



THE
POWER
TO KNOW.

SAS[®] Visual Analytics 7.3

Getting Started with Data Preparation

The correct bibliographic citation for this manual is as follows: SAS Institute Inc. 2015. *SAS® Visual Analytics 7.3: Getting Started with Data Preparation*. Cary, NC: SAS Institute Inc.

SAS® Visual Analytics 7.3: Getting Started with Data Preparation

Copyright © 2015, 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, North Carolina 27513-2414.

October 2015

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.

Contents

<i>Using This Book</i>	v
<i>Accessibility</i>	vii
Chapter 1 • Introduction	1
About Data Preparation	1
Accessing Data	2
Chapter 2 • Basic Tasks	3
Introduction	3
Create a Data Query	3
Create a LASR Star Schema	4
Append Data to a Table	5
Chapter 3 • Specific Tasks	7
Introduction	7
Cheat Sheet for Data Queries	8
Cheat Sheet for LASR Star Schemas	9
Chapter 4 • Quick Reference	11
The Data Builder	11
Where to Find Additional Resources	12

Using This Book

Audience

This book covers the basics of how to prepare data in SAS Visual Analytics. The emphasis is on introductory instructions, how-to hints, and quick reference information.

Requirements

Prerequisites

If you choose to perform the tasks in this book, you need the following software, information, and privileges:

- access to SAS Visual Analytics 7.3
- a supported web browser (see the [SAS support site](#) for supported browsers and versions)
- a supported version of the Adobe Flash Player (see the [SAS support site](#) for supported versions)
- an account that can log on to SAS Visual Analytics
- access to the data source that you want to use (for example, a SAS data set, a Microsoft Excel file, an Oracle database, and so on)
- Write permission to at least one library (without the necessary permission, you cannot save tables, data queries, and so on)

- data preparation capabilities (without the necessary capabilities, you cannot see certain user interface elements)

System Requirements

Detailed system requirements and additional web browsers that are supported are documented on the [SAS support site](#).

Key Terms

This book makes specialized use of the following terms:

delimited file	a text file that contains data that is separated by commas, tabs, or custom delimiters.
LASR	the SAS in-memory analytic and data functionality that SAS Visual Analytics uses. LASR is an adjective, not an acronym.
SAS data set	a table that is in the SAS7BDAT file format.
spreadsheet	Microsoft Excel workbook (XLSX, XLSM, and XLSB) file or Microsoft Excel 97 to 2003 workbook (XLS) file.
data query	the primary method for preparing data in SAS Visual Analytics.
LASR star schema	similar to a typical star schema in a relational database. Usually includes a fact table that is surrounded by dimension tables.

Accessibility

For information about the accessibility of this product, see [Accessibility Features of SAS Visual Analytics 7.3](#) at support.sas.com.

1

Introduction

<i>About Data Preparation</i>	1
<i>Accessing Data</i>	2

About Data Preparation

Data preparation involves getting data ready for use in reports and explorations. In SAS Visual Analytics, you can prepare data using the SAS Visual Data Builder (the data builder).

Here are some of the tasks that you can perform using the data builder:

- join data from multiple tables
- append tables (data union)
- create aggregations
- sort data
- filter data
- create calculated columns
- pivot data

Data preparation is needed only if the data that you want to use is not ready as is. If the data is ready as is, you can start working with it immediately in the designer or the explorer.

Accessing Data

If the data that you want to prepare is not already loaded in SAS Visual Analytics, then you can import it into the SAS Visual Analytics environment. Here are the types of data that you can import:

- a file from your local machine (for example, a Microsoft Excel spreadsheet, a delimited text file (CSV), or a SAS data set)
- a SAS data set on a server
- a table from a database (for example, a server database or a Hadoop database)
- data from social media and other sources (for example, Facebook, Google Analytics, or Twitter)

To import data, click **File** ► **Import Data**. In the Import Data window, click the type of data that you want to import.

For more information about accessing data, see [SAS Visual Analytics: Getting Started with Data on Windows](#).

2

Basic Tasks

<i>Introduction</i>	3
<i>Create a Data Query</i>	3
<i>Create a LASR Star Schema</i>	4
<i>Append Data to a Table</i>	5

Introduction



This chapter provides step-by-step instructions to guide you through basic data preparation tasks in the data builder. The purpose of this chapter is to provide a brief, hands-on orientation to some of the workflows for preparing data.

Before you begin, access the data builder by clicking **Data Preparation** from SAS Home.

Create a Data Query

To create a data query:



- 1 On the menu bar, click **File** ► **New Data Query**.

- 2 Drag and drop tables from the SAS Folders tree onto the workspace.
- 3 Select the columns that you want to include in the data query by clicking on the column names in the original tables in the workspace.
- 4 Make other optional changes. For more information about optional changes, see [“Cheat Sheet for Data Queries” on page 8](#).
- 5 Click the **Outputs** tab, and specify the following:
 - In the **Table** field, specify a unique name for the output table or choose an existing output table to overwrite. This is the name of the file that you will use to create reports and explorations.
 - In the **Location** field, specify a folder where you want the data query to appear. The **My Folder** folder is private. Only you can use the data that you save to **My Folder**. If you want the data query to be accessible to others, save it to the **Shared Data** folder.
 - In the **Library** field, specify the library where the data query will be loaded. Be sure to choose a LASR library if you want to use the data query in reports or explorations. For more information about libraries, see [SAS Visual Analytics 7.3: User’s Guide](#).
- 6 Click  to save. The Save As dialog box appears.
- 7 Specify a name and location for the data query, and then click **Save**.
- 8 Click , and select **Run** to run the data query.

Create a LASR Star Schema


To create a LASR star schema:



- 1 On the menu bar, click **LASR** ► **Create a Star Schema**.


- 2 Drag and drop tables from the SAS Folders tree onto the workspace.
Note: You must drag and drop the fact table first and then the dimension tables.
- 3 Make other optional changes. For more information about optional changes, see [“Cheat Sheet for LASR Star Schemas” on page 9](#).
- 4 Click the **Output** tab, and specify the following:
 - In the **Table** field, specify a unique name for the output table or choose an existing output table to overwrite. This is the name of the file that you will use to create reports and explorations.
 - In the **Location** field, specify a folder where you want the LASR star schema to appear. The **My Folder** folder is private. Only you can use the data that you save to **My Folder**. If you want the LASR star schema to be accessible to others, save it to the **Shared Data** folder.
 - In the **Library** field, specify the library where the LASR star schema will be loaded. Be sure to choose a LASR library if you want to use the LASR star schema in reports or explorations. For more information about libraries, see [SAS Visual Analytics 7.3: User’s Guide](#).
- 5 Click  to save. The Save As dialog box appears.
- 6 Specify a name and location for the LASR star schema, and then click **Save**.
- 7 Click  to run the LASR star schema.

Append Data to a Table

To append data to a table:

- 1 On the menu bar, click **LASR** ► **Append Tables**.
- 2 Select the table to which you want to append data from the **Available tables** list, and click  to move it to the **Base table** field.

If the table that you want to use is not listed, click  to view a different source library. If a table is added to the source library *after* you click , then the **Available tables** list does not automatically refresh to show the new table. Select the same source library again to refresh its contents.

3 Select the tables to append to the base table from the **Available tables** list, and click  to move them to the **Source tables** list.

4 (Optional) If you select the **Unload source tables after appending them to the base table** check box, then the source tables are removed from memory after the append completes.

Select this option only if you do not want to access the source tables individually.

5 Click  to save. The Save As dialog box appears.

6 Specify a name and location, and then click **Save**.

7 Click  to run.

3

Specific Tasks





<i>Introduction</i>	7
<i>Cheat Sheet for Data Queries</i>	8
<i>Cheat Sheet for LASR Star Schemas</i>	9



Introduction

This chapter highlights core functionality in the data builder. For most tasks, there are multiple ways to begin. In this book, one way to begin each task is documented. If you ever get stuck or have questions, some of the Help text in the interface might help. For comprehensive information, access the [SAS Visual Analytics: User's Guide](#) from the Help menu or visit the [SAS support site](#).



Complete the following tasks using the data builder. To access the data builder, click **Data Preparation** from SAS Home.


Cheat Sheet for Data Queries

Task	How to Begin in the Data Builder
Add another data source.	Drag and drop the table or data query from the SAS Folders tree onto the workspace.
Add an aggregation to an existing column.	On the Column Editor tab, click to highlight the Aggregations cell for the column to which you want to add an aggregation, and click the ellipsis button.
Create a calculated column.	On the Column Editor tab, click the plus sign (+) to the left of the last entry in the Column Name list. A new line appears. Click to highlight the Expression cell of the new line, and click  .
Filter data.	Click the Where or Having tab, depending on how you want to filter the data.
Change or add a join condition.	On the Joins tab, click the plus sign (+).
Delete a column.	On the Column Editor tab, right-click the name of the column, and select Remove Column .
Pivot data.	On the Column Editor tab, click to highlight the Pivot By cell for the column that you want to pivot values, and then click the ellipsis button.
Sort data.	On the Column Editor tab, click to highlight the Sort cell for the column by which you want to sort the data. Select the sort preference from the drop-down list.
Validate a data query.	On the toolbar, click  .
Preview the results of a data query.	On the toolbar, click  , and select Preview to preview the results.
Save a data query.	On the toolbar, click  .

Task	How to Begin in the Data Builder
Run a data query.	After saving the data query, on the toolbar, click  , and select Run to run the data query.
Schedule the data query to run.	After saving the data query, on the toolbar, click  . In the Schedule dialog box, specify information.
View the results of a data query.	After you preview or run the data query, click the Results tab.

Cheat Sheet for LASR Star Schemas

Task	How to Begin in the Data Builder
Add a data source.	Drag and drop the tables from the SAS Folders tree onto the workspace. You must drag and drop the fact table first and then the dimension tables.
Change the join condition of a dimension column.	On the Joins tab, double-click the Dimension Column cell.
Change the join condition of the fact column.	On the Joins tab, double-click the Fact Column cell.
Specify a different column prefix.	On the Inputs tab on the right pane, choose a different table name from the drop-down menu.
Delete a column from a dimension table.	Right-click on the column in the table in the workspace, and select Remove Column .
Save a LASR star schema.	On the toolbar, click  .
Run a LASR star schema.	After saving the LASR star schema, on the toolbar, click  , and select Run to run the LASR star schema.

Task	How to Begin in the Data Builder
Schedule the LASR star schema to run.	After saving the LASR star schema, on the toolbar, click  . In the Schedule dialog box, specify information.
View the results of a LASR star schema.	After you preview or run the LASR star schema, click the Results tab.

4

Quick Reference

The Data Builder 11

Where to Find Additional Resources 12

The Data Builder

The screenshot displays the SAS Data Builder interface. At the top, the title bar shows 'Prepare Data' and the SAS logo. Below it is a menu bar with 'File', 'LASR', 'Tools', and 'Help'. A progress indicator shows '16%' completion. The main workspace is divided into several sections:

- File Explorer (Left):** Shows a tree view of folders including 'My Folder', 'Products', 'SAS Projects', 'Shared Data', 'System', and 'User Folders'.
- Design Canvas (Center):** Displays three data tables: 'ORGANIZATION_DIM', 'ORDER_FACT', and 'PRODUCT_DIM'. Lines indicate relationships between columns in these tables.
- Column Editor (Bottom):** A table showing the selected columns and their expressions.

	Column Name	Expression	Type	Form
1	Order_ID	ORDER_FACT.Order_ID	NUMERIC	12.
2	Product_Name	PRODUCT_DIMProduct_Name	CHARACTER(45)	
3	Product_Group	PRODUCT_DIMProduct_Group	CHARACTER(25)	
4	Product_Line	PRODUCT_DIMProduct_Line	CHARACTER(20)	
- Properties Panel (Right):** Shows the 'Inputs' tab for the selected 'PRODUCT_DIM' table. It lists properties such as 'Alias: PRODUCT_DIM', 'Name: PRODUCT_DIM', 'Location: /Shared Data', and 'Library: /Shared Data/SASApp - SASDATA'. It also includes 'Default Aggregations' with 'Auto-aggregate' set to 'Disable'.

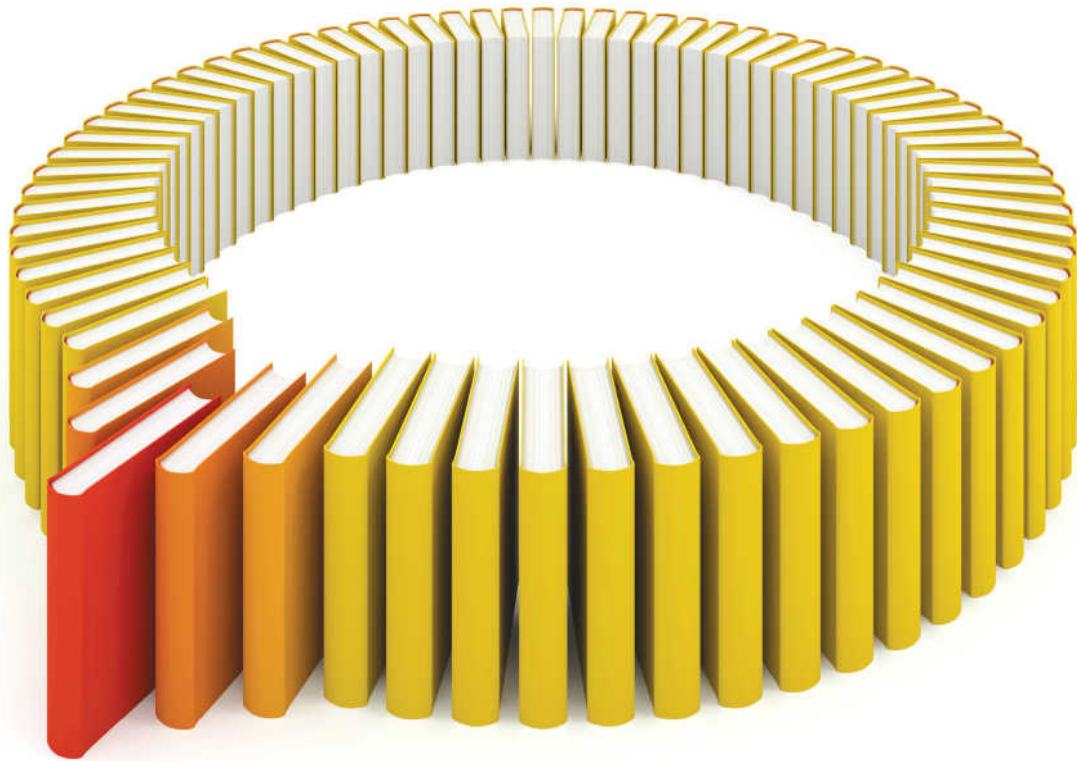
- 1 The application bar enables you to return to the home page and to access other parts of SAS Visual Analytics and other SAS applications that integrate with the home page.
- 2 The menu bar contains menus that enable you to perform tasks such as creating new data queries and LASR star schemas. The right side of the menu bar has a memory gauge that displays the memory utilization for a distributed SAS LASR Analytic Server. You can also sign out of SAS Visual Analytics.
- 3 The navigation pane displays a tree (the SAS Folders tree) of tables and data queries.
- 4 The center of the screen contains the workspace. When you create a new object, such as a data query, it is represented as a tab on the top of the workspace.

The bottom of the workspace contains a series of tabs that enable you to create column expressions, joins, and filter data.
- 5 The right pane enables you to manage the properties of the item that is selected in the workspace.
- 6 The toolbar contains icons that enable you to manage, run, and schedule data queries.

Where to Find Additional Resources

The most current technical resources for SAS Visual Analytics are available on the [SAS Visual Analytics](#) page on the SAS support site. You can view SAS Tutorials from the [SAS Visual Analytics: Video Library](#)

Your experience with SAS software should be as smooth as possible. Please submit your [feedback](#).



Gain Greater Insight into Your SAS® Software with SAS Books.

Discover all that you need on your journey to knowledge and empowerment.

