What is the SAS environment and how can it be utilized for data analysis?

Authored by

stats writer

June 24, 2024

The SAS environment is a comprehensive software suite used for data analysis. It provides a user-friendly interface and a wide range of tools and techniques for manipulating, analyzing, and visualizing data. This environment includes various modules such as data management, statistical analysis, predictive modeling, and machine learning. It also offers a programming language, SAS programming, for more advanced and customized data analysis. With its powerful features and advanced capabilities, the SAS environment can be utilized by individuals, businesses, and organizations for a variety of data analysis tasks, including data mining, forecasting, and decision-making. It allows users to efficiently manage and analyze large datasets, obtain insights and patterns, and make data-driven decisions. Overall, the SAS environment is a valuable tool for data analysis that can assist in improving productivity, accuracy, and efficiency in various industries and fields.

Part 1: The SAS Environment

SAS is a command-driven program used for statistical analysis. Because of this, there is a steep learning curve for new users, especially those new to programming. This part of the tutorial will familiarize you with basics of the SAS environment.

The SAS 9.4 User EnvironmentDescribes the window and toolbar components of SAS's graphical user interface (GUI) and how they are navigated.Basic Programming GuidelinesThis guide covers how to write syntax of SAS, how to run a program, and how to recognize a data step versus a procedure step.SAS LibrariesSAS libraries allow users to safely store data sets and user-defined formats so that they can be accessed without having to reload them every time SAS is started. This tutorial discusses how to create and access a library in SAS, as well as the special WORK library, where temporary data sets and user-created formats are stored for the duration of the SAS session.

Part 2: Data Creation and Import

Before you can do any data manipulation and analysis, you need to have access to a SAS dataset file in your current session. You can do that in a number of ways. This tutorial will start off with a discussion about SAS Libraries and then discuss how to import or create SAS datasets.

The Data StepThe data step is the cornerstone of data management. This tutorial will introduce you to the basic components of the data step: the SET statement, the DROP statement, and the KEEP statement.Informats and FormatsInformats tell SAS how to read data. Formats tell SAS how to print data. This tutorial covers the ground rules of informats and formats.User-Defined FormatsThere are a number of built-in formats for numeric and date variables in SAS, but you may need to create your own formats from time to time, especially for categorical variables. This tutorial covers how create new formats using PROC FORMAT, and how to apply them to variables in a

data step. Also discussed is how to save user-defined formats to a library so that they can be re-used.

Part 3: Working with Variables

Now that you have a SAS dataset that you can work with, let's talk about what you can do with your variables. Much of this tutorial will cover data management standards and practices that, when followed, will make the rest of your work in SAS a lot easier. This tutorial will cover variable formats and labeling, missing values, manipulating variables, and calculating variables.

Defining VariablesThis SAS software tutorial covers the defining properties of variables: types, names, and labels.Missing Values in SASMissing values are observations where the information for a specific variable has not been recorded. SAS handles missing values for numeric and character variables differently. This tutorial describes working with missing values in SAS.

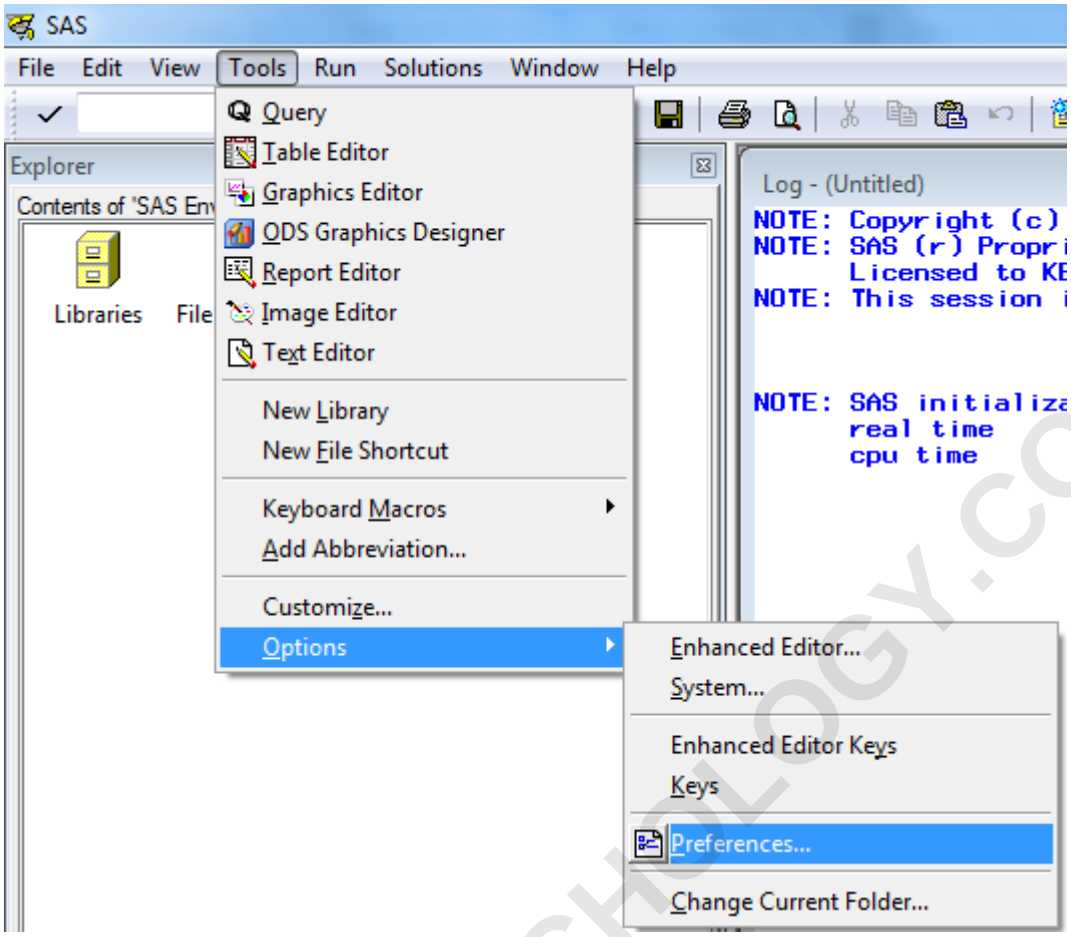
Changes to the default output in SAS 9.3

SAS produces high-quality graphics and formatting using its Output Delivery System (ODS). While ODS output is attractive, it requires a large amount of computing power. In SAS 9.2 and earlier, ODS was turned off by default, and had to be specifically enabled and disabled with a set of commands. In SAS 9.3, ODS is turned on by default. Because ODS can sometimes cause older computers to run more slowly, you may wish to disable ODS until you have a need to use it.

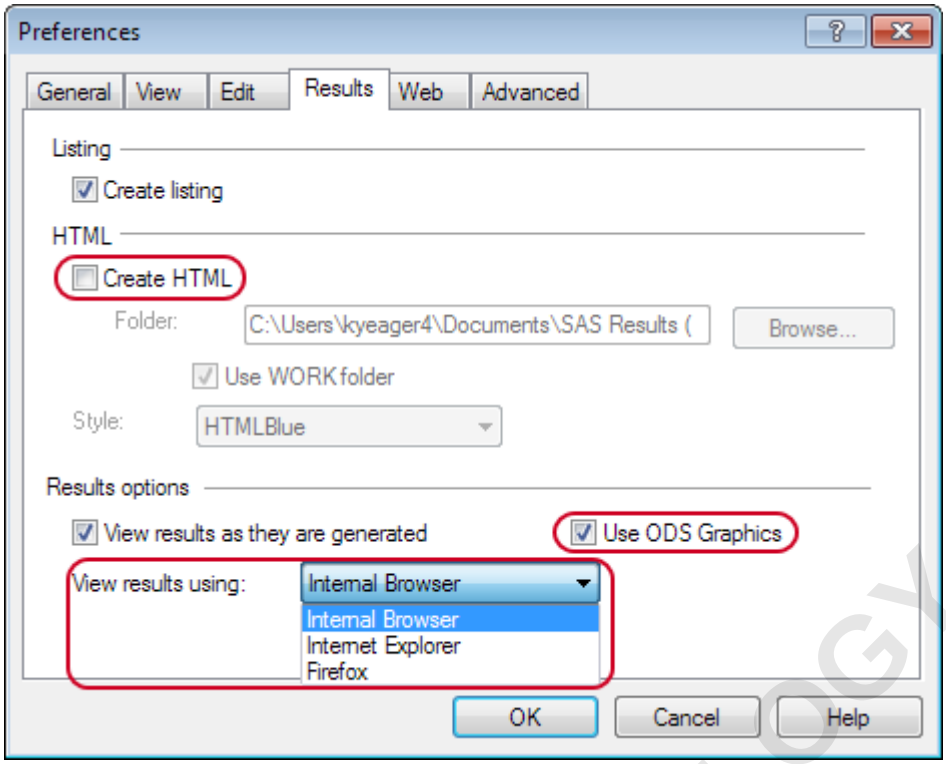
Additionally, our tutorials were produced using SAS 9.3, but all of our screencaps show our results going to the *Output* window. This was the default method of display in earlier versions of SAS. In SAS 9.3, the default option is to produce the output as an HTML file, and not send any text to the *Output* window. One advantage of using the HTML output is that copying-and-pasting the results into a word processing program is much less error prone. However, you may want to change this setting in your copy of SAS so that it is easier to compare your results to those in the tutorials. (Additionally, the majority of SAS tutorials online still use the *Output* window to display results.)

You can make changes to both of these settings by doing the following:

1. From the drop-down menus, select **Tools > Options > Preferences**.



2. In the **Preferences** window, click on the **Results** tab.



To disable the results being saved as an HTML file, un-check the box next to **Create HTML**.**To turn off ODS Graphics**, uncheck the box next to **Use ODS Graphics**.**To have the results printed in the *Output* window,** click on the drop-down menu next to *View results using*, and select **Internal Browser**. Alternatively, if you choose to save your output as an HTML file, you can use this drop-down menu to specify which browser to view the HTML document in.