

Creating Client-Friendly SAS Output

Arthur C. Singer, Wyeth-Ayerst Research, Philadelphia, PA
Caroline M. Bobik, Bobik Consulting, Inc., Paoli, PA

The most important goal of creating computer output is presenting concise communication in the shortest time, remembering that 'Time is money.' This poster presentation will emphasize readable annotation in tabular output. Appropriate documentation decreases the time spent with the client explaining the computer-generated information. SAS® programmers/users may find this approach helpful in dealing with clients who are not versed in computer-coded arguments.

Sorting of data by meaningful categories is often a neglected concept. In our experience, we have frequently received final listings of clinical results that were output in no apparent order and were virtually unusable. A user/programmer and the client should agree on the end format before data processing begins.

SAS® Software provides the tools necessary to produce such output. This presentation will display the transition from Microsoft Excel® output through SAS default output to easily read SAS output.

Displays of SAS software output will be used to illustrate the concepts relevant to this topic. The emphasis is on the content of the output rather than the syntax used to produce it. Programmers/users have different levels of expertise with some writing syntax more elegantly and efficiently than others. A sampling of 25 representative observations is used to show a progression of improvements, with possibly a blunder or two, to a data set.

We were given a Microsoft Excel file made up of 4 sheets (tables). The sheets contained demographic, chemistry, hematology, and urinalysis data, all constructed based on typical observations.

Microsoft Excel output displaying partial demographics and hematology data are headed with those terms. The demographic data are displayed by patients' initials in a row format, while the patients' initials are in a column format in the hematology data. The hematological variables are represented by abbreviated labels. Other laboratory data, though not shown, were received in the same format as that of the hematology data.

Format consistency makes computer output easier to read for most clients. Transposition of the laboratory files to row format was done using an Excel function. The dates were reformatted in Excel so that they would be easily read by SAS software. Blank lines were eliminated before creating 'comma separated values' Excel tables to be read by SAS software.

A SAS software user can input data any convenient way. We have demonstrated this feature in Tables 1 and 2, using the supplied data. It is not pretty and should never be distributed to the client in this form.

Table 1 shows the demographic data labeled as it was input by SAS. There is duplication of information that can be used by the programmer to verify the input. The data were input at character variables. The variable BD is the birth date expressed as an internal SAS value. BDATE is the same information that has been formatted as a date. DX1 is the code for gender while GENDER is labeled in the English. DX2 through DX4, input as character variables have been converted to numeric variables, WGT, HGT, and BSURF, respectively.

Table 2 displays variables input for convenience of the programmer, all as character variables. HDATE and HTIME were created using arithmetic statements. All the variables that have numeric values will have to be converted to numeric variables for arithmetic operations.

All data that is to be printed and distributed to a client should be labeled in words rather than variable names, as displayed in the resulting tables.

The last two tables at the end of this presentation each show at least one sorting category. Table 3 displays the data sorted by gender, and Table 4 shows the data sorted on an individual patient basis. Note that the patients listed in Table 3 are sorted by the initials of their last names.

In conclusion, we propose that users/programmers produce output that is clear and easy to understand. Your clients will thank you.

The authors can be contacted at:

Arthur C. Singer
Wyeth-Ayerst Research
PO Box 8299
Philadelphia, PA 19101
(610) 341-2077
e-mail: singera@war.wyeth.com

Caroline M. Bobik, President
Bobik Consulting, Inc.
PO Box 545
Paoli, PA 19301-0545
(610) 408-0661
e-mail: bobik.consulting@worldnet.att.net

References

SAS is a registered trademark or the trademark of SAS Institute Inc. in the USA and other countries. © indicates USA registration.

Microsoft Excel Version 5.0c, Copyright © 1985-1994
Microsoft Corporation

Demographics

	DOB	Sex	BW	H	BSA
NDA	13025	2	70.1	142.3	1.6
POT	32950	2	69.9	153.1	1.67
CHL	32831	2	59.3	151.2	1.55
BIC	82742	2	68.1	163.7	1.72
CAI	52638	2	69.23	170	1.81
PHO	42549	2	47.1	158.9	1.43
WBC	52952	1	70.4	154.9	1.7
PLT	112116	1	74.54	173	1.89
NET	12535	1	67.9	167.6	1.76
MON	7841	1	68	171	1.8
HCT	42026	1	82.3	181.2	2.11
RBC	111323	1	55.4	160	1.58
EOS	52622	1	98.1	182.4	2.31
BAS	62437	1	78	170	1.89
CEL	11230	1	73.6	179.5	1.92
RET	12614	1	58.6	167.6	1.68
TOP	4423	1	80.5	172.4	1.92

1=male
2=female

Hematology

Inits	NDA	POT	CHL	BIC	CAI	PHO	BUN
LD	72395	62195	72095	61395	61495	61695	71395
LT	7:00	11:20	10:10	8:10	2:55	7:30	18:20
Hgb	8.2	8.8	7.8	15.4	14.2	12.1	11
Hct	23.7	25.4	23.1	44.8	41.1	34.8	32.9
Plt	52	26	59	73	67	46	32
RBC	2.81	3.03	2.74	5.26	4.88	4.15	3.93
Ret	ND	ND	ND	ND	ND	ND	ND
WBC	5.1	2.3	1.6	13.7	11	10.6	8.8
bands	0	ND	24	4	3	ND	0
Segs	93	ND	72	87	94	ND	93
Lymphs	4	ND	4	4	2	ND	4
monos	3	ND	0	5	1	ND	3
Eos	0	ND	0	0	0	ND	0
Basos	0	ND	0	0	0	ND	0
OC	0	ND	0	0	0	ND	0

Protocol 123A-456-SUGI22
 Table 1 Demographic Data

PATID	BD	DX1	DX2	DX3	DX4	BDATE	GENDER	WGT	HGT	BSURF
NDA	-12754	2	70.10	142.3	1.6	30JAN25	Female	70.1	142.3	1.60
POT	-3565	2	69.90	153.1	1.67	29MAR50	Female	69.9	153.1	1.67
CHL	-10506	2	59.30	151.2	1.55	28MAR31	Female	59.3	151.2	1.55
BIC	-6336	2	68.10	163.7	1.72	27AUG42	Female	68.1	163.7	1.72
CAI	-7890	2	69.23	170.0	1.81	26MAY38	Female	69.2	170.0	1.81
PHO	-3903	2	47.10	158.9	1.43	25APR49	Female	47.1	158.9	1.43
BUN	-7783	2	54.30	161.2	1.53	10SEP38	Female	54.3	161.2	1.53
URA	-6750	2	77.30	168.4	1.85	09JUL41	Female	77.3	168.4	1.85
SGO	-11774	2	62.90	170.5	1.74	07OCT27	Female	62.9	170.5	1.74
SPT	-5539	2	48.12	166.7	1.56	01NOV44	Female	48.1	166.7	1.56
ALP	12390	2	79.50	151.3	1.76	03DEC93	Female	79.5	151.3	1.76
LDH	-12356	2	73.90	150.4	1.68	04MAR26	Female	73.9	150.4	1.68
GIU	-14877	2	71.30	152.8	1.71	09APR19	Female	71.3	152.8	1.71
WBC	-2773	1	70.40	154.9	1.7	29MAY52	Male	70.4	154.9	1.70
PLT	-15746	1	74.54	173.0	1.89	21NOV16	Male	74.5	173.0	1.89
NET	-9107	1	67.90	167.6	1.76	25JAN35	Male	67.9	167.6	1.76
MON	-6751	1	68.00	171.0	1.8	08JUL41	Male	68.0	171.0	1.80
HCT	-12309	1	82.30	181.2	2.11	20APR26	Male	82.3	181.2	2.11
RBC	-13198	1	55.40	160.0	1.58	13NOV23	Male	55.4	160.0	1.58
EOS	-13734	1	98.10	182.4	2.31	26MAY22	Male	98.1	182.4	2.31
BAS	-8226	1	78.00	170.0	1.89	24JUN37	Male	78.0	170.0	1.89
CEL	-10946	1	73.60	179.5	1.92	12JAN30	Male	73.6	179.5	1.92
RET	-16462	1	58.60	167.6	1.68	06DEC14	Male	58.6	167.6	1.68
TOP	-13421	1	80.50	172.4	1.92	04APR23	Male	80.5	172.4	1.92

Protocol 123A-456-SUGI22
Table 2 Hematology Data

PATID	HD	HT	HX1	HX2	HX3	HX4	HX5	HX6	HX7
NDA	12987	25200	8.2	23.7	52	2.81	ND	5.1	0
POT	12955	40800	8.8	25.4	26	3.03	ND	2.3	ND
CHL	12984	36600	7.8	23.1	59	2.74	ND	1.6	24
BIC	12947	29400	15.4	44.8	73	5.26	ND	13.7	4
CAI	12948	10500	14.2	41.1	67	4.88	ND	11	3
PHO	12950	27000	12.1	34.8	46	4.15	ND	10.6	ND
BUN	12977	66000	11	32.9	32	3.93	ND	8.8	0
CRT	12949	26100	7.4	21.1	76	2.55	ND	0.9	QNS
URA	12934	3600	7.9	23.6	80	3.2	ND	4.2	0
SGO	12940	7200	6.7	20.2	290	249	ND	5.3	2
SPT	12942	37500	13.6	40.1	204	4.75	ND	4.3	1
ALP	12920	72600	5.8	17.2	40	2.07	ND	0.4	ND
LDH	12931	36900	9.4	27	24	3.21	ND	0.4	ND
GIU	12952	46200	9.3	27.5	18	3.23	ND	1.1	4
TOP	12908	39600	9.1	27.6	185	3.33	ND	2.1	18
WBC	12912	41100	8.8	26.1	324	3.13	ND	6.2	ND
PLT	12950	35100	7.5	22.5	105	2.71	ND	0.7	0
NET	12882	45000	8.8	26.4	46	3.28	ND	2.5	ND
MON	12922	42420	8.6	26.4	421	3.29	ND	2.4	4
HCT	12921	33900	9.4	28	578	3.48	ND	5	1
RBC	12899	47100	6.1	18.9	90	2.38	ND	1.7	11
EOS	12851	5400	9	26.8	575	3.31	ND	6.9	2
BAS	12887	45240	6.6	20.6	80	2.6	ND	0.7	0
CEL	12881	5820	8.8	26.4	46	3.28	ND	2.5	ND
RET	12858	25200	7.9	24.4	87	3	ND	3.5	24

PATID	HX8	HX9	HX10	HX11	HX12	HX13	HDATE	HTIME
NDA	93	4	3	0	0	0	23JUL95	7:00
POT	ND	ND	ND	ND	ND	ND	21JUN95	11:20
CHL	72	4	0	0	0	0	20JUL95	10:10
BIC	87	4	5	0	0	0	13JUN95	8:10
CAI	94	2	1	0	0	0	14JUN95	2:55
PHO	ND	ND	ND	ND	ND	ND	16JUN95	7:30
BUN	93	4	3	0	0	0	13JUL95	18:20
CRT	QNS	QNS	QNS	QNS	QNS	QNS	15JUN95	7:15
URA	41	30	10	0	0	0	31MAY95	1:00
SGO	69	11	17	1	0	0	06JUN95	2:00
SPT	68	17	12	2	0	0	08JUN95	10:25
ALP	ND	ND	ND	ND	ND	ND	17MAY95	20:10
LDH	ND	ND	ND	ND	ND	ND	28MAY95	10:15
GIU	4	84	4	4	0	0	18JUN95	12:50
TOP	53	22	6	0	1	0	05MAY95	11:00
WBC	ND	ND	ND	ND	ND	ND	09MAY95	11:25
PLT	67	32	0	1	0	0	16JUN95	9:45
NET	ND	ND	ND	ND	ND	ND	09APR95	12:30
MON	69	23	4	0	0	0	19MAY95	11:47
HCT	80	8	10	0	1	0	18MAY95	9:25
RBC	39	49	1	0	0	0	26APR95	13:05
EOS	81	8	6	0	2	1	09MAR95	1:30
BAS	44	54	1	0	0	1	14APR95	12:34
CEL	ND	ND	ND	ND	ND	ND	08APR95	1:37
RET	37	24	13	0	0	0	16MAR95	7:00

Protocol 123A-456-SUGI22
Table 3 Hematological Data

----- Gender=Male -----

Patient's Initials	Birth Date	Age at Lab Date	Body Weight	Height	Body Surface Area	Laboratory Data Date
RBC	13NOV23	71	55.4	160.0	1.58	26APR95
WBC	29MAY52	42	70.4	154.9	1.70	09MAY95
CEL	12JAN30	65	73.6	179.5	1.92	08APR95
MON	08JUL41	53	68.0	171.0	1.80	19MAY95
TOP	04APR23	72	80.5	172.4	1.92	05MAY95
BAS	24JUN37	57	78.0	170.0	1.89	14APR95
EOS	26MAY22	72	98.1	182.4	2.31	09MAR95
HCT	20APR26	69	82.3	181.2	2.11	18MAY95

Patient's Initials	Laboratory Data Time	Hemoglobin	Hematocrit	Platelet Count	RBC Count
RBC	13:05	6.1	18.9	90	2.38
WBC	11:25	8.8	26.1	324	3.13
CEL	1:37	8.8	26.4	46	3.28
MON	11:47	8.6	26.4	421	3.29
TOP	11:00	9.1	27.6	185	3.33
BAS	12:34	6.6	20.6	80	2.6
EOS	1:30	9	26.8	575	3.31
HCT	9:25	9.4	28	578	3.48

Patient's Initials	Reticulocytes	WBC Count	Band Neutrophils	Segmented Neutrophils	Lymphocytes
RBC	ND	1.7	11	39	49
WBC	ND	6.2	ND	ND	ND
CEL	ND	2.5	ND	ND	ND
MON	ND	2.4	4	69	23
TOP	ND	2.1	18	53	22
BAS	ND	0.7	0	44	54
EOS	ND	6.9	2	81	8
HCT	ND	5	1	80	8

Patient's Initials	Monocytes	Eosinophils	Basophils	Other Cells
RBC	1	0	0	0
WBC	ND	ND	ND	ND
CEL	ND	ND	ND	ND
MON	4	0	0	0
TOP	6	0	1	0
BAS	1	0	0	1
EOS	6	0	2	1
HCT	10	0	1	0

ND means "Not Done", and QNS means "Quantity Not Sufficient"

Protocol 123A-456-SUGI22
 Table 4 Hematological Data

----- Patient's Initials=BAS -----

Gender	Birth Date	Age at Lab Date	Body Weight	Height	Body Surface Area	Laboratory Data Date	Laboratory Data Time
Male	24JUN37	57	78.0	170	1.89	14APR95	12:34

Hemoglobin	Hematocrit	Platelet Count	RBC Count	Reticulocytes	WBC Count	Band Neutrophils
6.6	20.6	80	2.6	ND	0.7	0

Segmented Neutrophils	Lymphocytes	Monocytes	Eosinophils	Basophils	Other Cells
44	54	1	0	0	1

----- Patient's Initials=BIC -----

Gender	Birth Date	Age at Lab Date	Body Weight	Height	Body Surface Area	Laboratory Data Date	Laboratory Data Time
Female	27AUG42	52	68.1	163.7	1.72	13JUN95	8:10

Hemoglobin	Hematocrit	Platelet Count	RBC Count	Reticulocytes	WBC Count	Band Neutrophils
15.4	44.8	73	5.26	ND	13.7	4

Segmented Neutrophils	Lymphocytes	Monocytes	Eosinophils	Basophils	Other Cells
87	4	5	0	0	0

ND means "Not Done", and QNS means "Quantity Not Sufficient"