

Contents

Preface	xi
Introduction	1
1 Accessing Data	5
1.1 Using FORMATTED, LIST, and COLUMN input to read raw data files	5
1.1.1 LIST input	6
1.1.2 COLUMN input	9
1.1.3 FORMATTED input	10
1.1.4 NAMED input	11
1.2 Using INFILE statement options to control processing when reading raw data files	12
1.2.1 MISSOVER	12
1.2.2 PAD	13
1.2.3 TRUNCOVER	14
1.3 Using various components of an INPUT statement	14
1.3.1 Column pointer	14
1.3.2 Line pointer	16
1.4 Accessing existing SAS data sets with DATA step statements	17
1.5 Combining SAS data sets using DATA step statements	18
1.5.1 Concatenating	18
1.5.2 Interleaving	19
1.5.3 One-to-one reading	21
1.5.4 One-to-merging	22
1.5.5 Match-merging	23
1.5.6 Updating datasets	26
Chapter summary	29
Two-minute drill	30
Assessment exam	30
Assessment exam answers	31
Practice exam	32
Practice exam answers	32
Problems	51
Problem solutions	52
2 Creating Data Structures	55
2.1 Creating temporary and permanent SAS datasets	55
2.1.1 Two-level names	55
2.1.2 One-level names	56
2.2 Explaining how the DATA step is compiled and executed	56
2.2.1 Compilation phase	56
2.2.2 Execution phase	57
2.2.3 DATA step boundary and type	58
2.2.4 When variable values are set to "missing" automatically	58

2.3 Using the DATA step to export data to standard and comma-delimited raw data file	61
2.3.1 Creating an external file with column-aligned data	61
2.4 Creating and manipulating the SAS date values	62
2.4.1 Processing time data	63
2.4.2 Processing datetime values	64
2.5 Understanding the role of a BY statement in the DATA step	65
2.5.1 Understanding the BY group	65
2.5.2 How the DATA step identifies the BY group	65
2.6 Controlling which variables and observations in a SAS dataset are processed	66
2.6.1 Differences between IF and WHERE statements	68
2.6.2 Incorrect use of IF statement	68
2.6.3 When as an alternative to IF-THEN/ELSE coding	68
Chapter summary	70
Two-minute drill	71
Assessment exam	73
Assessment exam answers	73
Practice exam	78
Practice exam answers	
Problems	79
Problem solutions	80
3 Managing Data	85
3.1 Investigating SAS data libraries using Base SAS utility procedures	85
3.1.1 PROC DATASETS	85
3.1.2 PROC CONTENTS	89
3.1.3 PROC SORT	91
3.1.4 PROC COPY	91
3.1.5 PROC FORMAT	
3.2 Using SAS procedures to investigate and evaluate the quality of data	94
3.2.1 PROC FREQ	94
3.2.2 PROC MEANS	96
3.2.3 PROC SUMMARY	101
3.3 Sorting observations in a SAS dataset	102
3.3.1 BY statement	102
3.3.3 Other options	103
3.4 Modifying variable attributes using options and statements in the DATA step	104
3.5 Conditionally executing SAS statements	107
3.5.1 WHERE statement and comparison with IF statement	107
3.5.2 IF-THEN/ELSE statement	110
3.5.3 SELECT statement	111
3.6 Using assignment statements in the DATA step	112
3.7 Accumulating subtotals and totals using DATA step statements	112
3.7.1 SUM statement	112
3.7.2 SUM function	114
3.8 Using SAS functions to manipulate character data, numeric data, and SAS date values	114
3.9 Using SAS functions to convert character data to numeric data and vice versa	119

3.10 Processing data using DO Loops	119
3.10.1 DO	120
3.10.2 Iterative DO	120
3.10.3 DO UNTIL	123
3.10.4 DO WHILE	123
3.11 Processing data using SAS ARRAYS	124
Chapter summary	126
Two-minute drill	127
Assessment exam	131
Assessment exam answers	132
Practice exam	134
Practice exam answers	139
Problems	141
Problem solutions	142
4 Generating Reports	147
4.1 Generating list reports using the PRINT and REPORT procedures	147
4.1.1 PRINT procedure	147
4.1.2 REPORT procedure	152
4.2 Generating summary reports and frequency tables using Base SAS procedures	157
4.2.1 PROC TABULATE	157
4.3 Enhancing reports through the use of titles, footnotes, labels, SAS formats, user-defined formats, and SAS system reporting options	176
4.3.1 TITLE	176
4.3.2 FOOTNOTE	176
4.3.3 LABEL	177
4.3.4 FORMAT	177
4.3.5 SAS reporting options	178
4.4 Generating HTML reports using ODS statements	178
4.4.1 The ODS statement	179
4.4.2 Other literature on ODS	180
Chapter summary	180
Two-minute drill	181
Assessment exam	183
Assessment exam answers	184
Practice exam	186
Practice exam answers	189
Problems	191
Problems solutions	192
5 Handling Errors	195
5.1 Recognizing and correcting syntax errors	195
5.2 Identifying and correcting semantic errors	198
5.3 Examining and resolve execution-time errors	200
5.4 Identifying and correcting data errors	
5.5 Examining and resolving macro-related errors	205
5.6 Identifying and resolving programming logic errors	205

x Contents

5.6.1 DATA step debugger	207
5.6.2 PUT statement	208
Chapter summary	211
Two-minute drill	212
Assessment exam	213
Assessment exam answers	213
Practice exam	215
Practice exam answers	221
Problems	223
Problem solutions	229
Index	231