## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Programs</td>
<td>xi</td>
</tr>
<tr>
<td>About This Book</td>
<td>xvii</td>
</tr>
<tr>
<td>About the Author</td>
<td>xxi</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>xxiii</td>
</tr>
<tr>
<td><strong>Part 1: Getting Acquainted with the SAS Studio Environment</strong></td>
<td></td>
</tr>
<tr>
<td>Chapter 1: Introduction to the SAS University Edition</td>
<td>1</td>
</tr>
<tr>
<td>Introduction: An Overview of SAS and the SAS University Edition</td>
<td>3</td>
</tr>
<tr>
<td>Obtaining Your Free Copy of SAS</td>
<td>3</td>
</tr>
<tr>
<td>Installing the SAS University Edition</td>
<td>6</td>
</tr>
<tr>
<td>Conclusion</td>
<td>10</td>
</tr>
<tr>
<td>Chapter 2: The SAS Studio Interface</td>
<td>11</td>
</tr>
<tr>
<td>Introduction</td>
<td>11</td>
</tr>
<tr>
<td>Exploring the Built-In Data Sets</td>
<td>13</td>
</tr>
<tr>
<td>Sorting Your Data</td>
<td>17</td>
</tr>
<tr>
<td>Switching Between Column Names and Column Labels</td>
<td>19</td>
</tr>
<tr>
<td>Resizing Tables</td>
<td>20</td>
</tr>
<tr>
<td>Creating Filters</td>
<td>20</td>
</tr>
<tr>
<td>Conclusion</td>
<td>23</td>
</tr>
<tr>
<td>Chapter 3: Importing Your Own Data</td>
<td>25</td>
</tr>
<tr>
<td>Introduction</td>
<td>25</td>
</tr>
<tr>
<td>Exploring the Utilities Tab</td>
<td>25</td>
</tr>
<tr>
<td>Importing Data from an Excel Workbook</td>
<td>27</td>
</tr>
<tr>
<td>Listing the SAS Data Set</td>
<td>32</td>
</tr>
<tr>
<td>Importing an Excel Workbook with Invalid SAS Variable Names</td>
<td>33</td>
</tr>
<tr>
<td>Importing an Excel Workbook That Does Not Have Variable Names</td>
<td>34</td>
</tr>
</tbody>
</table>
# Table of Contents

How SAS Works: A Look Inside the “Black Box” ................................................................. 106
Conclusion ......................................................................................................................... 106

**Chapter 8: Reading Data from External Files** ............................................................... 107
Introduction ....................................................................................................................... 107
Reading Data Values Separated by Delimiters ............................................................... 107
  Reading Comma-Separated Values Files ........................................................................ 110
  Reading Data Separated by Other Delimiters ............................................................... 111
Reading Data in Fixed Columns ...................................................................................... 112
  Column Input .................................................................................................................. 112
  Formatted Input ............................................................................................................. 114
Reading Excel Files ......................................................................................................... 116
  Reading from an Excel Workbook where Column Headings Are Invalid SAS Variable Names ......................................................................................................................... 118
Conclusion ....................................................................................................................... 119
Problems ........................................................................................................................... 120

**Chapter 9: Reading and Writing SAS Data Sets** ......................................................... 123
What’s a SAS Data Set? .................................................................................................... 123
Temporary versus Permanent SAS Data Sets ............................................................... 125
Shared Folders: Communicating between Your Virtual Computer and Your Real Computer ......................................................................................................................... 126
Creating a Shared Folder and Reading Data from It ...................................................... 128
Creating a Permanent SAS Data Set ............................................................................... 133
Reading from a Permanent SAS Data Set ..................................................................... 134
Conclusion ....................................................................................................................... 135
Problems ........................................................................................................................... 136

**Chapter 10: Creating Formats and Labels** ................................................................. 137
What Is a SAS Format and Why Is It Useful? ................................................................. 137
Using SAS Built-in Formats .......................................................................................... 144
More Examples to Demonstrate How to Write Formats ............................................... 145
Describing the Difference between a FORMAT Statement in a Procedure and a FORMAT Statement in a DATA Step ................................................................. 146
Making Your Formats Permanent .................................................................................. 147
Creating Variable Labels ............................................................................................... 150
Conclusion ....................................................................................................................... 151
Problems ........................................................................................................................... 151
Describing Some Useful Character Functions

Function Names: LENGTHN and LENGTHC ................................................................. 231
Function Names: TRIMN and STRIP ........................................................................... 233
Function Names: UPCASE, LOWCASE, and PROPCASE (Functions That Change Case) ................................................................. 234
Function Name: PUT .................................................................................................. 235
Function Name: SUBSTRN (Newer Version of the SUBSTR Function) ................. 236
Function Names: FIND and FINDC ............................................................................ 237
Function Names: CAT, CATS, and CATX ................................................................. 238
Function Names: COUNT and COUNTC ................................................................. 240
Function Name: COMPRESS ..................................................................................... 241
Function Name: SCAN ............................................................................................. 244
CALL Routine: CALL MISSING .................................................................................. 245
Function Names: NOTDIGIT, NOTALPHA, and NOTALNUM .............................. 246
Function Names: ANYDIGIT, ANYALPHA, and ANYALNUM .............................. 246
Function Name: TRANWRD ...................................................................................... 247
Conclusion .................................................................................................................. 247
Problems ..................................................................................................................... 248

Chapter 16: Working with Multiple Observations per Subject.............. 251
Introduction ............................................................................................................... 251
Useful Tools for Working with Longitudinal Data ............................................. 251
Describing First. and Last. Variables ................................................................. 252
Computing Visit-to-Visit Differences ................................................................. 253
Computing Differences Between the First and Last Visits ............................ 255
Counting the Number of Visits for Each Patient ............................................. 256
Conclusion ............................................................................................................... 258
Problems .................................................................................................................. 258

Chapter 17: Describing Arrays ............................................................... 261
Introduction ............................................................................................................... 261
What Is an Array? ................................................................................................. 261
Describing a Character Array ............................................................................. 264
Performing an Operation on Every Numeric Variable in a Data Set ............ 265
Performing an Operation on Every Character Variable in a Data Set .......... 265
Converting a Data Set with One Observation per Subject into a Data Set with Multiple Observations per Subject.............................. 266
Chapter 18: Displaying Your Data ........................................................... 273
Introduction .......................................................................................... 273
Producing a Simple Report Using PROC PRINT ........................................... 273
Using Labels Instead of Variable Names as Column Headings ..................... 277
Including a BY Variable in a Listing ....................................................... 279
Including the Number of Observations in a Listing ..................................... 280
Conclusion ............................................................................................ 280
Problems ............................................................................................... 281

Chapter 19: Summarizing Data with SAS Procedures ............................. 283
Introduction .......................................................................................... 283
Using PROC MEANS (with the Default Options) ......................................... 283
Using PROC MEANS Options to Customize the Summary Report ................. 286
Computing Statistics for Each Value of a BY Variable ................................. 288
Using a CLASS Statement Instead of a BY Statement ................................. 289
Including Multiple CLASS Variables with PROC MEANS ........................... 290
Statistics Broken Down Every Way ............................................................ 291
Using PROC MEANS to Create a Summary Data Set .................................... 293
Letting PROC MEANS Name the Variables in the Output Data Set .................. 294
Creating a Summary Data Set with CLASS Variables ................................. 295
Using a Formatted CLASS Variable .......................................................... 297
Demonstrating PROC UNIVARIATE ........................................................ 298
Conclusion ............................................................................................ 302
Problems ............................................................................................... 303

Chapter 20: Computing Frequencies ...................................................... 305
Introduction .......................................................................................... 305
Creating a Data Set to Demonstrate Features of PROC FREQ ....................... 305
Using PROC FREQ to Generate One-Way Frequency Tables ....................... 307
Creating Two-Way Frequency Tables ....................................................... 309
Creating Three-Way Frequency Tables ..................................................... 311
Using Formats to Create Groups for Numeric Variables ............................... 313