Alain Smits
Phuse SDE Belgium

SUPERCHARGING PROC CONTENTS

Introduction

- Combines a structure and contents overview in a dynamic web page
- Easier to explore new data sets, especially complex ones
- Pick up data issues quickly

The macro

```
%dscont(inds      = ,
        limit     = ,
        path      = ,
        file      = );
```

Example

```
%dscont(inds = sds.lb);
```

Default output to work folder, file opens automatically in SAS DMS

Sample output

ADAM.LBAD

Laboratory Results Analysis Dataset
Created: 19OCT10:15:50:19
Obs : 9750
Key :

#	VARIABLE	LABEL	TYPE	LENGTH	CONTENTS
1	**STUDYID**	STUDY IDENTIFIER	CHAR	40	1 distinct values, 0 missing values
2	**SITEID**	STUDY SITE IDENTIFIER	CHAR	10	11 distinct values, 0 missing values
3	**INVNAM**	INVESTIGATOR NAME	CHAR	120	11 distinct values, 0 missing values
4	**SUBJID**	SUBJECT IDENTIFIER FOR THE STUDY	CHAR	20	27 distinct values, 0 missing values
5	**USUBJID**	UNIQUE SUBJECT IDENTIFIER	CHAR	40	27 distinct values, 0 missing values
6	**SUBJINIT**	SUBJECT INITIALS	CHAR	5	27 distinct values, 0 missing values
7	**AGE**	AGE (years) AT REFERENCE DATE/TIME	NUM	8	3 distinct values, 0 missing values Mean: 3.968 Minimum: 3 Maximum: 5
8	**SEX**	SEX	CHAR	1	2 distinct values, 0 missing values
9	**RACE**	RACE	CHAR	80	4 distinct values, 0 missing values
10	**ARMCD**	PLANNED ARM CODE	CHAR	8	1 distinct values, 0 missing values
11	**ARM**	DESCRIPTION OF PLANNED ARM	CHAR	200	1 distinct values, 0 missing values
12	**ARMN**	PLANNED ARM NUM. CODE	NUM	8	1 distinct values, 0 missing values Mean: 1 Minimum: 1 Maximum: 1 Mis
13	**LBGRPID**	GROUP ID	CHAR	120	214 distinct values, 227 missing values
14	**PHASE**	PHASE	CHAR	80	3 distinct values, 0 missing values
15	**PHASEN**	PHASE (NUM CODE)	NUM	8	3 distinct values, 0 missing values Mean: 0.864 Minimum: 0 Maximum: 2
16	**VISITNUM**	VISIT NUMBER	NUM	8	16 distinct values, 0 missing values Mean: 17.618 Minimum: 1 Maximum: 1
17	**VISIT**	VISIT NAME	CHAR	120	16 distinct values, 0 missing values
18	**ANALTPT**	ANALYSIS TIMEPOINT	CHAR	100	10 distinct values, 0 missing values

Sample output

44	**LBSPEC**	SPECIMEN TYPE	CHAR	40	+	4 distinct values, 1760 missing values
45	**LBFAST**	FASTING STATUS	CHAR	1	+	3 distinct values, 227 missing values
46	**LBTOX**	TOXICITY	CHAR	40	+	9 distinct values, 440 missing values
47	**LBTOXN**	TOXICITY NUMBER	NUM	8	+	9 distinct values, 440 missing values Mean: 2.61 Minimum: 0 Maximum: 7 Missing
48	**LBTOXNM**	TOXICITY NAME	CHAR	60	–	49 distinct values, 385 missing values **Only approx. 30 values shown** 385x **MISSING** 224x ALANINE AMINO TRANSFERASE 225x ALKALINE PHOSPHATASE 135x ALPHA-1 ACID GLYCOPROTEIN 219x AMYLASE 218x ASPARTATE AMINO TRANSFERASE 174x BASOPHILS 221x BASOPHILS (%) 225x BICARBONATE 218x BLOOD UREA NITROGEN 227x CALCIUM 223x CHLORIDE 227x CREATININE 219x DIRECT BILIRUBIN 174x EOSINOPHILS 219x EOSINOPHILS (%) 220x HAEMATOCRIT 221x HAEMOGLOBIN 128x HIGH DENSITY LIPOPROTEIN 225x HYPERBILIRUBINEMIA

Under the hood

- Programming challenges
  - Performance when processing large data sets
  - Handling user defined formats
  - Avoiding information overload in the output

Under the hood

- Performance
  - Proc tabulate provides incredible processing efficiency to create statistics on all variables.
  - Requires a fair amount of data tweaking to get a useful data format for further processing due to the special format of the output data set of proc tabulate

Under the hood

- User defined formats
  - The macro needs to collect the user defined formats to ensure the table correctly displays the values, although the format itself is currently not displayed
  - In our organization, the use of user defined formats is limited
  - Code can be simplified through the recently introduced vvalue() function that returns the formatted value of a variable

Under the hood

- Avoiding information overload
  - The use of dynamic HTML (JavaScript) allows information to be collapsed to allow a clean structure view and expanded for a contents overview

Extras

- Saving as an excel file
  - Useful for copy and pasting.

Future enhancements

- (User Defined) Formats are currently not shown because they are not used in SDTM or ADaM.

Contact details

Alain Smits
Associate Director Clinical Programming
Janssen R&D
Beerse, Belgium
+32 14 64 1325 or +32 495 26 90 08
asmits@its.jnj.com