PharmaSUG 2017 - Paper QT16

Common Mistakes by Programmers & Remedies

Venkata Sairam Veeramalla, Technical Manager, GCE Solutions Pvt. Ltd, India

ABSTRACT

A SAS® programmer is a critical entity of any company. It is of our knowledge that a programmer is responsible for providing support for statisticians and statistical programming expertise for the company. Simultaneously, programmers are also occupied with varied tasks, for instance, writing programs to create SDTM, ADaM and TLFs, weekly and monthly meetings, one to ones, completing time sheets and meeting the tight timelines. In the midst of this busy schedule, it is usual for programmers to overlook few basic concepts and perform mistakes.

I would now share similar mistakes which I have observed during my experience working as a programmer, validator and as reviewer and would also like to provide easy solutions to rectify those mistakes. By doing so, we can ultimately deliver good quality deliverables and it also increases First Time Right (FTR) numbers which helps boost the performance of both the company and the individual as well.

INTRODUCTION

The job of a SAS® programmer can be a challenging and rewarding equally providing opportunities for technically minded people that would like to work within drug development process. As a SAS® programmer, we would typically work on development of SAS® code which creates Analysis Datasets, Tables, Listings and Figures (TLFs) and electronic submission packages, dealing with differing personalities and opinions, client requests and delivering to stretched timelines on an on-going basis is the most challenging part of our profession. It is very natural to commit certain mistakes during this process.

In this presentation, I would like to focus on these mistakes and would also come up with certain remedies that would help to reduce the mistakes count and save time.

1. COMMON PRESENTATION MISTAKES IN TLFS

Example 1: Presentation of Titles, Headings, Labels ought to be in an appropriate manner for accuracy and for better interpretation.

The labels of columns should be displayed correctly as “Inclusion” instead of INCLUSION, Dlt should be “DLT” and YES/NO values should be “Yes/No”:

Study 003 OPEN LABEL

Listing 16.03.01-12 (Page 1 of 1)
Listing of Dose Limiting Toxicities (DLTs) evaluation
Safety Set

Schedule: 1/ Dose Level: 50 mg/m² TID

Subject (Age/Sex/ Race)	INCLUSION Date	Date of First Dose (day)	Date of Last Dose (day)	Dlt Evaluable? (YES/NO)	Reason why DLT is not Evaluable	Patient Experienced Dlt? (Yes/No)
0004-052-/ 30001 (62/F)	17DEC2010 (-3)	20DEC2010 (1)	26MAY2011 (158)	Yes		No
0004-052-/ 30002 (61/F)	28DEC2010 (-6)	03JAN2011 (1)	14JUL2011 (193)	Yes		No
0004-052-/ 30003 (56/F)	07JAN2011 (-3)	10JAN2011 (1)	14JUL2011 (186)	Yes		No

Output 1. Dose Limiting Toxicities Evaluation Report

Example 2. No Subscript was found in the below example. Having a footnote for any given subscript is very essential for development, validation and review:

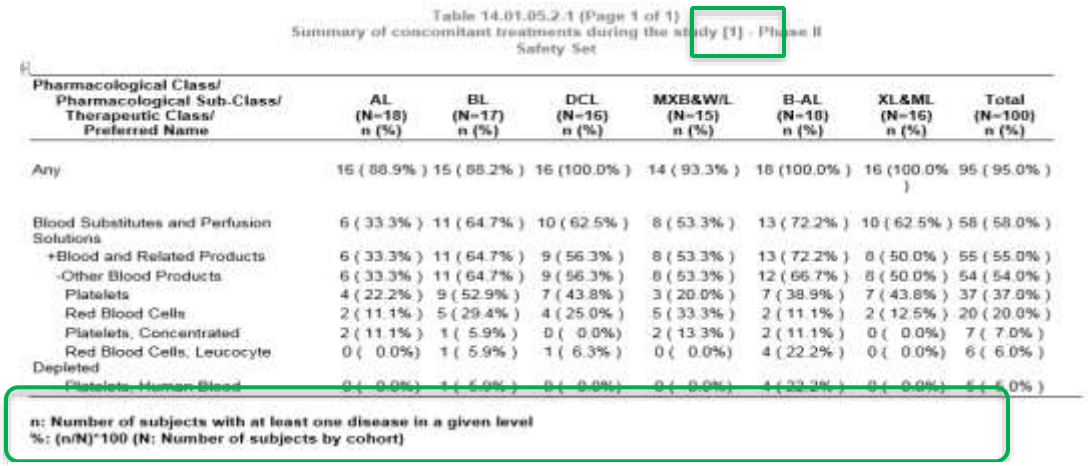


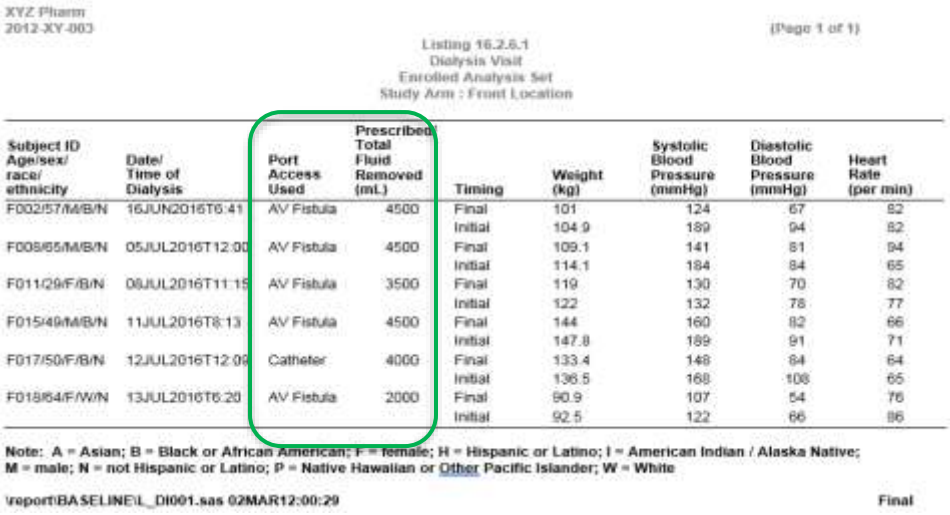
Table 14.01.05.2.1 (Page 1 of 1)
Summary of concomitant treatments during the study [1] - Phase II
Safety Set

Pharmacological Class/ Pharmacological Sub-Class/ Therapeutic Class/ Preferred Name	AL (N=18) n (%)	BL (N=17) n (%)	DCL (N=16) n (%)	MXB&W/L (N=15) n (%)	B-AL (N=18) n (%)	XL&ML (N=16) n (%)	Total (N=100) n (%)
Any	16 (88.9%)	15 (88.2%)	16 (100.0%)	14 (93.3%)	18 (100.0%)	16 (100.0%)	95 (95.0%)
Blood Substitutes and Perfusion Solutions	6 (33.3%)	11 (64.7%)	10 (62.5%)	8 (53.3%)	13 (72.2%)	10 (62.5%)	58 (58.0%)
+Blood and Related Products	6 (33.3%)	11 (64.7%)	9 (56.3%)	8 (53.3%)	13 (72.2%)	8 (50.0%)	55 (55.0%)
-Other Blood Products	6 (33.3%)	11 (64.7%)	9 (56.3%)	8 (53.3%)	12 (66.7%)	8 (50.0%)	54 (54.0%)
Platelets	4 (22.2%)	9 (52.9%)	7 (43.8%)	3 (20.0%)	7 (38.9%)	7 (43.8%)	37 (37.0%)
Red Blood Cells	2 (11.1%)	5 (29.4%)	4 (25.0%)	5 (33.3%)	2 (11.1%)	2 (12.5%)	20 (20.0%)
Platelets, Concentrated	2 (11.1%)	1 (5.9%)	0 (0.0%)	2 (13.3%)	2 (11.1%)	0 (0.0%)	7 (7.0%)
Red Blood Cells, Leucocyte Depleted	0 (0.0%)	1 (5.9%)	1 (6.3%)	0 (0.0%)	4 (22.2%)	0 (0.0%)	6 (6.0%)
Platelets, Human Blood	0 (0.0%)	1 (5.9%)	0 (0.0%)	0 (0.0%)	4 (22.2%)	0 (0.0%)	5 (5.0%)

n: Number of subjects with at least one disease in a given level
%: (n/N)*100 (N: Number of subjects by cohort)

Output 2. Summary of Concomitant Treatment Report

Example 3. Highlighted part of the below example shows that there are two extra columns. This may happen if a programmer is using an old version of SAP instead of the latest version where the columns have been eliminated lately:



XYZ Pharm
2012-XY-003

(Page 1 of 1)

Listing 16.2.6.1
Dialysis Visit
Enrolled Analysis Set
Study Arm : Front Location

Subject ID Age/sex/ race/ ethnicity	Date/ Time of Dialysis	Port Access Used	Prescribed Total Fluid Removed (mL)	Timing	Weight (kg)	Systolic Blood Pressure (mmHg)	Diastolic Blood Pressure (mmHg)	Heart Rate (per min)
F002/57/M/B/N	16JUN2016T6:41	AV Fistula	4500	Final	101	124	67	82
				Initial	104.9	189	94	82
F008/65/M/B/N	05JUL2016T12:00	AV Fistula	4500	Final	109.1	141	81	94
				Initial	114.1	184	84	65
F011/29/F/B/N	08JUL2016T11:15	AV Fistula	3500	Final	119	130	70	82
				Initial	122	132	78	77
F015/49/M/B/N	11JUL2016T8:13	AV Fistula	4500	Final	144	160	82	66
				Initial	147.8	189	91	71
F017/50/F/B/N	12JUL2016T12:09	Catheter	4000	Final	133.4	148	84	64
				Initial	136.5	168	108	65
F018/64/F/W/N	13JUL2016T6:20	AV Fistula	2000	Final	90.9	107	54	76
				Initial	92.5	122	66	86

Note: A = Asian; B = Black or African American; F = female; H = Hispanic or Latino; I = American Indian / Alaska Native; M = male; N = not Hispanic or Latino; P = Native Hawaiian or Other Pacific Islander; W = White

\report\BASELINE\L_DI001.sas 02MAR12:00:29

Final

Output 3. Dialysis Visit Report

Example 4: It can be observed that the age value 15yrs has been included though the inclusion criteria requires age above 18 yrs

Reason : Data entry error or Wrong age derivation

Study D1020C00009 Final TFL

Listing 12.2.4.1 Demographic characteristics (all subjects)

Centre	ECode	Treatment	Age at randomization (Years)	Sex	Weight (kg)	Height (cm)	BMI (kg/m²)	Race
582	E0102001	Lower dose AZD1656	68	Male	75	158	30.0	White
583	E0103002	Higher dose AZD1656	48	Female	68	155	28.3	White
584	E0104001	Glipizide	51	Female	96	175	31.3	White
	E0104002	Glipizide	56	Female	60	157	24.3	White
	E0104004	Higher dose AZD1656	63	Male	90	169	31.5	White
	E0104008	Fixed dose AZD1656 40 mg	61	Female	92	157	37.3	White
	E0104013	Placebo	15	Male	90	185	26.3	White
586	E0106001	Glipizide	50	Male	105	178	33.1	White
	E0106002	Lower dose AZD1656	62	Male		166	31.9	White
	E0106005	Fixed dose AZD1656 40 mg	63	Ma				
587	E0107002	Higher dose AZD1656	47	Ma				

Age values are very different from other values

Output 4. Demographic Report

Example 5. In the example below, the order of visit values are incorrect.

Reason: Programmer used character variable instead of numeric variable while sorting the data.

Listing 11 Clinical Chemistry: By Visit Comparisons Part I

Patient ID/Site	Sex/ Age	Lab ID	Visit	Date:Time	Albumin (g/L) (CTC Grade)	Alkaline Phosphatase (U/L) (CTC Grade)	ALT (U/L) (CTC Grade)	AST (U/L) (CTC Grade)	Calcium (mmol/L) (CTC Grade)
1001/1101	F/63	1101LAB01	SCREENING (DAYS -14 TO 0)	22JAN2007 :10:47	ND	139 H (1)	43	35 H (1)	2.32
			CYCLE 1 DAY 1	29JAN2007 :10:29	39.00	129 H (1)	46	40 H (1)	ND
			CYCLE 1 DAY 15	12FEB2007 :10:43	38.00	123 H (1)	65 H (1)	61 H (1)	2.27
			CYCLE 1 DAY 8	05FEB2007 :11:42	37.00 L (1)	118 H (1)	38	36 H (1)	2.20
			CYCLE 2 DAY 1	19FEB2007 :11:42	39.00	132 H (1)	28	29	2.25
			CYCLE 2 DAY 15	05MAR2007 :09:26	39.00	116 H (1)	ND	ND	2.27
			CYCLE 2 DAY 8	26FEB2007 :12:25	41.00	122 H (1)	42	42 H (1)	2.25
			CYCLE 3 DAY 1	12MAR2007 :09:22	40.00	112 H (1)	23	28	2.25
			CYCLE 3 DAY 8	19MAR2007 :09:22	ND	ND	ND	ND	ND

Output 5. Clinical Chemistry Report

Example 6: In this example, the table shows extraction date as 02JAN2014 and date of run as 19JUN2014 which indicates that Extraction date was not updated:

Subject Disposition	Statistic(s)	ITT: N = 455	SP: N = 451
Not Treated	n (%)	4 (0.9)	0
Treatment phase			
Ongoing	n (%)	187 (41.1)	187 (41.5)
Discontinued	n (%)	264 (58.0)	264 (58.5)
Primary reason for discontinuation from study treatment			
Death	n (%)	40 (8.8)	40 (8.9)
Adverse event	n (%)	29 (6.4)	29 (6.4)
Pregnancy	n (%)	0	0
Progressive disease	n (%)	165 (36.3)	165 (36.6)
Lack of efficacy	n (%)	1 (0.2)	1 (0.2)
Recovery	n (%)	0	0
Withdrew consent	n (%)	14 (3.1)	14 (3.1)
Non-compliance with study drug	n (%)	0	0
Lost to follow-up	n (%)	1 (0.2)	1 (0.2)
Study terminated by sponsor	n (%)	0	0
Transition to commercially available treatment	n (%)	1 (0.2)	1 (0.2)
Protocol violation	n (%)	0	0
Other	n (%)	12 (2.6)	12 (2.7)

ITT = Intention to treat population; SP = Safety population.

Safety\t_disp.SAS
Data source: ADSL
Data extraction date: 02JAN2014
Run time: 19JUN2014/12:4

Output 6. Disposition Report

2. COMMON PROGRAMMING MISTAKES IN TLFS

Example 1: In the below listing, under the column, "Treatment Given" a truncation is noted.

Reason: Programmer forgot to supply length statement while creating this Treatment Given variable with No/Yes values:

Listing 6 Adverse Events Part II

Patient ID/Site	System Organ Class	Preferred Term	Onset Cycle	Start Date/ Stop Date	SAE	Study Drug Action Taken	Treatment Given
1001/1101	BLOOD AND LYMPHATIC SYSTEM DISORDERS	*NEUTROPENIA	1	05FEB2007/ 19FEB2007	No	Dose Delayed	No
	BLOOD AND LYMPHATIC SYSTEM DISORDERS	*LEUKOPENIA	1	05FEB2007/ 12FEB2007	No	No Action Taken/Dose Not Changed	No
	BLOOD AND LYMPHATIC SYSTEM DISORDERS	*MONOCYTOPENIA	1	05FEB2007/ 12FEB2007	No	No Action Taken/Dose Not Changed	No
	NERVOUS SYSTEM DISORDERS	*PARAESTHESIA	3	12MAR2007/ 30JUL2007	No	No Action Taken/Dose Not Changed	No
	GENERAL DISORDERS AND ADMINISTRATION SITE CONDITIONS	*INFLUENZA LIKE ILLNESS	3	13MAR2007/ 19MAR2007	No	No Action Taken/Dose Not Changed	Ye
	GENERAL DISORDERS AND ADMINISTRATION SITE CONDITIONS	*FATIGUE	4	10APR2007/ 30JUL2007	No	No Action Taken/Dose Not Changed	No

Output 7. Adverse Events Report

Example 2: Based on the title, only adverse events (PT =>5%) should display in the table

Reason: Programmer did not subset the dataset correctly:

Table 5.6.28
Adverse Events Reported (PT =>5%) During the Treatment Period
All Treated Subjects

System Organ Class (SOC) (%) Preferred Term (PT) (%)	Treatment A N=104
Total Subjects with AE	89 (21.2)
Infections and Infestations	19(18.3)
Urinary Tract Infection	2(1.9)
Bronchitis	2(1.9)
Bursitis Infective	1(1.0)
Herpes Zoster	5(4.8)
Influenza	1(1.0)
Nasopharyngitis	3(2.9)
Respiratory Tract Infection	1(1.0)
Rhinitis	1(1.0)
Tonsillitis	3(2.9)
Vaginal Infection	1(1.0)
Vulvovaginal Mycotic Infection	1(1.0)
Gastrointestinal Disorders	25 (24.0)
Nausea	18 (17.3)
Abdominal Pain Upper	5 (4.8)
Haematochezia	12 (11.5)
Paraesthesia Oral	4 (3.8)
Nervous System Disorders	4 (3.8)
Headache	2 (1.9)
Burning Sensation	1 (1.0)
Dizziness	1 (1.0)

Output 8. Adverse Events with >= 5% PTs Report

Example 3: There are 18 patients under Placebo, but under Sex column it can be observed that there are 9 Males and 10 Females, which is more than 18 patients and improper percentages:

Study XYZ

Table 14.01.2-1 (Page 1 of 1)
Summary of demographics
Included Set

		Placebo (N=18)	Trt A (N=17)	Trt B (N=16)	Total (N=51)
Remaining report is not displayed					
Sex	Total	18	17	16	
Male	n (%)	9 (50.0%)	7 (41.2%)	10 (62.5%)	26 (51.0%)
Female	n (%)	10 (55.5%)	10 (58.8%)	6 (37.5%)	26 (51.0%)

Output 9. Summary of Demographics Report

Example 4: In the below listing the first two columns were not grouped:

Listing 3 Prior and Concomitant Medications

Patient ID/Site(1)	Prior Med or Conmed	WHO Preferred Term	Drug Name	Start Date/ Stop Date	Total Daily Dose	Route	Reason for Use
1001/1101	Con	AMOXICILLIN	AMOXICILLINE	27FEB2007/ 08MAR2007	3 Other	Oral	Adverse experience: DENTAL EXTRACTION
1001/1101	Con	IBUPROFEN	IBUPROFEN	27FEB2007/ 08MAR2007	600 mg	Oral	Adverse experience: DENTAL EXTRACTION
1001/1101	Con	PARACETAMOL	DAFALGAN	13MAR2007/ 19MAR2007	500 mg	Oral	Adverse experience: FLU LIKE SYNDROM
				19MAR2007/ ONGOING	500 mg	Oral	Adverse experience: BREAST PAIN
			PARACETAMOL	27FEB2007/ 08MAR2007	500 mg	Oral	Adverse experience: DENTAL EXTRACTION
1001/1101	Con	POVIDONE-IODINE	ISOBETADINE BUCCAL	27FEB2007/ 08MAR2007	15 cc/mL	Oral	Adverse experience: DENTAL EXTRACTION
1001/1101	Con	TRAMADOL HYDROCHLORIDE	CONTRAMAL	10APR2007/ ONGOING	10 drops	Oral	Adverse experience: BREAST PAIN

Output 10. Prior and Concomitant Medication Report

Example 5: When we check the below population table, the numbers look weird. ITT set has less number of patients when compared to PP set, which is not the case usually.

Page 1 of 1

Table 14.1.1
Subject Population Summary
All subjects

	xyz 5mg	xyz 10 mg	xyz 15 mg	xyz 20 mg	xyz 25 mg	Total
Number of Subjects Screened						835
Number of Subjects Randomised	87	93	91	89	88	448
Number of Subjects in Safety Set	87 (100.0%)	93 (100.0%)	91 (100.0%)	88 (98.9%)	88 (100.0%)	447 (99.8%)
Number of Subjects in ITT Set	77 (88.5%)	82 (88.2%)	81 (89.0%)	86 (97.7%)	85 (96.6%)	411 (91.7%)
Number of Subjects in PP Set	78 (89.7%)	81 (87.1%)	86 (94.5%)	87 (98.9%)	69 (78.4%)	401 (89.5%)

Source: Listing 16.2.1.4

Output 11. Population Report

Example 6: Total number of patients in ABC treatment are 18 and age category shows 14+5=19 patients.

Reason: The programmer has forgotten to write defensive coding like if age greater than dot (.) and Median and Max values were interchanged

Study 001 OPEN LABEL

Table 14.01.1-3 (Page 1 of 1)
Summary of demographics
Safety Set

		ABC (N=18)	BCL (N=17)	XCL (N=16)	BCL&M/LL (N=15)	L-CL (N=18)	CL&LT (N=16)	Total (N=100)
Age (years)	Total	18	17	16	15	18	16	100
	Mean ± SD	59.7 ± 14.2	61.8 ± 13.4	70.9 ± 6.6	66.8 ± 8.8	60.7 ± 9.4	68.2 ± 8.8	64.5 ± 11.2
	Median	83.0	85.0	83.0	78.0	78.0	80.0	85.5
	Min ; Max	32 ; 61	38 ; 63	59 ; 69	52 ; 69	47 ; 61	46 ; 68	32 ; 66
<=65 years	n (%)	14 (72.2%)	9 (52.9%)	3 (18.8%)	5 (33.3%)	12 (66.7%)	6 (37.5%)	48 (48.0%)
> 65 years	n (%)	5 (27.8%)	8 (47.1%)	13 (81.3%)	10 (66.7%)	6 (33.3%)	10 (62.5%)	52 (52.0%)

Output 12. Demographic Report

Example 7: The below adverse event table shows that total number of patients are 18 under disease category but we can observe from the 1st row, the number of subjects affected with at least 1 adverse event is 20, which is an error.

Table 14.03.05.2-3 (Page 1 of 1)
Treatment-emergent adverse events (TEAEs) by system organ class and preferred term - Phase II
Safety Set

Disease category: AL (N= 18)

System Organ Class Preferred Term	# Subjects Affected	# of Subjects At Risk	# of Non-serious Events	# Subjects Affected by Related Non-serious Events	# of Related Non-serious Events
Subjects having at least 1 adverse event	20	18	267	18	219
Blood and lymphatic system disorders	13	18	177	13	164
Thrombocytopenia	12	18	128	12	117
Anaemia	5	18	19	4	18
Neutropenia	4	18	30	3	29

Adverse events that started or worsened from the date of the first dose of study drug up to 30 days after the date last dose of study drug and all possibly related or related AEs are considered Treatment Emergent (TEAE). Subjects with multiple events for a given preferred term or system organ class are counted once only under each referred term or system organ class, respectively. If a subject had multiple events for a given preferred term or system organ class, all events will be counted for the number of events system organ class, respectively. Adverse events are coded by MedDRA Version 18.0.

Output 13. Treatment Emergent Adverse Event Report

Example 8: In the below listing, we can note that the "Last dose date" is 10DEC2015 and "Study Termination date" is 02NOV2015. Study Termination date should always be later than Last dose date.

Reason: It could be data issue or the programmer might have picked the wrong variable while working on this report:

Study BAC1023 OPEN LABEL

Listing 16.03-1.1 (Page 1 of 1)
Safety Set

Subject/ (Age/Sex)	Histology	Last dose date (Study day)	Study termination date (Study day)	Death date (Study day)
00039/ (68/F)	Extranodal Marginal Zone B-Cell Lymphoma of Malt Type	14JAN2016/ (1746)		
00201/ (49/M)	Follicular Lymphoma	10DEC2015/ (2350)	02NOV2015/ (2291)	
01115/ (53/F)	Follicular Lymphoma	19MAR2015/ (1240)	30MAR2015/ (1241)	
01118/ (41/F)	Follicular Lymphoma	12NOV2015/ (1443)	23NOV2015/ (1451)	
01147/ (46/F)	Follicular Lymphoma	20SEP2015/ (1332)	18NOV2015/ (1383)	
01160/ (56/F)	Follicular Lymphoma	23JUL2015/ (1216)	16AUG2015/ (1240)	
01161/ (75/M)	Marginal Zone B Cell Lymphoma	11OCT2015/ (1282)		

Output 14. Last dose and Study Termination Report

REMEDIES:

1. Cross verify whether all the labels are displayed correctly or not before submitting for validation.
2. Check for superscripts & subscripts and corresponding information, also make sure no extra spaces before submitting for validation
3. Be clear in communication, always use the final version of study documents
4. Be clear with all documents (i.e. Protocol, SAP, CRF, all specs) and requirements before deriving the variables
5. Sort and group the data with correct numeric variables
6. Update the date of extraction macro variable on daily basis (when we got new data form CDM)
7. Need to confirm the display of decimal places for derived variables with stat if not mentioned in SAP or study docs.
8. Always use length statement while creating new variables
9. Review the report to check if any column values are truncating
10. Read and understand titles & footnotes in each report before subset with right variables
11. Give a glance and check manually to see the percentages are matching correctly before submitting for validation or client
12. Double check the sub-setting condition.

13. Check for “Are we using correct **macro variables** to create percentages”
14. Use **defensive** coding to avoid losing the data.
15. Check the study documents and raise issue with DM:

CONCLUSION

I hope at least some of the programmers could relate to this presentation as it is very common for most of us to commit these above said mistakes during our routine programming, by taking extra and being a little vigilant during programming could avoid most of these mistakes and can help us achieve our desired results with more precision and accuracy.

Finally, I would like to say one important mantra for all our programmers

“Self-Validation is the best way to avoid 50 to 60% of mistakes

Try to work towards “First Time Right Approach (FTRA)”

ACKNOWLEDGMENTS

The author is very grateful to his mentor Frank Senk, and manager Amy Rosen & Pinku Hooda, CEO Neha Aggarwal and Director Nalin Gupta, the author would also like to acknowledge Vura K Pallayya Gupta for making available all the time to share his experiences

CONTACT INFORMATION

Venkata Sairam Veeramalla
Technical Manager
GCE Solutions India Pvt. Ltd
Hyderabad, India
Sairam.Veeramalla@gcesolutions.com
Sairam.sas@gmail.com
www.gcesolutions.com