

## Table of Contents

---

List of Programs	ix
Preface	xvii
Acknowledgments	xix

### 1

## The RETAIN Statement

---

Introduction	1
Demonstrating a DATA Step with and without a RETAIN Statement	1
Generating Sequential SUBJECT Numbers Using a Retained Variable	7
Using a SUM Statement to Create SUBJECT Numbers	9
Demonstrating That Variables Read with a SET Statement Are Retained	10
A Caution When Using a RETAIN Statement	11

### 2

## The LAG and DIF Functions

---

Introduction	13
Using the LAG Function to Compute Differences	13
Demonstrating Some Related Functions: LAG2, LAG3, and So Forth	16
Demonstrating the DIF Function	17

### 3

## FIRST. and LAST. Temporary Variables

---

Introduction	19
How to Create FIRST. and LAST. Temporary Variables	19
Using More Than One BY Variable	22
A Simple Application Using FIRST. and LAST. Variables	24

**4****Flags and Counters**

---

Introduction	27
Using a Flag Variable to Determine If a Particular Event Ever Occurred in Any One of Several Observations for Each Subject	27
Counting the Number of Positive Outcomes for Each Patient	29

**5****Summarizing Data Using PROC MEANS and PROC FREQ**

---

Introduction	33
Using PROC MEANS to Output Means to a Data Set	34
Comparing CLASS and BY Statements with PROC MEANS	37
Computing Other Descriptive Statistics	38
Automatically Naming the Variables in the Output Data Set	40
Demonstrating an Alternative Way to Select Specific Descriptive Statistics for Selected Variables	41
Adding Additional Variables to the Summary Data Set Using an ID Statement	42
Specifying More Than One CLASS Variable	44
Selecting Multi-Way Breakdowns Using the TYPES Statement	47
Using the PROC MEANS CHARTYPE Option to Simplify the _TYPE_ Interpretation	49
Comparing PROC MEANS and PROC FREQ for Creating an Output Data Set Containing Counts	50
Counting Frequencies for a Two-Way Table	52

**6****Using PROC SQL with Longitudinal Data**

---

Introduction	55
Creating a Demonstration Data Set	55
A Simple SQL Query	57
Using PROC SQL to Count Observations within a BY Group	58
Demonstrating a HAVING Clause	59
Using PROC SQL to Create a Macro Variable	60
Using a Summary Function to Compute Group Means	62

**7****Restructuring SAS Data Sets Using Arrays**

---

Introduction	65
Creating a New Data Set with Several Observations per Subject from a Data Set with One Observation per Subject	66
Another Example of Creating Multiple Observations from a Single Observation	69
Going from One Observation per Subject to Many Observations per Subject Using Multidimensional Arrays	72
Demonstrating the Use of a Multidimensional Array	74
An Alternative Program	77
Another Example of a Multidimensional Array	78

**8****Restructuring SAS Data Sets Using PROC TRANSPOSE**

---

Introduction	81
Going from One Observation to Several Observations	81
Another Example of Creating Multiple Observations from a Single Observation	84
Going from One Observation per Subject to Many Observations per Subject	86
Creating a Data Set with One Observation per Subject from a Data Set with Multiple Observations per Subject	88

**9****Study One: Operations on a Clinical Database**

---

Introduction	94
Description of the Clinical Data Set	94
Selecting the First or Last Visit for Each Patient	95
Computing Differences between the First and Last Visits	97
Another Method of Computing Differences between the First and Last Visits	99
Computing Differences between Every Visit	100
Counting the Number of Visits for Each Patient (DATA Step Approach)	101
Counting the Number of Visits for Each Patient (PROC FREQ)	103
Counting the Number of Visits for Each Patient (PROC MEANS)	103
Counting the Number of Visits for Each Patient (PROC SQL)	104
Selecting All Patients with $n$ Visits (DATA Step Approach)	105
Selecting All Patients with $n$ Visits (PROC FREQ Approach)	106
Selecting All Patients with Two Visits (Using PROC SQL)	107
Selecting All Patients with Two Visits (Using SQL in One Step)	107
Using PROC SQL to Create a Macro Variable	108
Computing Summary Statistics for Each Patient (Using PROC MEANS)	109

Computing Summary Statistics for Each Patient (Using PROC SQL)	110
Adding a Value from the First Visit to Each Subsequent Visit	111
Looking Ahead: Making a Decision about the Current Observation Based on Information in the Next Observation	114
Using Flags to Ascertain Vitamin Use	117
Using PROC FREQ to Ascertain Vitamin Use	118
Counting the Number of Routine Visits for Each Patient	119

**10****Study Two: Operations on Daily Weather Data and Ozone Levels**


---

Introduction	121
The OZONE Data Set	121
Computing Weekly Averages	122
Using the MOD Function to Group Data Values	125
Computing a Moving Average for a Single Variable	127

**11****Study Three: Producing Summary Reports on a Library Data Set**


---

Introduction	129
Computing the Number of Books per Patron Visit and by Library	130
Computing the Number of Patrons by Day of Week and Library	134
Generating a Table of LC Categories by Age Group and Overall	135

**12****Useful Macros**

---

Introduction	141
Listing All or Part of a Data Set	141
Computing Differences between Successive Observations	143
Computing Differences between the First and Last Observations per Subject	145
Computing a Moving Average	147
Computing Cell Means and Counts	149
Counting the Number of Observations per Subject	151

**Appendix****List of Data Files and SAS Data Sets**

---

The TEST_SCORES Data Set	153
The CLINICAL Data Set	154
The CLIN_FIRST Data Set	156
The OZONE Data Set	157
The LIBRARY Data Set	160

**Index**

165