SCstatisticalprogramming.com

How to Read Data in Fixed Columns Using Formatted Input Method Part 1

Formatted Input Method

- Reads data from fixed columns
- Reads character and numeric data including standard and nonstandard numerical values, such as numbers with dollar signs and commas
- Reads dates in different formats, not like column input which can only read dates as character values
- It is the most common and powerful of all the input methods
- Any time you have nonstandard data in fixed columns, you should consider using formatted input to read the file.

Raw Data

12345678901234567890123456789012 ---- Columns indicator, not part of the data

Tim M14510/21/1978

Sara 13009/20/1964

Mike M18011/23/1965

Laura F13011/06/1980

Sean M16704/07/2000

I use the same raw data file (data_column.txt) that was used in the tutorial on column input. The data is shown as above. The variables information is listed as below.

There are 4 variables in the data.

- Name: occupies total five columns, column 1-5
- Gender: occupies total one column, column 6
- Weight: occupies three columns, column 7-9
- DOB: occupies ten columns, column 10 - 19

SAS Informats

- Informats are built-in instructions that tell SAS how to read a data value. The choice of which informat to use is determined by the data.
- For standard numeric and character values, two basic informats are *w.d* and *$w.*
  - The w.d format: reads standard numeric values. The w tells SAS how many columns to read. The optional d tells SAS that there is a decimal point in the value, the decimal point counts as one of the columns to be read. For example, the variable Weight containing 3 digits values (such as 145 lbs) has an informat of 3. ; Another example, value 2.1 takes up 3 column spaces, 2 digits + 1 decimal point, so it has an informat of 3.1
  - The $*w*. informat: tells SAS to read *w* columns of character data. In this tutorial, the variable Name is read as character data and takes up five columns, and has informat of $5.; values of Gender take up a single column, and has informat of $1.

- For date: the MMDDYY10. informat tells SAS that the date you are reading is in the *mm/dd/yyyy* form, 10 means this form takes up 10 columns including 8 digits and 2 slash signs. For example, the variable DOB has values like 10/21/1978. SAS reads the date and converts the value into a SAS date. SAS stores dates as numeric values equal to the number of days from January 1, 1960. So, if you read the value **01/01/1960** with the MMDDYY10. informat, SAS stores a date value of 0. The date **01/02/1960** read with the same informat would result in a date value of 1, and so forth.
- For complete list of SAS informats, Please see SAS language, reference :

http://support.sas.com/documentation/cdl/en/lrdict/64316/HTML/default/viewer.htm#a001239776.htm