



# SAS® GLOBAL FORUM 2016



IMAGINE. CREATE. INNOVATE.

## 3N Validation to Validate PROC COMPARE Output

---

Amarnath Vijayarangan  
Emmes Services Pvt Ltd, India

#SASGF







# 3N Validation to Validate PROC COMPARE Output

Amarnath Vijayarangan

Emmes Services Pvt Ltd, India

## Abstract

- In the clinical research world, various validation methods are available to confirm data accuracy.
- PROC COMPARE is one of the SAS® procedures used to demonstrate a perfect match between production and validation data sets.
- In the current practice, the programmer must manually review the output file to ensure an exact match.
- The proposed approach programmatically validates the output of PROC COMPARE in all the programs
- Generates an HTML output with a Pass/Fail status flag for each output file

## PROC COMPARE: Expected Output

Data Set Summary					
Dataset	Created	Modified	NVar	NObs	Label
SASHELP.CLASS	24MAY11:13:52:29	24MAY11:13:52:29	5	19	Student Data
SASHELP.CLASS	24MAY11:13:52:29	24MAY11:13:52:29	5	19	Student Data

Variables Summary	
Number of Variables in Common: 5.	

Observation Summary		
Observation	Base	Compare
First Obs	1	1
Last Obs	19	19

Number of Observations in Common: 19.  
Total Number of Observations Read from SASHELP.CLASS: 19.  
Total Number of Observations Read from SASHELP.CLASS: 19.

Number of Observations with Some Compared Variables Unequal: 0.  
Number of Observations with All Compared Variables Equal: 19.

NOTE: No unequal values were found. All values compared are exactly equal.

## PROC COMPARE: Misleading Output

Data Set Summary					
Dataset	Created	Modified	NVar	NObs	Label
SASHELP.CLASS	24MAY11:13:52:29	24MAY11:13:52:29	4	19	Student Data
SASHELP.CLASS	24MAY11:13:52:29	24MAY11:13:52:29	5	19	Student Data

Variables Summary	
Number of Variables in Common: 4. Number of Variables in SASHELP.CLASS but not in SASHELP.CLASS: 1.	

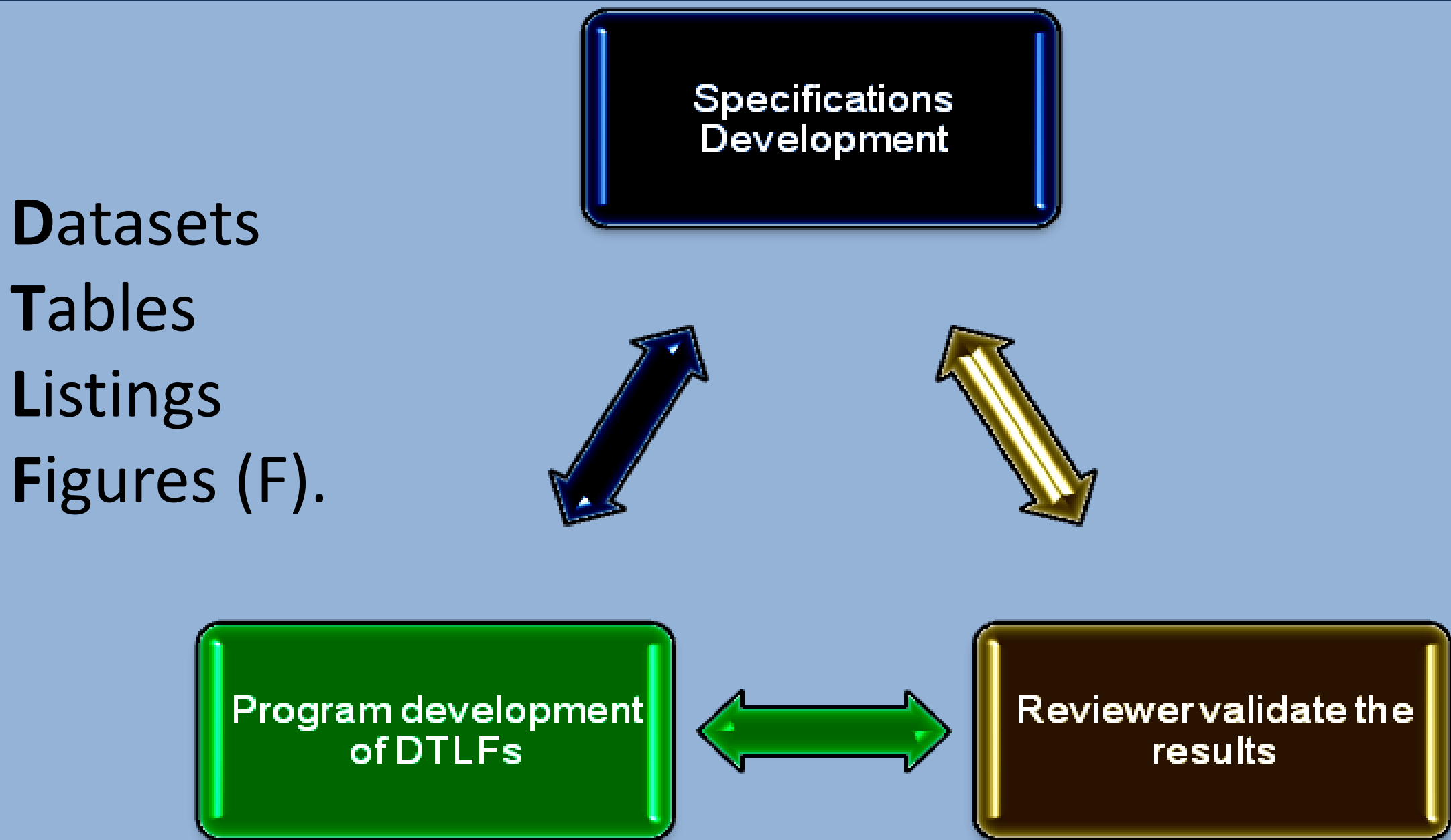
Observation Summary		
Observation	Base	Compare
First Obs	1	1
Last Obs	19	19

Number of Observations in Common: 19.  
Total Number of Observations Read from SASHELP.CLASS: 19.  
Total Number of Observations Read from SASHELP.CLASS: 19.

Number of Observations with Some Compared Variables Unequal: 0.  
Number of Observations with All Compared Variables Equal: 19.

NOTE: No unequal values were found. All values compared are exactly equal.

## Program Life Cycle



## Challenges

Manual review of 100s of outputs is tedious, time consuming and error prone	PROC COMPARE can also be misleading if programmer is not cautious	Supervisor should await for the status update from the programmers
---	---	--

## Solution : %threeN

- Programmatically validates PROC COMPARE output
- Generates a html output file with ‘Pass / Fail’ status flag for each output file along with important base and compared datasets attributes
- The status is flagged as Pass whenever the output file meets the following 3N criteria.
- NOTE: No unequal values were found. All values compared are exactly equal.
- Number of observations in Base and Compared datasets are equal.
- Number of variables in Base and Compared datasets are equal.

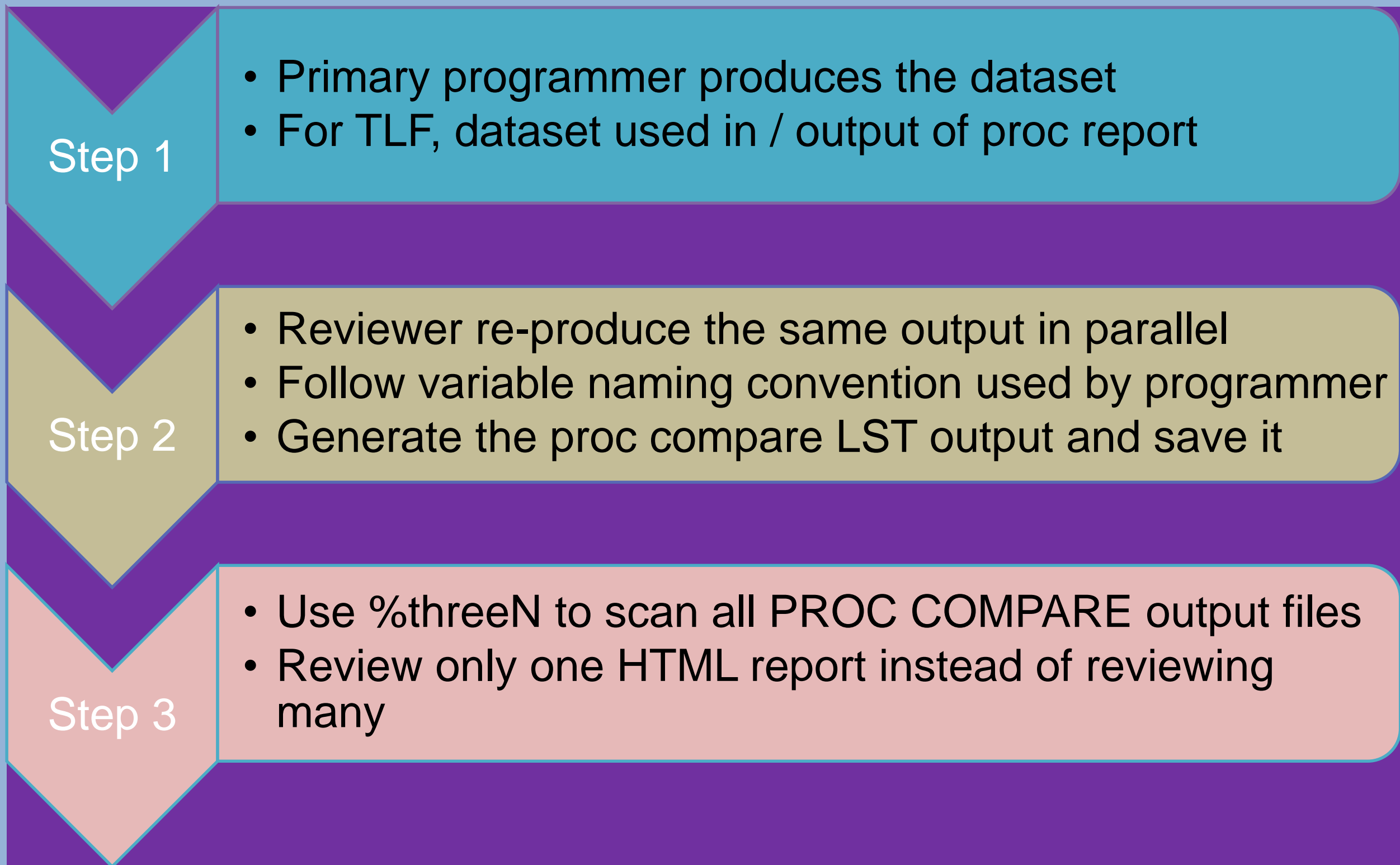


# 3N Validation to Validate PROC COMPARE Output

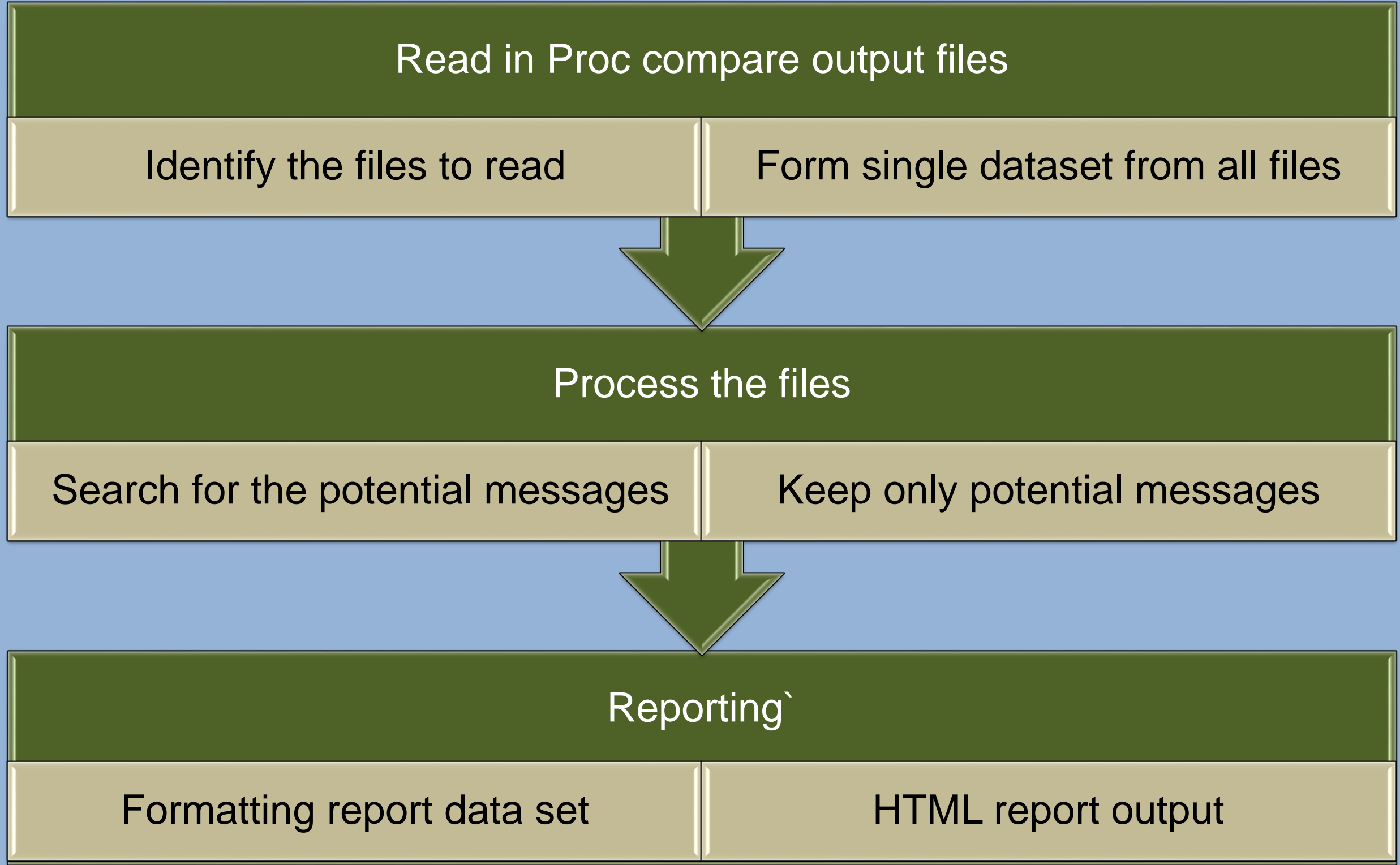
Amarnath Vijayarangan

Emmes Services Pvt Ltd, India

## Process Flow



## %threeN Macro



## Conclusion

### Advantages

- Efficient approach to review numerous outputs
- Very useful while re-running of programs
- 100% accuracy in reviewing Proc compare LST output
- Can be plugged with other programs
- Almost 90% of review time is saved

### Limitations

- Requires double programming
- Programmer and reviewer should follow the same variable naming convention
- Does not check the cosmetics related items

## Output

C:\3N Validation Report

3N Validation Report															
		Common Attributes		Base Data Set Attributes						Compare Data Set Attributes					
Output File Name	Status	# of Vars	# of Obs	Name	# of Vars	# of Obs	Label	# of Vars Not In Compare	# of Obs Not In Compare	Name	# of Vars	# of Obs	Label	# of Vars Not In Base	# of Obs Not In Base
class	Fail	3	9	SASHELP.CLASS	4	19	Student Data	1		WORK.CLASS	4	10		1	1
class1	Pass	5	19	SASHELP.CLASS	5	19	Student Data			WORK.CLASSFULL	5	19			
class2	Fail	4	19	SASHELP.CLASS	4	19	Student Data			WORK.CLASSFULL	5	19		1	





# SAS<sup>®</sup> GLOBAL FORUM 2016

IMAGINE. CREATE. INNOVATE.

LAS VEGAS | APRIL 18-21

#SASGF