

**2018**

**WESTERN USERS OF  
SAS SOFTWARE**

**SACRAMENTO, CA.**



# How SAS Thinks: SAS Basics I

---

Susan J. Slaughter, Avocet Solutions

# What is SAS Essentials?

Section for people new to SAS

Core presentations

1. How SAS Thinks
2. Introduction to DATA Step Programming
3. Introduction to SAS Procedures

We'll go fast

Slides are on my website

Pay attention because there will be a test

Do you have the handout?



# What is SAS?

Originally stood for Statistical Analysis System

Since 1980s officially no longer an acronym

Pronounced "sass" not "S. A. S."

Better answers:

- Company named SAS Institute
- Family of software products
- Programming language



# What is the SAS language?

SAS is a fourth generation language with features for data manipulation and analysis.

Generations of computer languages

- First—binary
- Second—machine code
- Third—"high-level" languages
- Fourth—built-in features such as query language

One generation is not better than another.



# Susan says

This is an overview

- Keep in mind that there are many exceptions

There are always at least 3 ways to do anything in SAS

- Don't worry about it

SAS is a big language

- No one knows it all

SAS is a language of defaults

- Once you know what the default is, you can override it



# Ways to run SAS

You have a choice of environments

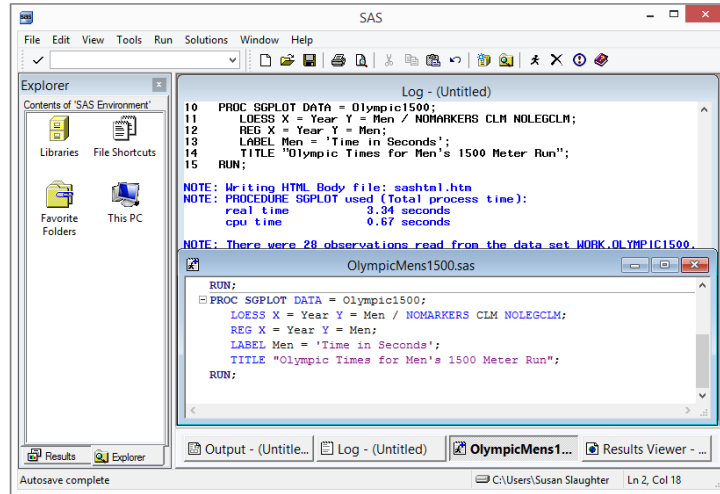
These 3 are included with Base SAS:

- SAS windowing environment (“Display Manager”)
- SAS Enterprise Guide
- SAS Studio

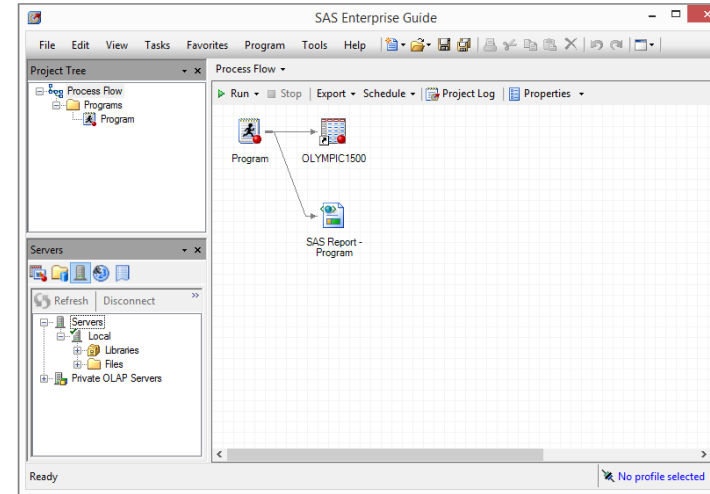


# Ways to run SAS

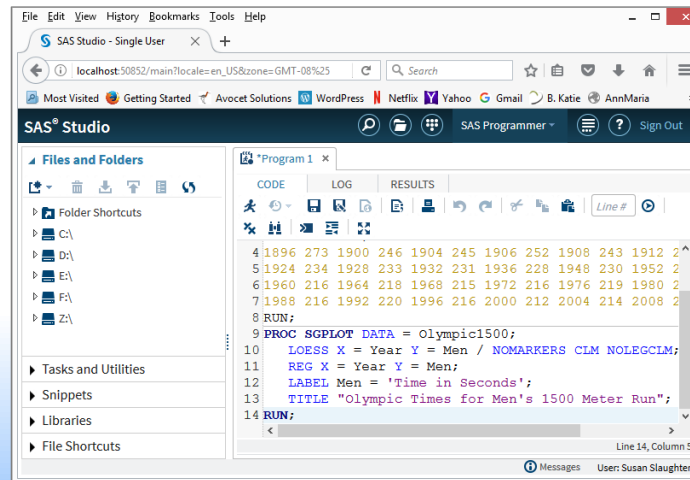
Display Manager



Enterprise Guide



SAS Studio



# Ways to run SAS

|                  | Write programs? | Point-and-click?<br>(SAS writes program) | Batch?           | Other considerations   |
|------------------|-----------------|--|------------------|--|
| Display Manager  | Yes             | No                                       | Yes              | Traditional<br>Most widely used<br>All operating systems<br>Easiest to learn |
| Enterprise Guide | Yes             | Yes                                      | Yes              | Windows only<br>Harder to learn<br>Can run slow                              |
| SAS Studio       | Yes             | Yes<br>(sort of)                         | Yes<br>(sort of) | New<br>Runs in a browser<br>Nice, but still pretty "raw"                     |





# SAS Studio

Many different ways to get it

- SAS Studio (Single-User, Basic, and Enterprise versions)
  - Part of standard SAS installation
  - Free with Base SAS license
- SAS University Edition
  - Download
  - Free for learning!
- SAS OnDemand for Academics
  - Use online
  - Free for learning!



# SAS data sets

Before SAS can use data, must be in a SAS data set

SAS data sets are comprised of two parts

- Data
- Descriptor (self-documenting)

When you run SAS programs, they are compiled and then executed

- What does this have to do with SAS data sets?
- Descriptor is constructed during compilation
- Data are read during execution



# SAS data sets

Descriptor portion contains:

Information about data set

- Name of data set
- Date created
- Number of observations and variables

Information about variables

- Name of variable
- Type (character or numeric)
- Length (in bytes)
- Label (if any)
- Informat and format (if any)



# SAS data sets

SAS data sets are rectangular

- Records = rows = observations
- Fields = columns = variables

Two types of variables

## Numeric

Numbers

Add, subtract

Period (.) for missing

## Character

Letters, numerals, \$#@!

Cannot add, subtract

Blank for missing

Some data values can be numeric or character

- Example: ID numbers

|   | ID    | Name   | Age | Major |
|---|-------|--------|-----|-------|
| 1 | 78374 | Thomas | 21  |       |
| 2 | 75638 | Cathy  | .   | STA   |
| 3 | 78634 | David  | 20  | ENG   |



# Dates in SAS

SAS date values are number of days since January 1, 1960

| <u>Date</u>       | <u>SAS date value</u> |
|-------------------|-----------------------|
| December 31, 1959 | -1                    |
| January 1, 1960   | 0                     |
| January 1, 1961   | 366                   |
| September 5, 2018 | 21432                 |

Stored as ordinary numeric data values

How many days old are you?

```
DaysOld = TodayDate - BirthDate;
```



# SAS variable names

## Rules for names of variables

- Start with letter or underscore
- Contain only letters, numerals and underscores
- Up to 32 characters long

## Can be upper or lower case

- SAS doesn't care
- Age, age, AGE and AgE all refer to the same variable
- But SAS remembers the case of first occurrence of variables



# SAS data set names

SAS data set names always have two levels:  
`libref.membername`

## Libref

- SAS data library reference
- Specifies location (disk, folder, path)

## Member name

- Individual data set in that library

Separated by a period



# SAS data libraries

SAS comes with built-in data libraries

- WORK
  - Temporary = erased when you exit SAS
- SASHELP
  - Permanent = not erased
  - Read-only, sample data sets
- SASUSER
  - Permanent = not erased
  - Read-write, for your data sets

Define your own SAS data libraries





# SAS data set names

## MySASLib.students

- two level name
- library= MySASLib
- member name = students
- will be permanent
- not erased by SAS

## students

- one level name
- library = WORK
- member name = students
- will be temporary
- erased by SAS
- real name WORK.students



# SAS data libraries

Many ways to create permanent SAS data libraries

- Depend on operating system

LIBNAME statement always works

General form (on Windows):

```
LIBNAME libref 'drive:\directory-path';
```

Example:

```
LIBNAME mysaslib 'c:\MySASData';
```



# Getting data into SAS

There are many ways

- Type it in yourself
  - Viewtable window in DM, Data Grid in EG
- DATA step with INPUT statement
- Import wizard
- PROC IMPORT
- SAS LIBNAME engines such as EXCEL
- SAS/ACCESS for files such as Oracle



# SAS programs

## First rule of SAS programming

Every SAS statement ends with a semicolon;

## No rules for formatting of code

- Upper or lowercase
  - I will use uppercase for keywords, mixed case for variables
- More than one statement per line
- Statements can continue on next line
- Any indention or none
- Can be completely unreadable—not recommended



# SAS programs

Always include comments in your programs!

Two styles

```
* This is a comment;
```

```
/* This is another comment */
```

Comments might include

- Your name
- Date
- Purpose
- Describe any non-obvious bits of code



# DATA versus PROC steps

Two basic parts of SAS programs

## DATA step

Begin with DATA statement

Input and modify data

Create SAS data set

Flexibility of programming

## PROC step

Begin with PROC statement

Perform analysis or task

Produce report

Like filling out a form

This is a simplification, but good guideline

Common mistake made by beginners is to use statements in wrong kind of step



# DATA versus PROC steps

A simple example:

```
DATA temps;  
    Fahrenheit = 68;  
    Celsius = (Fahrenheit - 32) * 0.5556;  
  
PROC PRINT DATA = temps;  
    TITLE 'Temperature Conversions';  
RUN;
```

DATA  
step

PROC  
step

A step ends when SAS encounters a DATA, PROC, RUN, QUIT, STOP or ABORT statement



# Global SAS statements

Global statements are not part of DATA or PROC steps

- Stay in effect until changed
- Only a few: OPTIONS, TITLE, FOOTNOTE, LIBNAME

```
OPTIONS NODATE;
```

```
TITLE 'Temperature Conversions';
```

```
DATA temps;
```

```
    Fahrenheit = 68;
```

```
    Celsius = (Fahrenheit - 32) * 0.5556;
```

```
PROC PRINT DATA = temps;
```

```
RUN;
```

Global

DATA  
step

PROC  
step





# SAS logs

When you run SAS programs, you get SAS log with

- SAS statements that you submitted
- Error messages
  - Program will not run!
  - Usually syntax or spelling problem
- Warnings
  - There may be a problem
- Notes
  - Data sets read and created
  - Number of observations and variables
  - May indicate a problem



# SAS log: Be sure to check!

```
1 DATA temps;  
2     Fahrenheit = 68;  
3     Celsius = (Fahrenheit - 32) * 0.5556;
```

NOTE: The data set WORK.TEMPS has 1 observations and 2 variables.

NOTE: DATA statement used (Total process time):

|           |              |
|-----------|--------------|
| real time | 0.01 seconds |
| cpu time  | 0.00 seconds |

```
4 PROC PRIN DATA = temps;
```

**ERROR: Procedure PRIN not found.**

```
5     TITLE 'Temperature Conversions';
```

```
6 RUN;
```

NOTE: The SAS System stopped processing this step because of errors.

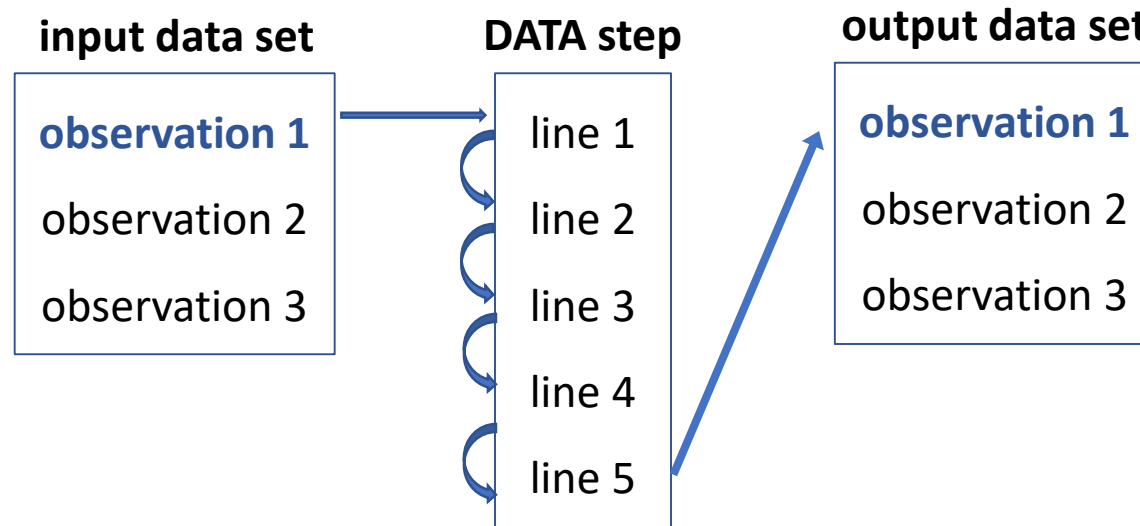
NOTE: PROCEDURE PRIN used (Total process time):

|           |              |
|-----------|--------------|
| real time | 0.00 seconds |
| cpu time  | 0.00 seconds |



# DATA step's built-in loop

DATA steps execute line-by-line and observation-by-observation



# Output Delivery System

ODS handles all procedure output

Susan says: You always use ODS!

Output formats are called destinations

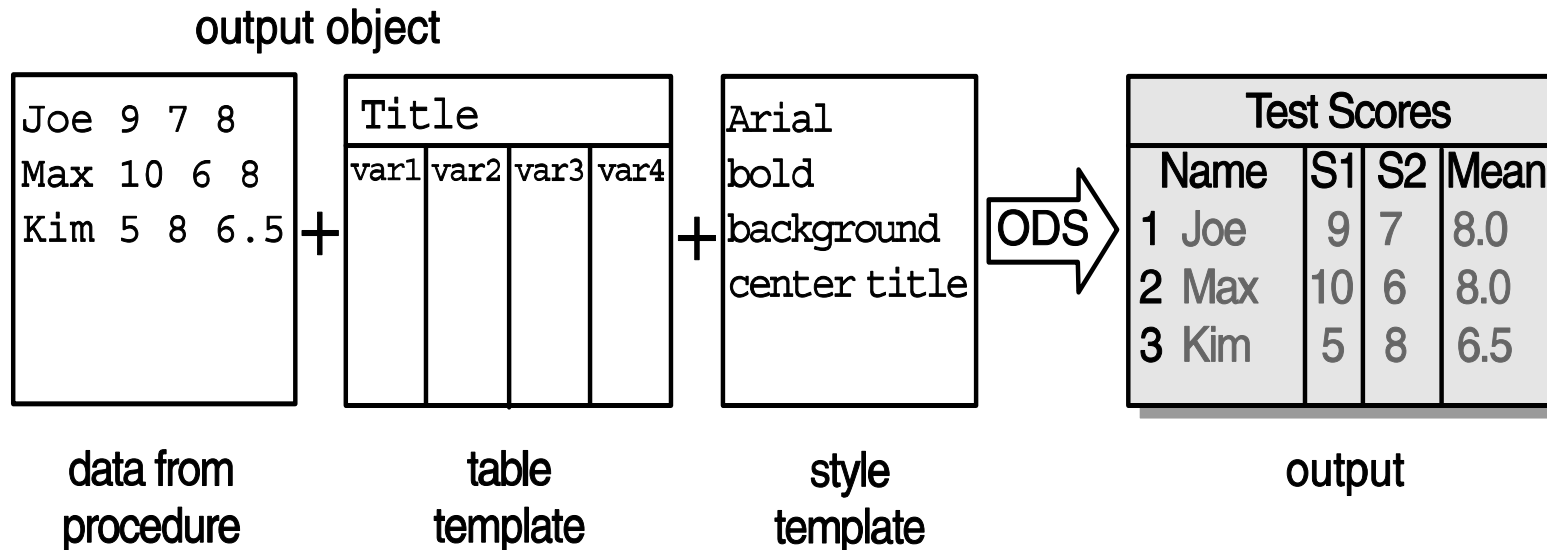
Many destinations

- HTML (default starting SAS 9.3)
- LISTING (text, default SAS 9.2 and earlier)
- PDF
- RTF
- POWERPOINT
- OUTPUT (SAS data set)



# Output Delivery System

How ODS works:



# Resources: Software

## Free software!

- Free for purposes of learning
- SAS University Edition
  - SAS Studio interface
  - Download and install virtual machine
- SAS OnDemand for Academics
  - SAS Studio interface is default
  - Use online, zero footprint



# Resources: Training

## Free training!

- [support.sas.com/training/](https://support.sas.com/training/)
- SAS Programming 1: Essentials online self-paced course
- Statistics 1: Introduction to ANOVA, Regression, and Logistic Regression online self-paced course
- More than 450 free tutorials



# Resources: Help

## Free help!

- [communities.sas.com](https://communities.sas.com)
- [blogs.sas.com](https://blogs.sas.com)

## Free SAS conference papers!

- [www.wuss.org](http://www.wuss.org)
- [www.lexjansen.com](http://www.lexjansen.com) (every SAS conference paper ever published)





# Pop quiz

1) Officially, what do the letters SAS stand for?

Nothing

2) What three interfaces are included with Base SAS?

Display Manager, Enterprise Guide, and SAS Studio

3) SAS data sets are comprised of what two basic parts?

Data and descriptor portions

4) What are the two types of variables in SAS?

Numeric and character



# Pop quiz

5) Someone tells you that the SAS date value for his date of birth is 17. On what date was he born?

January 18, 1960

6) Is this a valid SAS variable name?  
ABCDEFGHIJKLMNOPQRSTUVWXYZ

Yes

7) Which SAS data library is temporary?

WORK



# Pop quiz

8) Every SAS statement ends with what?

A semicolon;

9) What are the two basic parts of SAS programs?

DATA steps and PROC steps

10) What is the current default output format in SAS?

HTML



# Other presentations

Next up in this room

- Introduction to DATA Step Programming: SAS Basics II
- Introduction to SAS Procedures: SAS Basics III

Beginner's Techniques

Thursday 8:30-2:30 in Big Sur



# Thank you!

I hope you can stay for the next presentation.

Susan Slaughter

Avocet Solutions

Can download slides from [www.avocetsolutions.com](http://www.avocetsolutions.com)

