
Introduction to SAS and SAS Studio

Welcome to the SAS statistical labs. In these labs, you learn how to use statistics to analyze and draw conclusions from real data by writing programs in SAS

What Is SAS?

SAS is an integrated software suite for advanced analytics, business intelligence, data management, and predictive analysis. You can use SAS software through both a programming language (called Base SAS) and a graphical interface, such as SAS Studio.

With SAS software, you can do the following:

- access data in almost any format, including SAS tables, Microsoft Excel tables, comma-separated values (CSV) files, and database files.
- manage and manipulate your existing data to get the data that you need. For example, you can subset your data, combine it with other data, and create new columns.
- analyze your data using statistical techniques ranging from descriptive measures such as correlations to logistic regression and mixed models to sophisticated methods such as modern model selection and Bayesian hierarchical models.
- present the results of your analyses in a meaningful report that you can share with others. The reports that you create can be saved in a wide variety of formats, including HTML, PDF, and RTF.

To complete these labs, you are using the SAS University Edition vApp. A *vApp*, or virtual application, is just a convenient way to use software without the complexity of having to install and configure the software products yourself. The SAS University Edition is a *vApp* that includes both Base SAS and SAS Studio.

You can download the SAS University Edition from http://www.sas.com/en_us/software/university-edition/download.html.

What Is SAS Studio?

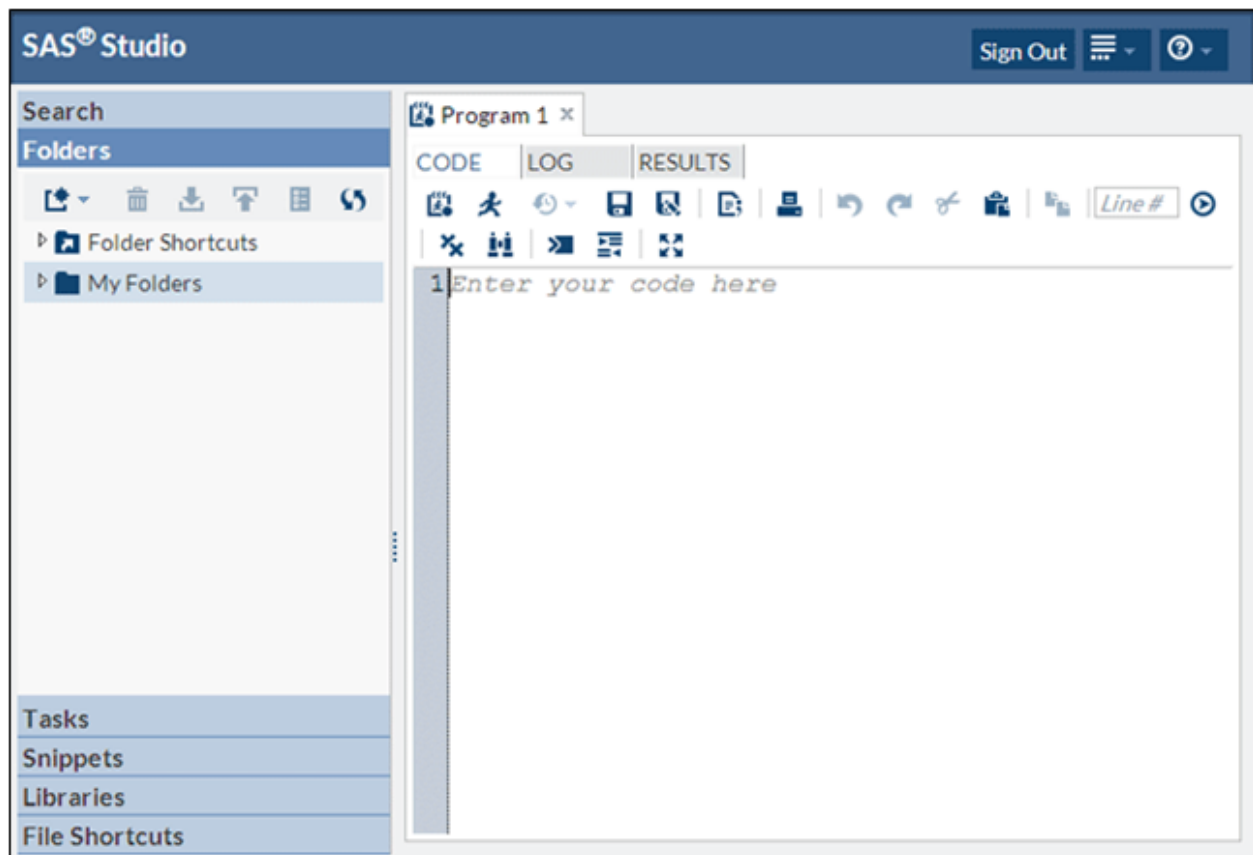
SAS Studio is a tool that you can use to write and run SAS code through your web browser. With SAS Studio, you can access your data files, libraries, and existing programs and write new programs. When you use SAS Studio, you are also using other SAS products behind the scenes. SAS Studio connects to a SAS server in order to process SAS commands. After the code is processed by the SAS server, the results are returned to SAS Studio in your browser.

¹ 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) and are not included under the CC-BY-SA license.



Getting Started in SAS Studio

When you run the SAS University Edition and open SAS Studio, the main SAS Studio window appears.



The main window of SAS Studio consists of a navigation pane on the left and a work area on the right. The navigation pane provides access to the search feature, your folders and folder shortcuts, your tasks and code snippets, the libraries that you have access to, and your file shortcuts. The Folders section is displayed by default.

The work area is used to display your data, code, logs, and results. When you first open SAS Studio, the work area displays a new program window. When you open data and run programs, other windows open in the work area with a tabbed interface.

Accessing Data in SAS

Before you start programming in SAS, you need to know how to access data in SAS. SAS tables are stored in SAS libraries. A SAS library is a collection of one or more SAS files that are recognized by SAS. In a Microsoft Windows environment, for example, a SAS library is typically a group of SAS files in the same folder or directory.

To access a SAS library, you assign it a name. The name is known as the *libref*, or library reference, and it refers to the physical location of the files. The libref is the nickname that you use to identify libraries in a SAS program.

All SAS tables have a two-level name. The first name is the libref, which tells SAS which library contains the table. The second name is the name of the table. The two names are separated by a period.

For example, **sashelp.class** refers to the table named **class** that is stored in the **sashelp** library.

To create a SAS library, you use the LIBNAME statement. The LIBNAME statement associates the name of the library, or libref, with the physical location of the library. In the SAS University Edition, you can create a libref to your shared folder. The following LIBNAME statement creates the **mydata** libref to your shared folder.

```
libname mydata '/folders/myfolders/';
```

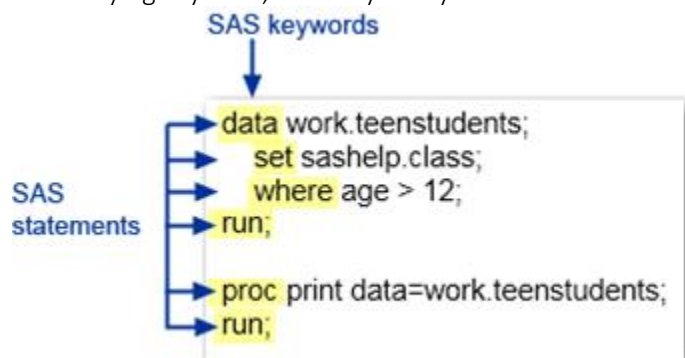
Writing a SAS Program

SAS programs consist of SAS statements. A SAS statement has two important characteristics:

- It usually begins with a SAS keyword.
- It always ends with a semicolon.


SAS programs include two types of statements: DATA steps and PROC steps. DATA steps begin with the keyword DATA and enable you to manage and manipulate your data. PROC steps begin with the keyword PROC and enable you to analyze and present your data. Generally, a step ends with a RUN statement or when a new DATA or PROC step begins. A RUN statement tells SAS to process all of the preceding lines of the step.

Each step in a SAS program consists of a sequence of statements. SAS statements usually begin with an identifying keyword, and they always end with a semicolon.



SAS statements are free format. This means that

- they can begin and end anywhere on a line
- one statement can continue over several lines
- several statements can be on one line.

To create a new program in SAS Studio, click **Folders** in the navigation pane. Then click  and select **SAS Program**.

Next Steps

This was a quick introduction to the basics of accessing data and writing a program in SAS. As you work through the lab, you learn more about specific analytical procedures and some different ways that SAS can help you better understand your data. If you want to learn more, look at the free tutorials that are available at <http://support.sas.com/training/tutorial/>.

Notes:

This lab was adapted for *OpenIntro* by Andrew Bray and mine Cetinkaya-Rundel from a lab written by Mark Hansen of UCLA Statistics.