

Sasshiato© Documentation

By Izabella Peszek

Document Issue: 1.01

Issue Date: 2011/06/24

Sasshiato release 2.03 documentation

Copyright 2009, Izabella Peszek

THIS MATERIAL IS PROPERTY OF IZABELLA PESZEK.
Released under GNU General Public License v3.0 as part of Sasshiato project.

Table of Contents

Document Overview	4
Version Control	4
References and Trademarks	4
Dataset description	5
Variables Required by Sasshiato	6
Restricted Literals.....	14
Formatting Instructions	15
Troubleshooting.....	17

Document Overview

Version Control

Version	Major Changes
1.01	Initial release

References and Trademarks

SAS® software is a registered trademark of SAS Institute Inc. Henceforth this document references SAS® as simply SAS.

Dataset description

The XML file passed to Sasshiato has the following structure.

The first record (`__datatype='RINFO'`) contains the information about general formatting properties of the report. There should be one and only one such record and it should be the first record in the dataset.

This record is followed by one or more group of records (each group corresponding to the value(s) of “PAGE=Y” grouping variables). Within each group, there should be one or more record corresponding to the table header, followed by the records corresponding to the table body. Records are printed in order they appear in the dataset.

Each time Sasshiato encounters a table header record (`__datatype="HEAD"`), a new page is started (except when table header is encountered for the first time). If such record has `__varbylab` and `__varbygrp` specified, `__varbylab` is printed above the table headers. All “adjacent” header records must have the same values of `__varbylab` and `__varbygrp`.

Sasshiato does not use SAS® datasets directly. The dataset passed to Sasshiato must be converted to xml file. Note that xml descriptor for the first record must include descriptions of some variables described below, even if their values are null. For that reason, built-in xml SAS XML engine can not be generally used and xml file has to be created using put statements.

Variables Required by Sasshiato

The following variables must appear in the first record (RINFO record). They do not need to be populated on the subsequent records.

Variable Name	Allowed Values	Description/Notes
__dest	RTF PDF	Output format. One or both can be specified.
__layouttype	STD, INTEXT	Applicable to RTF reports only. If __layouttype=INTEXT, the rtf report generates a single table (possibly spanning multiple pages) with footnotes printed after the table. If vertical page splits are necessary, the table with first group of columns is generated, followed by the table with the second group of columns etc. Many advanced formatting options may not be available with INTEXT option. If __layouttype=STD, each page of the rtf report is a separate table. The footnotes are printed on every page. If vertical page splits are necessary, they are created for every page of the report.
__filename		The name of the output file, without extension. Depending on the value of __outformat, __filename.rtf and/or __filename.pdf files are generated.
__path		Full path to the output folder where report will be generated.
__nodatamsg		The text to display if there is no data to show in the report (the only record in the XML file is the RINFO record)
__orient	L P	Orientation (Landscape or Portrait)
__margins		Size of margins (with unit) in order top, right, bottom, left. Unit can be inch (in) or centimeter (cm). Examples: 1in 1in 1.5in 1in 2cm 2.5cm 2cm 2cm
__papersize	LETTER A4	Paper size
__font	TIMESROMAN HELVETICA COURIER	Font Face. For RTF output, Times Roman is replaced with Times New Roman, Helvetica is replaced with Arial, and Courier with Courier New

Variable Name	Allowed Values	Description/Notes
fontsize		Font size
sfoot fs		Font size for system footnote.
__bookmarks_rtf		Comma delimited string of tokens, where each token takes the form of bkN=titleX or bkN=titleX-titleY. Bookmarks in PDF files are created differently than bookmarks in RTF files. In particular, RTF does not support titles bookmarked in reversed order. If the range is provided (titleX-titleY), all titles in the range are assigned the same heading style.
__bookmarks_pdf		Comma delimited string of tokens, where each token takes the form of bkN=titleX or bkN=titleX-titleY. Bookmarks in PDF files are created differently than bookmarks in RTF files. In particular, if the range is provided (titleX-titleY) then this creates a bookmark with concatenated text of all titles in the range.
__title1, ..., title6		The titles. Each title is printed in a new paragraph. Blank titles are ignored.
__footnot1, ..., __footnot8		The footnotes. Each footnote is printed in a new paragraph. Blank footnotes are ignored. See “XXX” for description of special formatting for footnotes.
shead l		System header (left side) ^(a)
shead m		System header (center) ^(a)
shead r		System header (right side) ^(a)
sfoot l		System footer (left side) ^(a)
sfoot m		System footer (center) ^(a)
sfoot r		System footer (right side) ^(a)
__breakokat		String of space delimited integers indicating after which columns vertical page break can occur. For example, breakokat = 3 6 9 indicates that page breaks can occur after __col_3, __col_6 and __col_9. The numbers specified must have corresponding columns in the dataset. If __breakokat is not specified the vertical breaks are fully decided by sasshiato.
__colwidths		String of tokens indicating the width of each column. If specified, the number of tokens should be the same as the number of columns. The following values can be

Variable Name	Allowed Values	Description/Notes
		<p>used for each token: A - auto N - no wrapping, NH - no wrapping including header, LW – fit longest word, LWH – fit longest word including header <number> - fixed number in points, (1in=72 points) <number>in - fixed number in inches <number>cm - fixed number in cm</p> <p>Note: for best results it is recommended to always use NH and/or LWH.</p> <p>If not all columns are provided with instructions the following defaults are set: - if no instructions are provided then 72 points is set for <code>__col_0</code> if it is a header column (otherwise 15 points is set) and N is set for all remaining columns. - if only first column is provided then all other columns are assumed as N. - if more than one column is provided then the last <code>__colwidths</code> instruction is repeated for the remaining columns. (Robert verify this)</p>
<code>__indentsize</code>		The size in points of the single indentation.
<code>__lasttheadid</code>		This variable only takes effect when vertical page splits are necessary and indicates the number of the last “header” column. All columns \leq <code>__lasttheadid</code> are repeated on each page of the report. By default, <code>__lasttheadid=0</code> .
<code>__gcols</code>		String of integers indicating which columns contain “grouping” variables. Grouping variables are printed only when their value changes, and in the first record per page. (Iza verify this)
<code>__stretch</code>		String of space delimited tokens, one for each report column, indicating whether the column is allowed to be stretched (Y or N). If specified, the number of tokens should be the same as the number of columns.
<code>dist2next</code>		String of space delimited integers, one

Variable Name	Allowed Values	Description/Notes
		for each report column, indicating distance in points between this and the next column. If specified, the number of tokens should be the same as the number of columns. If not specified, the value of <i>__col_sp</i> specified in <i>__sprops</i> is used for each column.

- (a) The following substitutions are applied: *__DATE__* is replaced with the date and *__PAGE__* is replaced with Page X of Y

Iza Move sprops here

The following variables must appear in each record, including RINFO record.

Variable Name	Allowed Values	Description/Notes
<i>__col_0</i> , <i>__col_1</i> , ... <i>__col_X</i>		Variables corresponding to the “columns” of the report. These variables must form a consecutive string (e.g. <i>__col_0</i> , <i>__col_1</i> , <i>__col_2</i> etc.). The “gaps” (e.g. <i>__col_0 __col_1 __col_3</i>) are not allowed.
<i>__datatype</i>	RINFO, HEAD, TBODY	Indicates the type of a record: general report info (1 st record), table headers, or table body

The following variables are optional variables used when there are two or more parts of the report with different headers. Whenever a new header record (*__datatype*=‘HEAD’) is encountered, the value of *__varbylab* is printed above the table header. Note that *__varbygrp* and *__varbylab* must have the same values for each record of the same header (if there is more than one).

Variable Name	Allowed Values	Description/Notes
<i>varbygrp</i>	integer	Numeric identifier of the header (integer)
<i>varbylab</i>		Text to print above table header

The following variables must appear on each record except RINFO record (they are ignored in RINFO record if they appear there).

Variable Name	Allowed Values	Description/Notes
__align		String of space delimited tokens, one for each report column, indicating alignment to be applied to the cell in this column. Each token can take the following values: C, L, R, D or RD. (center, left, right. Decimal, right-decimal). The RD alignment is not supported if the text contains line breaks (/). In such cases the line break is printed as "/". The number of tokens should be the same as the number of columns. The values of this variable can vary between rows. To calculate the positions of decimal points, all rows with D alignments are analyzed as a group and the best placement of decimal point is determined; and all rows with RD alignments are analyzed as a group (independently from D group) and the best placement of decimal points is determined.
__suffix	// blank (null)	If __suffix="//", a blank row is printed after this record (unless this is last record per page)
__indentlev		Applies to __col_0 only and specifies indentation level. The actual indentation size is equal to $\text{__indentlev} * \text{__indentsize}$
__keepn	0 1	Indicates whether page break can occur on this record. If __keepn=1, this record must be kept on the same page as the record that follows. If several subsequent records have all __keepn=1, and they can not be all fitted on one page, __keepn is ignored.
rowid	integer	Row identifier. Should be unique for each row.
__tcol		A text spanning all columns of the report. When non-blank value of __tcol is encountered, Sasshiato first prints this value spanning all columns of the reports, and after that resumes regular printing of report columns.
__sprops		See below

The following are the component of `__sprops` parameter. The components are passed to Sasshiato as comma-delimited string of component=value tokens.

appendable	If appendable=true then Page numbers are not being printed (awaiting next part of the report)
append	If append=true then the report is being appended to the existing report with the same <i>path</i> and <i>filename</i>
col_sp	Default spacing between columns (in points). Can be overridden for individual columns using <i>dist2next</i> parameter
rtf_extlns	Extra lines used for PDF/RTF compatibility if system headers, system footers or footnotes are placed inside the table. The default is 2.
esc_char	A single character indicating an “escape” character, that is, a character followed by formatting instructions. (See Formatting Instructions section). The default is “/”.
gen_size_info	If <code>__font=COURIER</code> and <code>__gen_size_info=simple</code> then Sasshiato generates three additional files in the output folder, containing the information about column widths (in number of characters), the sas statement for spanned headers (columns parameter of PROC REPORT) and the rtf statement defining page size and margins. These files are named <i>__filename_info.txt</i> , <i>__filename_info2.txt</i> , and <i>__filename_info3.txt</i> .
splitchars	A string of characters indicating on which characters line breaks can occur (blank is always added to this list). Note that if other formatting instruction forces the column width to be smaller than some words in a cell, a line break may occur in the middle of the word.
rtf_linesplit	calc or hyphen (default). Allows users to control where words are split if they can not fit on one line. If <i>rtf_linesplit</i> = calc then sasshiato calculates on which <i>splitchars</i> character the line break will occur and inserts a space after this character, which should result in rtf rendering a line break in this place. However, MS Word rendering is sometimes random and this calculation may not always work as designed. If <i>rtf_linesplit</i> = hyphen then sasshiato inserts a very small white hyphen after each <i>splitchars</i> character. In both cases nothing is inserted after standard <i>splitchars</i> (space and dash).
date_fmt	Any format supported by java.text.SimpleDateFormat, e.g. ddMMMyyyy HH:mm
date_fmt_uc	true/false: if true, the date is printed in upper case
body_pd	Size (in points) of the vertical space between last table row on a page and the solid line above the footnotes. It affects only the case when (a) the page is completely "filled" by table rows - in that case it ensures esthetically pleasing separation between table body and footnotes or (b) when “running” footnotes are used (appearing right under table body). In other cases, this vertical space is adjusted so that footnotes always appear at the bottom of the page
body_ld	size (in points) of leading for table body ^(a)

body_ff	font face for table body ^(b)
title_al	alignment of regular titles: L (left) R(right) or C (center)
title_pd	size (in points) of the vertical space between last title and a solid line on top of the table
title_ld	size (in points) of leading for titles ^(a)
title_ff	font face for titles ^(b)
head_pd	the space (in points) before each line of table header
head_ld	size (in points) of leading for table header ^(a)
head_ff	font face for table header ^(b)
headu_sp	width (in points) of the space separating lines under adjacent spanned table headers
foot_pd	size (in points) of the vertical space between the solid line above the footnotes and the 1st footnote
foot_ld	size (in points) of leading for footnotes ^(a)
foot_ff	font face for footnotes ^(b)
rtfpl_foot	USED FOR RTF FILE ONLY: placement of footnotes. t=as part of the table on each page, hf=in the footer area of the page
foot_pos	position of the footnotes: B (bottom of the page) or R (running, that is right under table body)
foot_ftsp	size (in points) of the horizontal space between "tabs" in the footnotes (defined using /ftl or /ftr control words)
bseparator	Position of the horizontal line separating table body and footnotes: f=just above footnotes, t=just under table, n = none
shead_pd	size (in points) of the vertical space between the last line of the system header line and the first title
shead_ld	size (in points) of leading for system headers ^(a)
shead_ff	font face for system headers ^(b)
rtfpl_shead	USED FOR RTF FILE ONLY: placement of system headers: t=as part of the table on each page hf=in the footer area of the page
sfoot_pd	size (in points) of the vertical space between the last footnote and the system footers
sfoot_ld	size (in points) of leading for system footers ^(a)
sfoot_ff	font face for system footers ^(b)
rtfpl_sfoot	USED FOR RTF FILE ONLY: placement of system footers: t=as part of the table on each page hf=in the footer area of the page
super_rs	size (in points) of the raise for superscript
super_fs	font size for superscript. If not specified (recommended), font size of table body-2 is used

^(a) Leading = height of the capital letter plus space above it used to separate lines
If left blank (recommended), the default is 1.2 times font size)

^(b) bf=bold, it=it, bf it = bold italic. If not specified then font face is set to normal

Restricted Literals

The following characters have special meaning for Sasshiato and should not be used inside any of the text variables in the xml file. If necessary, they must be encoded as shown in the Formatting Instructions section.

\
{
}
/ (this is a restricted escape character but can be changed to another escape character using *esc_char* component of *__sprops*)

In addition,

~{super
is a restricted word to indicate superscript (~{super *mytext*} prints *mytext*)

Formatting Instructions

The following in-text formatting instructions can be passed to Sasshiato.

/t<number>	line break with tab. <number> indicates number of tabs
/i<number>	initial tab (recognized only at the beginning of the text)
//	line break
/#<number><space>	Prints a character with ASCII decimal encoding = <number>
/s#<number><space>	Prints a character with ASCII decimal encoding = <number> using Symbol font
/s#le<space>	Prints ≤
/s#ge<space>	Prints ≥
/s#pm<space>	Prints ±

Notes:

/t<number> and /i<number> are supported only for __col_0 - __col_X variables.

The escape character / (shown) can be replaced using *esc_char* component of *__sprops*.

How to print the restricted characters:

- to print “/”: either use /#47 or change escape to a different character
- to print “\”: use /#92 or \\
- to print “{”: use /#123
- to print “}”: use /#125

The following syntax can be used in footnotes to create tabbed footnotes:

Text1 /ftr Text2

Text1 /ftl Text2

This will put a footnote in a table row containing three cells. The first cell will contain Text1 (right-aligned if followed by /ftr and left-aligned if followed by /ftl). The second is a very narrow empty cell with no text. The third cell will contain Text2.

Example:

<__footnot1>Note: /ftr Creatinine Clearance was calculated using the subject's actual body weight, the serum creatinine level measured at the local laboratory and the Cockcroft-Gault formula. </__footnot1>

<__footnot2>(a)/ftr Obtained prior to enrollment </__footnot2>

<__footnot3>(b)/ftl at Informed Consent </__footnot3>

This will result in the footnotes as follows:

Note: Creatinine Clearance was calculated using the subject's actual body weight, the serum creatinine level measured at the local laboratory and the Cockcroft-Gault formula.

- (a) Obtained prior to enrollment
- (b) at Informed Consent

Note that the first two footnotes have right-aligned text in the first cell while the last footnote has this text left-aligned.

Troubleshooting

Sasshiato will fail to generate the report in the following situations:

- Illegal characters are encountered in XML file, preventing parsing of this file.
- All header columns (columns with numbers \leq `__lasttheadid`) can not fit on one page with the column widths instructions provided.
- `__breakokat` instructions do not allow to fit on one page all the columns requested to be kept on the same page

There may be other scenarios when the instructions provided to sasshiato are internally incompatible (e.g., when combined width of the left and the right margins exceeds the page width). Generally, the log will provide the reason why report could not have been generated.