

Paper 068-29

Dating SAS® and MS Excel

Erik W. Tilanus, independent consultant, Vinkeveen, the Netherlands

ABSTRACT

Exchanging formatted date and time values between SAS and Excel may appear problematic, since the formats do not always match. Using local language versions of Excel can aggravate the problems. Internally both SAS and Excel use a numeric representation of dates and times. However they differ in anchor point (day 0) and in method. Using the internal structure, problems can be avoided, although a relatively simple transformation is required.

INTRODUCTION

There are several options to exchange data between SAS and Excel. The PC-File formats interface can read and create spreadsheets directly. If you do not have that option installed, a simple and effective method is to exchange data using tab-separated or comma separated values files. In general this works fine, except that dates, times and date/time combinations can cause troubles because of the different format specifications used by SAS and Excel. This is more so if you are using localized versions in different languages.

In this paper we show how to avoid those problems, by exchanging non-formatted (internal) dates, times and date/time values.

INTERNAL REPRESENTATION OF DATES AND TIMES

As generally known, a SAS date is a simple numeric value internally: the number of days since 1 January 1960. **1 January 1960** is **day zero**. Dates before this reference date have a negative internal value and the calendar is correct back to the 16th century and into the future for many centuries to come.

Times are counted internally in SAS as seconds since midnight and date/time combinations are calculated as the number of seconds since midnight 1 January 1960.

Excel also uses simple numerical values for dates and times internally. For the date values the difference with the SAS date is only the anchor point. Excel uses **1 January 1900** as **day one**.

Times are represented somewhat differently in Excel: a time in Excel is a fraction of a day. So for instance 12:00 noon in SAS is 43200 (seconds since midnight), in Excel it is 0.5 (half day).

This fraction approach is also used in date/time combinations. The integer part of the date/time combination is equal to the single date value. The fraction adds the time on the day.

CONVERSION FROM EXCEL TO SAS

With the knowledge of the internal values of Excel and SAS dates conversion is simple.

If you want to convert from an Excel date to a SAS date, subtract 21916: the difference in the starting points of the calendars.

Conversion of an Excel time value into a SAS time value is a question of multiplying by 86400, the number of seconds in a day.

Conversion of a date/time value is hardly more complicated: correct the date part by subtracting 21916 and then multiply the results by 86400 to convert it to the seconds used in SAS date/time values.

EXAMPLES:

```
SAS_date = Excel_date - 21916;
SAS_time = Excel_time * 86400;
SAS_date_time = (Excel_date_time - 21916) * 86400;
```

CONVERSION FROM SAS TO EXCEL

The other way around is obvious: reverse the calculations used in the previous paragraph. The examples are self-explanatory.

EXAMPLES:

```
Excel_date = SAS_date + 21916;
Excel_time = SAS_time / 86400;
Excel_date_time = SAS_date_time / 86400 + 21916;
```

CAVEATS

NEGATIVE DATE VALUES

There is one major difference between SAS and Excel dates and times: a SAS date, time or date/time value can also be negative. In Excel this is impossible. If you enter day 0, it will display as "0 January 1900"! A negative value creates an error display (#####)

The same happens with time values.

TIME VALUES HIGHER THAN 24 HOURS

Also note that the default format for times in Excel is HH:MM<:SS>. But HH does not go over 24 by default. So 15:00 + 16:00 displays as 7:00, unless you change the cell format!

DEMONSTRATION

Figure 1 shows a part of an Excel spreadsheet with various date and date/time values, according to default formatting (column A,D,G), their unformatted values (column B,E) and their conversion to the corresponding SAS value (column C,F). Column H shows the difference of D2 and G2: a positive value, formatted as time, with next to it the unformatted value. Note that the real difference should be 46:40:00! Column J and K show what happens if you subtract D2 from G2: a negative value.

	A	B	C	D	E	F	G	H	I	J	K
			Converted unformatted date (=B2-21916)	Standard Date/time value	Unformatted date/time value	Converted date time value (=E2-21916)*86400	Second date/time value	Difference between date/time values (=D2-G2), HH:MM:SS format)	Unformatted difference between date time values	Difference between date/time values =D2-G2, HH:MM:SS format)	Unformatted difference between date time values
1	Standard Excel date	Same - unformatted	16200	12/5/2004 8:10	38119.3403	1399968600	10/5/2004 9:30	22:40:00	1.94444444	#####	-1.94444
2	9-May-04	38116	16200	12/5/2004 8:10	38119.3403	1399968600	10/5/2004 9:30	22:40:00	1.94444444	#####	-1.94444

Figure 1: an Excel spreadsheet with a number of date and date time values, formatted and unformatted and their conversion to SAS values.

Figure 2 shows what the result is when you save this spreadsheet as a comma separated values (CSV) file.

```
Standard Excel date,Same - unformatted,Converted unformatted date (=B2-21916),Standard Date/time value,Unformatted date/time value,Converted date time value (=E2-21916)*86400,Second date/time value,"Difference between date/time values (=D2-G2), HH:MM:SS format)",Unformatted difference between date time values,"Difference between date/time values =D2-G2), HH:MM:SS format)",Unformatted difference between date time values
9-May-04,38116,16200,12/5/2004 8:10,38119.34028,1399968600,10/5/2004
9:30,22:40:00,1.944444444,#####
#####
#####
#####,-1.944444444
```

Figure 2: A CSV file version of the spreadsheet.

The easy way to read this CSV file is by using the IMPORT DATA wizard from the file menu in the display manager. However the result will not be satisfactory: it will not recognize several of the formats and it will create duplicate variable names for columns H to K. But still it is useful to run the wizard and then recall the generated source to adapt it to your own needs. This may be changing the format or informat specifications, the variable names or any other change or addition to the generated DATA step. In this example the modified source looks as follows:

```
data WORK.date_time_values;
  infile 'C:\SUGI29\date-time examples.csv'
    delimiter = ',' MISSOVER DSD lrecl=32767 firstobs=2 ;
  informat Standard_Excel_date anydtdte8.; * <-- New SAS 9 informat!;
  informat Same_unformatted best32. ;
  informat Converted_unformatted_date best32. ;
  informat Standard_Date_time_value anydtdtm. ; * <-- New SAS 9 informat!;
  informat Unformatted_date_time_value best32. ;
  informat Converted_date_time_value best32. ;
  informat Second_date_time_value anydtdtm. ; * <-- New SAS 9 informat!;
  informat Difference_between_DT_pos anydttme. ; * <-- New SAS 9 informat!;
  informat Unformatted_difference_DT_pos best32. ;
  informat VAR10 $255. ;
  informat Unformatted_difference_DT_neg best32. ;
```

```

format Standard_Excel_date date9. ;
format Same_unformatted best12. ;
format Converted_unformatted_date date9. ;
format Standard_Date_time_value datetime. ;
format Unformatted_date_time_value best12. ;
format Converted_date_time_value datetime. ;
format Second_date_time_value datetime. ;
format Difference_between_DT_pos time. ;
format Unformatted_difference_DT_pos best. ;
format Converted_difference_DT_pos time.;
format VAR10 $20.;
format Unformatted_difference_DT_neg best. ;
format Converted_difference_DT_neg time.;
input      Standard_Excel_date
           Same_unformatted
           Converted_unformatted_date
           Standard_Date_time_value
           Unformatted_date_time_value
           Converted_date_time_value
           Second_date_time_value
           Difference_between_DT_pos
           Unformatted_difference_DT_pos
           VAR10
           Unformatted_difference_DT_neg
;
Converted_difference_DT_pos = Unformatted_difference_DT_pos*86400;
Converted_difference_DT_neg = Unformatted_difference_DT_neg*86400;

put      Standard_Excel_date=;
put      Same_unformatted=;
put      Converted_unformatted_date=;
put      Standard_Date_time_value=;
put      Unformatted_date_time_value=;
put      Converted_date_time_value=;
put      Second_date_time_value=;
put      Difference_between_DT_pos=;
put      Unformatted_difference_DT_pos=;
put      VAR10 =;
put      Unformatted_difference_DT_neg=;
put      Converted_difference_DT_pos=;
put      Converted_difference_DT_neg=;
run;

```

The PUT statements generate the following information in the SAS LOG:

```

Standard_Excel_date=09MAY2004
Same_unformatted=38116
Converted_unformatted_date=09MAY2004
Standard_Date_time_value=.
Unformatted_date_time_value=38119.34028
Converted_date_time_value=12MAY04:08:10:00
Second_date_time_value=.
Difference_between_DT_pos=22:40:00
Unformatted_difference_DT_pos=1.944444444
VAR10=#####
Unformatted_difference_DT_neg=-1.944444444
Converted_difference_DT_pos=46:40:00
Converted_difference_DT_neg=-46:40

```

From these results you can derive that the formatted date/time values are not read in correctly. The other input fields are read smoothly and with the proper conversion on the Excel side or on the SAS side the resulting values are correct.

CONCLUSION

Converting SAS date, time and date/time values into the Excel equivalent or vice versa is easy. However be aware of some limitations in the date and time handling in Excel.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Erik W. Tilanus
PO Box 77
3645 ZK Vinkeveen
the Netherlands
Phone: +31 297 263936
Fax: +31 297 261712
Email: erik_tilanus@compuserve.com
Web: www.eriktilanus.com

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.