THE POWER TO KNOW.

SAS® Contextual Analysis 14.2: Administrator's Guide

SAS® Documentation

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

SAS® Contextual Analysis 14.2: Administrator'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.

14.2-P1:utaag

Contents

About This Book . . . . . *v*
Accessibility . . . . . *vii*

Chapter 1 • SAS Contextual Analysis Administrator's Guide Overview **1**
Overview of SAS Contextual Analysis 1
SAS Information Online 1

Chapter 2 • Quick Start for Running SAS Contextual Analysis **3**
Quick Start Overview 3
Quick Start Steps 3

Chapter 3 • SAS Contextual Analysis Architecture **13**
Overview of the SAS Contextual Analysis Architecture 13
SAS Contextual Analysis Components 14

Chapter 4 • Updating to SAS Contextual Analysis 14.2 . . . . . **15**
Update Overview 15
Updating an Existing Deployment 15
Updating on a Target Machine and Migrating Project Data 15

Chapter 5 • Preparing Your Data for Use with SAS Contextual Analysis **17**
Data Access Requirements 17
Text and Category Variable Names 17
Registering Data (Including Start, Stop, and Synonym Lists) 18

Chapter 6 • SAS Contextual Analysis Installation and Configuration **19**
Installation Overview 19
Pre-Installation Tasks 19
Install and Configure SAS Contextual Analysis 20

Chapter 7 • Post-Installation Information **25**
Configure SAS Contextual Analysis in SAS Management Console 25
Define Additional SAS Contextual Analysis Users 25
Establish a Lock-Down Path List 25
Set Up Integrated Windows Authentication 25
Naming Project Files in SAS Contextual Analysis 26
Configure Secure Sockets Layer (SSL) 26

Chapter 8 • Administrator Tasks and Troubleshooting **27**
Troubleshooting 27

Recommended Reading **29**

About This Book

Audience

This book is designed for administrators of the SAS Contextual Analysis application. It guides administrators through the process of configuring the software for use at your site. It is assumed that you have some familiarity with SAS programs and installation processes.

Requirements

For a full list of software requirements for SAS Contextual Analysis, go to System Requirements.

Accessibility

For information about the accessibility of this product, see Accessibility Features of SAS Contextual Analysis 14.2 at support.sas.com.

Chapter 1

SAS Contextual Analysis Administrator's Guide Overview

Overview of SAS Contextual Analysis 1

SAS Information Online 1

Overview of SAS Contextual Analysis

SAS Contextual Analysis is a web-based text analytics application that uses contextual analysis to provide a comprehensive solution to the challenge of identifying and categorizing key textual data. Using this application, you can build models (based on training documents) that automatically analyze and categorize a set of documents. You can then customize your models in order to realize the value of your text-based data. In addition, you can create rules to extract concepts, which are properties such as a book title, last name, city, gender, and so on.

The SAS Contextual Analysis architecture is designed for single-tier or three-tier deployment.

SAS Information Online

The following resources are available to you online:

- The product page for SAS documentation for SAS Contextual Analysis, available at http://support.sas.com/software/products/ca/index.html
- SAS Technical Support, available at support.sas.com
- Help within the SAS Contextual Analysis application

Chapter 2

Quick Start for Running SAS Contextual Analysis

Quick Start Overview 3
Quick Start Steps 3

Quick Start Overview

This chapter is designed to help you start your SAS Contextual Analysis installation quickly. It provides basic steps for completing these tasks:

- signing in to SAS Management Console and creating a SAS library
- registering a data set for use in SAS Contextual Analysis
- adding a user profile in SAS Management Console
- signing in to the SAS Contextual Analysis application

Note: Before you perform the quick start steps, complete the installation from the SAS Install Center (available at support.sas.com) and complete the configuration steps. For information about installation and configuration, see "Install and Configure SAS Contextual Analysis" on page 20.

Quick Start Steps

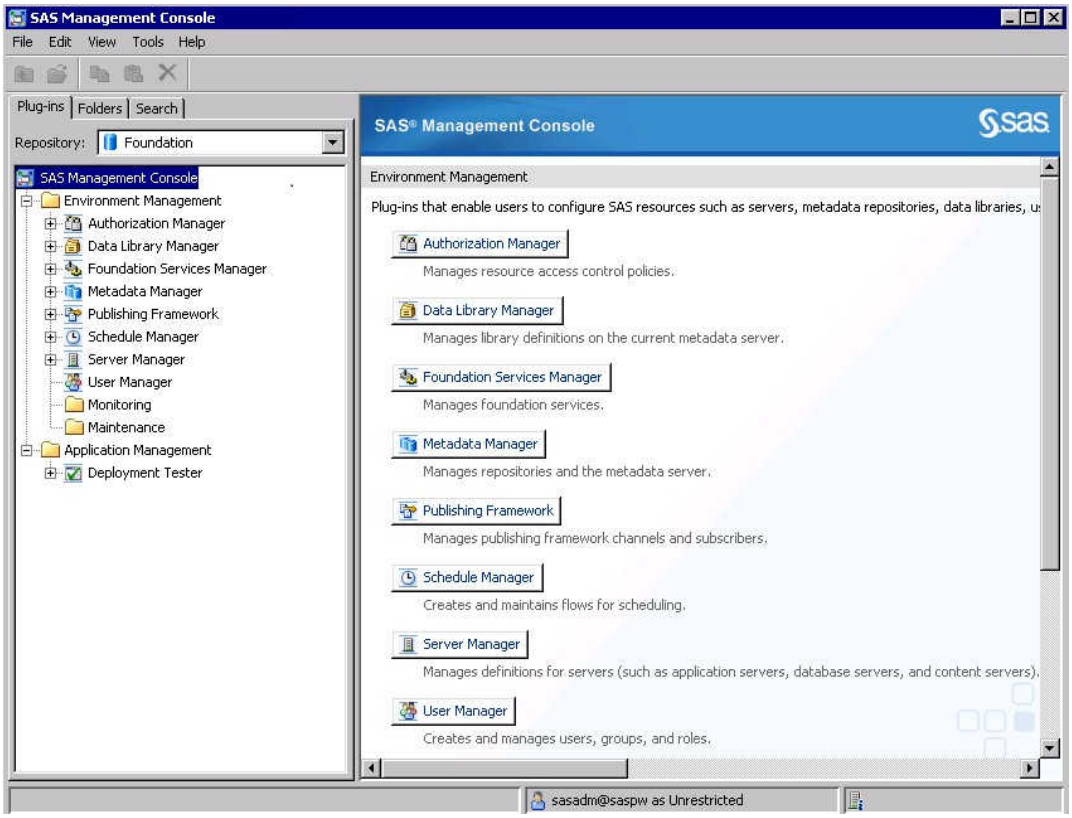
After you install and configure SAS Contextual Analysis, complete the following steps:

1. Open SAS Management Console. Connect to the metadata repository using these credentials:

   Metadata server host name: the host name that you entered during installation and configuration

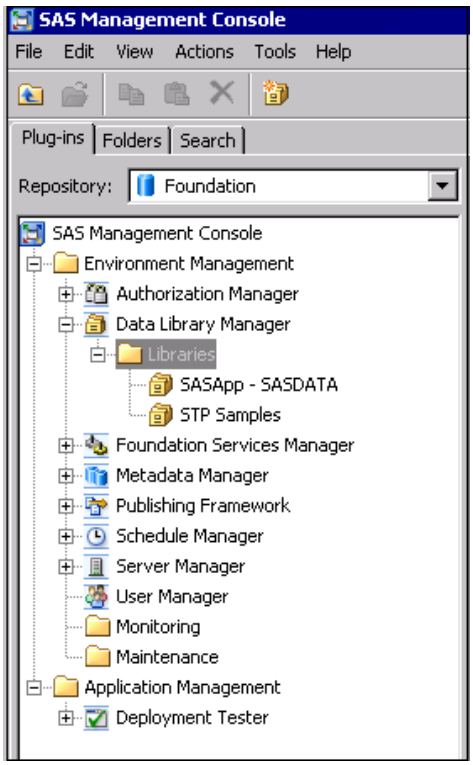
   User name: the default user name (**`sasadm@saspw`**) or other established user name

   Password: the password that you specified during installation and configuration

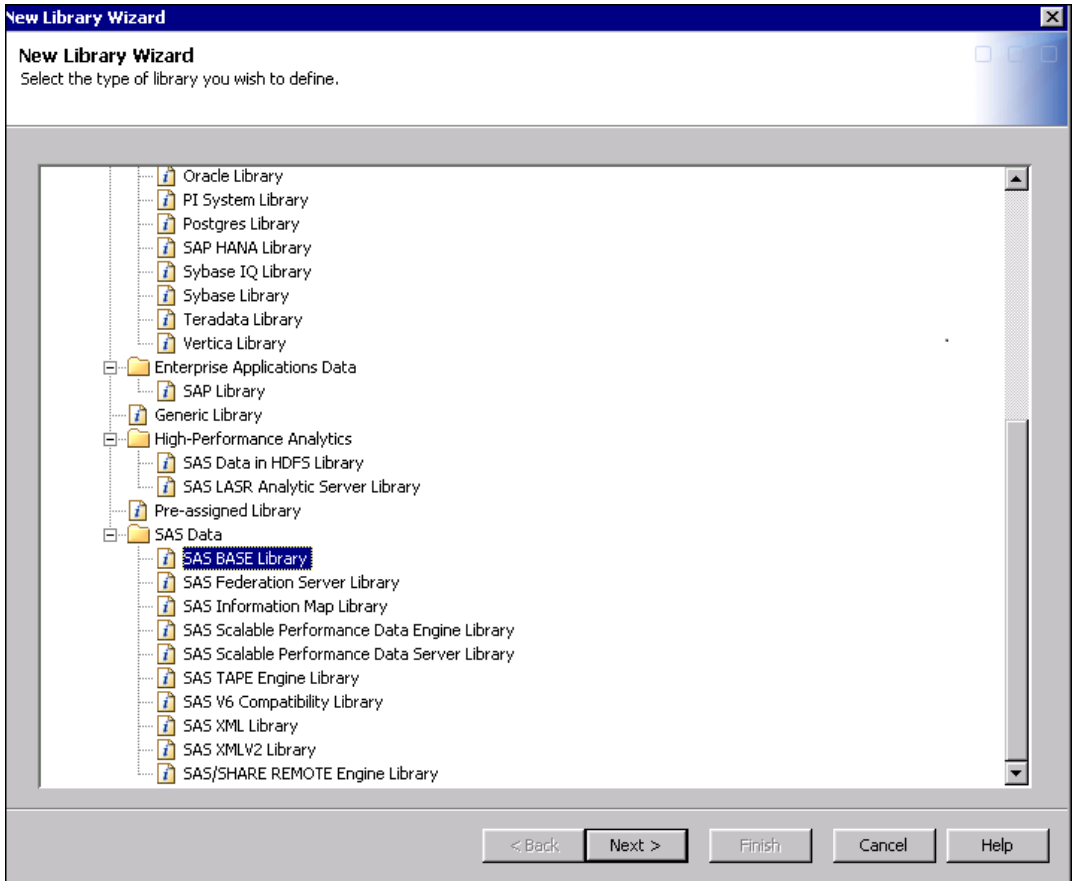


Note: If you are installing SAS Contextual Analysis on a UNIX host, you must use SAS Management Console on a Windows computer (that has SAS Management Console installed on it).

2. In the left pane under **Environment Manager**, navigate to the **Data Library Manager** plug-in and right-click **Libraries**. The New Library wizard appears.

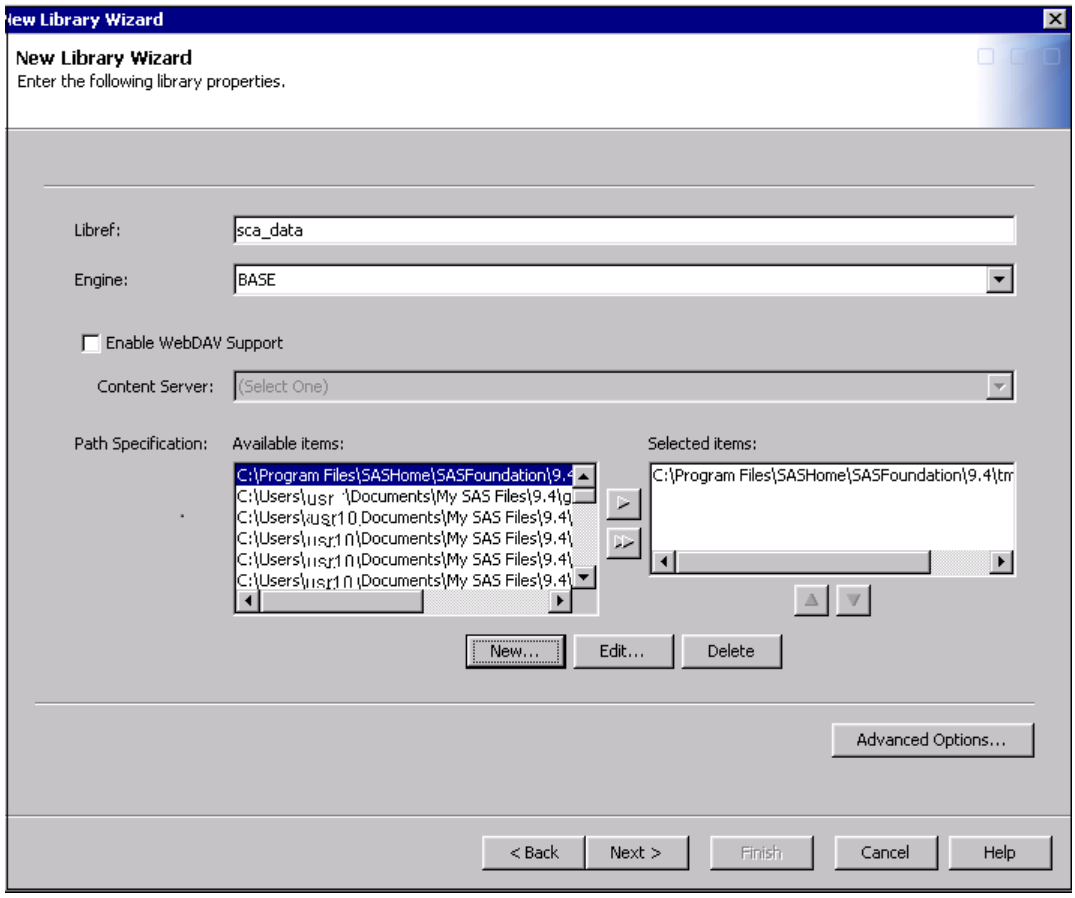


3. Select **SAS BASE Library**, and click **Next**.



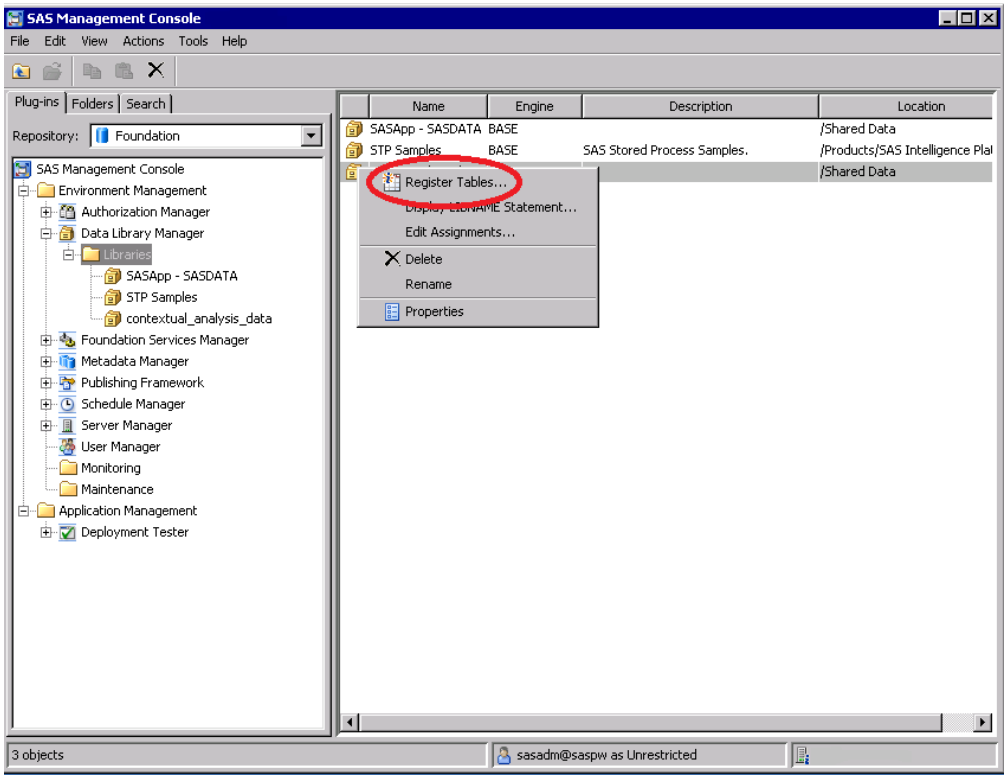
4. Enter a name for the library in the **Name** field (for example, **`sca_data`**), and click **Next**.
5. Select **SASApp** as the application server, and click **Next**.
6. In the New Library Wizard, enter a libref name in the **Libref** field.

   Click **New** to specify the path to the sample data sets. Sample data sets are provided in the **`[SAS_HOME]\SASFoundation\9.4\tmine\sample`** directory. In a Windows environment, the default location is **`C:\Program Files\SASHome\SASFoundation\9.4\tmine\sample`**. Select the directory from the dialog box.

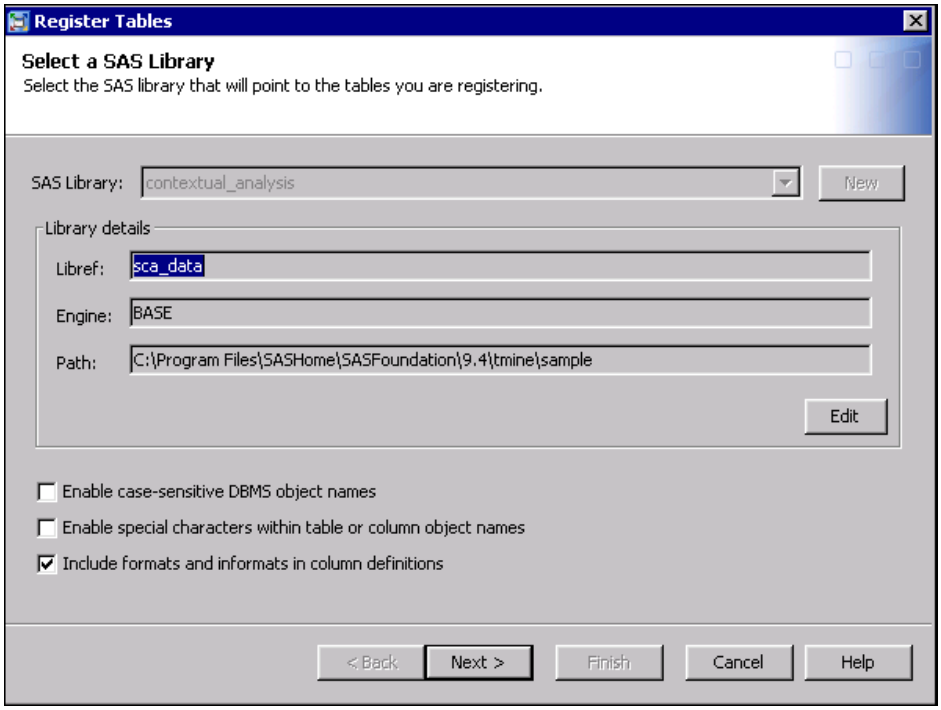


Click **Next**, and then click **Finish** to create the data library.

7. Register the tables from the data library into metadata. With the library selected, right-click to access the menu and then select **Register Tables**.



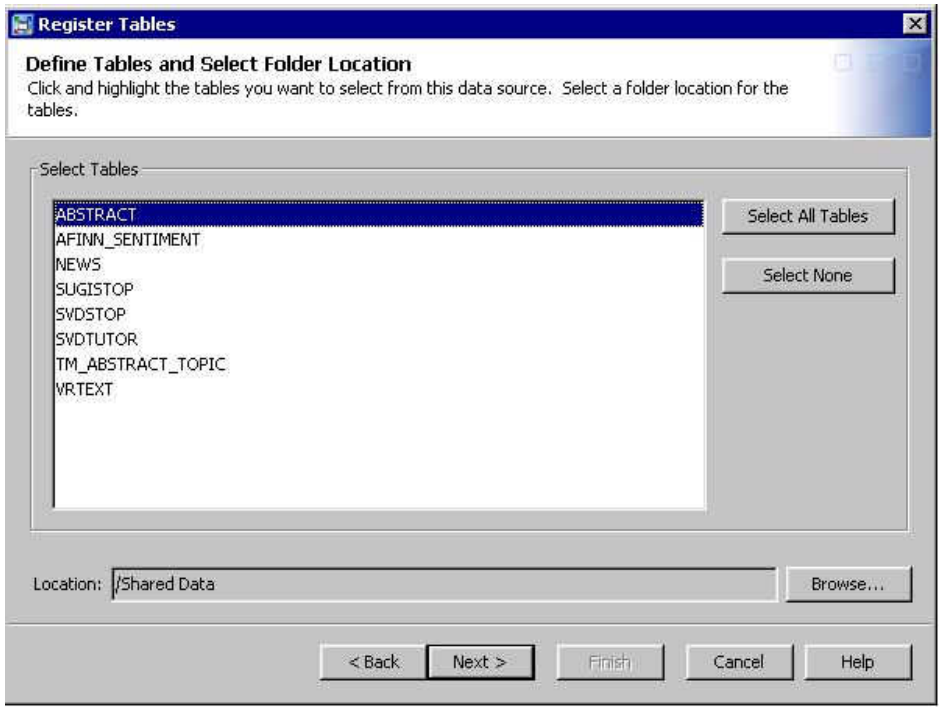
8. Enter the **Libref** name, and then click **Next**.



The Log On window for the application server appears.

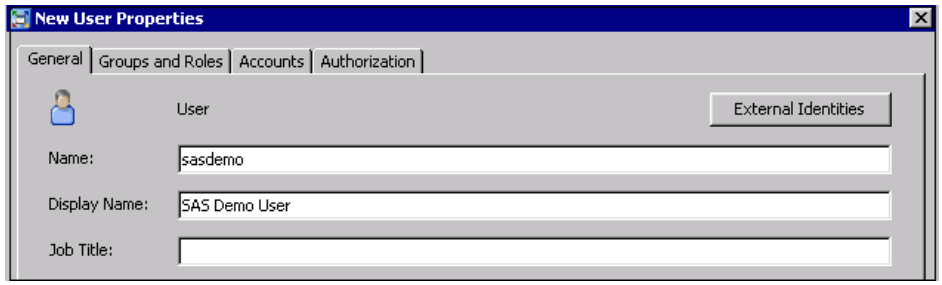
9. In the Log On window, enter credentials for an administrator account on the application server.

10. Select the tables that you want to register. Click **Next**, and then click **Finish** to register the tables in metadata.

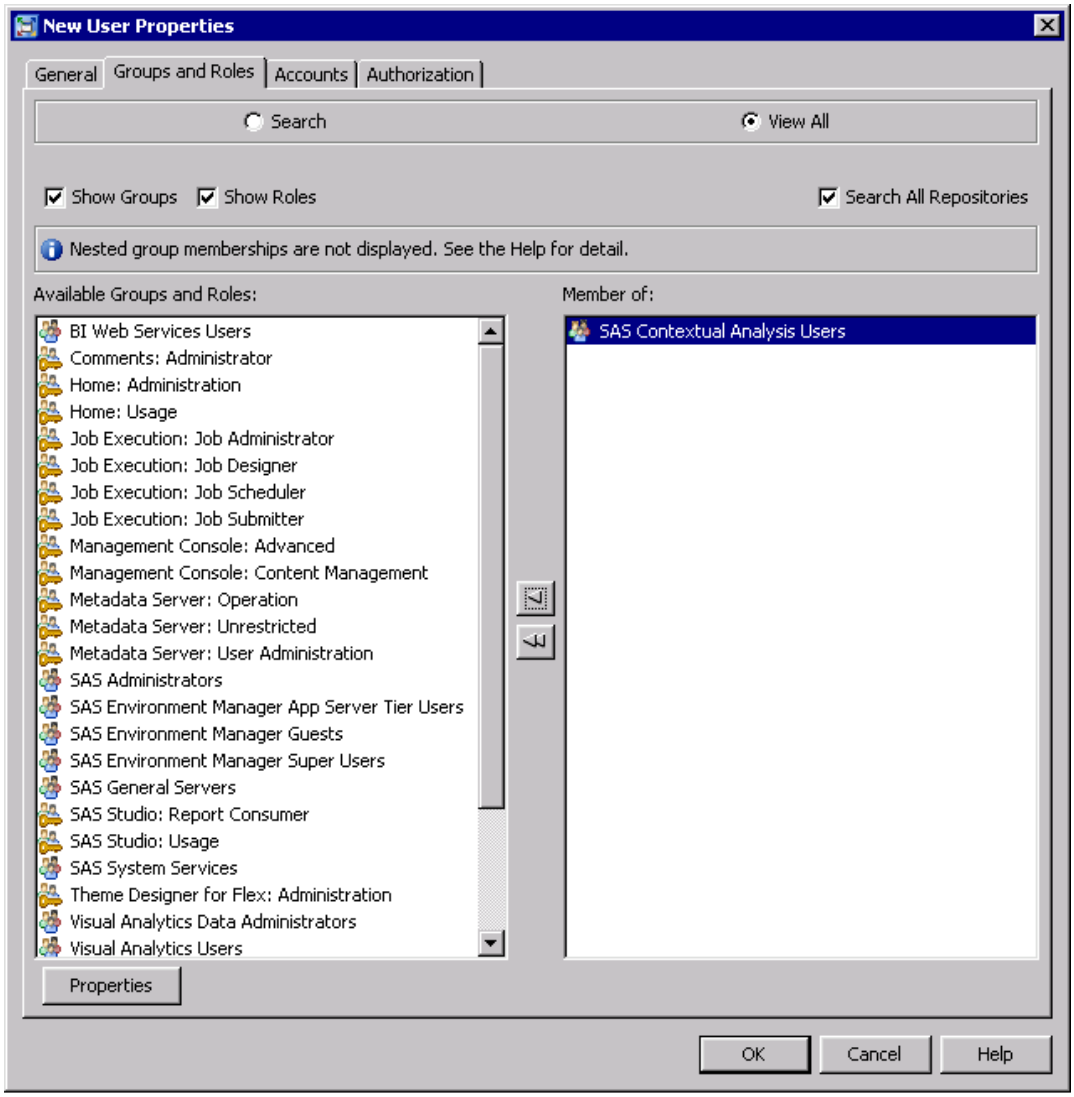


The tables are now accessible from within SAS Contextual Analysis.

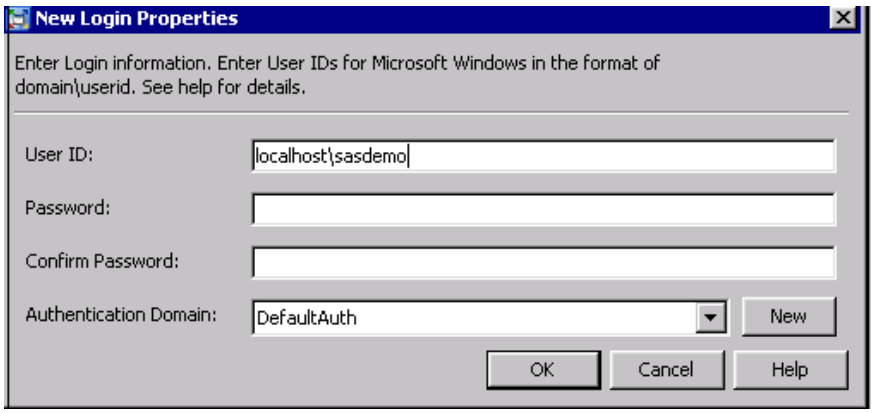
11. In SAS Management Console, use the User Manager to create a user profile.



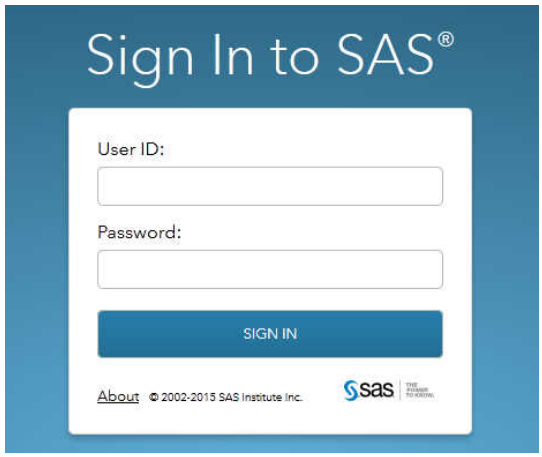
12. Click the **Groups and Roles** tab. Grant the user membership to the following group: **SAS Contextual Analysis Users**.



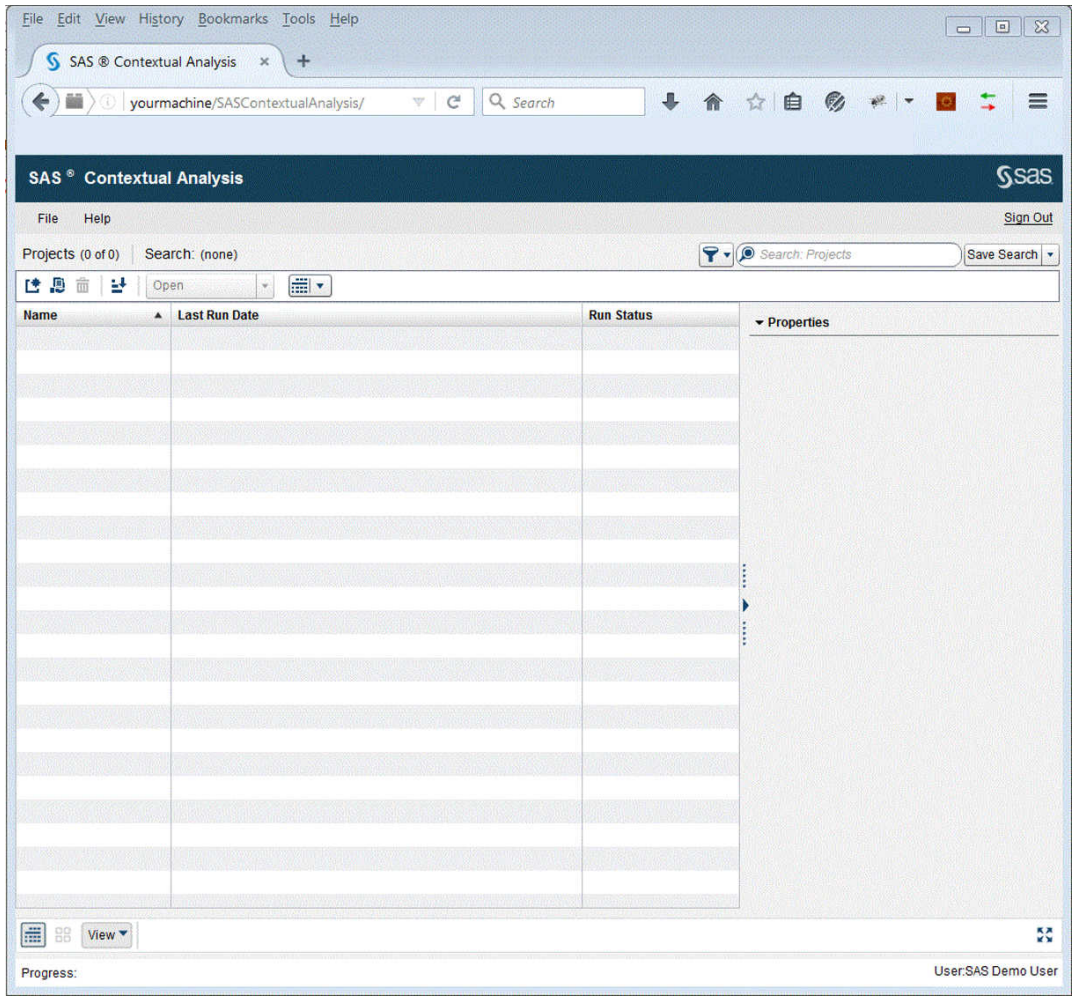
13. On the **Accounts** tab, click **New**. Enter the login for the user. If you leave the **Password** field blank, SAS uses the system's password (the domain password) for the credential.



14. Start SAS Contextual Analysis by using the following URL, where *localhost* is the name of the machine where the application is installed: http://localhost/SASContextualAnalysis. The Sign In window appears.



15. Sign in using your user credentials. The SAS Contextual Analysis application appears.



See Also

- For more information about SAS Management Console connection profiles, see "Change, Create, Delete, or Edit a Connection Profile" in *SAS Intelligence Platform: Desktop Application Administration Guide*

- For more information about creating a data library, see “Establishing Connectivity to a Library of SAS Data Sets” in *SAS Intelligence Platform: Data Administration Guide*
- For more information about registering tables in metadata, see “Registering and Verifying Tables” in *SAS Intelligence Platform: Data Administration Guide*
- For information about creating a project in SAS Contextual Analysis, see *SAS Contextual Analysis: User's Guide*

Chapter 3

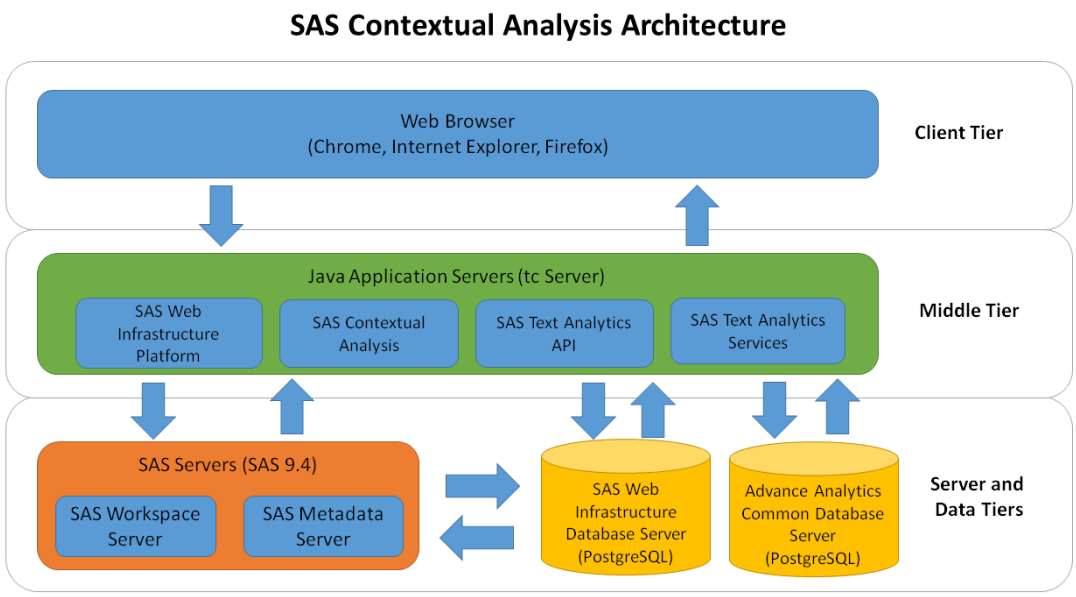
SAS Contextual Analysis Architecture

Overview of the SAS Contextual Analysis Architecture 13
SAS Contextual Analysis Components 14

Overview of the SAS Contextual Analysis Architecture

The SAS Contextual Analysis architecture is designed for single-tier or three-tier deployment. Figure 3.1 depicts a three-tier deployment model, which includes a client tier, a middle tier, and a server and data tier.

Figure 3.1 *SAS Contextual Analysis 14.2 Architecture*



SAS Contextual Analysis Components

The SAS Web Infrastructure Platform (WIP)
: provides the primary middle-tier components for the SAS platform. It supports *n*-tier architectures that are built around HTTP-based technologies. The middle-tier components in the SAS WIP not only provide for common infrastructure but also support distributed deployments and configurations that support enterprise requirements.

SAS Contextual Analysis middle tier
: is the web application that is responsible for serving the user interface to a client's web browser. After a client's browser has loaded the Adobe Flex user interface for SAS Contextual Analysis, Flex uses a REST API to communicate back to the SAS middle tier. The SAS Contextual Analysis middle tier also serves as the conduit through which a user's requests are routed to the SAS Text Analytics API middle tier.

SAS Text Analytics API middle tier
: is the web application that hosts the REST API that is used for communication between the Flex client and the SAS Text Analytics Service middle tier.

SAS Text Analytics Services middle-tier
: is the home of the majority of the "brains" for SAS Contextual Analysis. The SAS Text Analytics Services middle tier often communicates with the Advanced Analytics Common Database Server to store objects that are related to a SAS Contextual Analysis project (concepts, categories, and so on). The SAS Text Analytics Services middle tier also communicates through the SAS Web Infrastructure Platform to SAS Workspace Servers to process user requests that require a SAS session.

SAS Workspace Server
: enables client applications to submit SAS code to a SAS session by using an application programming interface (API). SAS Contextual Analysis generates the SAS code necessary to perform a project's analytical tasks and submits it to a SAS Workspace Server. You can run as many instances of SAS Workspace Servers as are needed to support your workload.

SAS Metadata Server
: controls access to a central repository of metadata that is shared by all of the SAS applications in the deployment. The SAS Metadata Server enables centralized control so that all users access consistent and accurate data. Administrators use SAS Management Console to administer the SAS Metadata Server, including SAS server configurations. SAS Contextual Analysis also uses the SAS Metadata Server to obtain metadata about SAS libraries, tables, and SAS Workspace Servers, and to authenticate users.

SAS Web Infrastructure Database Server
: is a database server that is used by the SAS WIP as the default location for middle tier data. These data include alerts, comments, and workflows.

Advanced Analytics Common Database Server
: is a PostgreSQL database that is used to store SAS Contextual Analysis project-related data such as projects, concepts, categories, and so on.

Chapter 4

Updating to SAS Contextual Analysis 14.2

Update Overview 15
Updating an Existing Deployment 15
Updating on a Target Machine and Migrating Project Data 15

Update Overview

There are two ways to update to SAS Contextual Analysis 14.2:

- Update SAS Contextual Analysis software on an existing deployment.
- Install SAS Contextual Analysis software on a target machine and migrate your settings and project data from an existing deployment.

Updating an Existing Deployment

To update SAS Contextual Analysis software on an existing deployment, you must use the SAS Deployment Wizard. There you can choose to upgrade your current SASHOME directory. For more information, see *SAS Guide to Software Updates*.

Updating on a Target Machine and Migrating Project Data

To install SAS Contextual Analysis software on a target machine, you must use the SAS Deployment Wizard. For more information, see *SAS Guide to Software Updates*.

To upgrade SAS Contextual Analysis on a target machine that is not the existing deployment, use the SAS Migration Utility. For more information, see *SAS Intelligence Platform: Migration Guide*. The section “Product-Specific SAS Migration Utility Properties” includes specific steps for SAS Contextual Analysis.

After the migration of the software is complete, you must copy the data that you used in your projects (in the folder that you registered to the data library during post-installation) from the source machine to the target machine.

You must copy all the project folders from the source machine to the same location on the target machine. For example, if you have a project folder in **C:\Users\myuserid\Documents\My SAS Files\9.4** on the source machine, then you must store the project folder in the same location on the target machine.

Note: During project migration, category names might be appended with an underscore and a number (**_n**). SAS Contextual Analysis appends category names only when name duplications occur.

Chapter 5

Preparing Your Data for Use with SAS Contextual Analysis

Data Access Requirements 17
Text and Category Variable Names 17
Registering Data (Including Start, Stop, and Synonym Lists) 18

Data Access Requirements

When you are accessing data in SAS Contextual Analysis, keep in mind that SAS Contextual Analysis must follow the rules of the user accounts on the SAS Workspace Server machine. For example, if you are a standard user, then you cannot access data that are stored in another user's personal folder. You also cannot access data that are stored in commonly restricted folders such as `C:\Windows\system32`.

Text and Category Variable Names

The names of text variables (columns that you are analyzing) and category variables must follow these standard SAS variable naming conventions:

- Names can be up to 32 characters in length.
- The first letter must begin with an English letter or an underscore. Subsequent characters can be English letters, numeric digits, or underscores.
- The name cannot contain blanks or any special characters (other than an underscore).

Note: Variables that do not follow these naming conventions could cause errors during processing.

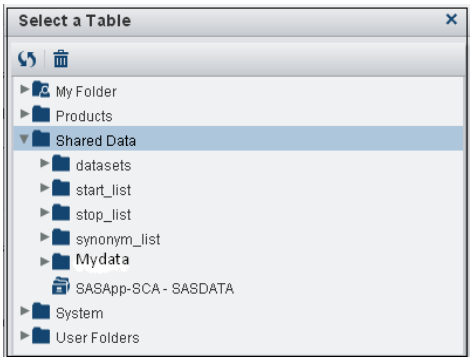
For more information about naming conventions for SAS variables, see *SAS Language Reference: Concepts*.

Registering Data (Including Start, Stop, and Synonym Lists)

You must register any data that you are using for text analysis, including synonym lists, stop lists, and start lists. Follow the steps for registering data in the section "Quick Start Steps" on page 3.

- For more information about creating a data library, see the section "Establishing Connectivity to a Library of SAS Data Sets" in *SAS Intelligence Platform: Data Administration Guide*.
- For more information about registering tables in metadata, see the section "Registering and Verifying Tables" in *SAS Intelligence Platform: Data Administration Guide*.

TIP You can organize your metadata so that you can easily find data. For example, you can create metadata folders within the Shared Data folder for data sets, synonym lists, start lists, and stop lists. Register the respective data sets to each library so that you can see the structure within the browse window. An example is shown in the following graphic.



Chapter 6

SAS Contextual Analysis Installation and Configuration

Installation Overview **19**

Pre-Installation Tasks **19**
Overview 19
Shared Memory Requirement for UNIX Systems 19
Create a SAS Software Depot 20

Install and Configure SAS Contextual Analysis **20**
Installation and Configuration Process 20
SAS Contextual Analysis Server Standard Deployment Plans 20

Installation Overview

Here are the major steps for installing SAS Contextual Analysis:

1. Make any system preparations that are necessary.
2. Create a SAS Software Depot.
3. Run the SAS Deployment Wizard.
4. Select from the standard deployment plans to install and configure SAS Contextual Analysis.

Pre-Installation Tasks

Overview

Before you install SAS Contextual Analysis, make any preparations to your system that you deem necessary. You might need to increase the shared memory of your machine, and you must create a SAS Software Depot on your machine.

Shared Memory Requirement for UNIX Systems

In order to ensure enough shared memory for operating the application's database on the SAS Advanced Analytics Common Data Server, it is recommended that you extend the size of your shared memory to 6 GB.

Refer to the documentation for your operating system and machine to determine the current shared memory settings. For example, on a 64-bit Linux machine, you can run the command `ipcs -l` on the command line and note the value in the results for **max total shared memory**.

To extend the shared memory to 6 GB, refer to the documentation for your operating system.

Create a SAS Software Depot

A SAS Software Depot is a file system that is used by the SAS Deployment Wizard, which is the tool used to install and initially configure SAS. The depot contains the SAS Deployment Wizard executable, one or more deployment plans, a SAS installation data file for each order, order data, and product data.

To create a SAS Software Depot, follow the instructions in "Creating a Software Depot" in *SAS Intelligence Platform: Installation and Configuration Guide*. During the creation of the depot, you identify on the target machine where you want to download the depot.

Note: You need to have access to your software order details in order for SAS Contextual Analysis to create a depot.

Install and Configure SAS Contextual Analysis

Installation and Configuration Process

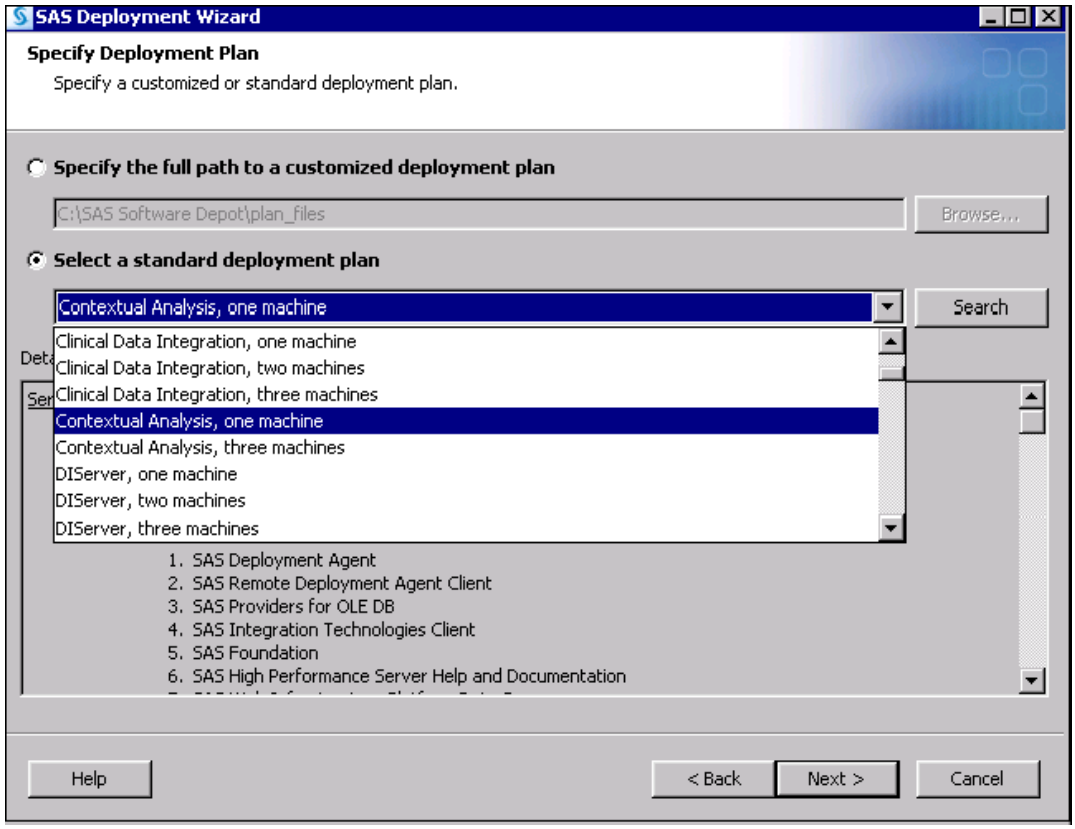
To complete the installation and configuration process, you must have completed the steps to download the SAS Software Depot. For more information, see "Create a SAS Software Depot" on page 20.

SAS Contextual Analysis Server Standard Deployment Plans

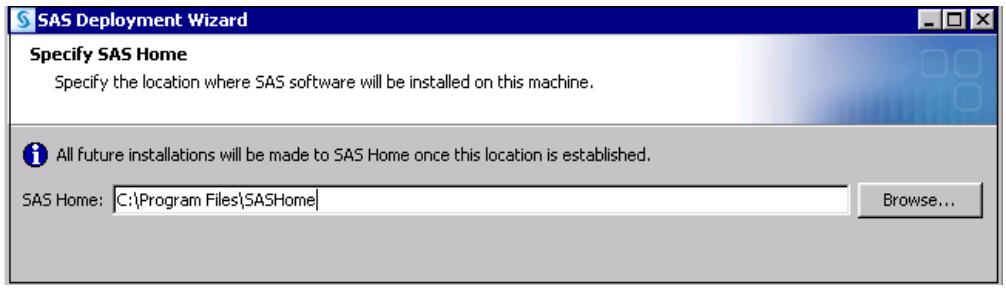
Deployment plans are blueprints for installing and configuring your software. SAS Contextual Analysis provides you with two standard plans in the SAS Software Depot: a one-machine deployment and a three-machine deployment.

To access the standard plan that is delivered with your SAS Software Depot, open the plan file. Select **Select a standard deployment plan**. Select **Contextual Analysis, one machine** or **Contextual Analysis, three machine**, depending on your desired configuration. For more information about the SAS Deployment Wizard, see *SAS Deployment Wizard and SAS Deployment Manager: User's Guide*.

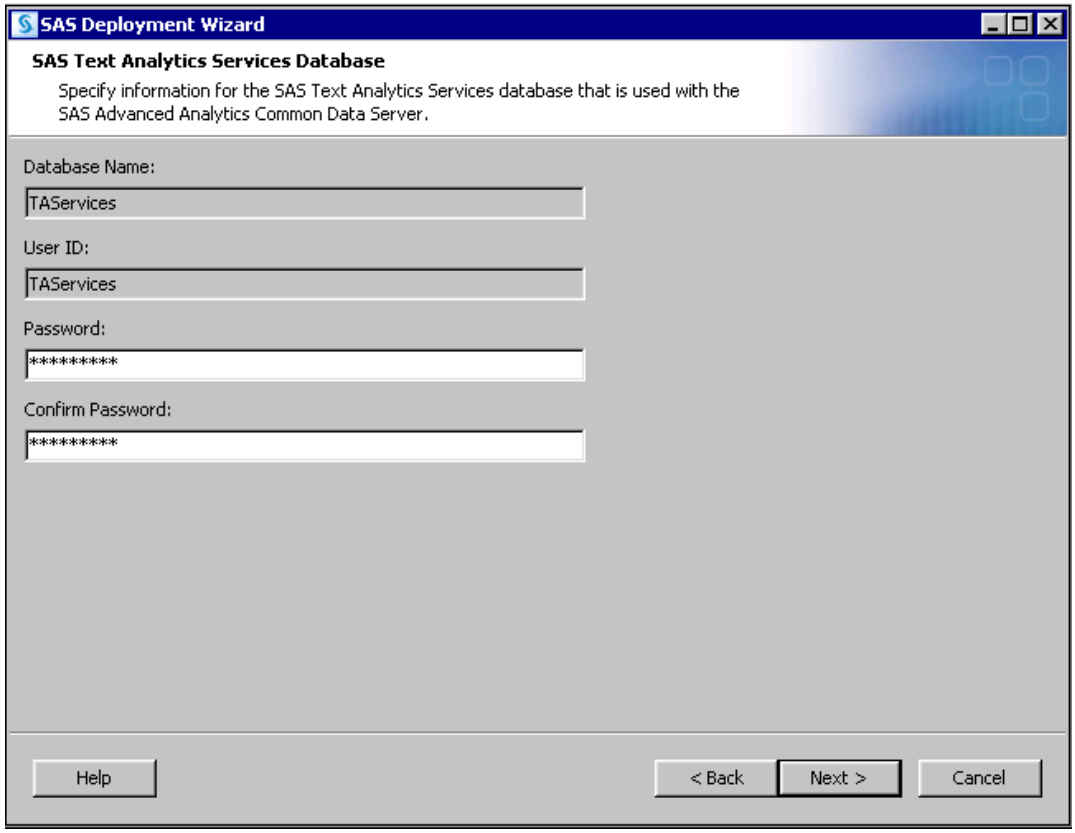
Note: If you cannot access your SAS Software Depot for some reason, you can download deployment plans from the SAS Install Center, available at http://support.sas.com/documentation/installcenter/index.html. In the navigation pane, select **Standard Plans**. Under **SAS 9.4**, select **9.4 Standard Deployment Plans**, and then select **SAS Contextual Analysis**. Download the appropriate plan.



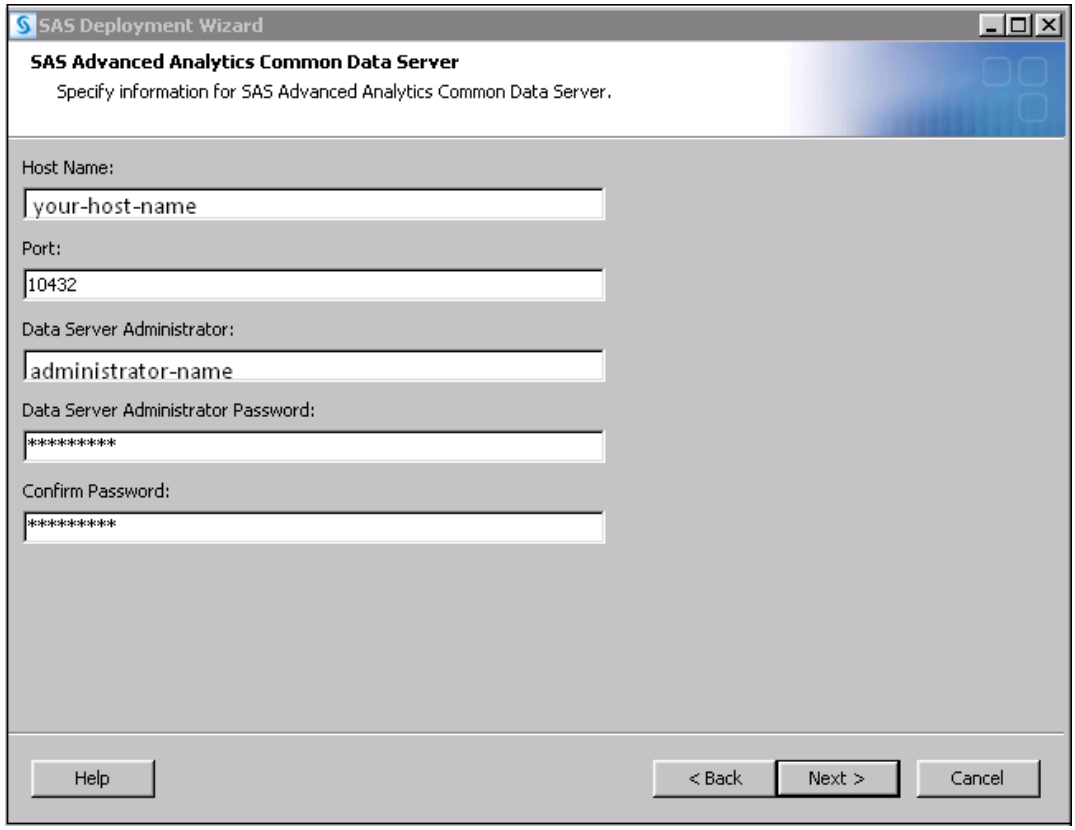
In the SAS Deployment Wizard, a prompt for the location of **SAS Home** appears. A default value is provided, based on the operating system on which are deploying. For example, on a 64-bit Windows machine, the default location is **`C:\Program Files\SASHome`**. If you choose to enter different location, be sure to note the location for later use.



Enter a password for the database.



You must also supply your host name, port, and administrator name (and password) for the SAS Advanced Analytics Common Data Server.



Continue through the wizard as described in *SAS Deployment Wizard and SAS Deployment Manager: User's Guide*.

Note: If your data are encoded in UTF-8, you must select the box that sets up the SAS server as a Unicode server. Setting up the server as a Unicode server ensures that your UTF-8 is processed properly. For more information about setting up a Unicode server, see "Locale and Encoding Considerations" in *SAS Intelligence Platform: Installation and Configuration Guide*.

Chapter 7

Post-Installation Information

Configure SAS Contextual Analysis in SAS Management Console 25
Define Additional SAS Contextual Analysis Users 25
Establish a Lock-Down Path List 25
Set Up Integrated Windows Authentication 25
Naming Project Files in SAS Contextual Analysis 26
Configure Secure Sockets Layer (SSL) 26

Configure SAS Contextual Analysis in SAS Management Console

To start SAS Management Console, you must log on to the metadata repository using a specific set of credentials. For more information, see Step 1 on page 3.

Define Additional SAS Contextual Analysis Users

You can define additional SAS Contextual Analysis users by creating user profiles in SAS Management Console. For more information, see Step 11 on page 8.

Establish a Lock-Down Path List

If you are using SAS on a server that is in a locked-down state, you must establish a path list that can be accessed. For more information about servers in a locked-down state, see "Locked-Down Servers" in *SAS Intelligence Platform: Security Administration Guide*.

Set Up Integrated Windows Authentication

SAS Contextual Analysis supports Integrated Windows authentication (IWA). IWA enables the system to verify credentials through the logged-in Windows accounts,

thereby eliminating the need for users to separately sign in to SAS Contextual Analysis. For information about setting up IWA, see "Support for Integrated Windows Authentication" in *SAS Intelligence Platform: Middle-Tier Administration Guide*.

Naming Project Files in SAS Contextual Analysis

SAS Contextual Analysis projects can be updated by different people, but projects are not meant to be shared concurrently. It is important that you understand the file system so that you can guide users appropriately.

When SAS Contextual Analysis projects are created, libraries are created in the Library Manager of SAS Management Console and are assigned the project name. Each project name is associated with the user who created it.

Note: Users are prevented from using a project name more than once. However, multiple users can create projects that have the same name. You must sign in as an unrestricted user to see duplicate project names in the Library Manager of SAS Management Console.

Configure Secure Sockets Layer (SSL)

Secure Sockets Layer (SSL) configuration enables clients to authenticate with the SAS middle tier by using a client certificate that is installed in their web browser. When this type of client certificate is used for authentication, you are not required to provide a user name and password to log on. For more information, see "Support for SSL with Client Certificate Authentication" in *SAS Intelligence Platform: Middle-Tier Administration Guide*.

Chapter 8

Administrator Tasks and Troubleshooting

Troubleshooting 27
Log Files 27
Log Information 27
Locating the Rules Files (LI and MCO) 28
Garbage Characters When Testing Uploaded Files 28

Troubleshooting

Log Files

The location of the SAS log files for the SAS Contextual Analysis middle tier varies from system to system, depending on the system configuration. In general, SAS Contextual Analysis middle tier log files are located here: **[*configuration-directory*]\Config\Lev1\Web\Logs\SASServer1_1**

Log Information

You can adjust the amount of information that is output to the SAS Contextual Analysis middle-tier log. To increase the information that is output to the log, do the following:

1. In the file system, navigate to **[*configuration-directory*]\Config\Lev1\Web\Common\LogConfig**.
2. Open the TextAnalyticsServices-log4j.xml file in a text editor.
3. Locate this section of code:

```
<category
 additivity="false"
    name="com.sas.ta">
<priority  value="WARN"/>
<appender-ref ref="SAS_CONSOLE"/>
<appender-ref ref="SAS_FILE"/>
</category>

<category
 additivity="false"
    name="com.sas.ta.service">
<priority  value="WARN"/>
```

```
<appender-ref ref="SAS_CONSOLE"/>
<appender-ref ref="SAS_FILE"/>
</category>
```

Change the value

```
WARN
```

to

```
INFO
```

4. Restart the web application server.

TIP When enhanced logging is used, the SAS code that is output can be copied and pasted into a SAS session on the server machine.

Locating the Rules Files (LI and MCO)

The LI files (used for concept rules) and MCO files (used for category rules) are located within the configuration folder of the project directory. For example, the location of the folder on a Windows machine is **C:\Users\[*user ID*]\Documents\My SAS Files\9.4\sca_project_1\config**. The folder contains the concepts.li file and the categories.mco file for the sca_project_1 project. These files can be used with the SAS Enterprise Content Categorization Server for scoring other documents. For more information, see *SAS Enterprise Content Categorization Studio: User's Guide*.

Garbage Characters When Testing Uploaded Files

If you use an uploaded text file to test concepts or categories, you might see garbage characters displayed in SAS Contextual Analysis. Garbage characters can occur when the encoding of the test file does not match the encoding on your workspace server.

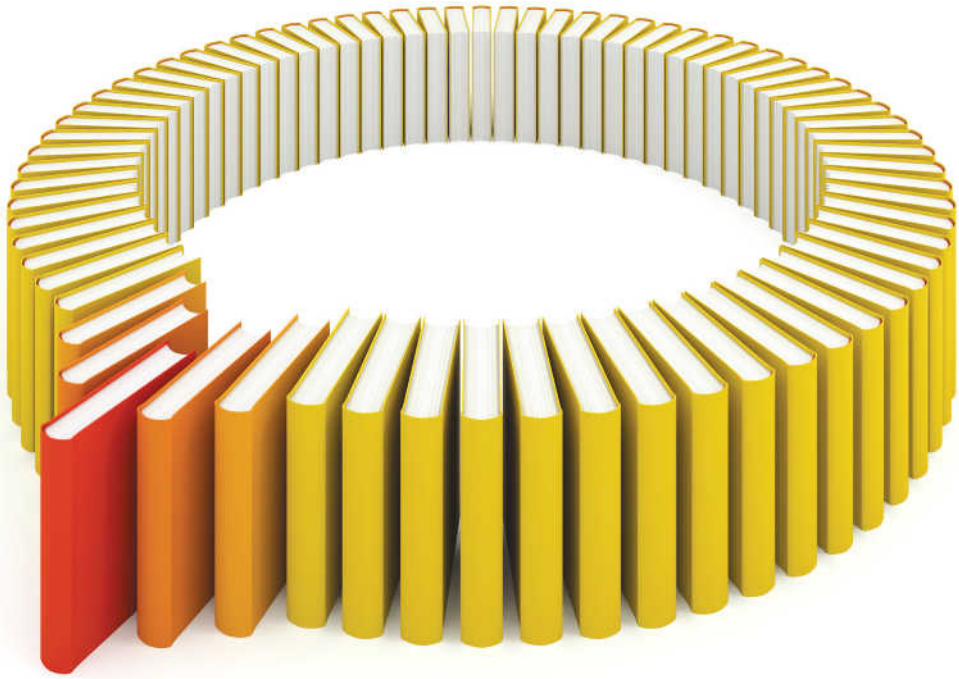
Recommended Reading

Here is the recommended reading list for this title:

- *SAS Content Categorization Single User Servers: Administrator's Guide*
- *SAS Deployment Wizard and SAS Deployment Manager 9.4: User's Guide*
- SAS Encoding: Understanding the Details
- *SAS Enterprise Content Categorization Studio: User's Guide*
- *SAS Intelligence Platform: Data Administration Guide*
- *SAS Intelligence Platform: Desktop Application Administration Guide*
- *SAS Intelligence Platform: Installation and Configuration Guide*
- *SAS Web Applications: Tuning for Performance and Scalability*
- Technical Paper, "Multilingual Computing with SAS 9.4"
- Text Mining and Analysis: Practical Methods, Examples, and Case Studies Using SAS

For a complete list of SAS publications, go to sas.com/store/books. If you have questions about which titles you need, please contact a SAS Representative:

SAS Books
SAS Campus Drive
Cary, NC 27513-2414
Phone: 1-800-727-0025
Fax: 1-919-677-4444
Email: sasbook@sas.com
Web address: sas.com/store/books



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

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

support.sas.com/bookstore
for additional books and resources.

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. © 2013 SAS Institute Inc. All rights reserved. S107969US.0613