

大数据分析工具SAS的 代码智能填充及代码 审查的可视辅助 (Make SAS do all the dirty work for you)

The roads ...

- **Before** we got into the market of big data analysis we heard that this market was the land of opportunity– the roads to the market were paved with gold.

But **after we got here, we learned three things.**

- ◆ 1. The roads were not paved with gold.
- ◆ 2. Some roads were not paved at all.
- ◆ 3. We are hired to pave them.

SAS coding vs paving.

- SAS coding tasks are not equal to paving but there are also **unattractive/unpopular** parts.
- Unfortunately, the unattractive parts are unavoidable, They are the nightmares for almost **any** programmer.

Definition of dirty/unattractive work?

- No one can precisely give a definition but you can smell the blood when some signs exist.

Code with plenty of variable names/values/formats

```
/*DSCOM*/  
if lnkseq = 1 then dscom = rdcdsreas sum long nm;  
else if lnkseq = 2 then dscom = rdcrsdcfin ltf long nm;  
else if lnkseq = 3 then dscom = rdcrsdcfin sf long nm;  
else if lnkseq = 4 then dscom = rdcdsreas_swov_long_nm;  
  
/*DTHDT*/  
if lnkseq = 1 and missing(dtmthdt sum dt) eq 0 then dthdt = input dtmthdt sum dt mmddy  
else if lnkseq = 4 and missing(dtmthdt_swov_dt) eq 0 then dthdt = input dtmthdt_swov_c  
  
/*DTHDTC*/  
if dthdt ne . then dthdtc=put(dthdt, yymmdd10.);  
  
/*AEID*/  
if lnkseq = 1 then do;  
  if ds = 2 then aeid = txtaeseq sum norm;  
  else if ds = 3 then aeid = txtaeseq sumd norm;  
end;
```

Structural but lengthy Proc Step

```
proc report data=ioqi.bg nowd headline;  
  column  sdytrtendtc sdytrtendt sdytrstddtc sdytrstddt subjrnd subjsafs subjpps subjfas birthdtmpuflg sex subrace  
  define sdytrtendtc / width=10 format=$10. "Study Treatment End Date Char" display flow;  
  define sdytrtendt / width=9 format=DATE9. "Study Treatment End Date" display flow;  
  define sdytrstddtc / width=10 format=$10. "Study Treatment Start Date Char" display flow;  
  define sdytrstddt / width=9 format=DATE9. "Study Treatment Start Date" display flow;  
  define subjrnd / width=8 format= 8. "Subject Randomized Set" display flow;  
  define subjsafs / width=8 format= 8. "Subject Safety Analysis Set" display flow;  
  define subjpps / width=8 format= 8. "Subj Per Protocol Set" display flow;  
  define subjfas / width=8 format= 8. "Subject Full Analysis Set" display flow;  
  define birthdtmpuflg / width=8 format= 8. "Birth Date Imputation Flag" display flow;  
  define sex / width=8 format= 8. "Sex" display flow;  
  define subrace / width=8 format= 8. "Subrace" display flow;  
  define race / width=8 format= 8. "Race" display flow;  
  define ageyr / width=8 format= 8. "Age in Years at Consent" display flow;  
  define trtpsort / width=8 format= 8. "Planned Treatment Sort Order" display flow;  
  define trtasort / width=8 format= 8. "Actual Treatment Sort Order" display flow;  
  define trtp / width=50 format=$50. "Planned Treatment" display flow;  
  define trta / width=50 format=$50. "Actual Treatment" display flow;
```


Peer view your co-worker's code

```
1 data adeg;
2   set eg;
3   keep TRTAN param TRTA subjid cohort trt cycle VIS AVISIT vis EGDTC EGDY AVALC
4   where param in ("Summary (Mean) QT Duration", "QTcB - Bazett's Correction Formula"
5   "Summary (Mean) Heart Rate"
6   , "Summary (Mean) PR Duration"
7   , "Summary (Mean) RR Duration"
8   , "Summary (Mean) QRS Duration"
9   , 'Interpretation'
10  )
11  ;
12  IF PARAM NE 'Interpretation' THEN avalc=%AHGputn(aval, 6.1);
13  cohort='Cohort' || %AHGputn(trtan);
14  trt=scan(trta, 2, '-');
15  cycle=visit;
16
17  /* cycle=left(tranwrd(SCAN(avisit,1), 'CYCLE', ''));*/
18  /* if upcase(cycle)=:'SCREEN' then cycle='0';*/
19  egdtc=trim(egdtc) || '/' || %AHGputn(EGDY);
```


Let's talk about something more practical

- Can we have a tool to generate some standard tables for us?
- By the way, Can I see the result **right now**? Before I decide to use Standard Code Package No. XXX ?

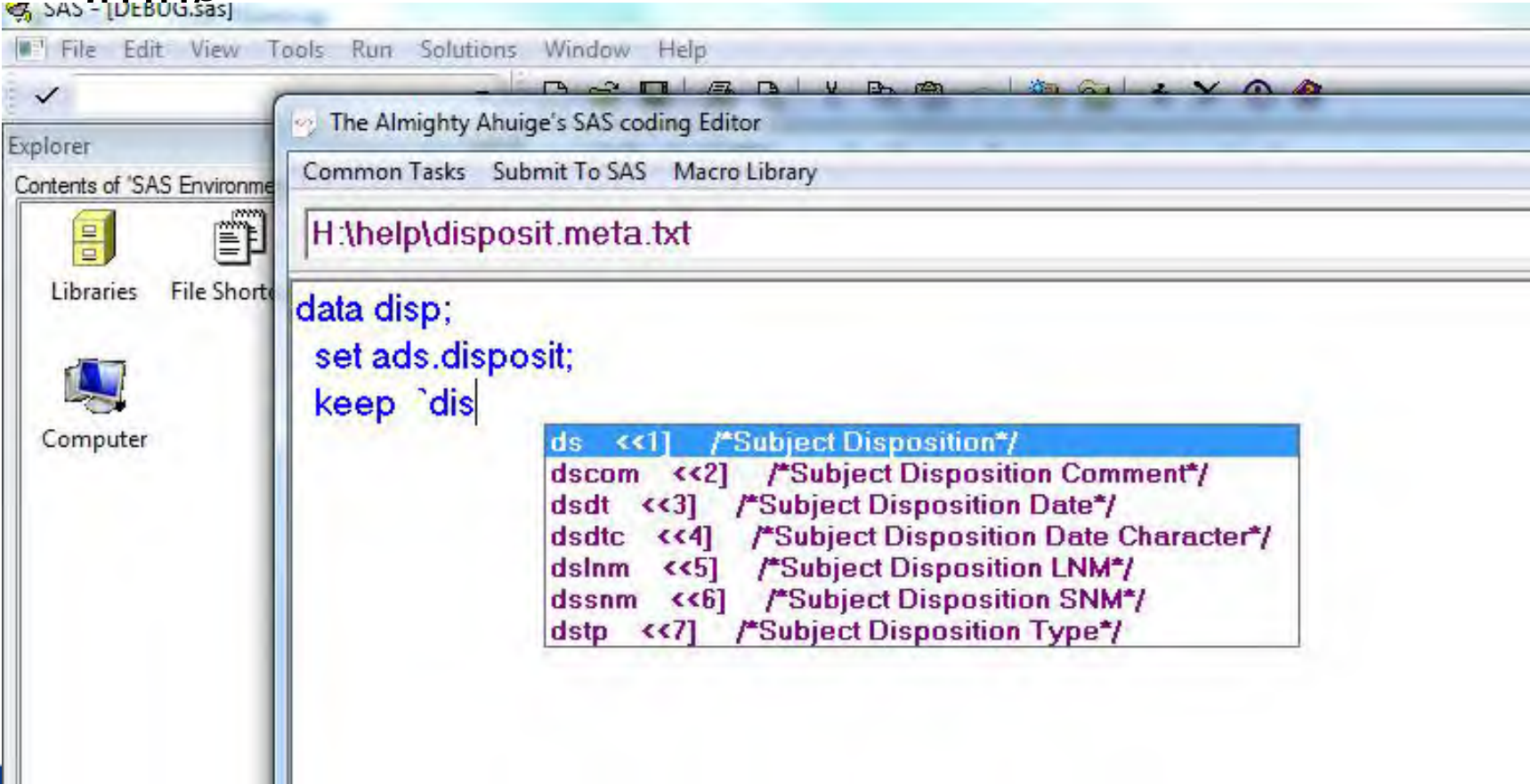
Here is a solution

- With a visual editor tool developed at Eli Lilly China, We can
- 1. Type code with pop-up hints and auto-fill if needed. The editor will suggest/complete the variable names/values/formats;
- 2. when peer-review code you can transparently SEE what is behind a mysterious variable.

- 3. Some complicated but structured code can be auto-generated such as PROC REPORT, it is just one-click away.
- 4. Within the editor, macros can be populated and executed instantly, results can be reviewed right away **EVEN BEFORE YOU REVIEW CODE** (What you see is what you get).

Information is at your fingertips

Auto popup/fill variable names by typing arbitrary
String



The screenshot shows the SAS coding editor window titled "The Almighty Ahuige's SAS coding Editor". The editor is open to the file "H:\help\disposit.meta.txt". The code in the editor is:

```
data disp;
  set ads.disposit;
  keep `dis|
```

An auto-popup menu is displayed, listing the following variables:

- ds <<1> /*Subject Disposition*/
- dscom <<2> /*Subject Disposition Comment*/
- dsdt <<3> /*Subject Disposition Date*/
- dsdtc <<4> /*Subject Disposition Date Character*/
- dslnm <<5> /*Subject Disposition LNM*/
- dssnm <<6> /*Subject Disposition SNM*/
- dstp <<7> /*Subject Disposition Type*/

- Auto popup/fill variable VALUE by searching string in the existing values of variables.

The Almighty Ahuige's SAS coding Editor

Common Tasks Submit To SAS Macro Library

H:\help\disposit.meta.txt

```
data disp;  
  set ads.disposit;  
  if dssnm = '.withd
```

.dscom	[1>>	"Withdrawal By Subject"
.dslnm	[2>>	"Withdrawal By Subject"
.dssnm	[3>>	"Withdrawal By Subject"

Review *INTO* your co-worker's code

```

The Almighty Ahuige's SAS coding Editor
Common Tasks Submit To SAS Macro Library
D:\lillycel\qally2835219\i3y_je_jpbcl\intrm2\data\shared\adam\adeg.meta.txt

data adeg;
  set eg;
  keep TRTAN param TRTA subjid cohort trt cycle VIS AVISIT vis EGDTC EGDY val ;
;
  cohort='Cohort'||%A
  trt=scan( TRTA?,2,
  cycle=visit;
run;

```

.TRTA	[1>>	"LY2835219-100mgQ12H"
.TRTA	[2>>	"LY2835219-150mgQ12H"
.TRTA	[3>>	"LY2835219-200mgQ12H"
.TRTA	[4>>	/*Actual Treatment*/

Any other typical tedious work?

- How about a Proc Report?
- It is just one-click away.

```
proc report data=adam.adtr nowd headline;  
  where not missing( LSPCGBST );  
  column LSPCGBST LSPERCHG LSSUMCHG LSSUMBL VISITNUM VISIT;  
  define LSPCGBST / width=18 format= 8. "Best Percentage Change of Sum of Target Lesion fro" display flow;  
  define LSPERCHG / width=8 format= 8. "Percentage Change of Sum of Target Lesion from Bas" display flow;  
  define LSSUMCHG / width=8 format= 8. "Change of Sum of Target Lesion from Baseline (cm)" display flow;  
  define LSSUMBL / width=8 format= 8. "Sum of Target Lesion at Baseline (cm)" display flow;  
  define VISITNUM / width=8 format= 8. "Visit Number" display flow;  
  define VISIT / width=50 format=$50. "Visit Name" display flow;  
run;
```


Auto generation of summary tables

review output before review code

```
Dataset: heart
Variable: Cholesterol
Treatment: Sex
By: Chol_Status [ Cholesterol Status ]
```

Chol_Status		Female	Male
s	label		
Borderline	n	959	902
Borderline	mean	219	219
Borderline	median	220.00	220.00
Borderline	min - max	200.00 - 239.00	200.00 - 239.00
Desirable	n	805	600
Desirable	mean	178	178
Desirable	median	180.00	180.00
Desirable	min - max	117.00 - 199.00	96.00 - 199.00
High	n	1010	781
High	mean	278	271
High	median	270.00	263.00

```
Dataset:  sashelp.heart  
Variable: MRW [ Metropolitan Relative Weight ]  
Treatment:  
By: BP_Status [ Blood Pressure Status ]
```

BP_Stat	n		mea		
us	n	n	median	min	max
High	2265	127	125.00	73.00	- 268.00
Normal	2141	116	115.00	67.00	- 246.00
Optimal	797	110	108.00	75.00	- 160.00

After you have chosen the output you can review the real code

```
The Almighty Ahuige's SAS coding Editor
Common Tasks Submit To SAS Macro Library debug Set Filter
sashelp\heart.meta.txt

data thedsn__2;
  set sashelp.heart;
  ;
run;

proc sort data = thedsn__2 out = thedsn__2 ;
  by Chol_Status Sex;
run;

;
proc means data=thedsn__2 noprint alpha=0.05;
  ;
  var Cholesterol;
  by Chol_Status Sex;
  output
  out=stat_Cholesterol n = n mean = mean median = median min = min max = max ;
run;
```

Q & A.