

# Five minutes - ISO 8601 Dates and Times

David Franklin, October 18, 2006

## Introduction

The ISO (International Organization for Standardization) has its roots with the International Electrotechnical Commission (IEC) in 1906 and the International Federation of the National Standardizing Associations (ISA) in 1926. In 1946 delegates from 25 countries met in London and as a result the ISO was began operations on 23 February 1947. There are ISO standards for so many things including what actually defines a "meter" (length) to international standards for the size of a seagoing cargo container.

ISO 8601 was first released in 1988 and dealt with the representation of dates, week numbers, times and durations. However, this paper will just look at some date and time representations of the standard.

It must be noted that dates and times have two formats, a basic and extended format however this paper will concentrate on the extended format that is more commonly used.

The main advantages of using the ISO 8601 format are:

- Easily readable and writable by systems
- Easily comparable and sortable
- Language independent
- Larger units are written in front of smaller units
- For most representations the notation is short and of constant length

## Dates

The standard representation for date is a character string in the format

YYYY-MM-DD

where YYYY is the year in four digit format, MM is the numeric representation of the month with '01' being for January and each month thereafter following in sequential order, and DD being the day of the month. The standard also allows for the less precision to be represented, i.e. YYYY-MM if no day is given or YYYY where a year is only available. The standard does not allow for cases where you have for example a year and day but no month, or where a month and day is present but no year.

## Times

The standard representation for time is a character string in the format

hh:mm:ss

where hh is the hour in two digit format using the 24 hour clock, mm is minutes and ss is seconds. Like in the case of dates, the standard does allow for less precision to be

displayed, i.e. where there is only hours and minutes the display only needs to be in hh:mm format. The standard also allows for the addition of a decimal fraction to the smallest time unit where higher precision is needed.

## Date and Time

The standard representation for a date and time, as a single variable, is a character string in the format

YYYY-MM-DDThh:mm:ss

where YYYY-MM-DD comes from the date, hh:mm:ss comes from the time and 'T' is used as a separator between date and time. As the case previously, the standard does allow for less precision than that available, e.g. if only the date is present then only the date will be represented. If decimal seconds are available then the standard does allow for that to be added at the end.

## What Does SAS Have?

For a number of years SAS programmers have had the functionality to read and write complete dates and complete hours using a combination of the YYMMDDxw format and the HHMMw.d format. SAS however in July 2004 released a hotfix (reference 82BC01, SAS Note SN-012361) that allowed for a number of new formats/informats dealing with ISO date/times, notably:

IS8601DA	date	YYYY-MM-DD
IS8601DN	datetime	w/date portion YYYY-MM-DD
IS8601DT	datetime	YYYY-MM-DDThh:mm:ss[.ffff]
IS8601TM	time	hh:mm:ss[.ffff]

The IS8601DA and IS8601TM have equivalent informats of YYMMDD10 and TIME8 (can have up to 24 decimal places after the second) and equivalent formats of YYMMDDD10 and TIME8. It must be noted that the IS8601TM informat requires a value format with a value for seconds.

IS8601DN and IS8601DT are interesting.

The IS8601DT format constructs an ISO datetime format and the equivalent informat will read it in.

```
34 usrdttm="16Oct2006:11:25"dt;  
35 put usrdttm IS8601DT.;  
2006-10-16T11:25:00
```

and

```
42 usrdttm=input("2006-10-16T11:27:00",IS8601DT.);  
43 put usrdttm datetime.;  
16OCT06:11:27:00
```

The format/informat IS8601DN can work with date/time to date conversions as the following examples show:

```

74 usrdttm=input("2006-10-16",IS8601DN.);
75 put usrdttm datetime.;
16OCT06:00:00:00

82 usrdttm="16OCT06:14:30:00"dt;
83 put usrdttm IS8601DN.;
2006-10-16

```

However, SAS does have its limitations and is very unforgiving when it comes to handling shortened forms of the date and/or time. Take the following example where the seconds is left off the time in the datetime:

```

50 usrdttm=input("2006-10-16T11:27",IS8601DT.);
51 put usrdttm datetime.;
NOTE: Invalid argument to function INPUT at line
50 column 13.

```

The second example below shows where there is a missing day off the date:

```

86 usrdttm=input("2006-10",IS8601DA.);
87 put usrdttm date9.;
NOTE: Invalid argument to function INPUT at line
86 column 13.

```

### Useful Hint

Joining a date and a time? Use the DHMS function, e.g.

```

1 data _null_;
2 StartDate='18OCT2006'd;
3 StartTime='09:06't;
4 StartDateTime=dhms(StartDate,0,0,StartTime);
5 put StartDateTime is8601DT.;
6 run;
2006-10-18T09:06:00

```

## What can be Used?

There is a macro called MakeISODate in the MACROS directory that creates a variable in the ISO8601 date standard using various input date time variables. The big advantage of this macro is that it will cater for the situations where the day and year may be present for a date but there is no month value. This situation is not ISO 8601 compliant but does fill the need for the data that is sometimes collected.

In almost every case we are dealing with complete time<sup>1</sup>, complete dates or partial dates. SAS can easily handle the cases where there is a complete date or complete time, using existing yymmddw and timew formats/informats as well as the IS8601DA and IS8601TM formats.

The problem still exists for partial dates, as it has always done. The advantage of having the date in ISO 8601 format is that it is easily sorted and a single variable can store the partial date. The disadvantage though is that the partial date will have to be transformed in some way to be useful for analysis and display purposes<sup>2</sup>. But this can be done easily as

<sup>1</sup> Must be in hh:mm:ss format to use IS8601TM, otherwise will have to use HHMMw format

<sup>2</sup> At present the current standard is to display dates in DATE9 format

the date and/time is just a character string that can be easily manipulated. The following SAS code demonstrates how different ISO date and time formats can be read and create a standard display for partial date and/or time fields:

```

data _null_;
length dx $15;
input dt $ 1-16;
if ^missing(dt) then do;
if substr(dt,3,1)=':' then dx=dt;
else do;
len=length(dt);
select(len);
when(16)
dx=put(input(scan(dt,1,'T'),
yymmdd10.),date9.)||
': '||substr(dt,12,5);
when(10)
dx=put(input(dt,yymmdd10.),date9.);
when(7)
dx='--'||put(input(compress(dt,'-'),
yymmn6.),monyy7.);
when(4) dx='----'||put(input(dt,4.),4.);
otherwise;
end;
end;
put @1 dt= @21 len= @23 dx=;
cards;
;
run;
dt=2006-10-12T14:14 len=16 dx=12OCT2006:14:14
dt=2006-10-12 len=10 dx=12OCT2006
dt=2006-10 len=7 dx=--OCT2006
dt=2006 len=4 dx=----2006
dt=15:23 len=. dx=15:23

```

This code can be easily adapted into a SAS macro for general use.

## Conclusion

ISO 8601 dates and times are relatively easy to use and understand. However the new IS8601 informats and formats, while useful, are identical to their YMMDD and TIME equivalents. The only informat/format that is new is IS8601DT where it is possible deal with a datetime in yyyy-mm-ddthh:mm:ss value but the following line will do it without the informat:

```

=dhms(input(scan(dt,1,'T'),yymmdd10.),0,0,
input(scan(dt,2,'T'),time8.));

```

and the format could easily be the following picture definition:

```

proc format;
picture isodt
other='%Y-%0m-%0dT%0H:%0M:%0S'
(datatype=datetime);
data _null_;
x=datetime(); put x isodt.;
run;
2006-10-17T14:45:18

```

*Cats regard people as warm blooded furniture.*

Jacquelyn Mitchard, The Deep End of the Ocean