

How to Recognize and Use Calendar Dates

Melvin Tooker, Animal Improvement Programs Laboratory, USDA, Beltsville, MD.

ABSTRACT

A frequent concern of beginning SAS® programmers is how to make SAS recognize different calendar date formats. For example, September, Sept., SEP, 9, 09, /9, and /09 are all valid representations for the month of September. Although a program may currently read a date correctly, a new challenge arises when the data provider alters date formats or new users remotely enter data.

SAS DATE INFORMATS

The SAS system stores date values as the number of days since January 1, 1960. A format is a layout specification for how a variable should be printed or displayed. An informat is a specification for how raw data should be read. SAS provides many pre-defined formats and informats. The appropriate informat to use depends on how the calendar dates have been stored in the input file.

Let's begin with a simple example of storing a SAS date.

```
data _null_;
  NESUG_day = "13Sep2005"d;
  put NESUG_day = ;
run;
```

We can then see in the log file that NESUG_day is now = 16692. This is what we want. The 'ddmmyyyy'd is a date constant and represents a single SAS date value. If we drop the 'Sep', and substitute a "09" the function does not work. To go any further, we'll need to use an informat to allow SAS to translate a 'real' date value to a usable SAS numeric variable.

```
data _null_;
  NESUG_day = "13092005";
  newd = INPUT(NESUG_day,
              DDMMYY10.);
  put newd = ;
run;
```

"DDMMYY10." is an example of an informat. SAS can determine the desired date value by using the "DDMMYY10." informat regardless of whether

input is changed to "13/09/2005" or "13.09.2005". If the month and day are switched, another informat "MMDDYY10." should be used.

```
data _null_;
  NESUG_day = "09132005";
  newd = INPUT(NESUG_day,
              MMDDYY10.);
  put newd = ;
run;
```

Switching the locations of the year "20050913" requires the use of a similar informat "YYMMDD10."

```
data _null_;
  NESUG_day = "20050913";
  newd = INPUT(NESUG_day,
              YYMMDD10.);
  put newd = ;
run;
```

Changing the month from '09' to '9' does not present any problems if the correct informat is used; i.e. the 'YYMMDD' is in the correct order. However, changing the year from '2005' to '05' brings up an interesting problem. Changing NESUG_day in the above example gives us the desired SAS date. But, what if '05' is supposed to be 1905? SAS uses a default 100-year range of 1920 to 2020. If we need to change the default date range for SAS to use when reading 2 digit years, the yearcutoff option is used.

```
options yearcutoff = 1900;
data _null_;
  NESUG_day = "91305";
  newd = INPUT(NESUG_day,
              MMDDYY10.);
  put newd = ;
run;
```

Good programming practice would be to always include the yearcutoff option when the year is ambiguous.

ANYDTDTE INFORMAT

New for SAS 9.x® is the anydtdte informat, which extracts the date part from a date or datetime format. This informat accepts a datestyle option.

By specifying the order of month, day and year, we can control how the date is interpreted by SAS.

```
options datestyle=mdy;
data _null_;
  x=input('09/13/2005',
         anydtdte10.);
  put x;
run;
```

If the date is unambiguous the datestyle option is not needed and is actually overridden by SAS common sense. For example:

```
options yearcutoff = 1900;
options datestyle=ymd;
data _null_;
  x=input('09/13/2005',
         anydtdte10.);
  put x;
run;
```

SAS finds the correct date (16692) and overrides the incorrect yearcutoff and datestyle option. Care must be taken because this 'common sense' cannot be relied upon; changing the input above to '09/12/2005' with the datestyle='ymd' yields 09DEC2005 instead of what may have been intended, September 12, 2005.

CHARACTER DATES

When SAS encounters a character date, what is to be done? Converting '13SEPTEMBER2005' forces us to use a longer way to interpret things:

```
data _null_;
  date = '13 September 2005';
  b = input(scan(date,1) ||
           substr(scan(date,2),1,3)
           ||scan(date,3),date9.);
  put b;
run;
```

Changing the pattern of Day, Month or Year requires a similar change in the order of the scan function. Luckily, SAS 9.1® has National Language Support (NLS) that allows the use of the nldate informat:

```
options locale=English_UnitedStates;
data _null_;
  b=input('September 13, 2005',
         , nldate17.);
```

```
  put b;
run;
```

The default for the locale option is English_United States. Other locale options include German_Germany and French_France. Canada has two official languages: English and French; two values can be specified for the locale = system option: English_Canada and French_Canada. Dates can have different representations. For example, the decimal separator is a dot (.) in some regions and a comma (,) in others. Using the appropriate locale option with the nldate informat will simplify obtaining the correct SAS date.

CONCLUSIONS

Dates have many representations. The month may be represented as a number or as a name. The name may be fully spelled or abbreviated. The order of the month, day, and year may differ. Different countries may use a variety of spellings. It is important to use the appropriate informat to allow SAS to interpret calendar dates properly.

REFERENCES

SAS Online Documentation (SAS V9.1)

Numerous postings on the SAS listserv: <http://www.listserv.uga.edu/archives/sas-l.html>

Gilsen, Bruce (2003), "Date handling in the SAS system," in Proceedings of the 2003 (SUGI 28) SAS Users Group International Conference.

CONTACT INFORMATION

Your comments and suggestions may be sent to:

Melvin Tooker, AIPL, USDA
 Room 306, Building 005, BARC-West
 10300 Baltimore Avenue, Beltsville, MD 20705
 Phone: 301-504-8680 Fax: 301-504-8092
 Web: <http://aipl.arsusda.gov>
 E-mail: mtookер@aipl.arsusda.gov

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 registered trademarks or trademarks of their respective companies.