

SAS/STAT[®] 14.2 User's Guide

Introduction

This document is an individual chapter from *SAS/STAT® 14.2 User's Guide*.

The correct bibliographic citation for this manual is as follows: SAS Institute Inc. 2016. *SAS/STAT® 14.2 User's Guide*. Cary, NC: SAS Institute Inc.

SAS/STAT® 14.2 User's Guide

Copyright © 2016, SAS Institute Inc., Cary, NC, USA

All Rights Reserved. Produced in the United States of America.

For a hard-copy book: No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute Inc.

For a web download or e-book: Your use of this publication shall be governed by the terms established by the vendor at the time you acquire this publication.

The scanning, uploading, and distribution of this book via the Internet or any other means without the permission of the publisher is illegal and punishable by law. Please purchase only authorized electronic editions and do not participate in or encourage electronic piracy of copyrighted materials. Your support of others' rights is appreciated.

U.S. Government License Rights; Restricted Rights: The Software and its documentation is commercial computer software developed at private expense and is provided with RESTRICTED RIGHTS to the United States Government. Use, duplication, or disclosure of the Software by the United States Government is subject to the license terms of this Agreement pursuant to, as applicable, FAR 12.212, DFAR 227.7202-1(a), DFAR 227.7202-3(a), and DFAR 227.7202-4, and, to the extent required under U.S. federal law, the minimum restricted rights as set out in FAR 52.227-19 (DEC 2007). If FAR 52.227-19 is applicable, this provision serves as notice under clause (c) thereof and no other notice is required to be affixed to the Software or documentation. The Government's rights in Software and documentation shall be only those set forth in this Agreement.

SAS Institute Inc., SAS Campus Drive, Cary, NC 27513-2414

November 2016

SAS® and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.

SAS software may be provided with certain third-party software, including but not limited to open-source software, which is licensed under its applicable third-party software license agreement. For license information about third-party software distributed with SAS software, refer to <http://support.sas.com/thirdpartylicenses>.

Chapter 2

Introduction

Contents

Overview of SAS/STAT Software	9
Experimental Software	9
About This Book	10
Chapter Organization	10
Typographical Conventions	11
Options Used in Examples	11
Where to Turn for More Information	12
Accessing the SAS/STAT Sample Library	12
Sashelp Data Sets	12
SAS Technical Support Services	13
Related SAS Software	13
SAS/IML Software	13
Base SAS Software	13
SAS/ETS Software	15
SAS/GRAPH Software	15
SAS/OR Software	15
SAS/QC Software	15
SAS/IML Studio	16

Overview of SAS/STAT Software

SAS/STAT software provides comprehensive statistical tools for a wide range of statistical analyses, including analysis of variance, categorical data analysis, cluster analysis, multiple imputation, multivariate analysis, nonparametric analysis, power and sample size computations, psychometric analysis, regression, survey data analysis, and survival analysis. A few examples include nonlinear mixed models, generalized linear models, correspondence analysis, and robust regression. The software is constantly being updated to reflect new methodology. SAS/STAT software provides more than 90 procedures for statistical analysis.

Experimental Software

Experimental software is sometimes included as part of a production-release product. It is provided to (sometimes targeted) customers in order to obtain feedback. All experimental uses are marked Experimental

in this document. Whenever an experimental procedure, statement, or option is used, a message is printed to the SAS log to indicate that it is experimental.

The design and syntax of experimental software might change before any production release. Experimental software has been tested prior to release, but it has not necessarily been tested to production-quality standards, and so should be used with care.

About This Book

Since SAS/STAT software is a part of the SAS System, this book assumes that you are familiar with Base SAS software and with the books *SAS Language Reference: Concepts* and the *Base SAS Procedures Guide*. It also assumes that you are familiar with basic SAS System concepts such as creating SAS data sets with the DATA step and manipulating SAS data sets with the procedures in Base SAS software (for example, the PRINT and SORT procedures).

Chapter Organization

This book is organized as follows.

Chapter 1, “[What’s New in SAS/STAT 14.2](#),” provides information about the changes and enhancements to SAS/STAT software in SAS 9.3.

Chapter 2, this chapter, provides an overview of SAS/STAT software and summarizes related information, products, and services. The remaining introductory chapters (Chapter 3, “[Introduction to Statistical Modeling with SAS/STAT Software](#),” through Chapter 17, “[Introduction to Structural Equation Modeling with Latent Variables](#)”) provide an introduction to the broad areas covered by SAS/STAT software.

Chapter 18, “[Introduction to Power and Sample Size Analysis](#),” provides documentation for the Power and Sample Size Application (PSS).

Chapter 19, “[Shared Concepts and Topics](#),” provides information about topics that are common to multiple procedures. Topics include parameterization of model effects, the EFFECT statement, and the NLOPTIONS statement. Starting in SAS/STAT 9.22, this chapter also documents the following statements that are used for postfitting analysis and are common across many modeling procedures: EFFECTPLOT, ESTIMATE, LSMEANS, LSMESTIMATE, SLICE, STORE, and TEST.

Chapter 20, “[Using the Output Delivery System](#),” explains the fundamentals of using the Output Delivery System (ODS) to manage your SAS output.

Chapter 21, “[Statistical Graphics Using ODS](#),” Chapter 22, “[ODS Graphics Template Modification](#),” and Chapter 23, “[Customizing the Kaplan-Meier Survival Plot](#),” describe the extension to ODS that enables many statistical procedures to create statistical graphics as easily as they create tables.

Subsequent chapters describe the SAS procedures that make up SAS/STAT software. These chapters appear in alphabetical order by procedure name and are organized as follows:

- The “Overview” section provides a brief description of the analysis provided by the procedure.
- The “Getting Started” section provides a quick introduction to the procedure through a simple example.

- The “Syntax” section describes the SAS statements and options that control the procedure.
- The “Details” section discusses methodology and miscellaneous details, such as ODS tables and ODS graphics.
- The “Examples” section contains examples that use the procedure.
- The “References” section contains references for the methodology and for examples of the procedure.

Following the chapters on the SAS/STAT procedures, Chapter A, “[Special SAS Data Sets](#),” documents the special SAS data sets that are associated with SAS/STAT procedures and Chapter B, “[Sashelp Data Sets](#),” documents Sashelp data sets that are used in this book.

Typographical Conventions

This book uses several type styles for presenting information. The following list explains the meaning of the typographical conventions used in this book:

roman	is the standard type style used for most text.
UPPERCASE ROMAN	is used for SAS statements, options, and other SAS language elements when they appear in the text. However, you can enter these elements in your own SAS programs in lowercase, uppercase, or a mixture of the two.
UPPERCASE BOLD	is used in the “Syntax” sections’ initial lists of SAS statements and options.
<i>oblique</i>	is used for user-supplied values for options in the syntax definitions. In the text, these values are written in <i>italic</i> .
VariableName	is used for the names of variables and data sets when they appear in the text.
bold	is used to refer to matrices and vectors.
<i>italic</i>	is used for terms that are defined in the text, for emphasis, and for references to publications.
monospace	is used for example code. In most cases, this book uses lowercase type for SAS code.

Options Used in Examples

The HTMLBLUE style is used to create the graphs and the HTML tables that appear in the online documentation. The PEARLJ style is used to create the PDF tables that appear in the documentation. A style template controls stylistic elements such as colors, fonts, and presentation attributes. You can specify a style template in an ODS destination statement as follows:

```
ods html style=HTMLBlue;
. . .
ods html close;

ods pdf style=PearlJ;
```

```
. . .
ods pdf close;
```

For more information about styles, see Chapter 21, “Statistical Graphics Using ODS.”

Most of the PDF tables are produced by using the following SAS System option:

```
options papersize=(6.5in 9in);
```

If you run the examples, you might get slightly different output. This is a function of the SAS System options that are used and the precision that your computer uses for floating-point calculations.

Where to Turn for More Information

This section describes other sources of information about SAS/STAT software.

Accessing the SAS/STAT Sample Library

The SAS/STAT sample library includes many examples that illustrate the use of SAS/STAT software, including the examples used in this documentation. To access these sample programs from the SAS windowing environment, select **Help ► SAS Help and Documentation** from the main menu. Then expand the **Learning to Use SAS, Sample SAS Programs**, and **SAS/STAT** items.

You can also access the SAS/STAT sample library at http://support.sas.com/documentation/onlinedoc/stat/ex_code/index.html.

Sashelp Data Sets

SAS provides over 200 data sets in the Sashelp library. These data sets are available for you to use for examples and for testing code. For example, the following step uses the Sashelp.Class data set:

```
proc reg data=sashelp.class;
    model weight = height;
run; quit;
```

You do not need to provide a DATA step to use Sashelp data sets.

The following steps list all of the data sets that are available in Sashelp:

```
ods select none;
proc contents data=sashelp._all_;
    ods output members=m;
run;
ods select all;

proc print;
    where memtype = 'DATA';
run;
```

The results of these steps (over 200 data set names) are not displayed.

The following steps provide detailed information about the Sashelp data sets:

```
proc contents data=sashelp._all_;  
run;
```

The results of this step (hundreds of pages of PROC CONTENTS information) are not displayed. See Chapter B, “[Sashelp Data Sets](#),” for more information about Sashelp data sets.

SAS Technical Support Services

As with all SAS products, the SAS Technical Support staff is available to respond to problems and answer technical questions regarding the use of SAS/STAT software. Go to <http://support.sas.com/techsup> for more information.

Related SAS Software

Many features not found in SAS/STAT software are available in other parts of the SAS System. If you do not find something you need in SAS/STAT software, try looking for the feature in the following SAS software products.

SAS/IML Software

SAS/IML software gives you access to a powerful and flexible programming language (Interactive Matrix Language) in a dynamic, interactive environment. The fundamental object of the language is a data matrix. You can use SAS/IML software interactively (at the statement level) to see results immediately, or you can store statements in a module and execute them later. The programming is dynamic because necessary activities such as memory allocation and dimensioning of matrices are done automatically. SAS/IML software is of interest to users of SAS/STAT software because it enables you to program your methods in the SAS System.

Base SAS Software

The features provided by SAS/STAT software are in addition to the features provided by Base SAS software. Many data management and reporting capabilities you will need are part of Base SAS software. See *SAS Language Reference: Concepts*, *SAS Language Reference: Dictionary*, and the *Base SAS Procedures Guide* for documentation of Base SAS software.

Base SAS software includes many useful SAS procedures. Base SAS procedures are documented in the *Base SAS Procedures Guide*. The following is a list of Base SAS procedures you might find useful:

CORR	computes correlations.
RANK	computes rankings or order statistics.
STANDARD	standardizes variables to a fixed mean and variance.
MEANS	computes descriptive statistics and summarizes or collapses data over cross sections.
TABULATE	prints descriptive statistics in tabular format.
UNIVARIATE	computes descriptive statistics.

ODS Graphics

Base SAS software provides the following:

- The SG family of procedures provides a simple syntax for creating stand-alone statistical graphics. These procedures include SGPLOT, SGSCATTER, and SGPANEL, which provide a simple and convenient syntax for producing many types of displays. They are particularly convenient for exploring and presenting data. See the *SAS ODS Graphics: Procedures Guide* for more information.
- The GTL (Graph Template Language) and the SGRENDER procedure provide a powerful syntax for creating customized graphs. See the *SAS Graph Template Language: User's Guide* and the *SAS Graph Template Language: Reference* for more information. You can also use the GTL to modify the SAS templates that are provided for use with SAS/STAT procedures. See Chapter 22, “[ODS Graphics Template Modification](#),” for more information about template modification.
- The ODS Graphics Editor enables you to make immediate changes to ODS Graphics by using a point-and-click interface. See the *SAS 9.4 ODS Graphics Editor: User's Guide* for more information.

See Chapter 21, “[Statistical Graphics Using ODS](#),” for more information about ODS Graphics.

SAS DATA Step

The DATA step is your primary tool for reading and processing data in the SAS System. The DATA step provides a powerful general purpose programming language that enables you to perform all kinds of data processing tasks. The DATA step is documented in *SAS Language Reference: Concepts*.

Base SAS Procedures

Base SAS software includes many useful SAS procedures. Base SAS procedures are documented in the *Base SAS Procedures Guide*. The following is a list of Base SAS procedures you might find useful:

CORR	computes correlations.
RANK	computes rankings or order statistics.
STANDARD	standardizes variables to a fixed mean and variance.
MEANS	computes descriptive statistics and summarizes or collapses data over cross sections.
TABULATE	prints descriptive statistics in tabular format.
UNIVARIATE	computes descriptive statistics.

SAS/ETS Software

SAS/ETS software provides SAS procedures for econometrics and time series analysis. It includes capabilities for forecasting, systems modeling and simulation, seasonal adjustment, and financial analysis and reporting. In addition, SAS/ETS software includes an interactive time series forecasting system.

SAS/GRAPH Software

SAS/GRAPH software includes procedures that create two- and three-dimensional plots and charts.

SAS/OR Software

SAS/OR software provides SAS procedures for operations research and project planning and includes a point-and-click interface to project management. Its capabilities include the following:

- solving transportation problems
- linear, integer, and mixed-integer programming
- nonlinear programming
- scheduling projects
- plotting Gantt charts
- drawing network diagrams
- solving optimal assignment problems
- network flow programming

SAS/OR software might be of interest to users of SAS/STAT software for its mathematical programming features. In particular, the NLP procedure in SAS/OR software solves nonlinear programming problems, and it can be used for constrained and unconstrained maximization of user-defined likelihood functions.

SAS/QC Software

SAS/QC software provides a variety of procedures for statistical quality control and quality improvement. SAS/QC software includes procedures for the following:

- Shewhart control charts
- cumulative sum control charts
- moving average control charts

- process capability analysis
- Ishikawa diagrams
- Pareto charts
- experimental design

SAS/QC software also includes the ADX interface for experimental design.

SAS/IML Studio

Many users of SAS/STAT software will be interested in SAS/IML Studio, which is new in SAS 9.2 software. Formerly known as SAS Stat Studio, SAS/IML Studio is a tool for data exploration and analysis; it provides a highly flexible programming environment in which you can run SAS/STAT or SAS/IML analyses and display the results with dynamically linked graphics and data tables. You can also move seamlessly between interactive analysis and programatically driven analysis. SAS/IML Studio is intended for data analysts who write SAS programs to solve statistical problems but need more versatility for data exploration and model building.

The programming language in SAS/IML Studio, which is called *IMLPlus*, is an enhanced version of the SAS/IML programming language. IMLPlus extends SAS/IML by providing features such as the ability to create and manipulate dynamically linked graphs and the ability to call SAS procedures.

SAS/IML Studio also includes an interface to the R language. The IMLPlus language provides functions that transfer data between SAS data sets and R data frames, and between SAS/IML matrices and R matrices.

SAS/IML Studio runs on a PC in the Microsoft Windows operating environment. For more information about SAS/IML Studio, see the *SAS/IML Studio User's Guide* and *SAS/IML Studio for SAS/STAT Users*.

Index

Base SAS software, [13](#)

control charts, [15](#)

CORR procedure, [14](#)

DATA step, [14](#)

descriptive statistics, *see also* UNIVARIATE procedure

experimental design, [15](#)

IMLPlus, [16](#)

MEANS procedure, [14](#)

operations research, [15](#)

order statistics, *see* RANK procedure

Pareto charts, [15](#)

RANK procedure, [14](#)

order statistics, [14](#)

SAS data set

DATA step, [14](#)

summarizing, [14](#)

SAS Stat Studio, [16](#)

SAS/ETS software, [15](#)

SAS/GRAPH software, [15](#)

SAS/IML software, [13](#)

SAS/IML Studio, [16](#)

SAS/OR software, [15](#)

SAS/QC software, [15](#)

Shewhart control charts, [15](#)

STANDARD procedure, [14](#)

standardized values, [14](#)

standardizing

values (STANDARD), [14](#)

statistical

quality control, [15](#)

TABULATE procedure, [14](#)

UNIVARIATE procedure, [14](#)