SYSTATDataConsolidator (SDC)

User Guide

# Definitions:

SYSTAT dataset A (strictly) rectangular array of data values; the columns are the values for a particular Variable and the rows the values for each SYSTAT Case; for SDC purposes, the SYSTAT dataset is conceptually divided into sub-arrays, each from a constituent FILMAN or CSV file

Individual SYSTAT column Contains the values for a single SYSTAT Variable

Individual SYSTAT row Contains the values for each of the Variables in a single SYSTAT Case

Collected column The values for a SYSTAT Variable Group; these values are selected from a set of files of a single type (FILMAN or CSV) that contain this information; the files that may be included in a Variable Group column must be of the same type; if they are CSV, they must have the same number of items in each row and the names of those items (in the first row) must be identical; if they are FILMAN files, they must have the same number of GVs, channels, and data points

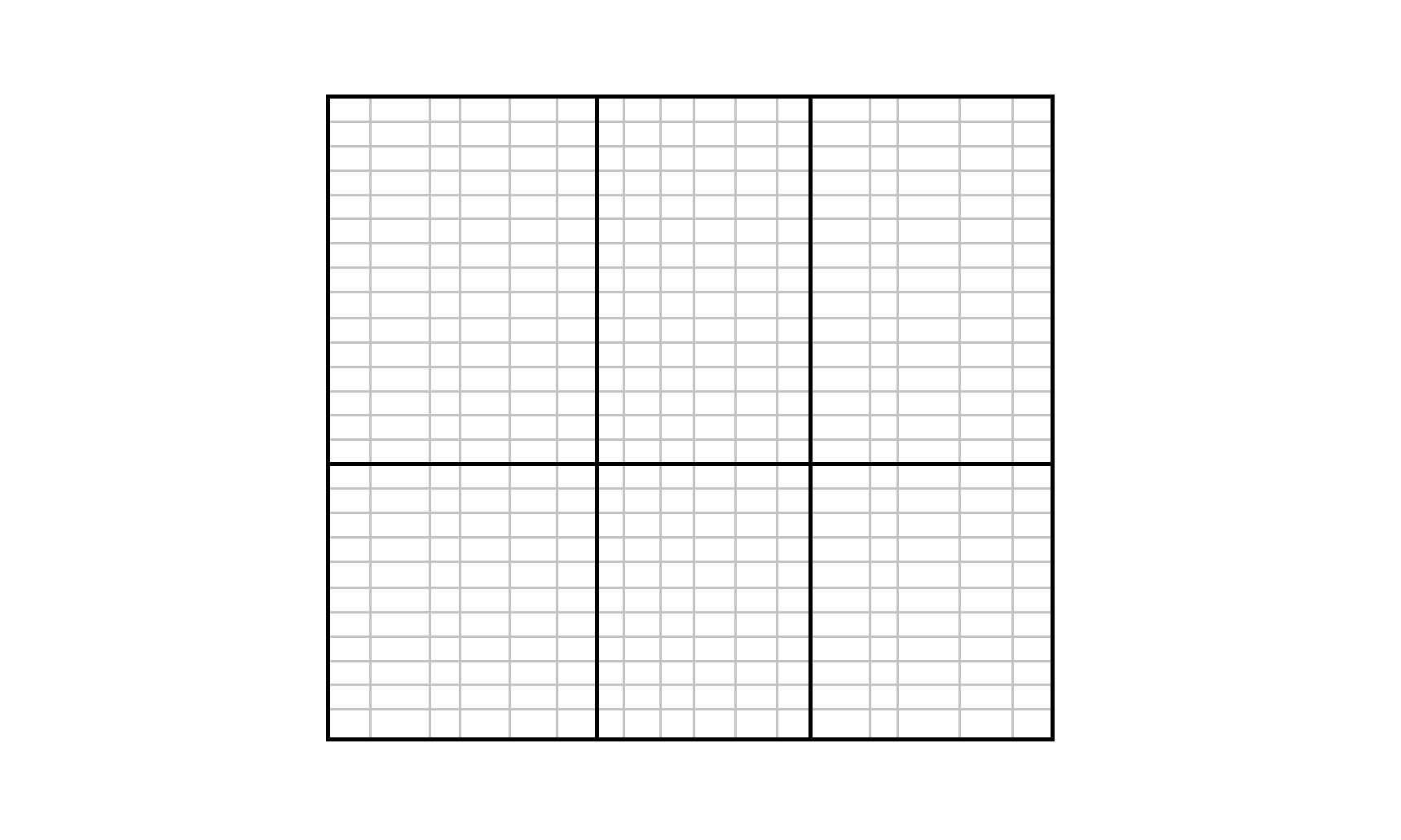
Collected row The Trials from a single experimental session; the files containing these the values may be of either FILMAN or CSV types (or both); all the files must have the same number of records (CSV) and/or recordsets (FILMAN), each representing a given SYSTAT Trial

CSV file Comma Separated Value file; the rows in this file consist of a series of values (either numbers or strings <= 12 characters in length); the first row contains the names of each of the Variables; these names must be <=12 characters in length, start with a letter and contain only letters, numbers and underscore; Variables which encode strings must end with a $ (which is included in the 12 character limit); each row ends in a CR/LF and the values in a row are separated by commas; Excel CSV files are in an acceptable format

FILMAN file Standard format FILMAN file created by FILMAN processing; a given FILMAN file is described by the number of GVs, number of channels, number of datapoints, and number of recordsets it contains. Sometimes FILMAN files entered into SYSTATDataConsolidator will be processed into a single channel so that a single record(set) represents a single trial. If more than one channel is present in the FILMAN file, the data from the various channels will be entered as separate individual SYSTAT columns

Missing data Indicated by the value -1E+36 in a file (SYSTAT, FILMAN or CSV); this value can be used to fill in “holes” in a data file in order to make files compatible

Below is a diagram of a SDC output file, a file that will be input into SYSTAT for analysis. The individual small gray rectangles each represent a single datum; each gray row is a SYSTAT trial; and each gray column a SYSTAT variable. Each of the larger black rectangles represents an individual FILMAN or CSV input file. Notice that **File 1** and **File 4** must have the same number, ordering, and type of columns but may have different number of trials. This assures that the SYSTAT variables “line up” in the consolidated output file. The same requirements are true for **File 2** and **File 5**, and **File 3** and **File 6**. Also note that **File 1**, **File 2**, and **File 3** must have the same number of trials, though they may (and almost certainly will) have different variables. This is analogously true for **File 4**, **File 5**, and **File 6**. SDC enforces these conditions and will proceed only if they are met. One mechanism for overcoming this “limitation” is the use of Missing data.

****

File 1

File 2

File 3

File 4

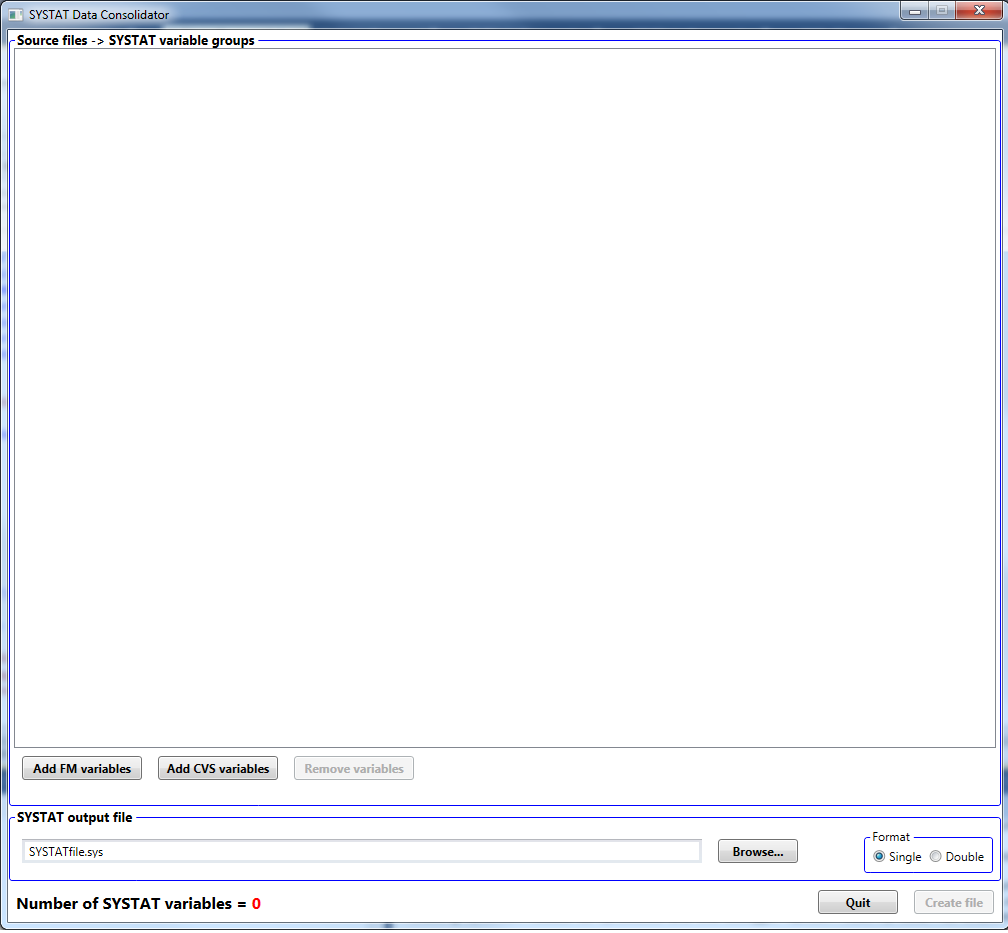
File 5

File 6

**SYSTAT variables**

**SYSTAT cases**

Now let’s look at SDC itself. Here’s the main window immediately after opening the program.

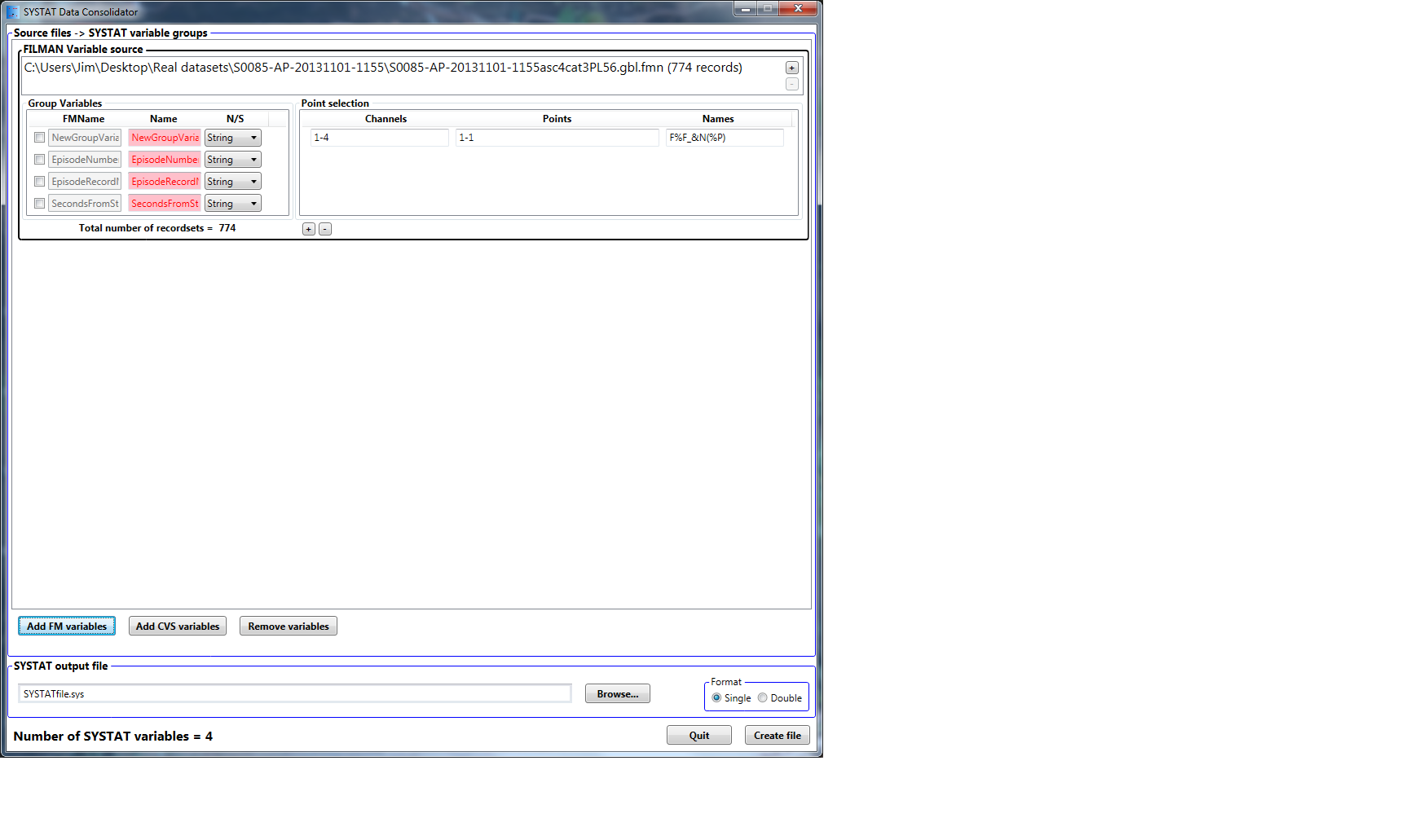


Click here to select output file location

SYSTAT output may be either single or double precision

Use these buttons to add FM or CSV files to create new SYSTAT output variables

After we’ve added a FILMAN file, the window looks like this.



See \*\* below

Here we select, name (no more than 11 characters if a string or 12 if a number), and format GVs to be included as SYSTAT columns. See below (\*) for naming convention

Variable blocks are removed by clicking here

SYSTAT variables are selected from the channels and points from each FILMAN recordset. SYSTAT variable names are also created. Names are encoded based on the filenumber (uniquely generated by the program) [%F], the channel number [%C] or name [&N], the point number [%P], etc. Other letters, numbers and \_() may be used as well. See below (\*). Remember that names may be no longer than 12 characters!

Use these buttons to add (or subtract) compatible files to this Collected Column. To successfully add a file it must have the same GVs, number of channels, and number of points.

Each block added to this list represents a set of Collected Columns (variables) in the SYSTAT output

\* Naming SYSTAT variables is somewhat problematic when there are many variables to be named. The system for naming points from FILMAN files attempts to simplify this problem by automatically generating names that “make sense”. Think of this as a “format” string for creating a name for each variable. The macros in the naming convention are two symbol codes as follows:

%F Each input file is assigned a unique number from 1 to the number of files; this symbol represents this number and thus permits creating unique variable names even though several files are involved which may otherwise have identical variables

%f Index assigned to each input file

%C The original channel number from the FILMAN file

%c Renumbered channel in the order selected in the channel selection field

%P Original point number from the FILMAN file

%p Renumbered point number in the order selected in the point selection field

&N The actual channel name from the FILMAN file

Group variables may have macros as well using these codes:

%G Original number of the GV in the FILMAN file

%g Renumbered (based on the GVs selected) GV

%F and %f As in the point naming convention above

Any of the numeric codes (those beginning with a “%”) can have an optional digit after the % to indicate a fixed field width, padded left with zeros. Thus the variable name for channel 16 point 123 would encode from “V%C(%P)” as “V16(123)”, while “V%3C(%4P)” would result in “V016(0123)”.

Valid SYSTAT names are less than 12 characters long (11 if the GV is naming a string-valued variable), begin with a letter or the symbol “\_” and contain only letters, numbers, and “\_”. A set of parentheses may be used to enclose a number at the end of the name (as in the above example). SDC cannot ensure that the name length is of appropriate, so **BE** **CAREFUL**!

\*\* These buttons are used to add additional naming conventions for channel/point subsets.