PharmaSUG 2024 - Paper SS-306

Leveraging SAS and Adobe Plug-in for CRF Bookmark Generation (Rave studies)

Swaroop Neelapu, Algorics

ABSTRACT

Annotated CRF (acrf.pdf) should be dual bookmarked: (1) bookmarks by chronology and (2) bookmarks by forms. This is usually done manually and is quite tedious, time consuming and an error prone process. This paper will discuss automation that can be leveraged to complete this process in record time with much higher precision and accuracy. This is possible irrespective of the number of pages in the CRF or the number of visits in the trial. It is a completely menu-driven process with very minimal use of keyboard involving only 4 major steps.

The resulting bookmarks in acrf.pdf are compliant with the standards mentioned in Metadata Submission Guidelines version 2.0 (CDISC-SDTM-MSG-v2.0). It also includes dual bookmarking by Visits/Forms including other features like nested levels, running records, inherit zoom magnification, etc.

INTRODUCTION

The prerequisites for using this process are:

a) The studies are built in Medidata Rave
b) SAS software
c) Adobe Acrobat Standard (or Pro version). It should include the capability to export and import bookmarks. There are many third-party tools available online for the purpose of exporting and importing bookmarks. For the purpose of this presentation, the plug-in software used is called "AutoBookmark Professional Plug-in for Adobe Acrobat" from Evermap Inc. This plug-in will install additional menu items in Acrobat for exporting and importing bookmarks from and to a csv file (see **Figure 2** below). (Note: A free version of the plug-in is available for 30-day trial download. Evermap Inc, USA is a certified Technology Partner for Adobe Inc.).

Software and versions used: Windows, SAS version 9.4, Adobe Acrobat Standard 2017 (along with Autobookmark Plug-in menu installed), MS Office 365

Skill level required for use: although a little SAS knowledge is preferred, anyone can deliver this even without any technical know-how. The process is majorly menu-driven apart from running a SAS program that has been provided in this paper.

Figure 1. Overall Process Flow Diagram

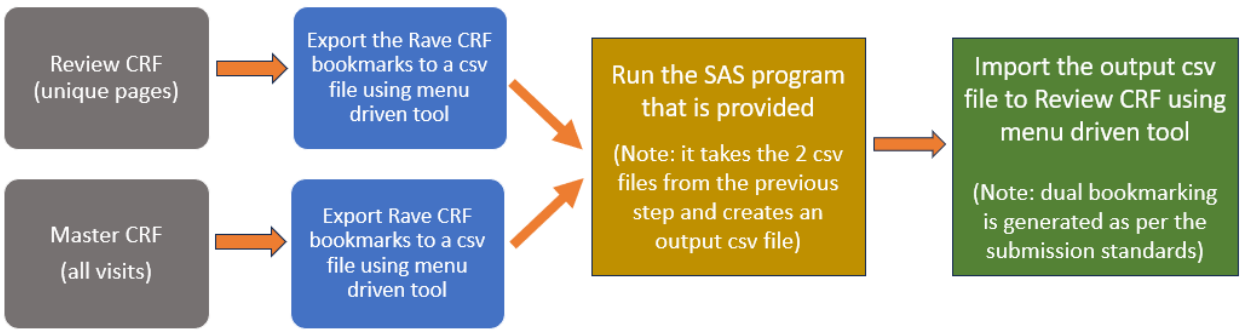
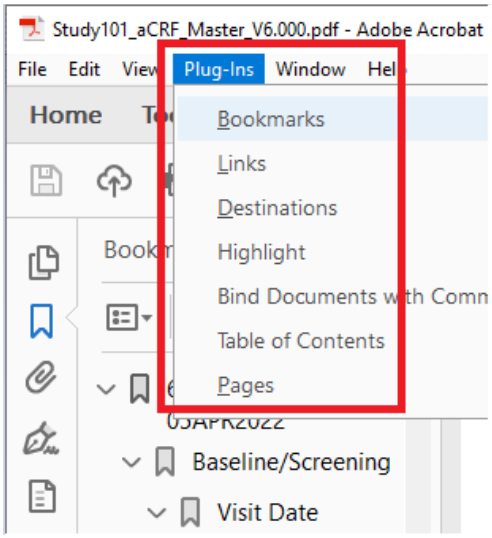


Figure 2. The Autobookmarks Plug-In when installed will add the menu items shown in red box below. It includes menu items for batch processing of Bookmarks apart from others.



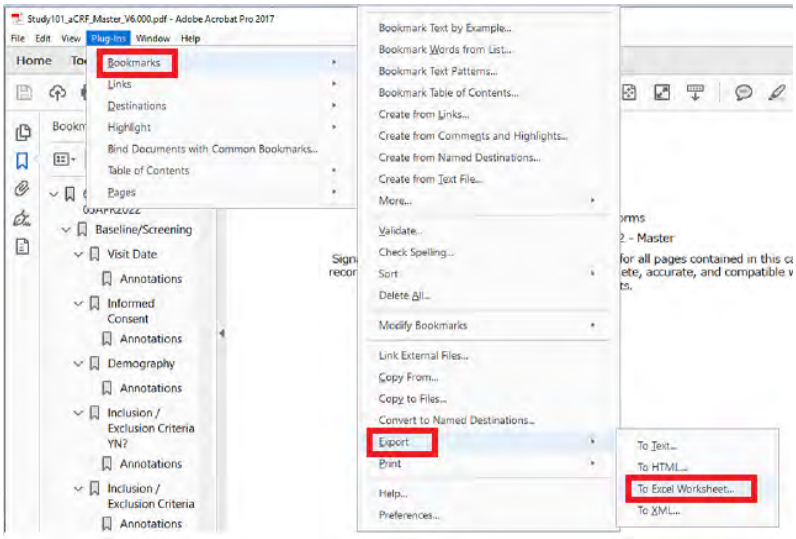
Process Steps:

a) The "EDC Developer" or "Database Programmer" can provide two different types of CRFs from Rave. "Review CRF" which contains only the unique pages from the study and "Master CRF" which contains all pages for all trial visits (i.e., if vital signs is occurring at multiple visits, then the same vitals page is present for all the visits with just the header being different).

b) Open the Master CRF pdf and export the already existing Rave created bookmarks to an Excel format file as below. (Note: these bookmarks are provided by Rave and are not conformant to health authority requirements. The CRF pdf file can be with or without Rave-generated annotation pages and it does not make a difference to the process).

Plug-in → Bookmarks → Export → To Excel Worksheet

Figure 3. Shows how to export bookmarks from Master CRF using the menu



c) Open the Excel spreadsheet and save the file as a csv file and name it as “input_Master.csv”. (NOTE: although the menu says “To Excel Worksheet…” the file is saved as an xml file but can be opened in MS Excel).

The XML file will look like below. Only 2 columns “Bookmark” and “Level” are relevant to the process from the Master CRF and are used in the SAS program.

Figure 4. shows the csv file that is exported from Master CRF.

Bookmark	Page Number	Named Destination	Web Link	JavaScript	Level
6.000 PROD SLF 05APR2022	1				1
Baseline/Screening	2				2
Visit Date	2				3
Annotations	4				4
Informed Consent	5				3
Annotations	7				4
Demography	9				3
Annotations	11				4
Inclusion / Exclusion Criteria YN?	13				3
Annotations	14				4
Inclusion / Exclusion Criteria	15				3
Annotations	16				4
Medical History YN?	17				3
Annotations	18				4

d) Similarly, for Review CRF (unique pages CRF) export the bookmarks to Excel and then save the file as “input_Review.csv”.

The XML file will look like below. Only 2 columns “Bookmark” and “Page Number” are relevant to the process from the Review CRF and are used in the SAS program.

Figure 5. shows the csv file that is exported from Review CRF

Bookmark	Page Number	Named Destination	Web Link	JavaScript	Level
6.000 PROD SLF 05APR2022	1				1
Enrollment	2				2
Annotations	3				3
Unscheduled Visit	4				2
Annotations	7				3
Visit Date	10				2
Annotations	12				3
Informed Consent	13				2
Annotations	15				3
Demography	17				2
Annotations	19				3
Inclusion / Exclusion Criteria YN?	21				2
Annotations	22				3
Inclusion / Exclusion Criteria	23				2
Annotations	24				3
Medical History YN?	25				2

e) Run the SAS code that has been provided in **Appendix 1**. The program will make use of the 2 above created csv files and generate an output file called “output_bookmarks.csv”.

(NOTE: A brief record of how the SAS program is written. Although both the input csv files have many columns, only 2 columns are relevant for us. Columns “bookmark” and “level” from “output_Master.csv”. Columns “bookmark” and “page_number” from “input_Review.csv”. Only these 4 columns are sufficient to generate a properly formatted output csv file.)

Figure 6. Below is Output CSV file for “VISIT” bookmarks. Notice the Page Numbers (4,8) after comma. Also, Running Records lists AEs, CMs

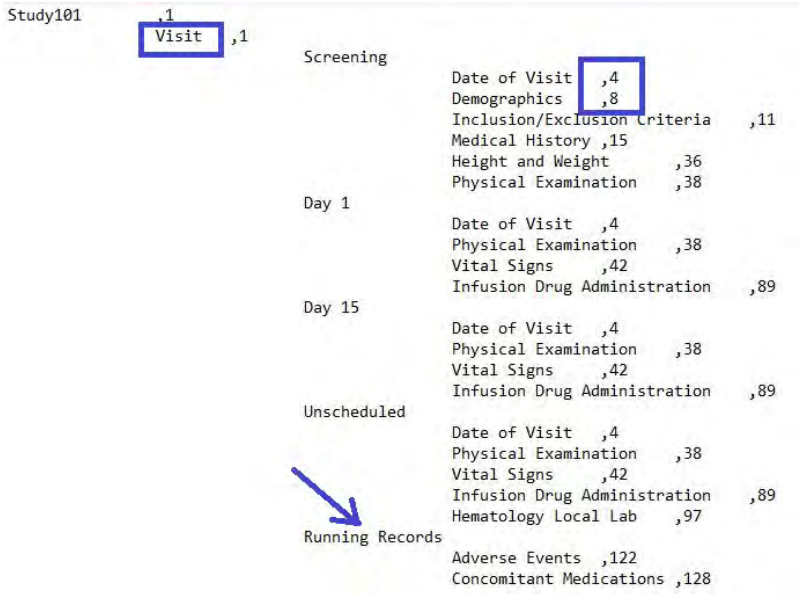
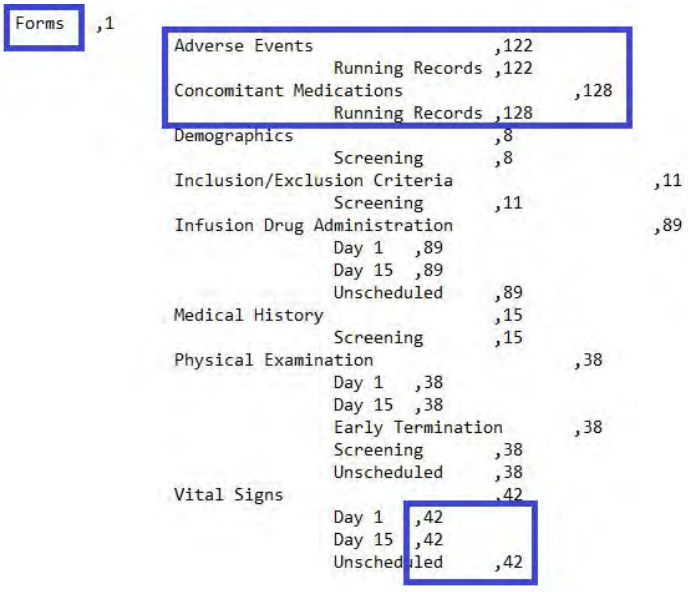


Figure 7. Below is Output CSV file under “FORMS” bookmarks. Notice the Page Number is the same (42) for all Vital Signs visits. Also, Running Records for AEs, CMs

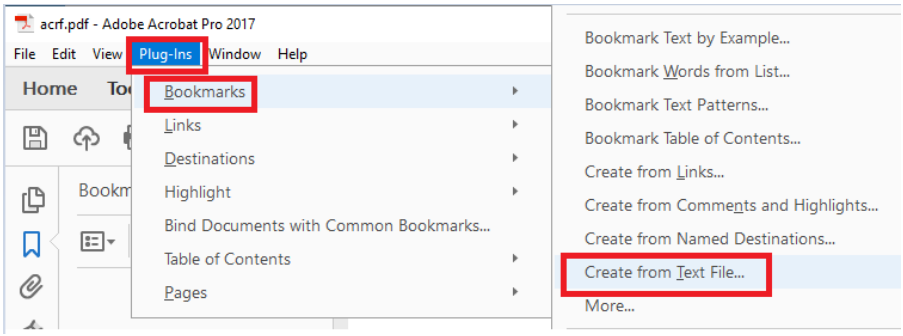


f) Make a duplicate copy of Review CRF and rename the file as acrf.pdf. Delete all existing Rave bookmarks from this file.

g) Open acrf.pdf and import the above generated file “output_bookmarks.csv” as below:

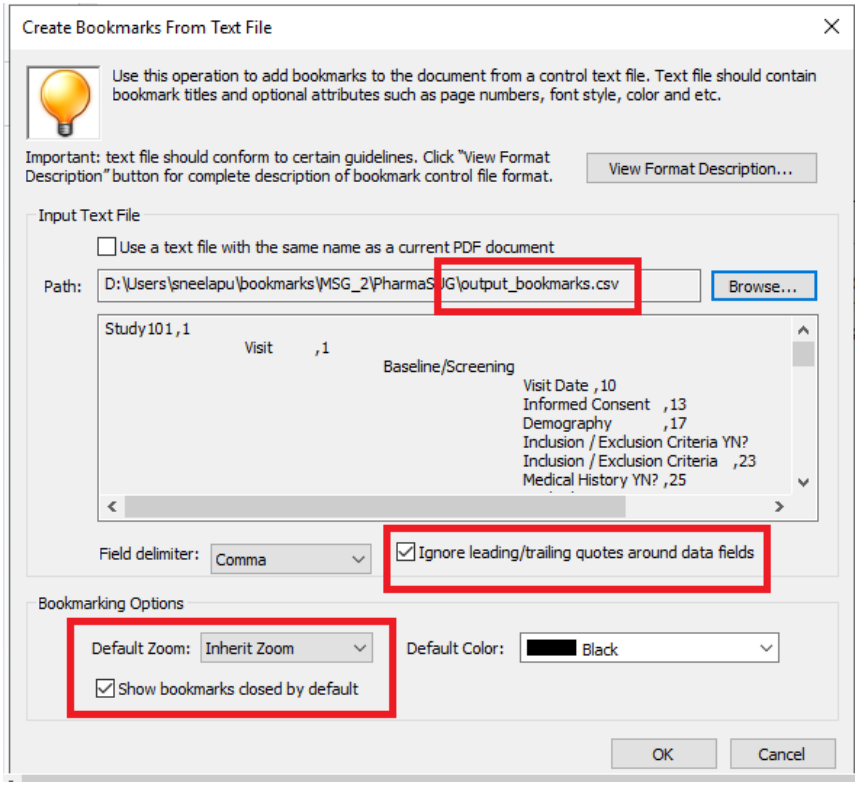
Plug-ins → Bookmarks → Create From Text File

Figure 8. shows how to import the csv file that is generated by the SAS program



The above menu item when clicked will open the below window and confirm these options are as marked:

Figure 9. shows the options to be chosen while importing the csv



Click on “OK” and the bookmarks will be created as per the regulatory requirement as in below 3 screenshots.

Figure 10. shows dual bookmarks created as per CDISC-MSG-v2.0

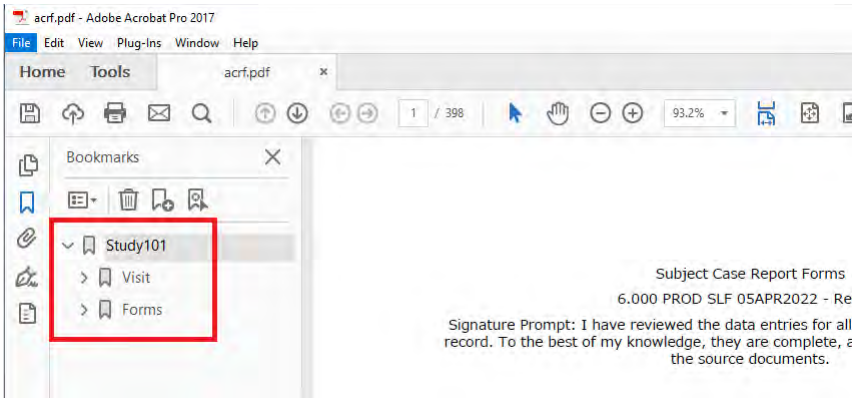


Figure 11. shows the "VISIT" bookmarks expanded. The sort order is by visit. Notice the "Adverse Events", "Concomitant Medications" and other menu items created under "Running Records".

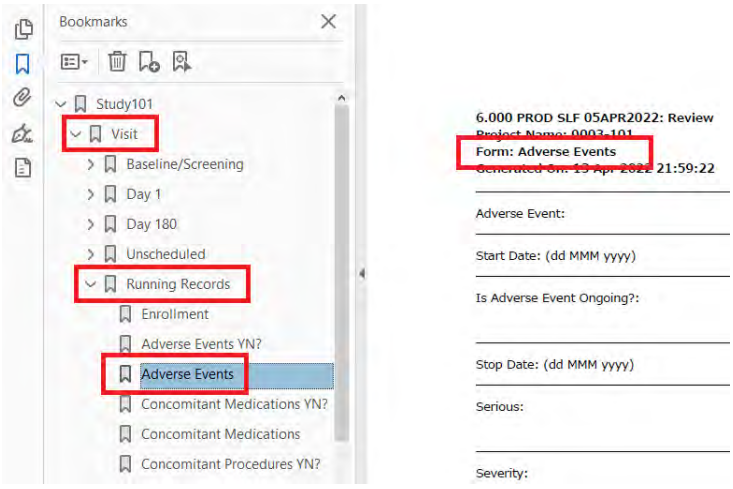
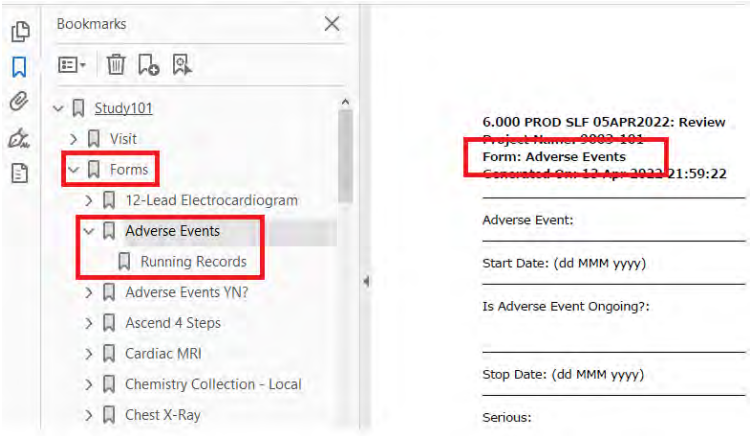


Figure 12. shows the "FORMS" bookmarks expanded. The sort order is alphabetical. Notice the "Running Records" created under the "Adverse Events" bookmark.



CONCLUSION

Dual bookmarking is a tedious task especially for studies that have a large number of CRF pages and visits. However, the special structure of CRF bookmarks in Medidata Rave system makes it possible to automate the process. Through this automation, we significantly reduce the manual burden and make it more consistent.

There is an additional benefit of this Autobookmarks plug-in. As per the CDISC guidelines, a “Table of Contents” is also needed in the annotated CRF. The plug-in menu can create it quite easily. After creating dual bookmarking as per above process, just click the menu item “Plug-In” -> “Table of Contents”. With this the “acrf.pdf” is ready and complete for submission to regulatory authorities.

REFERENCES

None

CONTACT INFORMATION

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

Swaroop Neelapu, Algorics, Hyderabad, India
neelapu.swaroop@gmail.com

APPENDIX 1:

This macro will output one CSV file named “"output_bookmarks.csv"”.

```
/*
Author: Swaroop Neelapu
Version: 1.0
Date: 15.Mar.2024

Purpose: this SAS Program creates CSV file for acrf.pdf bookmarks
      as per SDTM-MSG-v2.0 standards;

Assumptions: the program assumes that the structure of the Rave bookmarks in
            Master CRF and Review CRF are as per screenshots provided in the
            paper

Input: 2 CSV files that are created from the 2 CRFs;
Output: one CSV file is generated by the program (named
"output_bookmarks.csv");

eg: Sample macro call;

%bookmrk (  Study_Name = Study101,
                  Files_Path =
D:\Users\sneelapu\bookmarks\MSG_2\PharmaSUG\Study101,
                  Input_Master_CSV_FileName = input_Master,
```

```
                Input_Review_CSV_FileName = input_Review);
*/

%macro bookmrk
      (Study_Name = ,          /* the Trial name */
      Files_Path = ,  /* the file path where the 2 input CSV files are
located */
      Input_Master_CSV_FileName = , /* filename of input CSV file for Master
CRF */
      Input_Review_CSV_FileName = ); /* filename of input CSV file for Review
CRF */

* import the master csv file;
proc import datafile = "&Files_Path.\&Input_Master_CSV_FileName..csv"
  out = master_bk(keep=bookmark level)  replace;
  guessingrows = 32767;
  getnames = yes;
  datarow = 2;
run;

* import the review csv file;
proc import datafile = "&Files_Path.\&Input_Review_CSV_FileName..csv"
  out = review_bk(keep=bookmark page_number)  replace;
  guessingrows = 32767;
  getnames = yes;
  datarow = 2;
run;

* delete all Annotations;
data review_bk;
  set review_bk;
  if bookmark ^= 'Annotations';
run;

* configuring VISIT bookmarks structure;
data master_bk1(keep=bookmark level page_num1);
  set master_bk;
   length page_num1 8.;
   page_num1 = .;
   if bookmark = 'Annotations' then delete;
run;

proc sql noprint;
  update master_bk1 a set page_num1 = (select page_number
                from review_bk b
                where a.bookmark = b.bookmark
             /*   and bookmark not in ('Vital Signs' '12 Lead ECG')*/ );
quit;

data master_bk1a;
  set master_bk1;
  length nm2 $200.;
  if level = 1 then do; output; nm2 = 'Visit'; pg2 = ',1'; end;
  output;
run;

data master_bk1b;
```

```
  set master_bk1a;
   rownum = _N_;
run;

proc sort data = master_bk1b;
  by rownum page_num1;
run;

data master_bk1c;
  set master_bk1b;
  by rownum page_num1;
  level1 = lag(level);
run;

proc sql noprint;
  select min(rownum-1) into :rownum1 from master_bk1c
            where level = 2 and level1 = 2;
quit;

%put &rownum1;

* configure Running Records for VISITs;
data master_bk1d;
  set master_bk1c;

  if rownum = &rownum1 then nm2 = 'Running Records';
  if nm2 = 'Running Records' then do;  output; nm2 = ''; end;
  output;
run;

data master_bk2(keep=nm1 pg1 nm2 pg2 nm3 pg3 nm4 pg4);

  set master_bk1d;
   length nm1 nm3 nm4 $200. pg1 pg3 pg4 $10.;

   if level = 1 and nm2='' then do;
            nm1 = "&Study_Name."; pg1 = ',1';
      end;
   if level = 2 and page_num1 =. then do;
            nm3 = bookmark; pg3 = '';
      end;
   if level = 3 then do;
            nm4 = bookmark; pg4 = cats(',',put(page_num1,best.));
      end;
   if level = 2 and page_num1 ^=. and nm2^='Running Records' then do;
            nm4 = bookmark; pg4 = cats(',',put(page_num1,best.));
      end;
   if nm2='Running Records' then do;
            bookmark = ''; nm3 = nm2; nm2 = '';
      end;
run;

data master_bk3;
  retain nm1 pg1 nm2 pg2 nm3 pg3 nm4 pg4;
  set master_bk2;
run;
```

```
* configuring FORMS/DOMAINS;
data master_bk4(keep=nm3 nm4 pg4);
  set  master_bk3;
  if nm3 ^= '' or nm4^='';
  if pg3 = '';
run;

data master_bk5;
  set master_bk4;
  length _nm3 $200.;
  retain _nm3 ;
  if nm3^='' then _nm3 = nm3;
  else nm3 = _nm3;
  drop _nm3;
  if nm4 ^= '';
run;

proc sort data = master_bk5 out = master_bk6;
  by nm4 nm3 pg4;
run;

data master_bk7;
set master_bk6;
by nm4 nm3 pg4;
output;
if last.nm4 then do;
    call missing(of _all_);
    output;
    end;
run;

data master_bk8;
  set master_bk7;
  length _nm4 $200.;
  retain _nm4 ;
  if nm4^='' then _nm4 = nm4;
  else nm4 = _nm4;
  drop _nm4;
  length _pg4 $200.;
  retain _pg4 ;
  if pg4^='' then _pg4 = pg4;
  else pg4 = _pg4;
  drop _pg4;
run;

proc sort data = master_bk8 out = master_bk9;
  by nm4 nm3 pg4;
run;

data master_bk9a;
  set master_bk9;
  length nm5 $200.;
  nm5 = nm3;
  nm3 = nm4;
  nm4 = nm5;
  drop nm5;
```

```
  if nm4^='' then nm3 = '';
run;

* configuring Running Records for FORMS;
data master_bk9b;
  set master_bk3;
  if pg3 ^='' or nm2 ^= '' or nm1 ^= '';
  nm4 = nm3;
  pg4 = pg3;
  nm3 = '';
  pg3 = '';
  if nm2 ^= '' then nm2 = 'Forms';
  if nm1 ^= '' then do;
      nm3 = '_WITHOUT VISIT'; nm1 = ''; pg1='';
  end;
  if nm2 ^='' then ord = 1;
  if nm3 ^='' then ord = 2;
  if nm4 ^='' then do ord = 3; ord1 = 3; end;

run;

data master_bk9b;
  set master_bk9b;
    if nm3 = '_WITHOUT VISIT' then delete;
run;

proc sort data = master_bk9b;
  by ord ord1 nm4;
run;

data master_bk9c;
  set master_bk9b master_bk9a;
run;

* to order the visits;
data visit_ord;
  set master_bk3;
  if nm3 ^= '' and pg3 = '';
  vis_ord = _N_;
run;

data visit_ord (keep = nm3 vis_ord);
  set visit_ord;
  vis_ord = _N_ + 20;
run;

proc sql noprint;
  update master_bk9c a set ord1 = (select vis_ord from visit_ord b
                  where a.nm4 = b.nm3) where ord =.;
run;

data master_bk9d;
  set master_bk9c;
  by nm1 notsorted;
  if nm3 ^= '' then ord = 2;
  if ord in (3,.) then do; ord = 3; end;
  if first.nm1 then ord2 = 0.1;
```

```
  else if ord1=. then ord2 + 0.01;
run;

proc sort data = master_bk9d out = master_bk9e;
  by ord2 ord1;
run;

data master_bk9f;
  set master_bk9e;
  drop ord ord1 ord2;
run;

* finalizing the columns;
data master_final;
  retain nm1 pg1 nm2 pg2 nm3 pg3 nm4 pg4;
  set  master_bk3 master_bk9f;
  if pg3 = ',.' then pg3 = ',';
  if pg4 = ',.' then pg4 = ',';
  if index(nm3,',') then nm3 = tranwrd(nm3,',','');
  if index(nm4,',') then nm4 = tranwrd(nm4,',','');
  nm1 = strip(nm1);
  pg1 = strip(pg1);
  nm2 = strip(nm2);
  pg2 = strip(pg2);
  nm3 = strip(nm3);
  pg3 = strip(pg3);
  nm4 = strip(nm4);
  pg4 = strip(pg4);
run;

* exporting to CSV file;
proc export data=master_final
      outfile="&Files_Path.\output_bookmarks.csv" replace
      dbms=tab;
        putnames=no;
run;

%mend bookmrk;
******* End of Macro *******;

*** execute the macro ***;
option mprint mlogic symbolgen;

%bookmrk (  Study_Name = ,
            Files_Path = ,
            Input_Master_CSV_FileName = ,
            Input_Review_CSV_FileName = );
```