

Grooming of Coarse Data by Smart Listing Using SAS System

Soujanya Konda, inVentiv Health Clinical, Hyderabad, India

ABSTRACT

Organizations strive to deliver the quality database; one of the crucial and typical parts is data cleaning. Though word data cleanings sounds simpler it is very tedious to the clean entire data and provide the quality data. However; we can try new techniques to manage the data the effectively, efficiently to provide discrepant free data for further process. Data Management (DM) is to deliver a qualitative database to SAS Programming, Statistical Analysis teams for their analysis. DM team is to conduct reviews more often and repeatedly on same reports and to locate the discrepant data would take lots of time; find out the optimized process by creating smart listings in such a way that it identifies, newly entered, modified records on each and every fresh run. In turn; it avoids lot of manual effort, the various challenges, efficient techniques are thoroughly discussed further in this paper.

INTRODUCTION

Achieving a whole, better results out of clinical trials is very essential and difficult. Clinical trials include various teams and many efforts. Even a step of failure can fail the entire study.

Is collecting and entering the clinical data into database easy? Yes, it is. However, entering the accurate data and managing it efficiently to reach the goal of clean database is hard and time-consuming. It requires lots of manual effort because it is an ongoing process till the study ends.

Even after adopting new techniques—several programs or scheduled outputs by using various filtering options—using various filters and programmed reports, the reviewers need to clean the entire data. To overcome this manual intervention, you can use efficient techniques like Smart Listing that is generated using the SAS system.

INTRODUCTION TO SMART LISTINGS

Smart Listing plays a vital role in cleaning the data without using any filters. Reviewers can pull entire data and locate the newly entered, modified records along with reviewer comments on every fresh run of the program. This helps the reviewers to locate new records and their previous reviewed comments. In turn, it helps to track the same on each run of the report.

In this paper, let us discuss the following information on Smart Listing:

- Basic requirements
- Flow chart and clear description of smart listings
- Pictorial representation of sample outputs
- Restrictions
- Advancements
- Benefits

BASIC REQUIREMENTS

The following are the basic requirements to work with Smart Listing:

- A common drive
- Read and Write access to this drive for all required users, such as programmers and reviewers

Programmers can program the Smart Listing and reviewers can execute the program. Reviewers can then try to modify or update the report.

The output in excel format needs to be read into SAS system for next execution of the program.

FLOW CHART

Figure 1 illustrates the flow chart of smart listings, which explains the process right from programming to review.

Step 1: Connect to the Database from the SAS system.

Step 2: Copy the required panels from the database like AE, CM etc.

Step 3: Program panels as required and include all the variables along with adding desired columns to capture date, initials, comments, required formats, lengths, labels etc.

Step 4: Generate the output in desired format, preferably excel because it would be most convenient or desired output for reviewers.

Step 5: Output destination is directed to common drive.

Step 6: Reviewer completes the review and enters the comments, initials etc.

Step 7: Same output becomes one of the input sheets for the next run.

Step 8: Compare this excel input with the current database data to track the new records.

Step 9: Program in such a way that the comments and initials are carry forwarded for each and every run.

Figure 1: Smart Listing Flow Chart

Figure 2 illustrates the brief pictorial representation of flow that explains the flowchart.

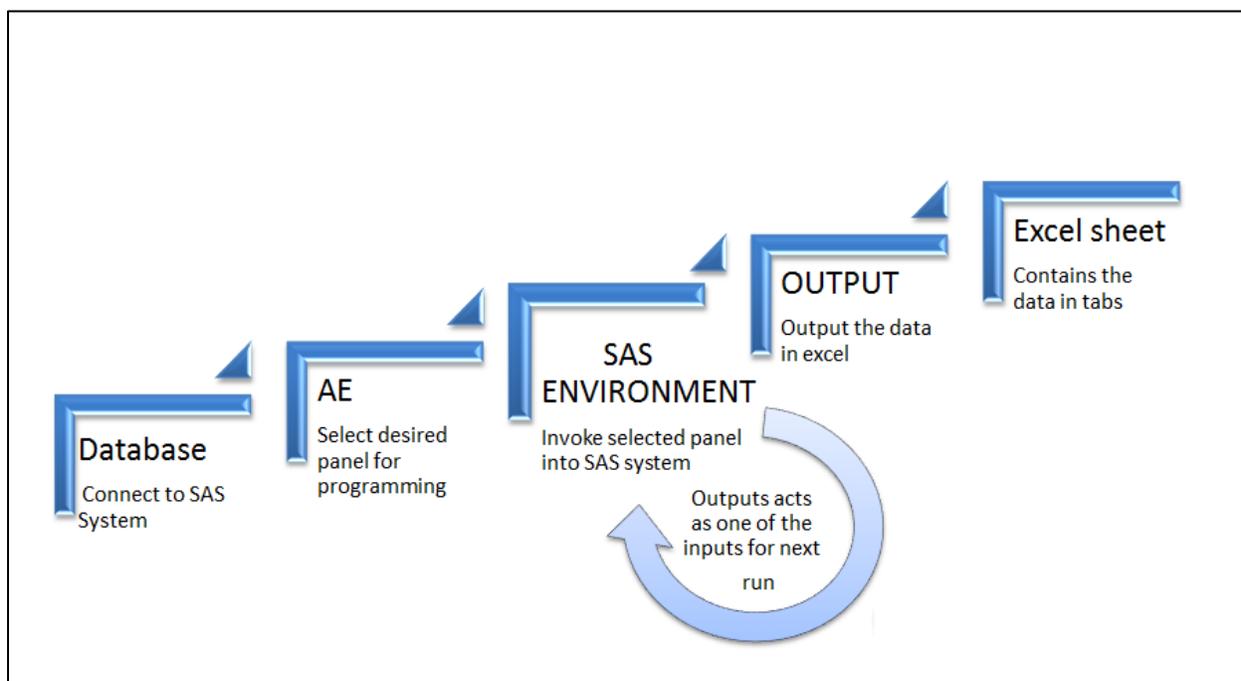


Figure 2: Pictorial Presentation of Smart Listing Process

PICTORIAL REPRESENTATION OF SAMPLE OUTPUTS

Display 1 shows the initial or first output samples.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	
1	STAT	INVID	SUBJVISID	PAGE	PAGE	TRTIC	TRTNAME	TRTTERM	DOSE	DOSEU_JAEID	AETE	IFU_D	TRTS	TRTE	TRTF	ENTRY_D	SIGN_DATE	COMMENTS					
2	NEW	100	1000	4	1	1												7/9/2007					
3	NEW	100	1001	4	1	1	1	TOPICRE/ANTIFUNGALSFOR	TOPICALUSE		1	Adver	28MAI	30MAR	2008			7/3/2009					
4	NEW	100	1001	4	1	1	2	XYLOCAINE/LIDOCAINE			2	Adver	13JUN	13JUN	2008			7/3/2009					
5	NEW	100	1001	4	1	1	3	CICATRYL/ CICATRYL/01314301/			2	Adver	20JUN	26JUN	2008			7/3/2009					
6	NEW	100	1001	5	1	2	1	DIASEPTYL/CHLORHEXIDINE			5	KAPO	Adver	19DEC	03JAN	2009			2/16/2009				
7	NEW	100	1001	5	1	2	2	COLD-CRE/ COLD/01040401/			6	ECZEI	Adver	20JAN	20APR	2009			2/16/2009				
8	NEW	100	1001	5	1	2	3	DEXERYL/ DEXERYL			6	ECZEI	Adver	20JAN	20APR	2009			2/16/2009				
9	NEW	100	1001	5	1	2	4	LUCOMIST/ ACETYLCYSTEINE			11	Adver	31AUC	09SEP	2009			11/29/2010					
10	NEW	100	1001	5	1	2	5	AMOXICILLI/ AMOXICILLIN			12	Adver	09OCT	17OCT	2009			11/29/2010					
11	NEW	100	1001	5	1	2	6	PREDNISO/ PREDNISOLONE			12	Adver	10OCT	15OCT	2009			11/29/2010					
12	NEW	100	1001	5	1	2	7	CELESTEN/ BETAMETHASONE			9	Adver	01DEC	13DEC	2009			11/29/2010					
13	NEW	100	1001	5	1	2	8	GENTALLIN/ GENTAMICIN			9	Adver	01DEC	13DEC	2009			11/29/2010					
14	NEW	100	1001	6	1	3	1	FLECTOR/ DICLOFENAC			3	Adver	26NOV	30NOV	2008			1/15/2010					
15	NEW	100	1001	6	1	3	2	AERILUS5M/ DESLOBATADINE			4	Adver	16DEC	19DEC	2008			1/15/2010					

Display 1: Initial or First Output

ADVANCEMENTS

The following steps of the Smart Listing process (illustrated in Figure 1) can be automated:

- Importing of excel sheets can be automated using macros.
- Report generation can be auto-scheduled.
- SAS reports can be uploaded into the J-Review tool.
 - Reviewers can directly retrieve the report in a single click, which avoids the lots of confusion.
 - Reviewers require less or no additional training.

RESTRICTIONS

- Naming conventions of excel sheet should be maintained.
- Data types of excel columns should not be modified because this would cause errors and warnings, such as:
 - Data types are different.
 - Formats are different.
 - Excel should be closed before importing the excel sheet into SAS system etc.
- Reviewers should modify only required fields and should not modify other fields, because it would populate as "NEW" records though review is performed.

CONCLUSION

Smart listing:

- Is beneficial, because of the reduced turnaround time, easy tracking, and good and clear documentation.
- Helps to track new, updated, modified records on each every run of the report
- Is easy to handle the huge report
- Helps the reviewers and increases the scope of the report
- Helps attain the quality database with very less turnaround time

You do not need to maintain separate documentation for the issues because information is present in single report. You can easily assess the metrics of report and calculate the percentage of the database cleaned and completed. In turn, you can plan and assign the resources because the initial round would take more time when compared to next rounds.

REFERENCES

- Web site: SAS Global Forum 2009, Multiple authors, Available at <http://support.sas.com/resources/papers/proceedings09/TOC.html>

ACKNOWLEDGMENTS

I take this opportunity to thank Naveen Kodavoor; Associate Director at inVentiv Health Clinical, whose support and guidance encouraged to write this paper. A special thanks to Srivalli Konda, Sudarshan Reddy, Suresh Kumar and Vishnu Bapunaidu who gave me helping hand in meticulously organizing these words.

RECOMMENDED READING

- Base SAS® Procedures Guide
- SAS® For Dummies®

CONTACT INFORMATION

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

Name: Soujanya Konda
Enterprise: inVentiv Health Clinical
Address: Kondapur
City, State ZIP: Hyderabad – 500084, Andhra Pradesh, India
Work Phone: +91-9177489341
E-mail: soujanya.konda@inventivhealth.com or soujitheuniq@gmail.com
Web: <http://www.inVentivHealthclinical.com>

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.