class</code> Variables	332
18.5	Saving Statistics in a Restructured Dataset	333
18.5.1	Statistics in the <code>RESULTS</code> Tab	333
18.5.2	Statistics in the Created Dataset	333
18.6	Listing Statistics to Compute	334
18.7	Displaying Global and Category Statistics using the <code>ways</code> Statement	336
18.7.1	Adding Global Statistics using the <code>ways</code> Statement	336
18.7.2	Dealing with Missing Values	337
18.7.3	Moving Global Statistics Below Statistics by Category	340
18.7.4	Changing Categories Order	341
18.8	Displaying Statistics for Several Groups whose Values are Grouped Together: <code>multilabel</code>	344
18.8.1	Statistics Produced without Format	344
18.8.2	Statistics Grouped with a Standard Format	345
18.8.3	Statistics Grouped with a <code>multilabel</code> Format	346
18.9	Displaying All Format Values	347
18.9.1	Format Used in the Example	347
18.9.2	Default Result	347
18.9.3	Result with <code>completetypes</code>	348
18.9.4	Result with <code>completetype</code> and <code>preloadfmt</code>	349
18.10	An Alternative to the <code>ways</code> Statement	350
18.10.1	<code>ways</code> Statement with Three Levels	350
18.10.2	<code>printalltypes</code> Option	351
18.10.3	<code>printalltypes</code> Option with the <code>type</code> Statement	351
18.11	Rounding the Statistics	352

18.12	An Old Method for Saving Statistics in a Dataset	353
18.12.1	Example 1	353
18.12.2	Example 2	356
18.13	Working with Weighted Values	359
18.14	<code>proc summary</code> as Alternative to <code>proc means</code>	360
18.14.1	Identifying the Syntax to Extract Statistics in a Dataset	360
18.14.2	Outputting the Statistics in a Dataset	361
19	Understanding Common Descriptive Statistics	362
19.1	Quartiles and Percentiles	362
19.1.1	What is a Quartile?	362
19.1.2	What is a Percentile?	363
19.1.3	Data	363
19.1.4	Computing Quartiles with <code>proc means</code>	364
19.1.5	Computing Percentiles with <code>proc means</code>	365
19.1.6	Mathematical Formula	365
19.1.7	Calculating Manually the Median	366
19.2	Variance and Standard Deviation	367
19.2.1	The Mean is Not Sufficient to Describe the Data . . .	367
19.2.2	What Statistics Can Be Added To Describe The Data?	367
19.2.3	Population Variance	367
19.2.4	Estimating the Population Mean/Variance from a Sample	369
19.2.5	Computing the Variance using <code>proc means</code>	370
20	Creating Your First Graph: Box-and-Whisper Plot	371
20.1	Components of a Box-and-Whisker Plot	371
20.2	Data	372
20.3	Creating a Box-and-Whisker Plot with <code>proc sgplot</code>	372
20.4	Removing Frames	373
20.5	Improving Both X-Axis and Y-Axis	374
20.6	Changing the Color of the Lines	375
20.7	Changing the Color and Size of the Box	376
20.8	Changing the Orientation	377

Module 7	379
21 Using <code>proc sql</code> with One Dataset	379
21.1 Introduction	379
21.1.1 On the Agenda	379
21.1.2 SQL Procedure vs. Data Step	380
21.2 Data	381
21.2.1 Dataset <code>one</code>	381
21.2.2 Dataset <code>two</code>	382
21.3 Creating a Dataset from Another Dataset	382
21.3.1 Creating a Dataset from Another Dataset	382
21.3.2 Creating Multiple Datasets in One SQL Procedure	383
21.3.3 The Impact of the <code>nolabel</code> Global Option on Variable Labels	384
21.3.4 Using the Same Input and Output Dataset Name	385
21.4 Displaying Data without First Creating a Dataset	388
21.4.1 Displaying Data by Starting the Statement with a <code>select</code> Clause	388
21.4.2 Displaying All Variables using an Asterisk	389
21.5 Displaying Metadata	389
21.5.1 Displaying the List of Dictionaries	389
21.5.2 Displaying the List of Datasets and Variables	390
21.5.3 Displaying the Metadata in the Log as a Statement of the SQL Procedure	393
21.5.4 Updating Datasets and Variables Attributes	394
21.6 Adding Constants	396
21.6.1 Adding Constants	396
21.6.2 Exercise	397
21.7 Extracting Summary Statistics	398
21.7.1 Extracting Summary Statistics of a Single Variable	398
21.7.2 Extracting Summary Statistics of Several Variables	401
21.8 Creating a Variable using a Conditional Structure	403
21.8.1 Creating a Variable using <code>case when</code>	403

21.8.2	Exercise	404
21.8.3	Using the <code>calculated</code> Keyword to Create a Variable based on a Newly Created Variable	405
21.9	Selecting Variables and Observations	406
21.9.1	Selecting Variables and Observations using Dataset Options (<code>keep=</code> / <code>drop=</code> and <code>where=</code>)	406
21.9.2	Selecting Observations using a <code>where</code> Clause	407
21.9.3	Selecting the First Few Observations	409
21.10	Displaying the Observation Numbers	411
21.10.1	Of the Whole Dataset	411
21.10.2	After Selecting a Subset of the Input Dataset	411
21.10.3	After Selecting a Subset of the Output Dataset	412
21.11	Extracting Summary Statistics by Group of Observations	413
21.11.1	<code>group by</code> Clause	413
21.11.2	<code>having</code> Clause	414
21.12	Sorting Data Explicitly using <code>order by</code>	415
21.12.1	Sorting Data in Ascending and Descending Order	415
21.12.2	Exercises	417
21.12.3	Sorting the Data using the Computed Values	418
21.13	Keeping a Single Observation among Duplicates Observations	419
21.13.1	Using the Keyword <code>distinct</code> with a List of Variables	419
21.13.2	Using the Keyword <code>distinct</code> with an Asterisk	420
21.13.3	Example of Implicit <code>distinct</code>	421
21.13.4	Using the Keyword <code>distinct</code> in a Function	421
21.14	Creating a Dataset and Adding Observations Manually	422
21.14.1	Creating an Empty Dataset from Scratch	422
21.14.2	Using <code>insert into . . . values . . .</code> to Add Data to a Dataset	424
21.14.3	Using <code>insert into . . . set . . .</code> to Add Data into a Dataset	424
22	Using <code>proc sql</code> with Multiple Datasets	426
22.1	Adding Observations to an Existing Set of Data	426

22.1.1 Data	426
22.1.2 <code>union all</code>	427
22.1.3 <code>union</code>	429
22.1.4 Creating an Empty Dataset from an Existing Dataset .	429
22.1.5 <code>insert into</code>	429
22.2 Selecting Observations Based on Data Available in Another Dataset	431
22.2.1 Combining the <code>where</code> Clause with a Sub-query: The <code>in</code> and <code>not in</code> Operators	431
22.2.2 Combining the <code>where</code> Clause with a Subquery: The <code>exists</code> and <code>not exists</code> Operators	434
22.3 Joining Data using a <code>left join</code> - Example in a <i>Many-to-One</i> Scenario	436
22.3.1 Data	436
22.3.2 <code>left join</code>	437
22.3.3 Exercise	439
22.4 Joining Data using a <code>full join</code> - Example in a <i>Many-to-One</i> Scenario	440
22.4.1 Data	440
22.4.2 <code>full join</code>	441
22.4.3 Exercise	442
22.5 Using the <code>coalesce</code> Function outside of <code>full join</code>	443
22.5.1 Data	443
22.5.2 <code>if/else if</code> Statements vs. <code>coalescec</code> Functions .	444
22.6 Joining Data in case of a <i>Many-To-Many</i> Scenario	446
22.6.1 Data	446
22.6.2 Can a Data Step (<code>merge</code>) Handle a <i>Many-To-Many</i> Scenario?	446
22.6.3 Handling a <i>Many-to-Many</i> Scenario with a <code>full join</code>	447
22.6.4 Handling a <i>Many-to-Many</i> Scenario with a <code>cross join</code>	448
22.7 Generating a Cartesian Product	449
22.7.1 Data	449

22.7.2	Generating a Cartesian Product using a <code>cross join</code> in the SQL Procedure	449
22.7.3	Generating a Cartesian Product using a Data Step . .	450
Module 8		451
23	Creating a Macro Variable and Displaying Its Value	451
23.1	Creating a Macro Variable	451
23.1.1	The <code>%let</code> Statement	451
23.1.2	The <code>call symputx</code> Routine	455
23.1.3	Saving the Number of Observations in a Macro Vari- able using <code>call symputx</code>	457
23.1.4	The <code>call symput</code> Routine (Old Syntax)	460
23.1.5	Deleting a Macro Variable	463
23.1.6	Creating a Macro Variable using the SQL Procedure .	465
23.1.7	Managing Blanks	468
23.1.8	Creating Several Macro Variables using the SQL Pro- cedure	470
23.1.9	Macro Parameters are Macro Variables	473
23.1.10	Creating a Macro Variable using a Loop	474
23.2	Local and Global Macro Variables	475
23.2.1	Displaying the List of All the Macro Variables	475
23.2.2	What's a Local Macro Variable?	477
23.2.3	When a Local Macro Variable Becomes Global	480
23.3	The Third Argument of the <code>call symputx</code> Routine	483
23.3.1	Definition	483
23.3.2	Outside a Macro	484
23.3.3	Inside a Macro	485
23.4	Solving Macro Variables in Several Steps	487
23.5	Case Study	488
23.5.1	Goal	488
23.5.2	Creating a Macro Variable using a Loop	490

23.5.3	Generating Macro Variables based on Dataset Values	491
23.5.4	Creating the Reference Dataset using Dictionaries . . .	492
23.5.5	Using Local Macro Variables	493
23.5.6	Alternative using a Data Step: the <code>call execute</code> Routine	494
23.5.7	Alternative using a Data Step: the <code>put</code> Statement (Ex- ercise)	495
24	Getting Familiar with a Few Macro Functions	496
24.1	Two Examples using the Macro Function	
	<code>%sysfunc</code>	496
24.1.1	Identifying the Path of the <code>work</code> Library	496
24.1.2	Displaying Today's Date	498
24.2	Performing Calculations using Macro Variables Values	499
24.2.1	A Macro Variable Value is Always Character	499
24.2.2	The <code>%eval</code> Macro Functions (1/2)	499
24.2.3	The <code>%eval</code> Macro Functions (2/2)	499
24.2.4	The <code>%sysevalf</code> Macro Functions (1/2)	500
24.2.5	The <code>%sysevalf</code> Macro Functions (2/2)	500
24.2.6	The <code>%datatyp</code> Macro Function	502
24.3	Using the <code>%str</code> and <code>%nrstr</code> Macro Functions	503
24.3.1	Using a <code>%</code> in Front of a Quote in the <code>%str</code> Macro Func- tion	503
24.3.2	Using the <code>%str</code> Macro Function	504
24.3.3	The <code>%nrstr</code> Macro Function	510
24.4	The <code>%cmpres</code> Macro Function	513
24.5	The <code>%upcase</code> Macro Function	514
25	Getting Started with Macros	515
25.1	Keyword Macro Parameters vs. Positional Macro Parameters	515
25.1.1	Keyword Macro Parameter	515
25.1.2	Positional Macro Parameter	517
25.2	A Few Global Options	518

25.2.1	General Considerations	518
25.2.2	The <code>mprint</code> Global Option	519
25.2.3	The <code>mlogic</code> Global Option	520
25.2.4	The <code>symbolgen</code> Global Option	521
25.2.5	The <code>mprint</code> and <code>mfile</code> Options used with the <code>filename</code> Statement	521
25.3	Basics on Conditions in Macro Language	524
25.3.1	<code>%else</code> and Comments	524
25.3.2	Skipping Part of the Program	525
25.3.3	Using the <code>in</code> Operator in a Macro	527
25.3.4	Using <code>%if</code> Outside a Macro	530
25.4	Including a Program Containing a Macro Definition	531
25.4.1	Macro Definition	531
25.4.2	Including a Macro and Call it	531
25.5	Using a Permanent Macro Catalog	533
25.5.1	Overview	533
25.5.2	Compiling a Macro and Saving It in a Permanent Macro Catalog	534
25.5.3	Displaying the List of Macros Available in a Macro Cat- alog	536
25.5.4	Calling a Macro Available in a Permanent Macro Catalog	538
Module 9		540
26	Using <code>proc report</code> with a Listing Destination	540
26.1	Data	541
26.2	A First Example	542
26.3	Managing Formats	543
26.3.1	Removing Formats	543
26.3.2	Adding or Replacing the Format Used	544
26.4	Managing Labels	544
26.4.1	Removing Labels from the Variables	544

26.4.2	Adding or Replacing Permanent Labels	545
26.4.3	Displaying a Label on Multiple Lines	546
26.4.4	Adding a Label Common to Several Variables	547
26.4.5	Adding a Label Common to Several Variables on Multiple Lines	548
26.4.6	Removing Variable Names	549
26.4.7	Adding a Label Common to Several Columns while Removing Variable Names	549
26.5	Grouping Data	550
26.5.1	<code>order order=formatted</code> (Default Value)	550
26.5.2	<code>order</code> and Missing Values	551
26.5.3	<code>order order=internal</code>	553
26.5.4	<code>order order=freq</code>	553
26.5.5	<code>order order=data</code>	554
26.5.6	Creating a Temporary Variable to Sort the Data	555
26.5.7	<code>group</code> vs <code>order</code>	556
26.5.8	<code>group</code> and <code>analysis</code>	557
26.5.9	<code>group order=data</code> and <code>preloadfmt</code>	558
26.5.10	<code>group, across</code> and <code>analysis</code>	559
26.5.11	Exercise	560
26.6	Updating the Display of Missing Numeric Values	562
26.7	Displaying the Output in a <i>listing</i> Destination and Updating the Global Options	563
26.7.1	Displaying the Output in a <i>listing</i> Destination	563
26.7.2	Updating Global Options: Example 1	564
26.7.3	Updating Global Options: Example 2	565
26.7.4	Updating Global Options: Example 3	565
26.8	Customizing the Table Header [<i>listing</i> Destination]	568
26.8.1	<code>proc report</code> Statement Options	568
26.8.2	Drawing a Line Below the Label Common to Several Columns	569
26.8.3	Drawing a Line Below Each Column Label	570

26.8.4	Drawing a Line Below the Label Common to Several Columns and Each Column Label	571
26.9	Adding Some Text Before and After the Data	571
26.9.1	<code>compute before</code> and <code>compute after</code>	571
26.9.2	<code>compute before</code> and <code>compute after</code> followed by a Variable Name	574
26.10	The <code>page</code> and <code>skip</code> Options of the <code>break</code> Statement	576
26.10.1	Starting a New Page for Each Group of Values (<code>page</code>)	576
26.10.2	Adding a Line Before/After Each Group of Values (<code>skip</code>)	578
26.11	Adding a Statistic for Each Group of Values	579
26.11.1	Adding a Statistic for Each Group of Values	579
26.11.2	Deleting the Text at the Front of a Statistics Row . . .	580
26.11.3	Customizing the Text at the Front of the Statistics Row	581
26.11.4	Exercise	582
26.11.5	Revision: Changing the Default Width of Format Labels	584
26.12	Adding and Customizing Lines Drawn Around Statistics [<i>listing Destination</i>]	585
26.12.1	Adding Lines Drawn Around Statistics	585
26.12.2	Customizing Lines Drawn Around Statistics	587
26.13	Adding Global Statistics	588
26.13.1	The <code>rbreak</code> Statement	588
26.13.2	The <code>rbreak</code> and <code>by</code> Statements vs. the <code>break</code> Statement with the <code>page</code> Option	590
26.14	Exercise	592
26.15	Displaying the Values of the Temporary <code>_BREAK_</code> Variable . .	594
26.15.1	Method 1: Extracting <code>proc report</code> Data in a New Dataset	594
26.15.2	Method 2: Creating a New Variable in <code>proc report</code>	595
26.16	Displaying <code>proc report</code> Options with their Value and Managing Columns Spacing [<i>listing Destination</i>]	597
26.16.1	The <code>list</code> Option: Displaying <code>proc report</code> Options with their Value	597

26.16.2	The <code>spacing=</code> Option: Managing Columns Spacing	599
26.17	Managing Columns Width [<i>listing</i> Destination]	600
26.17.1	Increasing Columns Width using the <code>width=</code> Option	600
26.17.2	Character Variable: Managing an Insufficient Column Width	602
26.17.3	Numeric Variable: Managing an Insufficient Column Width (1/2)	603
26.17.4	Numeric Variable: Managing an Insufficient Column Width (2/2)	604
26.18	Managing Text Alignment [<i>listing</i> Destination]	605
26.18.1	Aligning to the <code>left</code> , <code>right</code> or <code>center</code>	605
26.18.2	Creating Indentations	607
26.18.3	Aligning Numbers (Example 1)	608
26.18.4	Aligning Numbers (Example 2)	610
26.19	The <code>by</code> Statement and Titles	611
26.19.1	Data	611
26.19.2	The <code>by</code> Statement	612
26.19.3	The <code>nobyline</code> Option	613
26.19.4	Example 1: <code>#byline</code>	614
26.19.5	Example 2: <code>#byvar(...)</code> and <code>#byval(...)</code>	615
26.19.6	Example 3: <code>#byvar1</code> and <code>#byval1</code>	616
26.20	A Title on the Left, in the Middle and on the Right [<i>listing</i> Destination]	617
26.20.1	code <code>justify=</code> and the <i>listing</i> destination	617
26.20.2	Breaking down the steps for building the <code>title</code> state- ment	618
26.20.3	Save the Three Title Parts in Three Macro Variables	620
26.20.4	Compute the Number of Blanks to Add Between the Strings to Be Concatenated	620
26.20.5	Final Result	622
26.21	<code>proc report</code> and Alternatives	623

26.21.1 Structuring your Program and Limiting the Use of <code>proc</code>	
<code>report</code> to Reporting	623
26.21.2 Alternatives to <code>proc report</code>	623
Module 10	625
27 <code>proc transpose</code> by Example	625
27.1 Introduction	625
27.1.1 What Exactly Does It Mean to Transpose Data?	625
27.1.2 Why Transposing Data?	626
27.1.3 Data	627
27.2 Specifying the Variable Names in the Output Dataset (<code>_NAME_</code>	
and <code>COLx</code>)	629
27.2.1 A First <code>proc transpose</code>	629
27.2.2 Specifying the Variable Names in the Output Dataset .	630
27.2.3 Transposing Data by Group of Observations (<code>by</code> State-	
ment)	634
27.2.4 Rare Case: <code>copy</code> Statement	634
27.3 Managing Variable Labels	635
27.3.1 Using Labelled Variables in the <code>var</code> Statement	635
27.3.2 Displaying the Default Labels of the <code>_NAME_</code> and <code>_LABEL_</code>	
Variables	636
27.3.3 Adding a Label to <code>COLx</code> Variables using the <code>idlabel</code>	
Statement	637
27.3.4 Remarks	637
27.4 Transposing Formatted Variables	639
27.4.1 Transposing Numeric Variables with the Same Format	639
27.4.2 Transposing Character Variables with the Same Format	640
27.4.3 Transposing Variables Having Different Formats	641
27.5 Transposing Both Numeric and Character Variables	641
27.5.1 Transpose Both Numeric and Character Variables . .	641
27.5.2 Note: <code>_all_</code> , <code>_numeric_</code> and <code>_character_</code>	644

27.5.3	Be Careful! When No <code>var</code> Statement is Used	644
27.5.4	Challenges	645
27.6	Transposing Backwards!	648
27.6.1	Transposing Backwards!	648
27.6.2	Transposing Backwards! Yes, But...	649
27.7	Next Step	650
28	Discovering Common Functions, Series 3	651
28.1	Functions <code>length</code> , <code>lengthc</code> and <code>lengthn</code>	652
28.1.1	Data	652
28.1.2	Functions <code>length</code> , <code>lengthc</code> and <code>lengthn</code>	653
28.2	Functions <code>index</code> and <code>find</code> , and Functions Related	654
28.2.1	Data	654
28.2.2	Function <code>index</code>	656
28.2.3	Function <code>indexc</code>	657
28.2.4	Function <code>indexw</code>	658
28.2.5	Functions <code>index/c/w</code> vs. Functions <code>find/c/w</code> using Two Arguments	659
28.2.6	Function <code>findc</code>	660
28.2.7	Function <code>findw</code> using Three Arguments	661
28.2.8	Function <code>findw</code> using Five Arguments (1/2)	663
28.2.9	Function <code>findw</code> using Five Arguments (2/2)	664
28.2.10	Function <code>findw</code> and the Modifier <code>e</code> , <code>r</code> and <code>m</code>	665
28.2.11	Be careful! When Identical Values are in the 2nd and 3rd Arguments of <code>indexw</code> or <code>findw</code>	668
28.2.12	Selecting Observations using <code>contains</code>	669
28.2.13	Selecting Observations using Truncated String Com- parison Operators (<code>proc sql</code>)	670
28.3	Functions <code>lag</code> , <code>dif</code> and Related Computations	672
28.3.1	Data	672
28.3.2	Function <code>lag</code>	673
28.3.3	Function <code>lag2</code>	676
28.3.4	Function <code>dif</code> (1/2)	678

28.3.5	Function <code>dif</code> (2/2)	679
28.3.6	Function <code>dif2</code> (1/2)	679
28.3.7	Function <code>dif2</code> (2/2)	680
28.3.8	The Function <code>lead</code> Does Not Exist! Three-Step Alternative (1/2)	681
28.3.9	The Function <code>lead</code> Does Not Exist! Three-Step Alternative (2/2)	682
28.3.10	<code>proc expand</code> , an Alternative to Functions?	683
28.3.11	<code>proc sql</code> , an Alternative to Functions? (1/2)	684
28.3.12	<code>proc sql</code> , an Alternative to Functions? (2/2)	687
28.4	Functions <code>ifc</code> and <code>ifn</code>	688
28.4.1	Function <code>ifc</code>	688
28.4.2	Function <code>ifn</code>	689
28.4.3	The Fourth Argument of the Functions <code>ifc</code> and <code>ifn</code>	690
28.5	Functions <code>putc</code> , <code>putn</code> , <code>inputc</code> and <code>inputn</code>	691
28.5.1	Function <code>putc</code> (From Character To Character)	692
28.5.2	Function <code>putn</code> (From Numeric To Character)	693
28.5.3	Function <code>inputc</code> (From Character To Character)	694
28.5.4	Function <code>inputn</code> (From Character To Numeric)	695
28.6	Functions <code>week</code> and <code>weekday</code> and Formats of the Same Family	696
28.6.1	Data	696
28.6.2	Function <code>week</code> and Formats <code>weeku</code> , <code>weekv</code> and <code>weekw</code>	697
28.6.3	Function <code>weekday</code>	703
28.6.4	Formats to Display Dates in English	703
28.6.5	Formats to Display Dates in Local Language	706
28.7	Function <code>intnx</code> or How to Find a Date from Another Date	710
28.7.1	Using the Default Value for the 3rd Argument	712
28.7.2	Using <code>'sameday'</code> for the 4th Argument	721
28.7.3	Using Interval Multiples	723
28.7.4	Shifting by One or Several Intervals	725
28.7.5	Defining Your Own Intervals	727
28.7.6	Using the <code>intnx</code> Function with <code>%sysfunc</code>	728

28.7.7 Exercise	729
28.8 Function <code>intck</code> or How Much Time Has Elapsed Between Two Dates	730
28.8.1 Time Elapsed Between Two <i>SAS Dates</i>	731
28.8.2 Time Elapsed Between Two <i>SAS Times</i>	733
28.8.3 Time Elapsed Between Two <i>SAS Datetimes</i>	733
28.9 Function <code>yrdif</code>	736
28.9.1 Basis for Computation: 'age' (By Default)	736
28.9.2 Basis for Computation: 'act/act'	738
28.9.3 Basis for Computation: 'act/360'	739
28.9.4 Basis for Computation: 'act/365'	740
28.9.5 Basis for Computation: '30/360'	741
28.9.6 References	742
28.10 Function <code>datdif</code>	742
28.10.1 Basis for Computation: 'act/act'	742
28.10.2 Basis for Computation: Other Possibilities	743
28.10.3 Reference	743
Module 11	744
29 Generating a Text File using the <code>file</code> Statement	744
29.1 Writing to the Log or to the Default Destination	744
29.1.1 Using a Data Step Without Creating a Dataset	745
29.1.2 Writing to the Log	745
29.1.3 Writing to the Default Destination	745
29.1.4 Writing Information to the Default Destination and Oth- ers to the Log	746
29.1.5 Writing Multiple Lines with a Single <code>put</code> Statement	747
29.2 Writing a SAS Program and Submit It	747
29.2.1 Writing to a <code>.sas</code> File	747
29.2.2 Notes regarding the <code>%include</code> Statement	749
29.2.3 Writing to a Catalog and Managing Its Contents	751

29.2.4	Writing to a Temporary File with <code>filename...temp</code>	756
29.2.5	Writing a Program using Several Data Steps (<code>mod</code> Option)	756
29.3	Writing the Values of One or Several Variables Available in a Dataset	757
29.3.1	Data	757
29.3.2	Writing the Values of One or Several Variables	758
29.3.3	Alternating Between Variable Values and Text	758
29.3.4	Adding the Variable Name in Front of Each Value	759
29.3.5	Writing Formatted Values	759
29.3.6	Starting at a Specific Position to Align Values	759
29.3.7	Moving the Cursor Backwards to Remove a Blank	761
29.3.8	Writing All Values on a Single Line	761
29.4	Writing a Program using a Reference Dataset (<code>ref</code>)	762
29.4.1	Data	762
29.4.2	Writing a Program Having One Observation per Dataset in <code>ref</code>	762
29.4.3	Writing a Program Having One Observation per Variable in <code>ref</code>	770
29.4.4	Writing a <code>proc report</code> (Exercise)	773
29.5	Exploring the Options of the <code>put</code> Statement	774
29.5.1	Using Automatic Variables without <code>put</code> Statement Options	774
29.5.2	Using Automatic Variables with <code>put</code> Statement Options	775
29.5.3	Using Distinct Options of the <code>put</code> Statement	776
29.5.4	Using Together Several Options of the <code>put</code> Statement	778
29.5.5	Using Formats	779
29.6	Creating a CSV File	779
29.6.1	Data	780
29.6.2	Creating a CSV File with a Data Step (Exercise)	781
29.6.3	Creating a CSV File with <code>ods csv/ods csvall</code>	782
29.6.4	Creating a CSV File with <code>proc export</code>	783

30 Reading a Text File using the <code>infile</code> Statement	785
30.1 Introduction	785
30.1.1 Data Step: <code>output</code> Statement (Exercise)	786
30.1.2 <code>proc sql: insert into</code> Statement (Exercise)	786
30.1.3 Data Step: <code>infile</code> Statement	787
30.2 Managing Variable Length	789
30.2.1 What is the Default Length of Character Variables?	789
30.2.2 Setting the Length	791
30.3 Delimiting Values	792
30.3.1 Using a Comma Delimiter	792
30.3.2 Using a Semicolon Delimiter	792
30.3.3 Inputting Data having Semicolons	793
30.3.4 Using a Tab Delimiter	793
30.3.5 Reading Values Enclosed in Quotation Marks	794
30.4 Handling Special Missing Values	795
30.4.1 Inputting <code>.A</code> to <code>.Z</code> and <code>._</code> Values	795
30.4.2 Inputting <code>A</code> to <code>Z</code> and <code>_</code> Values using the <code>missing</code> Option	796
30.5 Interpreting Properly Input Data using Informats	796
30.5.1 Using an <code>informat</code> Statement	797
30.5.2 Using an informat in an <code>input</code> Statement	797
30.6 Managing Non-Standard Line Contents	798
30.6.1 Managing Multiple Observations on One Line	798
30.6.2 Managing Multiple Lines for One Observation	799
30.6.3 Managing a New Observation Every Other Line	799
30.6.4 Starting Inputting Data at Line 2	800
30.6.5 Increasing the Number of Bytes for Character Variables	800
30.6.6 Handling Carriage Returns and Line Feeds	801
30.7 Positioning the Pointer on the Line	802
30.7.1 Going to a Specific Column (<code>@n</code>) or Moving Forward/Back- ward from the Current Column on the Line (<code>+n</code>)	802
30.7.2 Reading the Value After a Given Text Introduced by <code>@</code>	803
30.7.3 Using <code>@n</code> , <code>+n</code> and <code>@</code> (Exercise)	804

30.8	Choosing Between Four Approaches	805
30.8.1	Overview of the Four Approaches	805
30.8.2	Inputting Data using <i>List Input</i>	805
30.8.3	Inputting Data using <i>Column Input</i>	807
30.8.4	Inputting Data using <i>Named Input</i>	808
30.9	Managing Lines Where Some Data is Non Available	808
30.9.1	Choosing Between <code>flowover</code> , <code>stopover</code> , <code>trunccover</code> and <code>missover</code>	809
30.9.2	Continuing the Search on the Next Line using the <code>scanover</code> Option	811
30.10	Reading the Log Contents	812
30.10.1	Creating and Saving a Log File	812
30.10.2	Reading the Entire Log	815
30.10.3	Reading a Subset of the Log File	816
30.11	Reading a <i>CSV file</i> with <code>proc import</code>	817
30.11.1	Data	817
30.11.2	Example with <code>class_procexport.csv</code>	817
30.11.3	Example with <code>class_procexport_dlm.csv</code>	820
30.11.4	Example with <code>class_ods_csvall.csv</code>	821
31	Working with a Transport File	825
31.1	Overview	825
31.1.1	Transport File vs. zip File	825
31.1.2	xpt Transport File vs. cpt Transport File	825
31.1.3	Pharmaceutical Industry	828
31.1.4	Disclaimer	828
31.1.5	On the Agenda	829
31.2	Testing the Different Approaches	829
31.2.1	Organizing Files in Directories	829
31.2.2	Preparing the Datasets to Transport	830
31.3	Approach 1: Using the Macros <code>%loc2xpt</code> and <code>%xpt2loc</code> (xpt File)	830
31.3.1	Creating a Transport File	831

31.3.2	Extracting the Datasets from the Transport File	832
31.4	Approach 2.1: Using the <code>xport</code> Engine with <code>proc copy</code> (<code>xpt</code> File)	833
31.4.1	Creating a Transport File	833
31.4.2	Extracting the Datasets from the Transport File	833
31.5	Approach 2.2: Using the <code>xport</code> Engine with a Data Step (<code>xpt</code> File)	834
31.5.1	Creating a Transport File	834
31.5.2	Extracting the Datasets from the Transport File	835
31.6	Can the Approaches Be Mixed? (<code>xpt</code>)	835
31.7	Approach 3: Using <code>proc cport/cimport</code> (<code>cpt</code> File)	835
31.7.1	Creating the Transport File	836
31.7.2	Extracting the Datasets from the Transport File	836
31.8	Communicating on the Transported Data	837
31.8.1	Choosing a Meaningful File Extension	837
31.8.2	Communicating the Approach Used	837
31.8.3	Providing Metadata Details	837
31.9	Extracting Information about the Creation of the Transport File	838
31.9.1	Using a Data step	838
31.9.2	Using Dictionaries	841
31.10	Comparing Data Before and After	841
31.10.1	Comparing the Data	842
31.10.2	Comparing the Dataset Attributes (<code>vtable</code>)	843
31.10.3	Comparing the Variable Attributes (<code>vcolumn</code>)	845
31.10.4	Understanding the <code>npos</code> and <code>obslen</code> Variables	845
31.11	Modifying the Dataset Creation Date	847
31.12	Summarizing the Results Based on Approaches and Data Used	848
31.12.1	Summary Table	848
31.12.2	Summary Table Explanations	850
31.13	Assignment	853
31.14	Reading Multiple <code>xpt</code> Transport Files	853

Module 12	855
32 Creating Reports using the Output Delivery System	855
32.1 General Information	855
32.1.1 Introduction	855
32.1.2 Data	857
32.1.3 Basic Concepts	858
32.1.4 Creating a New Page in a PDF File: <code>startpage=</code> . .	863
32.2 Merging Cells and Vertical Aligning Cells Content	866
32.2.1 Merging Cells in a PDF Destination: <code>spanrow</code>	866
32.2.2 Vertical Aligning Cells Content: <code>verticalalign=</code> . .	867
32.3 Customizing the Six Areas of a proc report Output	868
32.3.1 Area 1: Customizing the <code>column</code> Area	868
32.3.2 Area 2: Customizing the <code>header</code> Area	869
32.3.3 Area 3: Customizing the <code>report</code> Area	872
32.3.4 Area 4: Customizing the <code>lines</code> Area	874
32.3.5 Area 5: Customizing the <code>summary</code> Area	876
32.3.6 Area 6: The Three Arguments of the <code>call define</code> Routine	878
32.3.7 Area 6: The Second Argument of the <code>call define</code> Routine	879
32.3.8 Area 6: Reversing the Order of the Variables in the <code>column</code> Statement	883
32.3.9 Area 6: Using <code>style(calldef)</code>	885
32.3.10 Area 6: Using the <code>call define</code> Routine to Create Indentations	889
32.3.11 Area 6: Using the <code>call define</code> Routine to Alternate Rows Background Color	890
32.3.12 Area 6: The First Argument of the <code>call define</code> Rou- tine	891
32.3.13 Area 6: Using <code>'format'</code> as Second Argument of the <code>call define</code> Routine	893
32.3.14 Adding a Hyperlink	895

32.3.15	Updating Margins, Paper Size and Page Orientation	898
32.4	Style Templates	900
32.4.1	Using a Style Template of the SAS® system	900
32.4.2	Listing Currently Available Style Templates, Displaying and Saving Their Source Code	901
32.4.3	Creating and Using a Temporary Style Template	904
32.4.4	Listing and Deleting Temporary Style Templates	908
32.4.5	Creating, Using and Deleting a Permanent Style Tem- plate	914
32.5	Justifying Numeric Values	919
32.5.1	Previewing the Data (Before/After)	919
32.5.2	Using the <code>textalign=d just=d</code> Style Attribute	919
32.6	Colors	920
32.6.1	Changing the Title Color	920
32.6.2	Finding a Color Code using Either Excel or the SAS® System	920
32.6.3	Creating the <code>%rgb</code> Macro	922
32.6.4	Calling the <code>%rgb</code> Macro	922
32.6.5	Using a Format as Style Attribute Value	923
32.6.6	A Few External Resources	924
32.7	Titles and Lines of Text	924
32.7.1	Customizing the Style of Three Titles Displayed on the Same Line	924
32.7.2	Customizing the Title Style	927
32.7.3	Adding a Line Break in a String	930
32.7.4	Adding Blanks and Subscript/Superscript Values to a String	933
32.7.5	Using Unicode Values	935
32.7.6	Numbering Pages	940
32.8	<code>proc odstext</code> and Alternatives	944
32.8.1	Applying a Style to the <code>h</code> and <code>p</code> Statements of <code>proc</code> <code>odstext</code>	944

32.8.2	Creating as Many Lines of Text As Observations Available in a Dataset using <code>proc odstext</code>	951
32.8.3	Creating a Bulleted List using <code>proc odstext</code> : Basic	952
32.8.4	Creating a Bulleted List using <code>proc odstext</code> and Customizing the Bullet Points	953
32.8.5	Adding Data in a Bulleted List using <code>proc odstext</code>	957
32.8.6	Defining Bullet Points Style using the <code>list</code> Style Element	959
32.8.7	Creating Lists and Sub-lists using <code>proc odslist</code>	960
32.8.8	Customizing an <code>ods text=</code> Statement	961
32.9	Creating a Table of Contents in a PDF File (or RTF File)	963
32.9.1	Removing the Table of Contents of a PDF File: <code>notoc</code>	963
32.9.2	Using <code>ods proclabel</code> and the <code>contents=</code> Options of <code>proc report</code> to Apply New Labels to Table of Contents Elements	963
32.9.3	Using <code>proc document</code> to Apply New Labels to Table of Contents Elements	966
32.9.4	Customizing the Table of Contents using <code>proc document</code>	969
32.9.5	Inserting a Table of Contents in the First Page of a PDF File	973
32.9.6	Inserting a Table of Contents in the First Page of a RTF File	974
32.9.7	Excluding a Procedure from the Table of Contents	975
32.10	File Properties	977
32.10.1	Using the Options of the <code>ods pdf</code> Statement	977
32.10.2	Using the Options of the <code>ods rtf</code> Statement	978
32.10.3	Using the Native Language with a RTF Destination	979
32.11	Splitting Outputs Into Multiple Files: <code>newfile=</code>	981
32.11.1	Creating One File per Procedure: <code>newfile=proc</code>	981
32.11.2	Creating One File per Output: <code>newfile=output</code>	982
32.11.3	Creating One File per Page: <code>newfile=page</code>	983
32.11.4	Creating One File per Group of Values: <code>newfile=bygroup</code>	984

32.12	Saving the Same Output in Multiple Files	985
32.12.1	Example 1	985
32.12.2	Example 2	986
32.13	Displaying a Table on Several Pages	987
32.13.1	Keeping the Same Column Width on Every Page in a PDF File: <code>uniform</code>	987
32.13.2	Creating a Page Break When Displaying New Vari- ables in a RTF File: <code>keepn</code>	988
32.14	Options Which Don't Always Work in the PDF Destination . .	990
Module 13		991
33	Creating an Excel File using <code>ods excel</code>	991
33.1	Introduction	991
33.2	Data	992
33.3	A First Example using <code>ods excel</code>	993
33.4	Using the <code>sheet_name=</code> Option to Name Worksheets	994
33.4.1	Naming a Worksheet	994
33.4.2	Explicitly Naming Each Worksheet	995
33.4.3	Explicitly Naming the First Worksheet	996
33.4.4	Using the Default Excel Worksheet Names (<code>Sheet 1...</code>)	997
33.5	Displaying Outputs on Several Worksheets	998
33.5.1	Using the <code>sheet_interval='output'</code> Option to Create one Worksheet Per Output	998
33.5.2	Using the <code>sheet_interval='proc'</code> Option to Create one Worksheet Per Procedure	999
33.5.3	Using the <code>sheet_interval='none'</code> Option to Keep All Outputs in the Same Worksheet	1000
33.5.4	Using the <code>sheet_interval='now'</code> Option to Create a New Worksheet	1001
33.5.5	Using the <code>sheet_interval='bygroup'</code> Option to Display the Outputs of a given Procedure by Value . .	1002

33.5.6 Using the <code>sheet_interval='page'</code> Option to Display the Outputs on Several Worksheet Based on the <code>pageby</code> Statement of <code>proc print</code> . . .	1004
33.5.7 Using the <code>sheet_label=</code> Option to Specify the Worksheet Names Prefix	1005
33.6 Using the <code>blank_sheet=</code> Option to Add an Empty Worksheet	1006
33.7 Using the <code>tab_color=</code> Option to Color the Tabs	1007
33.8 Using the <code>autofilter=</code> Option to Add Filters	1008
33.8.1 Adding a Filter to All Columns	1008
33.8.2 Adding a Filter to Successive Columns	1009
33.8.3 Adding a Filter to a Specific Column	1010
33.9 Using the <code>frozen_headers=</code> and <code>frozen_rowheaders=</code> Options to Freeze Rows/Columns	1011
33.10 Using the <code>zoom=</code> Option to Zoom on the Data	1012
33.11 Using the <code>protect_worksheet=</code> Option to Protect a Worksheet	1013
33.12 Managing Titles/Footnotes and Subtitles	1014
33.12.1 Using the <code>embedded_titles='yes'</code> and <code>embedded_footnotes='yes'</code> Options to Insert Titles and Footnotes in the Worksheet	1014
33.12.2 Using the <code>title_footnote_width=</code> Option to Specify the Number of Columns to Merge	1016
33.12.3 Using the <code>title_footnote_nobreak='yes'</code> Option to Compute the Number of Columns to Merge in Order to Avoid Line Breaks	1017
33.12.4 Using the <code>suppress_bylines='yes'</code> Option to Suppress Subtitles Generated by a <code>by</code> Statement	1018
33.12.5 Using the <code>nobyline</code> Global Option to Suppress Subtitles Generated by a <code>by</code> Statement	1019

33.12.6	Using the <code>embed_titles_once='yes'</code> and <code>embed_footnotes_once='yes'</code> Options to Display Titles/Footnotes Only Once Per Worksheet	1020
33.13	Defining Rows Height and Columns Width	1021
33.13.1	Using the <code>row_heights=</code> Option to Set the Row Heights Based on the Row Type	1021
33.13.2	Using the <code>absolute_row_height=</code> and <code>absolute_column_width=</code> Options to Define the Row Heights and the Column Widths Within the Print Area	1025
33.13.3	Using the <code>height=</code> and <code>width=</code> Style Attributes to Define the Row Heights and the Column Widths of a Table	1026
33.14	Using the <code>hidden_rows=</code> and <code>hidden_columns=</code> Options to Hide Rows/Columns	1028
33.15	Using the <code>flow=</code> Option to Manage Line Breaks	1028
33.15.1	Possible Values for the <code>flow=</code> Options	1028
33.15.2	Program Used for the Demo	1029
33.15.3	<code>flow='<range>'</code>	1030
33.15.4	<code>flow='header'</code>	1030
33.15.5	<code>flow='rowheader'</code>	1031
33.15.6	<code>flow='data'</code>	1031
33.15.7	<code>flow='table'</code>	1032
33.16	<code>ods text=</code> Global Statement and <code>proc odstext</code>	1032
33.16.1	Impact of the <code>flow='text'</code> Option	1032
33.16.2	Displaying a Long String using <code>ods text=</code> or <code>proc odstext</code>	1034
33.16.3	Using <code>proc odstext</code> Before/After a Reporting Pro- cedure and Dealing with Titles - Basic Scenarios	1037

33.16.4	Using <code>proc odstext</code> Before/After a Reporting Procedure and Dealing with Titles - Advanced Scenarios	1041
33.17	Using <code>tagattr=</code> Style Attribute to Introduce Excel Features	1047
33.17.1	Using the <code>tagattr='mergeacross:...'</code> Style Attribute to Merge Several Cells	1047
33.17.2	Example of Limit of <code>tagattr='mergeacross:...'</code> Style Attribute	1050
33.17.3	Using <code>tagattr='rotate:...'</code> Style Attribute to Rotate a Text	1051
33.17.4	Using <code>tagattr='format:... type:'</code> Style Attribute or SAS Formats to Format Cells	1052
33.18	Using the <code>flyover=</code> Style Attribute to Add A Note to a Cell	1057
33.18.1	Using the <code>flyover=</code> Style Attribute to Add A Note to a Cell	1057
33.18.2	Exercise (<code>call define</code> Routine)	1058
33.19	Positioning Tables on the Worksheet	1059
33.19.1	Using the <code>start_at=</code> Option to Define the Starting Cell	1059
33.19.2	Juxtaposing Two Tables with PowerShell (Windows)	1060
33.20	Using Style Attributes to Customize Table Borders	1063
33.20.1	Overview	1063
33.20.2	Using the <code>borderstyle=</code> Style Attribute	1064
33.20.3	Using the <code>borderwidth=</code> Style Attribute	1065
33.20.4	Exercise	1066
33.20.5	Using a <i>CSS File</i>	1068
33.21	Creating a Table of Content	1072
33.21.1	Using the <code>formulas=</code> Option with the <code>hyperlink</code> Excel Function	1072
33.21.2	Using the <code>index=</code> Option to List the Worksheet Tabs	1073
33.21.3	Using the <code>contents=</code> Option to List the Outputs Displayed in the Worksheets	1074
33.21.4	Renaming Worksheets	1075
33.22	Inserting an Image or Graphic in a Worksheet	1075

33.22.1	Inserting an Unicode Character	1076
33.22.2	Inserting an Image	1079
33.22.3	Inserting a Graphic	1080
33.22.4	Inserting a Background Image	1082
33.23	Updating File Properties	1083
33.24	Page Setup: 'Page' Tab	1084
33.24.1	'Page' Tab: Default Result	1084
33.24.2	'Page' Tab (1): 'Orientation'	1084
33.24.3	'Page' Tab (2): 'Scaling' - 'Adapt to:'	1085
33.24.4	'Page' Tab (3a): 'Scaling' - 'Fit to:'	1086
33.24.5	'Page' Tab (3b): 'Scaling' - 'Fit to:' (Automatically) . . .	1086
33.24.6	'Page' Tab (4): 'Paper size:'	1087
33.24.7	'Page' Tab (5): 'Print quality:'	1087
33.25	Page Setup: 'Margins' Tab	1088
33.25.1	'Margins' Tab (1): 'Top', 'Right', 'Bottom' and 'Left' . . .	1088
33.25.2	'Margins' Tab (2): 'Header' and 'Footer'	1090
33.25.3	'Margins' Tab (3): 'Center on page'	1090
33.26	Page Setup: 'Header/Footer' Tab	1091
33.26.1	Using the <code>print_header=</code> and <code>print_footer=</code> Op- tions	1091
33.26.2	<code>print_header=</code> Option vs. <code>title</code> Statement	1092
33.26.3	Using Excel/VBA code to Customize Margins	1093
33.27	Page Setup: 'Sheet' Tab	1095
33.27.1	'Sheet' Tab (1): 'Print area:'	1096
33.27.2	'Sheet' Tab (2a): 'Rows to repeat at top:'	1097
33.27.3	'Sheet' Tab (2b): 'Columns to repeat at left:'	1098
33.27.4	'Sheet' Tab (3a): 'Print' - 'Gridlines'	1099
33.27.5	'Sheet' Tab (3b): 'Print' - 'Black and white'	1099
33.27.6	'Sheet' Tab (3c): 'Print' - 'Draft quality'	1100
33.27.7	'Sheet' Tab (3d): 'Print' - 'Row and column headings'	1100
33.27.8	'Sheet' Tab (4): 'Page order'	1100

33.28 Managing Page Breaks in Printout 1101

 33.28.1 Using the `rowbreaks_interval='none'` Option to
 Avoid Page Breaks 1101

 33.28.2 Using the `rowbreaks_interval='proc'` Option to
 Insert a Page Break After Each Procedure 1102

 33.28.3 Using the `rowbreaks_interval='output'` Option
 to Insert a Page Break After Each Output 1103

 33.28.4 Using the `rowbreaks_count=` Option to Insert a Page
 Break After x Observations 1104

33.29 Note on Tagsets.Excelxp Destination 1105

Module 14 **1107**

34 `proc freq` by Example **1107**

 34.1 Data 1107

 34.1.1 Dataset `grp_all` 1107

 34.1.2 Dataset `grp_cnt` 1109

 34.1.3 Dataset `grp_800` 1110

 34.2 Calculating Statistics for One Variable 1111

 34.2.1 Displaying Statistics in the Default Destination 1111

 34.2.2 Outputting Statistics in a Dataset using *ODS* 1112

 34.3 Calculating Statistics for a Combination of Two Variables 1113

 34.3.1 Displaying Statistics in the Default Destination 1113

 34.3.2 Outputting Statistics in a Dataset using the *ODS* 1115

 34.4 Calculating Statistics by Group of Observations 1116

 34.4.1 Combining More Than Two Variables 1116

 34.4.2 Using a Variable Having No Missing Values (`by` State-
 ment) 1118

 34.4.3 Using a Variable Having Missing Values (`by` Statement) 1120

 34.5 Using a Variable Having Missing Values in the `table` Statement 1121

 34.5.1 Including Missing Values in the Total `missing` Option 1121

 34.5.2 How does `missprint` differ from `missing`? 1123

34.6 Using Weights (<code>weight</code> Statement)	1124
34.6.1 Using Weights (<code>weight</code> Statement)	1124
34.6.2 Including Weights Equal to Zero (<code>zeros</code> Option)	1125
34.7 Extracting Statistics in a Dataset using <code>out=</code> (Old Method)	1126
34.7.1 Stop Printing in the Active Destination(s) (<code>noprint</code> Option)	1126
34.7.2 Adding Statistics to the Output (<code>outpct</code> and <code>outcum</code> Options)	1127
34.7.3 Listing All Possible Combinations (<code>sparse</code> Option)	1128
34.8 How Many Distinct Values are there Per Variable? (<code>nlevels</code> Option)	1129
34.8.1 Example	1129
34.8.2 Exercise	1130
34.9 Plotting Frequencies and Percentages on a Graph	1131
34.9.1 Including Missing Values on a Graph	1131
34.9.2 Plotting the Statistics of a Variable on a Graph	1132
34.9.3 Plotting the Statistics Related to the Combination of Two Variables on a Graph	1134
34.9.4 What are the Different Graph Types? (<code>plots=</code>)	1136
34.9.5 Customizing the Axes using Variable Formats and Labels	1138
34.9.6 Specifying the Categories' Order on the Axis	1140
34.9.7 Printing Only the Graph	1144
34.9.8 Updating the Graph Title (<code>proc template</code>)	1148
34.10 Modifying the Borders in a <i>listing</i> destination	1151
34.10.1 Example with One Variable (<code>table treat</code>)	1151
34.10.2 Example with the Combination of Two Variables (<code>table</code> <code>treat*ny</code>)	1151
34.11 Next	1152
35 Creating a Bar Chart using <code>proc sgplot</code>	1153
35.1 Data	1154
35.1.1 Overview	1154
35.1.2 Dataset <code>gender</code>	1154

35.1.3 Dataset <code>gender_freq</code> (Exercise)	1155
35.1.4 Dataset <code>gender_freq_multi</code>	1155
35.1.5 Dataset <code>gender_m</code>	1155
35.1.6 Dataset <code>gender_m_freq</code> (Exercise)	1156
35.1.7 Dataset <code>gender_mx_freq</code> (Exercise)	1157
35.1.8 Dataset <code>gender_age</code>	1159
35.1.9 Dataset <code>gender_age_freq</code>	1160
35.1.10 Dataset <code>gender_age_freq_15</code>	1160
35.1.11 Dataset <code>gender_age_freq0_17</code>	1161
35.1.12 Dataset <code>age_freq</code>	1162
35.1.13 Dataset <code>gender_age_result</code>	1163
35.1.14 Dataset <code>gender_age_yr_result</code>	1164
35.1.15 Dataset <code>gender_age_yr</code>	1164
35.1.16 Dataset <code>gender_result</code>	1164
35.1.17 Dataset <code>gender_result_mean</code>	1165
35.1.18 Dataset <code>gender_yr_result_mean</code>	1165
35.1.19 Dataset <code>age_result_mean</code>	1166
35.1.20 Dataset <code>gender_course_mark_freq</code>	1167
35.1.21 Dataset <code>shoes_sales_sum</code>	1168
35.2 Plotting the Statistics of Discrete and Continuous Variables	1169
35.2.1 Setup	1170
35.2.2 Plotting a Bar for Males and a Bar for Females	1172
35.2.3 Plotting an Extra Bar for Missing Gender	1175
35.2.4 Plotting Males and Females by Age	1176
35.3 Managing Spaces and Frames	1177
35.3.1 Visualizing Spaces	1178
35.3.2 Removing All the Frames	1178
35.4 Managing <i>offsets</i> , Bar Width and Cluster Width	1180
35.4.1 Defining the Axis Offsets	1180
35.4.2 Managing Bar Width	1183
35.4.3 Managing Cluster Width	1185
35.5 Modifying the Bar Order	1187

35.5.1	Using <code>discreteorder=</code> When No Format is Available	1188
35.5.2	Using One Format	1189
35.5.3	Using Two Formats	1193
35.5.4	Using <code>reverse</code> to Reverse the Order of the Bars	1193
35.5.5	Using <code>values=</code> to List the Axis Values	1194
35.5.6	Ordering Segments and Bar Clusters	1195
35.6	Managing the Axes Ticks and Values	1197
35.6.1	Overview	1197
35.6.2	Removing Axis Elements	1197
35.6.3	Overview of the Option <code>fitpolicy=</code>	1199
35.6.4	Add Line Feeds to Axis Values	1200
35.6.5	Displaying a Subset of the Axis Values	1202
35.6.6	Rotating Axis Values	1203
35.6.7	Stacking Axis Values (Vertical Alignment)	1205
35.6.8	Staggering Axis Values	1207
35.6.9	Rotating When Staggering Does Not Fit	1209
35.6.10	Rotating When Splitting Does Not Fit	1210
35.6.11	Values on Linear Axis: Defining Minimum and Maximum Values	1212
35.6.12	Adding Minor Ticks	1213
35.6.13	Changing Tick Style	1214
35.6.14	Values on Linear Axis: Defining a Range of Values	1215
35.6.15	Values on Linear Axis: Uncorrelating Minimum and Maximum Values from the Value Ranges	1217
35.6.16	Values on Linear Axis: Set Multiple Ranges of Values	1218
35.6.17	Defining the Style of Axis Values	1219
35.7	Managing Axis Labels and Lines Hidden by the Axis	1220
35.7.1	Overview	1220
35.7.2	Defining Axis Label Style	1220
35.7.3	Changing the Position of Axes on the Graph	1222
35.7.4	Moving the Baseline and Customizing its Style	1222
35.7.5	Taking the Axis Away in order to See the Baseline	1224

35.7.6 Taking the Baseline Away in order to See the Zero Frequency Bar	1225
35.7.7 Changing Axis Color and Ticks	1226
35.8 Adding Statistics on the <code>Wall</code> Space	1229
35.8.1 Adding Statistics Above the Bars	1229
35.8.2 Adding a Statistic to Each Segment	1237
35.8.3 Replacing the Statistics Above the Bars with Images .	1242
35.8.4 Replacing the Statistics Above the Bars with Symbols	1248
35.8.5 Exercise	1250
35.9 Adding Statistics Outside of the <code>Wall</code> Space	1251
35.9.1 Adding Statistic Table Outside of the <code>Wall</code> Space and Increasing the <code>Graph</code> Space Padding	1251
35.9.2 Defining the Spacing Before/After each Row Create with the <code>xaxistable</code> Statement using the <code>pad=</code> Option	1253
35.9.3 Specifying the Statistic Table Location	1253
35.9.4 Removing the Label Associated with the Statistics . .	1256
35.9.5 Customizing the Style of the Statistics	1257
35.9.6 Customize the Style of Values, Labels and Titles . . .	1258
35.9.7 Two Options for Title Alignment	1260
35.9.8 Creating Indentations	1261
35.9.9 Display Statistics by Class and Use the Default Color Scheme	1263
35.9.10 Using <code>colorgroup=</code> and <code>attrid=id</code> to Define Col- ors for Statistical Values	1265
35.9.11 Using <code>textgroup=</code> and <code>textgroup=id</code> to Define Col- ors for Statistical Values	1266
35.10 Overlaying Two Bar Charts	1267
35.10.1 Creating Two Bar Charts	1268
35.10.2 Modifying the Label	1269
35.10.3 Modifying the Legend	1269
35.10.4 Modifying the Position and Filling Color of the Bars . .	1271
35.11 Managing the Legend	1272

35.11.1	Plotting a Graph with a Legend	1273
35.11.2	Removing the Legend	1274
35.11.3	Removing Title and Some Labels from the Legend . .	1275
35.11.4	Modifying the Order of the Labels	1276
35.11.5	Modifying the Default Layout	1277
35.11.6	Modifying the Legend Position	1278
35.11.7	Customizing the Style of the Legend Title and Labels	1280
35.11.8	Modifying the Legend Symbol Scale (<code>scale=</code>)	1281
35.11.9	Modifying the Legend Symbol Height and Width . . .	1281
35.11.10	Modifying the Legend Symbols (<code>legenditem</code>)	1282
35.11.11	Changing the Legend Padding Width	1286
35.12	Managing Background Elements	1286
35.12.1	Managing the Background	1287
35.12.2	Adding Lines or Color Bands behind the Bars	1289
35.12.3	Managing the Bar Borders	1292
35.12.4	Managing the Bar Pattern	1294
35.12.5	Applying a Gradient to the Bars Fill Color	1296
35.12.6	Managing the Fill and Border Colors of Bar Segments	1297
35.12.7	Defining the Fill Color of Bars Based on Associated Statistics	1302
35.13	Uniforming Axes and Legends	1305
35.13.1	Plotting Several Graphics using the <code>by</code> Statement . .	1305
35.13.2	Defining the Ratio Between the Width of the x-Axis and the Height of the y-Axis (<code>aspect=</code>)	1309
35.13.3	Creating an Animated <code>gif</code> Image	1310
35.14	Adding Annotations using the <code>refline</code> , <code>dropline</code> and <code>inset</code> Statements	1311
35.14.1	Adding Reference Lines (<code>refline</code>)	1312
35.14.2	Drawing Lines Up to a Given Point on the Graph (<code>dropline</code>)	1318
35.14.3	Adding a Text Area (<code>inset</code>)	1320
35.15	Storing Annotations in a Dataset	1324
35.15.1	Adding Draft in the Middle of the Graph (Function <code>text</code>)	1324

35.15.2	Adding Frequencies in Each Bar Segment (Function <code>text</code>)	1326
35.15.3	Adding a Frame with Rounded Corners (Function <code>rectangle</code>)	1327
35.15.4	Adding Arrows (Function <code>arrow</code>)	1329
35.15.5	Linking and Annotating Multiple Bars (Function <code>polyline</code>)	1331
35.15.6	Adding a Logo to the Graph (Function <code>image</code>)	1336
35.16	Saving a Graph Image in a Directory	1337
35.16.1	Saving a Graph Image in a Directory	1337
35.16.2	Modifying the File Name and Extension	1339
35.17	Plotting a Graph for the HTML Destination	1340
35.17.1	Creating a HTML File and Specifying the Path to the Graphic Image	1340
35.17.2	Creating Interactions Between the Graph Bars and the HTML Pages	1344
35.17.3	Customizing the Style of the Bars (Fill Color, Border Color and Axis Color)	1345
35.17.4	Customizing the Style of the Bar Segments (Fill Colors and Border Colors)	1350
35.17.5	Displaying Information Bubbles (<code>tip=</code>)	1353

Module 15 **1359**

36 Exporting Data using Engine/`proc export` (`xlsx/excel`) **1359**

36.1	Introduction	1359
36.1.1	Overview	1359
36.1.2	Choosing an Engine/DBMS Export Type	1360
36.1.3	Two Methods: Engine and <code>proc export</code>	1361
36.1.4	<code>validvarname=any</code> and <code>validmemname=extend</code>	1361
36.1.5	Why Create an Excel File without <code>ods excel?</code>	1363
36.2	Data	1364
36.2.1	Dataset <code>xxcustomer</code>	1364
36.2.2	Datasets <code>male</code> and <code>female</code>	1365

36.2.3 Dataset <code>xxcustomer_lbl</code>	1365
36.2.4 Dataset <code>specialchar_lbl</code>	1366
36.2.5 Dataset <code>xxcustomer_fmt</code>	1367
36.2.6 Dataset <code>date_time</code>	1368
36.2.7 Dataset <code>xxcustomer_agec</code>	1369
36.2.8 Dataset <code>xxcustomer_null</code>	1369
36.3 Creating an <i>Excel File</i> Having One Worksheet	1369
36.3.1 With <code>xlsx</code> Engine	1369
36.3.2 With <code>proc export (dbms=xlsx)</code>	1372
36.4 Can 32 Characters be Used to Name a Worksheet?	1374
36.4.1 With <code>xlsx</code> Engine	1374
36.4.2 With <code>proc export (dbms=xlsx)</code>	1374
36.5 Creating an <i>Excel File</i> Having Multiple Worksheets	1375
36.5.1 With the <code>xlsx</code> Engine	1375
36.5.2 With <code>proc export (dbms=xlsx)</code>	1378
36.6 Displaying Library Details	1378
36.6.1 Using Dictionaries	1378
36.6.2 Using the <code>list</code> Option of the <code>libname</code> Statement	1379
36.7 Emptying the Worksheet Contents (<code>excel</code> Engine Only)	1379
36.7.1 Step 1: Creating an <i>Excel File</i>	1380
36.7.2 Step 2: Listing Worksheets	1380
36.7.3 Step 3: Emptying a Worksheet	1381
36.8 Replacing the Contents of a Worksheet	1382
36.8.1 With <code>xlsx</code> Engine	1382
36.8.2 With <code>proc export (dbms=xlsx)</code>	1382
36.8.3 With the <code>excel</code> Engine	1383
36.8.4 With <code>proc export (dbms=excel)</code>	1384
36.9 Deleting a Worksheet	1384
36.9.1 Step 1: Creating an <i>Excel File</i> with Two Worksheets	1384
36.9.2 Step 2: Setting	1385
36.9.3 Step 3: Opening the <i>Excel File</i>	1385
36.9.4 Step 4: Deleting the Worksheet	1386

36.9.5 Step 5: Closing the <i>Excel File</i>	1386
36.9.6 Step 6: Resetting	1387
36.9.7 From Unix to Windows with proc download and proc upload	1387
36.10 Deleting a File	1388
36.10.1 Using the <code>fdelete</code> Function and the <code>filename</code> Statement	1388
36.10.2 Using the Functions <code>fdelete</code> and <code>filename</code>	1389
36.10.3 Via a Macro	1389
36.11 Replacing an <i>Excel File</i> by Another One	1390
36.11.1 Generic Solution	1390
36.11.2 With proc export (<code>dbms=excel</code>) Only	1390
36.12 Exporting Data without Header using the <code>putnames=</code> Statement	1391
36.13 Displaying Variable Labels	1392
36.13.1 With <code>xlsx</code>	1392
36.13.2 With <code>excel</code>	1393
36.14 Displaying Formatted Values	1394
36.15 Displaying Dates and Times	1395
36.15.1 Data	1395
36.15.2 With Formatted Variables	1395
36.15.3 With <code>excel</code> Only	1397
36.16 Displaying Missing Numeric Values	1397
36.17 Displaying Character Values in a Numeric Form	1398
36.17.1 Generic Solution	1398
36.17.2 With <code>excel</code> Only	1398
36.18 Displaying Numeric Values in a Character Form	1399
36.18.1 Generic Solution	1399
36.18.2 With <code>excel</code> Only	1399
36.19 Replacing Missing Character Values with a Value of Your Choice	1400
36.19.1 Generic Solution	1400
36.19.2 With <code>excel</code> Only	1400
36.20 Keeping or Dropping Variables	1402

36.20.1	Generic Solution	1402
36.20.2	With <code>proc export</code> (<code>dbms=excel</code> Only)	1402
36.21	Creating a 2003 Version <i>Excel File</i> using the <code>excel</code> Engine .	1403
36.22	Adding Data to Specific Cells using the DDE Language . . .	1403
37	Importing Data using Engine/<code>proc import</code> (<code>xlsx/excel</code>)	1406
37.1	Preparing the <i>Excel Files</i>	1407
37.1.1	Datasets	1407
37.1.2	Creating <i>Excel Files</i> with <code>proc export</code>	1411
37.1.3	Creating an <i>Excel File</i> using <code>ods excel</code>	1414
37.1.4	Creating <i>Excel Files</i> manually	1415
37.2	Displaying the Metadata of <i>Excel File</i> Data	1416
37.2.1	With <code>xlsx</code> Engine	1416
37.2.2	With <code>excel</code> Engine	1418
37.3	Reading a Worksheet	1419
37.3.1	Displaying the Data of a Worksheet without Creating a Dataset	1420
37.3.2	Creating a Dataset from the Data of a Worksheet using a Data Step	1421
37.3.3	Copying the Data of One or Multiple Worksheets into a Given Library	1421
37.3.4	Creating a Dataset from the Data of Worksheet using <code>proc import</code>	1422
37.4	Note on <code>indsname=</code> , Option of the <code>set</code> Statement	1424
37.5	Accessing the Data Starting at Row 1 using <code>getnames=</code> . . .	1428
37.5.1	With <code>proc import</code> (<code>dbms=xlsx</code>)	1428
37.5.2	With <code>xlsx</code> Engine	1428
37.6	Defining Variable Names When More Than 32 Characters Are Used, Etc.	1429
37.7	Selecting a Subset of the Data using Dataset Options	1431
37.7.1	Selecting Specific Variables using <code>keep=</code> or <code>drop=</code> .	1431
37.7.2	Selecting Observations Based on their Position in the Dataset using <code>firstobs=</code> and <code>obs=</code>	1432

37.7.3	Selecting Observations Based on a Condition using <code>where=</code>	1433
37.8	Accessing Data Starting at a Specific Row using <code>datarow=</code> (<code>proc import</code> Only)	1434
37.9	Accessing Data of an Area of the Worksheet using <code>range=</code> and <code>Alike</code>	1434
37.9.1	With <code>xlsx</code> Engine	1435
37.9.2	With <code>proc import</code> (<code>dbms=xlsx</code>)	1436
37.10	Working with an Intermediate CSV File to Remove Characters 1437	
37.10.1	Step 1: Converting an <i>Excel File</i> to a CSV File using Visual Basic	1437
37.10.2	Step 2: Removing Bullet Points and Line Feeds from the CSV File	1440
37.10.3	Step 3: Importing the Updated CSV File	1441
37.11	Dealing with Variable Type and Length	1442
37.11.1	What is the Default Result?	1442
37.11.2	Updating the Input File to Specify Variable Type and Length	1443
37.11.3	Specifying the Variable Type and Length using <code>dbdsopts=</code> (<code>excel</code> Only)	1446
37.11.4	Specifying the Maximum Variable Length using <code>dbmax_text=</code> (<code>excel</code> Only)	1447
37.12	Reading Dates	1448
37.12.1	Reading <i>Excel Dates, Times and Datetimes</i>	1448
37.12.2	Reading a Mix of Numeric and Character Values	1448
37.13	Importing a Worksheet Having No Observation	1451
37.14	Importing an Empty Worksheet	1451
37.15	Adding Rows of Data to an <i>Excel File</i> (<code>excel</code> Only)	1451
37.15.1	Adding Rows of Data using <code>proc append</code>	1452
37.15.2	Adding Rows of Data using <code>insert into</code> of <code>proc sql</code> 1454	

Module 16	1456
38 Arrays by Example	1456
38.1 Introduction	1456
38.1.1 What Kind of Problem Does an Array Solve?	1456
38.1.2 Benefits	1456
38.1.3 Example	1457
38.1.4 On the Agenda	1458
38.2 Data	1459
38.2.1 Dataset <code>one_num</code>	1459
38.2.2 Dataset <code>one_char</code>	1459
38.2.3 Dataset <code>lb</code>	1460
38.2.4 Dataset <code>france</code>	1460
38.2.5 Dataset <code>client</code>	1461
38.2.6 Dataset <code>clientt</code>	1461
38.2.7 Dataset <code>yr</code>	1462
38.2.8 Dataset <code>result</code>	1462
38.3 Defining an Array: Three Main Components	1463
38.3.1 Overview of the Three Main Components	1463
38.3.2 Zooming on Array Name	1463
38.3.3 Zooming on Number and Arrangement of Array Elements	1464
38.3.4 Zooming on Array Elements	1465
38.4 Defining an Array : Defining the Type of New Variables	1468
38.5 Defining an Array: Initializing the Values of New Variables	1469
38.5.1 Listing Each Value	1469
38.5.2 Repeating the Same Value x Times	1470
38.5.3 Defining Values in a Macro Variable	1470
38.5.4 Defining a Series of Values on Explicitly Given Variables	1470
38.5.5 Defining a Series of Values on Variables Named by Deduction	1471
38.6 Using an Array with a First Set of Data	1471
38.6.1 Using No Array	1471
38.6.2 Using One Statement per New Variable	1472

38.6.3	Creating a Loop	1472
38.6.4	Creating a Loop using the <code>lbound</code> and <code>hbound</code> Function	1473
38.6.5	Creating a Loop using the <code>dim</code> Function	1473
38.6.6	Creating a Loop using <code>do over</code>	1473
38.7	Using an Array with a Second Set of Data	1474
38.7.1	Creating New Variables (Exercise)	1474
38.7.2	Updating Existing Variables using <code>do over</code> (Exercise)	1474
38.8	Use Case: Updating the Values of Temporary Variables . . .	1475
38.9	Use Case: Searching for a Value in a Series of Variables . .	1476
38.10	Use Case: Transposing (1/3)	1477
38.10.1	Goal	1477
38.10.2	Using <code>proc transpose</code> (Exercise)	1478
38.10.3	Using an Array	1478
38.11	Use Case: Transposing (2/3)	1480
38.11.1	Goal	1480
38.11.2	Using <code>proc transpose</code> (Exercise)	1480
38.11.3	Using an Array	1481
38.12	Use Case: Transposing (3/3)	1481
38.12.1	Goal	1481
38.12.2	Using <code>proc transpose</code> (Exercise)	1482
38.12.3	Using an Array to Name Variables (1/2)	1482
38.12.4	Using an Array to Name Variables (2/2)	1485
38.13	Use Case: Using an Array with the <code>lag</code> and <code>lag2</code> Functions	1486
38.14	Defining and Using a Multidimensional Array	1488
38.14.1	Goal	1488
38.14.2	Using an Array but no Loop	1489
38.14.3	Using an Array and a Loop	1490

Indexes		1491
Global Statement		1491
Global Statement: <code>ods pdf</code>		1495
Global Statement: <code>ods excel</code>		1496

Data Step	1498
Metadata	1500
Variable and Operator	1502
Function and Routine	1505
Format and Informat	1508
Value	1510
Procedure	1511
Procedure: proc means	1515
Procedure: proc freq	1516
Procedure: proc report	1517
Procedure: proc sql	1519
Procedure: Graph	1521
Styles	1534
IT	1537
SAS Institute Inc.	1540
Exercises and Other Resources	1542