SAS® GLOBAL FORUM 2017

April 2 – 5 | Orlando, FL

When ANY Function Will Just NOT Do

USERS PROGRAM

When ANY Function Will Just NOT Do

Richann Watson, Karl Miller

Experis, inVentiv Health

ANY AND NOT FUNCTIONS

String (required)*:*

- Character constant, variable or expression

Start (optional)***:***

- determine search direction
  - Start > 0, search is from left to right
  - Start < 0, search is from right to left
  - Start < negative length of string, search starts at the end of the string

Functions yield the position of the first encounter of the desired search. It returns a zero when one of the following is true:

- Search character is not found
- Start > length of the string
- Start = 0

Function Search Criteria	ANY Function	NOT Function
alpha-numeric character	ANYALNUM(*string, <start>*)	NOTALNUM(*string, <start>*)
alphabetic character	ANYALPHA(*string, <start>*)	NOTALPHA(*string, <start>*)
digit	ANYDIGIT(*string, <start>*)	NOTDIGIT(*string, <start>*)
character that is valid as the first character of a SAS variable*	ANYFIRST(*string, <start>*)	NOTFIRST(*string, <start>*)
lowercase letter	ANYLOWER(*string, <start>*)	NOTLOWER(*string, <start>*)
character that is valid in a SAS variable*	ANYNAME(*string, <start>*)	NOTNAME(*string, <start>*)
punctuation character	ANYPUNCT(*string, <start>*)	NOTPUNCT(*string, <start>*)
white-space character: blank, horizontal and vertical tab, carriage return, line feed, and form feed	ANYSPACE(*string, <start>*)	NOTSPACE(*string, <start>*)
uppercase letter	ANYUPPER(*string, <start>*)	NOTUPPER(*string, <start>*)

* SAS variable name under VALIDVARNAME=V7

Extracting Week Number from Text String

Data comes in a variety of formats but what is consistent is that there is usually a number that is preceded by non-number.

CLICK HERE

➢More commonly used approaches

```
visind=input(substr(visit,
      indexc(visit, '123456790')), best.);
viscomp=input(compress(visit,,'dk'),best.);
```

➢Another suggestion would be

```
visany=input(substr(visit,
      anydigit(visit)), best.);
```

Extracting Week Number and Day Number from Text String

Data comes in a variety of formats but what is consistent is that there is a number that represents the week that is preceded by a non-number which is followed by more non-numbers, with the last character in the string being a number which represents the day.

CLICK HERE

1. Find the location of the 1st number when searching from the left.

```
firstnumloc=anydigit(visit);
```

2. Find the location of 1st alpha character when searching from left. Starting at the position of 1st number.

```
secalploc=anyalpha(visit, firstnumloc);
```

3. Find the location of last number when searching from right

```
lastnum=anydigit(visit, -length(visit));
```

4. Extract the week portion using the number found in step 1 and step 2.

```
WEEK=input(substr(visit, firstnumloc, secalploc -
    firstnumloc), best.);
```

5. Extract the day portion using the number found in step 3.

```
DAY input(substr(visit, lastnum), best.);
```

Use in SDTM and ADaM Data set Creation

SDTM

➢Populating LBSTRESN in LB domain and determining data issues

```
anyalpha(lborres)
```

ADaM

CLICK HERE

➢Determine if AVALC is a partial numeric

➢Determine if AVAL is integer or float

```
not(anyalpha(AVALC))
anypunct(AVALC)
```

Convert Individual Date/Time Components to ISO 8601 Format

1. Determine which components has values other than a number. If it has a value other than a number, then it is assumed it is missing and denoted with a single dash.

CLICK HERE

```
if not(notdigit(&dttmvar.)) then _&dttmvar=&dttmvar;
else _&dttmvar='-';
```

2. Create date and time variables

```
isodt=catx("-", _year, _mon, _day);
isotm=catx(":", _hr, _min, _sec);
```

3. Determine if the time is completely missing (i.e., isotm = '-:-:-')

```
if notpunct(strip(isotm)) > 0 then _isotm=
   substr(isotm,1,notpunct(strip(isotm),-length(isotm)));
else _isotm=' ';
```

4. Combine date with the new time variable

```
newdttm=catx("T", isodt, _isotm);
```

5. Determine if the date is complete. If ANYALPHA returns a value greater than 0, then ISO 8601 date is complete and no further processing

```
if anyalpha(strip(newdttm)) > 0 then NEWDTC=newdttm;
```

6. ANYALPHA in step 5 returns 0, then there is no time so need report up to last non-missing date component

```
… if notpunct(strip(newdttm)) > 0 then NEWDTC=
  substr(newdttm,1,notpunct(strip(newdttm),-length(newdttm)));
```

When ANY Function Will Just NOT Do

Richann Watson, Karl Miller

Experis, inVentiv Health

CONCLUSION

- Simple yet powerful functions when used correctly
- Perform well across multiple data types, studies or compounds
- Applied structure to meet your specific needs
- Beneficial for efficiency in programming

CONTACT INFORMATION

Name: Richann Watson
Enterprise: Experis
Work Phone: 513.843.4081
E-mail: richann.watson@experis.com

Name: Karl Miller
Enterprise: inVentiv Health
Work Phone: 402.641.3089
E-mail: karl.miller@inventivhealth.com

REFERENCES

Dictionary of SAS Functions and CALL Routines. SAS Institute, Available at http://support.sas.com/documentation/cdl/en/lefunctionsref/67960/HTML/default/viewer.htm

RECOMMEND READING

SAS® Functions by Example by Ron Cody

SAS® GLOBAL FORUM 2017

April 2 – 5 | Orlando, FL

When ANY Function Will Just NOT Do

Richann Watson, Karl Miller

Experis, inVentiv Health

Use in SDTM Data set Creation

LBORRES	ANYALPHA	Handling
1.4	0	Convert to standardized unit and populate LBSTRESC and LBSTRESN
TRACE	1	No conversion required. Copy to LBSTRESC.
CANCELLED	1	Data should be queried or fixed so that LBSTAT = 'NOT DONE' and LBREASND = 'CANCELLED'

Use in ADaM Data set Creation

AVALC	ANYALPHA	NOTALPHA	ANYPUNCT	NOTDIGIT
TESTING	1	8	0	1
<1.0	0	1	1	1
1.2	0	1	2	2
354.12	0	1	4	4
	0	1	0	1
___123	0	1	1	1
124!	0	1	4	4
47821	0	1	0	6
1TEST	2	1	0	2

CAUTION

It might be tempting to use NOTALPHA for this, but it is also very important to keep in mind what is actually being searched by the functions. Since NOTALPHA will return the position of the first non-alphabetic character and due to the case that some character results can contain both alphabetic and numeric characters, the use of NOTALPHA would yield a non-zero value for results that are alphanumeric and not strictly numeric.

When ANY Function Will Just NOT Do

Richann Watson, Karl Miller

Experis, inVentiv Health

Extracting Week Number from Text String

VISIT	VISIND	VISCOMP	VISANY
w4	4	4	4
w 4	4	4	4
W-4	4	4	4
WK4	4	4	4
Wk: 4	4	4	4
Week 4	4	4	4
WEEK:4	4	4	4

Note: Although using the compress function for this particular example would be sufficient, the use of the ANYDIGIT approach allows us to build upon this example. See “Extracting Week Number and Day Number Text String” example.

Extracting Week Number and Day Number from Text String

VISIT	FIRSTNUMLOC	SECALPLOC	LASTNUM	WEEK	DAY
W4D1	2	3	4	4	1
w 4 d 1	3	5	7	4	1
W4 D1	2	4	5	4	1
w-4 d-1	3	5	7	4	1
wk4 d1	3	5	6	4	1
wk4 dy1	3	5	7	4	1
wk 4 dy 1	4	6	9	4	1
WK: 4 DY: 1	5	7	11	4	1
WEEK 4 DAY 1	6	8	12	4	1
week:4 day:1	6	8	12	4	1

When ANY Function Will Just NOT Do

Richann Watson, Karl Miller

Experis, inVentiv Health

Convert Individual Date/Time Components to ISO 8601 Format

MON	DAY	YEAR	HR	MIN	SEC	...	ISODT	_ISOTM	...	NEWDTC
01	01	2016	01	30	45		2016-01-01	01:30:45		2016-01-01T01:30:45
01	01	2016	01	30	--		2016-01-01	01:30		2016-01-01T01:30
01	01	2016	01	30	UN		2016-01-01	01:30		2016-01-01T01:30
01	01	2016	01	--	UN		2016-01-01	01		2016-01-01T01
01	01	2016	UN	30			2016-01-01	-:30		2016-01-01T-:30
01	01	2016	UN	UK	NA		2016-01-01			2016-01-01
01	--	2016	UN	UK	NA		2016-01--			2016-01
NA	--	2016	UN	UK	NA		2016----			2016
UK	01	2016	o1	30			2016---01	-:30		2016---01T-:30
NA	01	2016	UK	30			2016---01	-:30		2016---01T-:30
--	01	2016	NA	30			2016---01	-:30		2016---01T-:30
01	--	2016	O1	30			2016-01--	-:30		2016-01--T-:30
01	NA	2016	01	30			2016-01--	01:30		2016-01--T01:30
UK	UK	2016	01	30			2016----	01:30		2016----T01:30
01	01	UNK	01	30			--01-01	01:30		--01-01T01:30
0I	01	NA		30			----01	-:30		----01T-:30
--	UK	UNK	01	30	45		-----	01:30:45		-----T01:30:45
--	UK	UNK		NA			-----			
11	19	UNK	UK	NA	--		--11-19			--11-19