

# Using Google to Solve SAS Programming Problems

---

Michael Todd

Nth Analytics

418 Wall St Princeton NJ

[www.nthanalytics.com](http://www.nthanalytics.com)



# Google: Knowledge Base

---

- SAS has been around for 30 years
- Most, if not all, programming problems have been solved
  - Just a matter of finding the solution
  - That's where Google comes in
- This paper will present several examples and search strategies

# Why is this Useful?

---

- Programmers naturally rely heavily on code already in place at their companies
- If new situations come up, there may not be code to copy from
- New technologies in SAS may be completely unfamiliar
- With Google, there is always someone to ask for help!

# Sources

---

- Programming solutions come mainly from 4 sources:
  - Conference Papers
  - Books
  - Social Media
  - SAS Documentation

# Google Books

---

- Previously, we actually had to buy the books
- Now, we can read them online, for free
  - An outgrowth of Google's controversial book digitizing project
  - Usually want a little bit of information, a particular tip, or suggestion

# SAS Documentation

---

- Consists of manuals and sample programs
  - Years ago, SAS used to ship the manuals along with the software
  - We had copies in our offices, and would look in the index and try to find the information
- Now, you can easily search online and get exactly what you want
- Documentation is available at [support.SAS.com](http://support.SAS.com)

# Social Media

- SAS-L

An old listsrv, but still used

- SAS Blogs: [blogs.sas.com](http://blogs.sas.com)

- Twitter:

nthanalytics, swaynette, sgoodin,  
beckygraebe, sastips, saspublishing,  
sytruongus, annmariastat, cjdinger  
Hashtags: #sas, #sgf09, etc.

- LinkedIn groups:

Business Intelligence SAS,  
SAS Professionals

## blogs.sas

- » [A Shot in the Arm](#)  
Jason Burke on Health and Life Sciences
- » [BI and the Chicken Pot Pie](#)  
Tammi Kay George on Business Intelligence
- » [bLog-Normal Distribution](#)  
John Sall offers analytic lessons and random observations
- » [Closing the Intelligence Gap](#)  
Gary Cokins on Performance Management
- » [Conversations and Connections](#)  
David B. Thomas on Social Media at SAS
- » [DataFlux Community of Experts](#)  
Providing perspective and insight on issues surrounding data quality and data integration
- » [In Other Words](#)  
Jim Davis writes on the larger issues that affect a global business, and what it takes to be competitive in the 21st century
- » [In the Final Analysis](#)  
Mikael Hagström on the issues facing organizations today
- » [JMP Blog](#)  
Focusing on data visualization, visual Six Sigma, design of experiments and more
- » [Key Happenings at support.sas.com](#)  
Sharing updates and advances in SAS online support

# Conference Papers


---


- Major of source of programming techniques
- Often can lift code straight out
- Tend to be more step-by-step, as opposed to the SAS documentation
- Collected on:
  - [support.SAS.com](http://support.SAS.com)
  - [lexjansen.com](http://lexjansen.com)



# Lexjansen.com

- Has thousands of SAS papers


View my profile on  **Linked in**




**Lex Jansen**  
The Musical Blog

Sitemap

w Favorites To






**Pharmaceutical Programming: Journal**

[Call for Papers](#)


[More information](#)

**Featured new SAS titles**



**Lex Jansen**

Senior Consultant, Clinical Data Strategies at Octagon Research Solutions  
Greater New York City Area



Google™ Custom Search

This searches **9664** SAS papers from SAS Global Forum, SUGI, PharmaSUG, NESUG, SESUG, PhUSE, WUSS, MWSUG, PNWSUG and SCSUG. Alternatively, search for a word based on [title, author or keywords](#).

**4195** SAS papers presented at [SGF/SUGI 1996-2009](#).

**1191** SAS papers presented at [PharmaSUG 2000-2009](#).

**1666** SAS papers presented at [NESUG 1997-2008](#).

**1087** SAS papers presented at [SESUG 1999-2008](#).

**357** SAS papers presented at [Phuse 2005-2008](#).  
-- PhUSE 2009 preview.

**664** SAS papers presented at [WUSS 2003-2008](#).

**159** SAS papers presented at [MWSUG 2001, 2005-2007](#).

**147** SAS papers presented at [PNWSUG 2004-2008](#).

**197** SAS papers presented at [SCSUG 2003-2007](#).

**156** SAS papers related to [CDISC](#).  
Easy access to the [CDISC Forum](#).

# Google Search History

---

- Google stores all of your searches and the results
- This list itself can be searched
  - Pull all SAS searches
- [www.google.com/history](http://www.google.com/history)

# PROC SQL

# My PROC SQL Searches

[Search History](#)[Search the Web](#)

Web History for nthanalytics@gmail.com

## All History

- [Web](#)
- [Images](#)
- [News](#)
- [Products](#)
- [Sponsored Links](#)
- [Video](#)
- [Maps](#)
- [Blogs](#)
- [Books](#)
- [Pause](#)
- [Remove items](#)

## Trends

## Interesting Items

## Bookmarks

Search results 1 - 25 of 98 for SQL - [Browse all history](#)

Sorted by relevance | [Sort by date](#)

Searched for [proc sql like operator](#) - [Viewed 2 results](#)

★ <http://books.google.com/books?id=G8A1C9-17poC&pg=PA47&lpg=PA4...> - Mar 19



[PROC SQL: beyond the basics using SAS](#) - Mar 19

Kirk Paul Lafler

Viewed page [47](#)

Master the language of PROC SQL!

Searched for [SAS PROC SQL begins with operator](#) - [Viewed 1 result](#)

★ [PROC SQL for DATA Step Die-Hards](#) - [ucla.edu](#) - Dec 16

PROC SQL can be rather intimidating for those who have learned SAS data management techniques ... familiar, followed by SQL code that will accomplish the ...

Searched for [sas proc sql retain](#) - [Viewed 3 results](#)

★ [SUGI 24: Efficiency Techniques: SQL vs. Retain Variables](#) - [sas.com](#) - Apr 6

SQL vs. RETAIN. Table 1 contains a portion of 1997 March. Current Population Survey (CPS) survey data. ... result is obtained from SQL with less code and, ...

★ [The RETAIN Statement: One Window into the SAS Data Step Paul ...](#) - [ucla.edu](#) - Apr 6

discusses a PROC SQL alternative. The fourth section of the paper ..... alternative might be the MAX and MIN functions of PROC SQL. (22) proc sql; ...

★ [http://www.experts-exchange.com/Database/SAS/Q\\_24099179.html](http://www.experts-exchange.com/Database/SAS/Q_24099179.html) - Apr 6

Searched for [sas proc sql row difference](#) - [Viewed 1 result](#)

★ [Using SQL to Summarize Data Down Rows and Across Columns](#) - [lexjansen.com](#) - Apr 6

PROC SQL gives users the ability to summarize data ... function in PROC SQL is that it will generally compute ... Guide to the SQL

# SAS Searches: PROC SQL

- PROC SQL noprint
- SAS PROC SQL correlated subquery
- PROC SQL into
- SAS SQL coalesce
- PROC SQL insert
- SAS SQL between operator
- SAS PROC SQL union syntax
- PROC SQL except operator
- PROC SQL into separated by
- SAS PROC SQL intersect
- SAS PROC SQL begins with operator
- PROC SQL like operator
- SAS PROC SQL retain
- SAS PROC SQL row difference

# PROC SQL Problem: Lag Function

- Off the top of head, I can't think of any way to do this. Try Googling.

The screenshot shows a Google search interface with the query "sas proc sql lag" in the search bar. The search results are displayed under the "Web" tab. The first result is a PDF document titled "268-29: Introduction to PROC SQL" from SAS.com, with a description of PROC SQL capabilities and a link to the PDF. The second result is "SUGI 23: An Introduction to PROC SQL" from SAS.com, with a description of PROC SQL and a link to the document. A red box highlights the text "Looks promising!" with a red arrow pointing to the second search result. Below the search results, there is a link to "Show more results from www2.sas.com". The third result is an email from the SAS-L listserve archive, dated August 30, 2005, with the subject "In proc SQL, how to do something similar to lag function in >data step? ...". A red circle highlights the subject line of this email.

Google   [Adv](#)

Web

[PDF](#) [268-29: Introduction to PROC SQL](#)  
File Format: PDF/Adobe Acrobat  
**PROC SQL** can be used to retrieve, update, and report on information from **SAS** data ...  
**PROC SQL** can not only retrieve information without having to learn **SAS** ... can be used in an expression to create a new variable except **LAG**, **DIF**, and ...  
[www2.sas.com/proceedings/sugi29/268-29.pdf](http://www2.sas.com/proceedings/sugi29/268-29.pdf) - [Similar](#) -

[SUGI 23: An Introduction to PROC SQL](#)  
expression to create a new variable except **LAG**, **DIF**, and **SOUND**: ..... **SAS** and **PROC SQL** are registered trademarks of the **SAS** Institute Inc., .... NOTE: **PROC SQL SET OPTION NOEXEC** AND WILL CONTINUE TO CHECK THE SYNTAX OF STATEMENTS. ...  
[www2.sas.com/proceedings/sugi23/Handson/p130.pdf](http://www2.sas.com/proceedings/sugi23/Handson/p130.pdf) - [Similar](#) -   
by RK Gusinow - [Related articles](#) - [All 3 versions](#)

[SAS-L archives -- August 2005, week 5 \(#162\)](#) - 3:17pm  
Aug 30, 2005 ... COM> Sender: "**SAS**(r) Discussion" <[SAS-L@LISTSERV.UGA.EDU](mailto:SAS-L@LISTSERV.UGA.EDU)>Subject: In **proc SQL**, how to do something similar to **lag** function in >data step? ...  
[www.listserv.uga.edu/cgi-bin/wa?A2=ind0508e...sas-l...](http://www.listserv.uga.edu/cgi-bin/wa?A2=ind0508e...sas-l...) - [Cached](#) - [Similar](#) -

# Solution from SAS-L

Here's an example. First use the LAG function in a DATA step.

```
data age_lag;
set sashelp.class;
age_lag = lag(age);
keep name age age_lag;
run;
```

To emulate that in SQL, first add a column with serial numbers for the rows:

```
data numbered;
set sashelp.class;
sn + 1;
run;
```

Then code a subquery to get the AGE value from the "previous" row:

```
proc sql;
create table age_lag_sql as
select class.name,
       class.age,
       (select lag.age
        from numbered as lag
        where class.sn - lag.sn = 1) as age_lag
from numbered as class
order by class.sn
;
```

# Implementation

- I was able to use the code with only slight modification!

```
data work.OneWayFreqs;
set work.OneWayFreqs2;
retain sn 0;
sn+1;
run;

proc sql;
create table work.OneWayFreqs_lag as
select class.param,
class.window,
class.base,
class.cumFrequency,
class.cumPercent,
(select lag.CumPercent
 from work.OneWayFreqs as lag
 where class.sn - lag.sn = 1) as CumPct_lag,
(select lag2.Base
 from work.OneWayFreqs as lag2
 where class.sn - lag2.sn = 1) as base_lag
from work.OneWayFreqs as class
order by class.cumPercent
;
```

BASE	cumFrequency	cumPercent	CumPct_lag	base_lag
46	1	0.40322581		
48	2	0.80645161	0.40322581	46
49	6	2.41935484	0.80645161	48
51	12	4.83870968	2.41935484	49
52	20	8.06451613	4.83870968	51
53	25	10.0806452	8.06451613	52
54	30	12.0967742	10.0806452	53
55	35	14.1129032	12.0967742	54
56	37	14.9193548	14.1129032	55
57	41	16.5322581	14.9193548	56
58	46	18.5483871	16.5322581	57
59	53	21.3709677	18.5483871	58
60	58	23.3870968	21.3709677	59
61	61	24.5967742	23.3870968	60
62	68	27.4193548	24.5967742	61
63	71	28.6290323	27.4193548	62
64	77	31.0483871	28.6290323	63

25<sup>th</sup> percentile cutoff



RTF

# SAS Searches: RTF

- SAS RTF startpage
- SAS ODS PROC REPORT RTF image
- SAS RTF hidden text
- SAS RTF margins
- SAS ODS RTF text=
- SAS RTF tags
- SAS RTF different style header and body
- RTF underline
- SAS ODS RTF TOC
- RTF font code
- SAS RTF page x of y
- SAS ODS RTF page 1 of
- RTF escape character indent
- RTF \nofpages
- SAS ODS RTF ^page
- SAS ODS RTF title J=L
- ODS RTF J=L J=R title on same line
- SAS ODS RTF escapechar ^left

# SAS RTF Page x of y

[Advanced Search](#)

Web [+ Show options...](#)

Results 1 - 10 of about

[PDF] [Page X of Y with PROC REPORT](#) - 12:14pm

File Format: PDF/Adobe Acrobat - [View](#)

SAS Institute Inc. (b), "Q: In ODS RTF, can I get my **page** numbers in **page X of Y** format?"

SAS FAQ #. 4010. Product or Solution: Base SAS. ...

[changchung.com/download/pageXofY\\_draft.pdf](#) - [Similar](#) - [🗨](#) [📄](#) [🗕](#)

by CY Chung - 2005 - [Related articles](#)

Bad: macro-based

[15727 - Writing PAGE X OF Y in RTF does not work with BODYTITLE](#) - 12:14pm

Raw RTF code or the special function PAGEOF can be used on a TITLE or FOOTNOTE statement to write **PAGE X of Y** text in RTF output. However, if the BODYTITLE ...

[support.sas.com/kb/15727.html](#) - [Cached](#) - [Similar](#) - [🗨](#) [📄](#) [🗕](#)

Interesting, but not practical

[alternate \(.sas\) - SAS Customer Support](#) - 12:14pm

The length of the special text must be long enough to hold the generated **Page X of Y** text.

Does not apply to RTF files because the internal pagination is ...

[support.sas.com/rnd/base/ods/.../pageofpp\\_public.sas](#) - [Cached](#) - [Similar](#) - [🗨](#) [📄](#) [🗕](#)

Bad: too complicated

[+ Show more results from support.sas.com](#)

[PDF] [ODS RTF: Practical Tips](#) - 4 visits - 12:15pm

File Format: PDF/Adobe Acrobat - [View](#)

Since ODS RTF has been available in production in SAS .... "**Page x of y**" can be inserted into the footer using the following code: ...

[www.nesug.org/Proceedings/nesug03/at/at007.pdf](#) - [Similar](#) - [🗨](#) [📄](#) [🗕](#)

Ok, but old

[MeasureIT - Issue 5.11 - SAS 9 Tips - Part II](#)

In SAS 9 if you want to add **page** numbers to your RTF output in the form of "**page x of y**" then you can use the in-line formatting by specifying an escape ...

[www.cmg.org/measureit/issues/mit46/m\\_46\\_1.html](#) - [Cached](#) - [Similar](#) - [🗨](#) [📄](#) [🗕](#)

Winner!

# Bad: Macro-Based

- Forget it: way too complicated

```
/* example usage */
options linesize=64 nonumber nodate;

%pageXofY(report=%nrstr(

  proc report data=one nowd headline missing;
    column var;
    define var / display;

    compute after _page_;
      call execute('%let page = %eval(&page. + 1);');
      length _XofY $&len.;
      _XofY = symget('page') || ' of ' || symget('pages');
      line 'page ' _XofY $&len..;
    endcomp;
  run;

))
```

# Interesting, but not Practical

- PROC TEMPLATE would be a good solution, but implementing it would be a big change.

An alternative is to use PROC TEMPLATE to write Page X of Y, which will place the pageof information in the upper right corner of each page. The code below will apply to the default page number location only.

```
ods escapechar='^';

proc template;
  define style styles.test;
    parent=styles.rtf;
    style pageno from pageno /
      font=fonts("strongfont")
      posttext=" of ^(lastpage) ";
  end;
run;
```

# Bad: Too Complicated

- Macro goes on and on; not sure what it is doing.

```
%pageofpp(infile="test1.pdf",outfile="test2.pdf",text=&pagnumt,type=pdf)

* Example 2 - Centered in a TXT file ;
%let pagnumt=[Page_00000_of_99999];
footnote1 "%sysfunc(repeat(%str( ),50))&pagenumt";
...
%pageofpp(infile="t1.txt",outfile="t2.txt",text=&pagnumt,type=txt,justify=c)
-----*/

%local parmerr
      infilen outfilen
      numpages numpagel
      slength rtfcommand lpageof
;
%*-----
Validate parameters
-----;
%let parmerr=0;

%if (^%length(&infile)) %then %do;
  %put ERROR: INFILE not specified.;
  %let parmerr=1;
%end;

%if (^%length(&outfile)) %then %do;
  %put ERROR: OUTFILE not specified.;
  %let parmerr=1;
%end;

%else %if ("%infile" = "%outfile") %then %do;
  %put ERROR: OUTFILE same file as INFILE.;
  %let parmerr=1;
%end;
```

# Ok, But Old

- This solution works, and it's easy to copy, but SAS Version 9 has a simpler method.

"Page x of y" can be inserted into the footer using the following code:

```
footnote1 j=1 '{Page \field {\*\fldinst  
PAGE \*MERGEFORMAT}} { of \field  
\*\fldinst NUMPAGES \*MERGEFORMAT}}';
```

Similarly, "Page y" can be inserted into the footer using the following code:

```
footnote1 j=1 '{ Page \field {\*\fldinst  
NUMPAGES \*MERGEFORMAT}}';
```

When the document is opened the page numbers may initially appear wrong, especially in Word 97. This is because Word has not refreshed the fields. Doing "Print preview" can solve this problem.

# Winner!

- Newest, simplest solution: also has PDF method

## SAMPLE CODE FOR RTF

```
ods escapechar = '\';  
title 'This document will have page x of y '  
      j=r 'Page \{pageof}' ;  
ods rtf file='c:\test.rtf' ;  
proc print data=sashelp.prdsale;  
run;  
ods rtf close;
```

## SAMPLE CODE FOR PDF

```
ods escapechar = '\';  
title 'This document will have page x of y '  
      j=r 'Page \{thispage} of \{lastpage}' ;  
ods pdf file='c:\test.pdf' ;  
proc print data=sashelp.prdsale;  
run;  
ods pdf close;
```



# Age Calculation

# Age Calculation

---

- Most companies use this formula for age in years:
  - $\text{ageyear} = \text{floor}((\text{vdate} - \text{birthdt} + 1) / 365.25);$
- This can be extended to compute age in months
  - $\text{agemos} = \text{floor}((\text{vdate} - \text{birthdt} + 1) * 12 / 365.25);$
- Is this the most accurate method?
  - Try Googling “SAS age in months”

# Google Age Calculation

Web Images Videos Maps News Shopping Gmail more ▾

nthanalytics@gmail.com |



[Advanced Search](#)

Web [+ Show options...](#)

Results 1 - 10 of about 5,030,000 for sas

[Accurately calculating age in only one line](#) - 3 visits - 10/26/08

%macro **age**(date,birth); floor ((intck('month',&birth,&date) - (day(&date) < day(&birth))) / 12) %mend **age**; The macro is used in a **SAS** DATA step as follows: ...

[www.pauldickman.com/teaching/sas/age.php](http://www.pauldickman.com/teaching/sas/age.php) - [Cached](#) - [Similar](#) -

[24808 - Calculating Age with Only One Line of Code](#)

Aug 25, 2009 ... This sample shows how to easily determine a person's **age**, based on a given ... Often, **SAS** software users attempt to compute **age** using an expression such as: ... %macro **age**(date,birth); floor ((intck('month',&birth,&date) ...

[support.sas.com/kb/24808.html](http://support.sas.com/kb/24808.html) - [Cached](#) - [Similar](#) -

[24567 - Calculate a person's age](#)

Aug 25, 2009 ... Determine a person's current **age** using their date of birth. ... Note: This is Example 6.11 from Combining and Modifying **SAS** Data Sets - Examples. ... Divide the number of \*/ /\* **months** by 12 to produce the number of years ...

[support.sas.com/kb/24567.html](http://support.sas.com/kb/24567.html) - [Cached](#) - [Similar](#) -

[+ Show more results from support.sas.com](#)

[SAS FAQ - Date Processing with SAS Software](#)

DATA getage; birthday='25Dec1972'd; datevar=TODAY(); /\* Divide **months** by 12 to get years \*/ **age**=INT(INTCK('MONTH',birthday,datevar)/12); /\* Adjust **age** if ...

[www.hollandnumerics.co.uk/sasfaq/SASFAQ1.HTM](http://www.hollandnumerics.co.uk/sasfaq/SASFAQ1.HTM) - [Cached](#) - [Similar](#) -

[PDF] [DATES in SAS, an example of Calculating Age From Dates](#)

File Format: PDF/Adobe Acrobat - [View](#)

# SAS Knowledge Base

The screenshot shows the SAS Knowledge Base website interface. At the top left is the SAS logo with the tagline 'THE POWER TO KNOW.' Below the logo is a search bar. A navigation bar contains links for 'support.sas.com', 'KNOWLEDGE BASE', 'SUPPORT', 'LEARNING CENTER', and 'COMMUNITY'. The main content area is titled 'KNOWLEDGE BASE / SAMPLES & SAS NOTES'. On the left is a sidebar menu with categories like 'SAS Software', 'System Requirements', 'Install Center', 'Third-Party Software Reference', 'Documentation', 'Papers', 'Samples & SAS Notes' (with sub-links for 'Browse by Topic', 'Search Samples', 'Search Usage Notes', and 'Search Problem Notes'), and 'Focus Areas'. The main article is titled 'Sample 24808: Calculating Age with Only One Line of Code' and has tabs for 'Details', 'About', and 'Rate It'. The article text begins with a quote from William Kreuter, a senior computer specialist, and discusses the need for a one-line SAS solution to calculate age based on a birthday.

# Standard Algorithms for Age

## What Doesn't Work

Often, SAS software users attempt to compute age using an expression such as:

```
age = (somedate - birth) / 365.25;
```

where somedate and birth are SAS date variables (or constants or expressions). Clearly this usually doesn't return an integer and therefore it is not stating an age according to colloquial usage. That problem can be addressed by:

```
age = floor( (somedate - birth) / 365.25);
```

Now we're at least getting integers. In fact, for most dates in a given year this statement does produce the correct result. But in most years, age will increment on the wrong day. To account for the Gregorian calendar's idiosyncrasies, some users make attempts such as:

```
age = floor( (somedate - birth) / 365.2422);
```

However, extending the denominator to any number of significant digits doesn't help. Astronomers define several kinds of "years" for various technical uses, but the Gregorian calendar uses a different concept of "year" in which there are always either 365 or 366 days. No algorithm of this kind perfectly models such an interval.

# Solution Using SAS Date Functions

```
intck('month',birth,somedate)
```

returns the number of times the first day of a month is passed between birth and somedate. An enhancement is needed to alter this into the number of times the same day of the starting month is passed. This simply consists of subtracting one month if the day number of somedate is earlier than the day number of birth. Although one could program this concept using a separate if-then statement, it can be calculated more concisely as a logical expression returning a 0 or 1 value. The 0 or 1 is then subtracted from the result of intck, as in the following example.

```
intck('month',birth,somedate)  
- (day(somedate) < day(birth))
```

This now gives exactly the correct number of months for any pair of dates.

## **A one-line solution**

Converting months to years, we get:

```
age = floor  
((intck('month',birth,somedate)  
- (day(somedate) < day(birth))) / 12);
```

# Are the Results Consistent?

## ■ Old Method

$$\text{floor}((\text{VDATE} - \text{BIRTHDT} + 1) * 12 / 365.25)$$

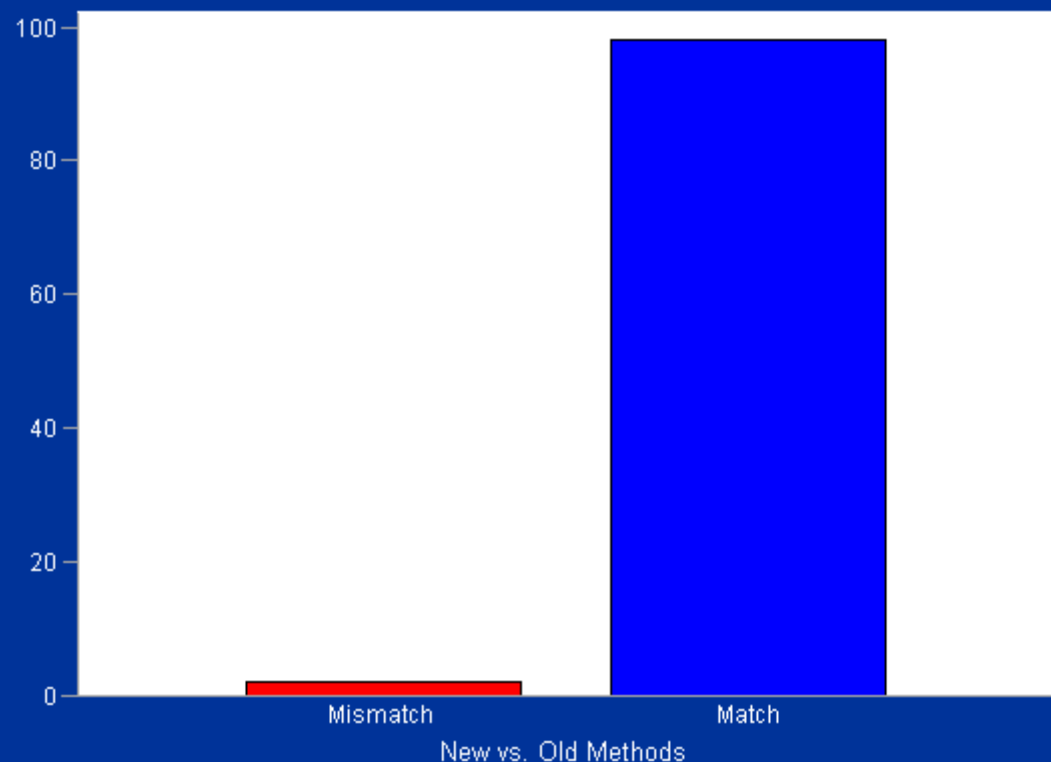
## ■ New Method

$$\text{intck}(\text{'month'}, \text{BIRTHDT}, \text{VDATE}) - (\text{day}(\text{VDATE}) < \text{day}(\text{BIRTHDT}))$$

## ■ Test Data: pediatric study (infants $\leq 2$ yrs), 3077 records

New vs. Old Methods		
new_vs_old	Frequency	Percent
Mismatch	68	2.08
Match	3209	97.92

Percent of Total Frequency



# Examine the Differences

Listing of Age (Months) Calculation Mismatches

Subject Number	Date of Birth	Visit Date	Age (Months) - New Method	Age (Months) - Old Method
101048	28APR2004	27NOV2006	✓ 30	31
101053	23JUN2005	22NOV2006	✓ 16	17
101054	09JAN2006	08FEB2007	✓ 12	13
101090	10DEC2004	10MAR2006	✓ 15	14
101094	29DEC2005	29MAR2006	✓ 3	2
101138	13APR2006	11JAN2007	✓ 8	9
101140	14APR2005	12JAN2006	✓ 8	9
101140	14APR2005	13JAN2006	✓ 8	9
101143	01APR2006	31JUL2007	✓ 15	16
101197	02MAY2005	31JAN2007	✓ 20	21
101232	06FEB2006	06SEP2006	✓ 7	6
101233	19FEB2005	19SEP2006	✓ 19	18
101236	03NOV2005	03APR2007	✓ 17	16
101238	15AUG2005	14DEC2006	✓ 15	16
300001	28DEC2004	27JAN2006	✓ 12	13
300001	28DEC2004	28MAR2006	✓ 15	14
300020	19AUG2003	18SEP2006	✓ 36	37
300063	07AUG2004	06FEB2006	✓ 17	18
300120	31OCT2005	31JUL2007	✓ 21	20

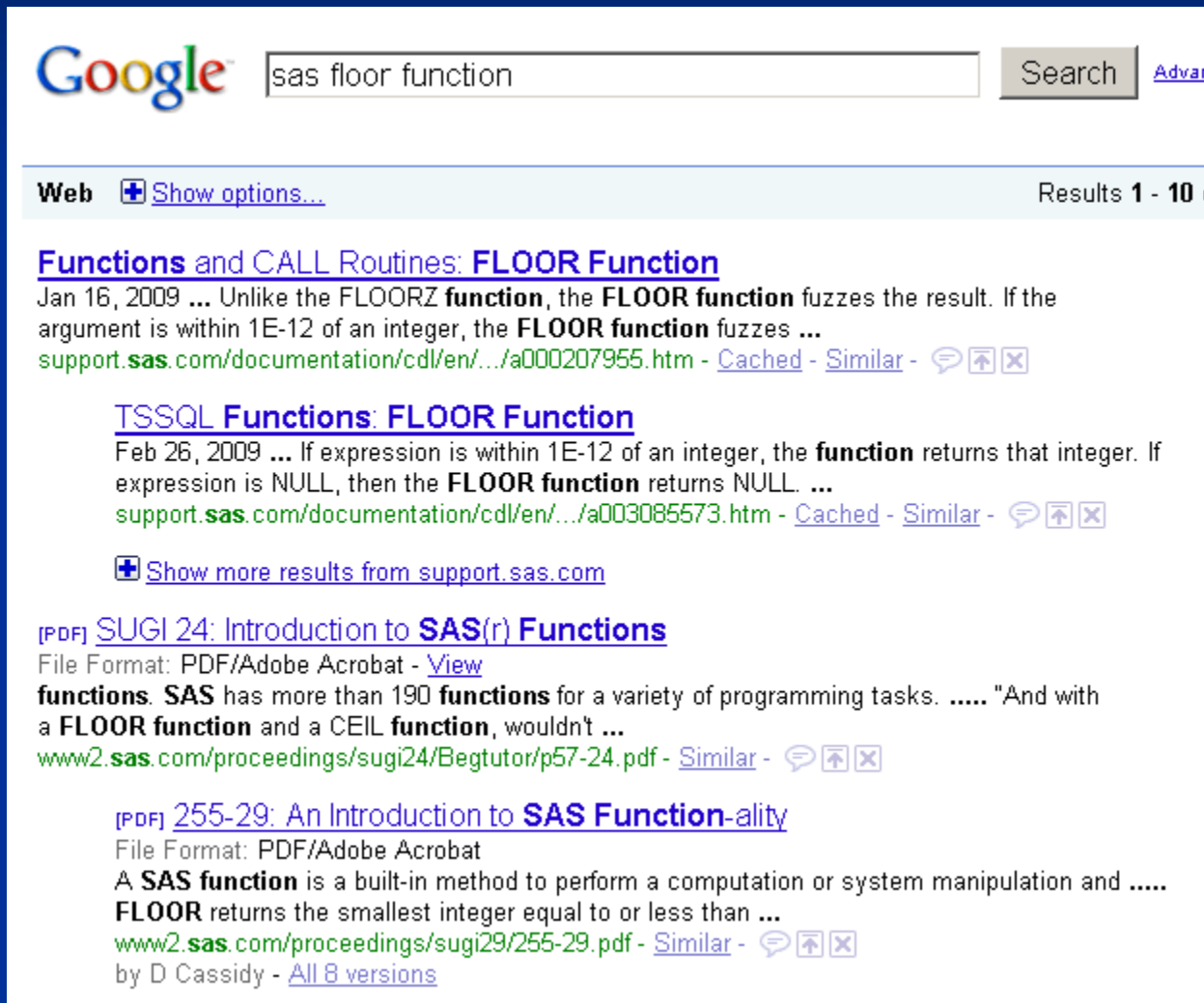
Does not reach 31 months  
until 28NOV2006

Reached 3 months on  
29MAR2006

Etc. for all 68 differences



# What does the FLOOR function do?



Google  Search [Advanced](#)

Web [+ Show options...](#) Results 1 - 10 of 10

[Functions and CALL Routines: FLOOR Function](#)  
Jan 16, 2009 ... Unlike the FLOORZ **function**, the **FLOOR function** fuzzes the result. If the argument is within 1E-12 of an integer, the **FLOOR function** fuzzes ...  
[support.sas.com/documentation/cdl/en/.../a000207955.htm](http://support.sas.com/documentation/cdl/en/.../a000207955.htm) - [Cached](#) - [Similar](#) - [Comment](#) [Share](#) [Close](#)

[TSSQL Functions: FLOOR Function](#)  
Feb 26, 2009 ... If expression is within 1E-12 of an integer, the **function** returns that integer. If expression is NULL, then the **FLOOR function** returns NULL. ...  
[support.sas.com/documentation/cdl/en/.../a003085573.htm](http://support.sas.com/documentation/cdl/en/.../a003085573.htm) - [Cached](#) - [Similar](#) - [Comment](#) [Share](#) [Close](#)

[+ Show more results from support.sas.com](#)

[\[PDF\] SUGI 24: Introduction to SAS\(r\) Functions](#)  
File Format: PDF/Adobe Acrobat - [View](#)  
**functions. SAS** has more than 190 **functions** for a variety of programming tasks. .... "And with a **FLOOR function** and a **CEIL function**, wouldn't ...  
[www2.sas.com/proceedings/sugi24/Begtutor/p57-24.pdf](http://www2.sas.com/proceedings/sugi24/Begtutor/p57-24.pdf) - [Similar](#) - [Comment](#) [Share](#) [Close](#)

[\[PDF\] 255-29: An Introduction to SAS Function-ality](#)  
File Format: PDF/Adobe Acrobat  
A **SAS function** is a built-in method to perform a computation or system manipulation and .... **FLOOR** returns the smallest integer equal to or less than ...  
[www2.sas.com/proceedings/sugi29/255-29.pdf](http://www2.sas.com/proceedings/sugi29/255-29.pdf) - [Similar](#) - [Comment](#) [Share](#) [Close](#)  
by D Cassidy - [All 8 versions](#)

CEIL returns the smallest integer equal to or greater than your value.

FLOOR returns the smallest integer equal to or less than your value.

INT returns the integer portion of your value. At first glance, someone might think that FLOOR and INT would always return the same value. However, if you look at the example with the negative number, you'll see that INT matches FLOOR for positive values but it matches CEIL for negative values.









# Confidence Interval for Median




# Confidence Interval for Median




Google  Search [Adva](#)

Web [+ Show options...](#) Results 1 - 10 of about 179,000

[PDF] [Confidence Intervals in Analysis and Reporting of Clinical Trials ...](#)  
File Format: PDF/Adobe Acrobat - [View](#)  
distribution-free **confidence intervals** for the **median** response within each treatment group using **SAS**. INTRODUCTION. **Confidence interval** estimation and ...  
[www.lexjansen.com/pharmasug/2003/.../sp050.pdf](http://www.lexjansen.com/pharmasug/2003/.../sp050.pdf) - [Similar](#) -   

[PDF] [Calculating a Nonparametric Estimate and Confidence Interval Using ...](#)  
File Format: PDF/Adobe Acrobat - [View](#)  
Calculating a Nonparametric Estimate and **Confidence Interval** Using **SAS** .... As described earlier, it is the **median** of the X\*Y differences calculated ...  
[www.lexjansen.com/pharmasug/2000/Coders/cc01.pdf](http://www.lexjansen.com/pharmasug/2000/Coders/cc01.pdf) - [Similar](#) -   

[The Mean, Median, and Confidence Intervals of the Kaplan-Meier ...](#)  
does not intersect a **confidence interval**. The **confidence intervals** in **SAS** Proc Lifetest for the **median**. (quartiles) are given by: 10:50 ¼ šť : šť ¼ Šťž ¼ ...  
[pubs.amstat.org/doi/pdf/10.1198/tast.2009.0015](http://pubs.amstat.org/doi/pdf/10.1198/tast.2009.0015) - [Similar](#) -     
by C Barker - 2009 - [Cited by 1](#) - [Related articles](#) - [All 2 versions](#)

[SAS-L archives -- January 2001, week 5 \(#142\)](#)  
Jan 31, 2001 ... Andrew Dixon wrote: > > Hi all, > How do I get **SAS** to return a **median** and put 95% **confidence intervals** around it? ...  
[www.listserv.uga.edu/cgi-bin/wa?A2=ind0101e...sas-l...](http://www.listserv.uga.edu/cgi-bin/wa?A2=ind0101e...sas-l...) - [Cached](#) - [Similar](#) -   

# CI for Median: First Search

Not promising ...

such cases, distribution-free confidence intervals should be constructed. A distribution-free interval sometimes may not exist, and its length is generally longer than the corresponding distribution-dependent interval for a particular distribution. This is the price that one pays for not making the distribution assumption (Hahn and Meeker, 1991). So, a distribution-dependent confidence interval should be chosen whenever there is solid evidence that the data follows a tractable distribution.

## DISTRIBUTION-DEPENDENT CONFIDENCE INTERVAL

If the assumption that the data are normally distributed is valid, one can construct confidence intervals for the mean treatment difference. The general form of a confidence interval for the mean difference between two treatment groups (Group A and Group B) is

$$\bar{Y}_a - \bar{Y}_b \pm t_{1-\alpha/2, df} * S(\bar{Y}_a - \bar{Y}_b) \quad (1)$$

code. The following code using SAS version 6.09 represents one way to obtain the 95% confidence interval for the mean treatment difference from an ANOVA model.

```
*-----*;  
* The following statements get output *;  
* datasets containing statistics *;  
* needed for calculation of confidence*;  
* interval *;  
*-----*;  
proc glm data=final outstat=glmdt  
  noprint;  
  class  &trt &str &invcd;  
  model  &dep=&indp;  
  lsmeans &trt/pdiff stderr tdiff  
         out=lsmeandt;  
  
*-----*;  
* The following statements get the *;  
* degree of freedom from the model *;  
*-----*;  
data _null_; set glmdt;  
  if _type_='ERROR';  
  call symput('df', df);  
  
*-----*;  
* The following statements get the *;  
* LSMEANS from the model *;  
*-----*;  
data _null_; set lsmeandt;  
  if &trt='&control' then call  
  symput('pLSM', lsmean);  
  if &trt='&trtl' then call  
  symput('tLSM', lsmean);  
*-----*;
```

Paper is from 2003,  
but they are using  
Version 6.09?

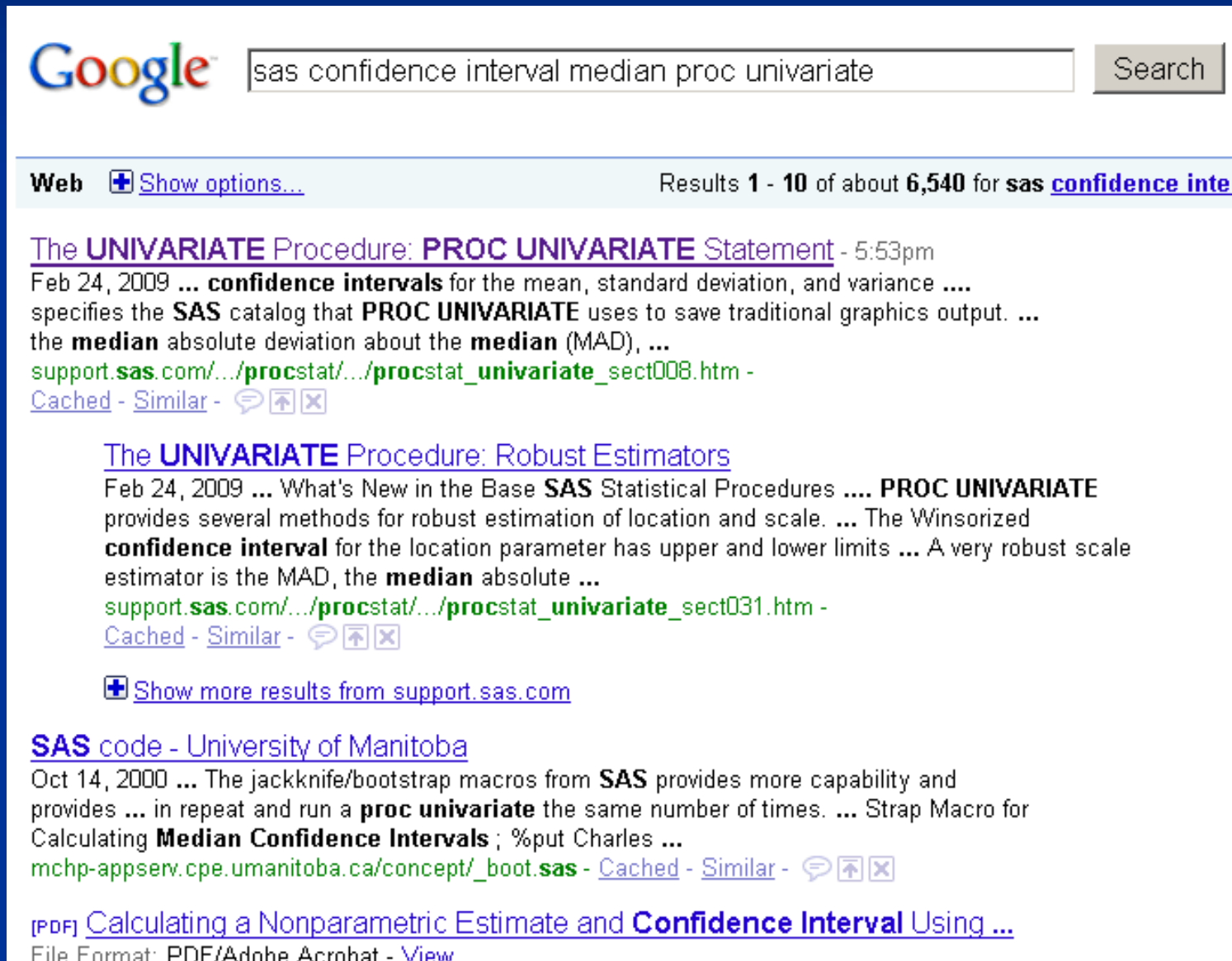
What does PROC  
GLM have to do  
with the median?

Want something  
quick. Don't want to  
deal with formulas.

Slide 36

# Try Again




Let's see if we can use PROC UNIVARIATE ...






The screenshot shows a Google search interface. The search bar contains the text "sas confidence interval median proc univariate". Below the search bar, the results are displayed. The first result is titled "The UNIVARIATE Procedure: PROC UNIVARIATE Statement" and includes a snippet about confidence intervals and the median absolute deviation. The second result is titled "The UNIVARIATE Procedure: Robust Estimators" and includes a snippet about robust estimation methods. The third result is titled "SAS code - University of Manitoba" and includes a snippet about jackknife/bootstrapping macros. The search results are for "sas confidence interval" and show 1-10 of about 6,540 results.




Google  Search

Web [+ Show options...](#) Results 1 - 10 of about 6,540 for **sas confidence inte**

[The UNIVARIATE Procedure: PROC UNIVARIATE Statement](#) - 5:53pm  
Feb 24, 2009 ... **confidence intervals** for the mean, standard deviation, and variance .... specifies the **SAS** catalog that **PROC UNIVARIATE** uses to save traditional graphics output. ... the **median** absolute deviation about the **median** (MAD), ...  
[support.sas.com/.../procstat/.../procstat\\_univariate\\_sect008.htm](http://support.sas.com/.../procstat/.../procstat_univariate_sect008.htm) -  
[Cached](#) - [Similar](#) -   

[The UNIVARIATE Procedure: Robust Estimators](#)  
Feb 24, 2009 ... What's New in the Base **SAS** Statistical Procedures .... **PROC UNIVARIATE** provides several methods for robust estimation of location and scale. ... The Winsorized **confidence interval** for the location parameter has upper and lower limits ... A very robust scale estimator is the MAD, the **median** absolute ...  
[support.sas.com/.../procstat/.../procstat\\_univariate\\_sect031.htm](http://support.sas.com/.../procstat/.../procstat_univariate_sect031.htm) -  
[Cached](#) - [Similar](#) -   

[+ Show more results from support.sas.com](#)

[SAS code - University of Manitoba](#)  
Oct 14, 2000 ... The jackknife/bootstrapping macros from **SAS** provides more capability and provides ... in repeat and run a **proc univariate** the same number of times. ... Strap Macro for Calculating **Median Confidence Intervals** ; %put Charles ...  
[mchp-appserv.cpe.umanitoba.ca/concept/\\_boot.sas](http://mchp-appserv.cpe.umanitoba.ca/concept/_boot.sas) - [Cached](#) - [Similar](#) -   

[Calculating a Nonparametric Estimate and Confidence Interval Using ...](#)  
File Format: PDF/Adobe Acrobat - [View](#)

# Proc Univariate Search Results

- The UNIVARIATE Procedure
  - Overview: UNIVARIATE Procedure
- Getting Started: UNIVARIATE Procedure
- Syntax: UNIVARIATE Procedure
- Details: UNIVARIATE Procedure
- Examples: UNIVARIATE Procedure
  - Computing Descriptive Statistics for Multiple Variables
  - Calculating Modes
  - Identifying Extreme Observations and Extreme Values
  - Creating a Frequency Table
  - Creating Plots for Line Printer Output
  - Analyzing a Data Set With a FREQ Variable
  - Saving Summary Statistics in an OUT= Output Data Set
  - Saving Percentiles in an Output Data Set
  - Computing Confidence Limits for the Mean, Standard Deviation, and Variance
  - Computing Confidence Limits for Quantiles and Percentiles

## PROC UNIVARIATE Statement

**PROC UNIVARIATE** <options> ;

The PROC UNIVARIATE statement is required to invoke the UNIVARIATE procedure. You can use the PROC UNIVARIATE statement by itself to request a variety of statistics for summarizing the data distribution of each analysis variable:

- sample moments
- basic measures of location and variability
- confidence intervals for the mean, standard deviation, and variance
- tests for location
- tests for normality
- trimmed and Winsorized means
- robust estimates of scale
- quantiles and related confidence intervals
- extreme observations and extreme values
- frequency counts for observations
- missing values

⇒ Yes!

# Proc Univariate Solution

The ODS SELECT statement restricts the output to the "Quantiles" table; see the section [ODS Table Names](#). The CIQUANTNORMAL option produces confidence limits for the quantiles. As noted in [Output 4.10.1](#), these limits assume that the data are normally distributed. You should check this assumption before using these confidence limits. See the section [Shapiro-Wilk Statistic](#) for information about the Shapiro-Wilk test for normality in PROC UNIVARIATE; see [Example 4.19](#) for an example that uses the test for normality.

```
title 'Analysis of Female Heights';
ods select Quantiles;
proc univariate data=Heights ciquantnormal(alpha=.1);
    var Height;
run;
```

## Analysis of Female Heights

The UNIVARIATE Procedure  
Variable: Height (Height (in))

Quantiles (Definition 5)			
Quantile	Estimate	90% Confidence Limits	
		Assuming Normality	
100% Max	70.0		
99%	70.0	68.94553	70.58228
95%	68.6	67.59184	68.89311
90%	67.5	66.85981	68.00273
75% Q3	66.0	65.60757	66.54262
50% Median	64.4	64.14564	64.98770

# Encoding Issues



# Language Encoding Problem

- We received a dataset from a Chinese affiliate we could not open.

```
1      options nofmterr;
2      libname in 'C:\test_drug\china';
NOTE: Libref IN was successfully assigned as follows:
      Engine:          V9
      Physical Name: C:\test_drug\china
3
4      data eg;
5      set in.eg;
6      run;

NOTE: Format VISF was not found or could not be loaded.
NOTE: Format ECGRSLF was not found or could not be loaded.
NOTE: Format YNF was not found or could not be loaded.
ERROR: Some character data was lost during transcoding in the dataset IN.EG.
ERROR: Some character data was lost during transcoding in the dataset IN.EG.
ERROR: Some character data was lost during transcoding in the dataset IN.EG.
NOTE: The data step has been abnormally terminated.
NOTE: The SAS System stopped processing this step because of errors.
NOTE: SAS set option OBS=0 and will continue to check statements. This may cause
NOTE: There were 226 observations read from the data set IN.EG.
WARNING: The data set WORK.EG may be incomplete.  When this step was stopped ther
NOTE: DATA statement used (Total process time):
      real time          0.18 seconds
      cpu time           0.00 seconds
```

# False Starts

---

- Step 1: Try Googling the error message:
  - “SAS Some character data was lost during transcoding”
  - This resulted in a series of system manual references that were not particularly helpful
  - Looking for some quick that just works

# Try Again with Another Search



sas chinese transcoding

Search

[Advanced Search](#)

Web [+ Show options...](#)

Results 1 - 10 of

[16520 - KCTV function does not \*\*transcode\*\* some Simplified \*\*Chinese\*\* ...](#)

The KCTV function does not **transcode** some Simplified **Chinese** characters correctly. Select the Hot Fix tab in this note to access the hot fix for this issue.

[support.sas.com/kb/16520](http://support.sas.com/kb/16520) - [Cached](#) - [Similar](#) -

Not  
promising

[Encoding Values in \*\*SAS\*\* Language Elements: SBCS, DBCS, and Unicode ...](#)

Feb 26, 2009 ... SBCS, DBCS, and Unicode Encoding Values Used to **Transcode** Data ...  
ms-936, zwin, Simplified **Chinese** PCMS. ms-949, kwin, Korean PCMS ...

[support.sas.com/documentation/cdl/en/.../a002607278.htm](http://support.sas.com/documentation/cdl/en/.../a002607278.htm) - [Cached](#) - [Similar](#) -

Maybe ...

[+ Show more results from support.sas.com](#)

[\[PDF\] SUGI 28: Multi-Lingual Computing with the 9.1 \*\*SAS\*\*\(r\) Unicode Server](#)

File Format: PDF/Adobe Acrobat - [View as HTML](#)

Japanese Characters, or **Chinese** characters etc. Until 9.1, however, a **SAS** session could not represent .... 8.2) to force **SAS** to **transcode** the data on input. ...

[www2.sas.com/proceedings/sugi28/281-28.pdf](http://www2.sas.com/proceedings/sugi28/281-28.pdf) - [Similar](#) -

[SAS 9.1 National language support \(NLS\): user's guide - Google Books Result](#)

by **SAS** Institute - 2004 - Computers - 448 pages

Difference between Encoding and **Transcoding** Encoding establishes the default working environment for your **SAS** session. For example, the Windows Latin1 ...

[books.google.com/books?isbn=1590471946...](http://books.google.com/books?isbn=1590471946...) -

# Chinese Encoding Options Fail

## Windows Double-Byte Encodings

PLATFORM	LANGUAGE	ENCODING=VALUE
Windows	Japanese	shift-jis or ms-932
Windows	Korean	euc-kr or ms-949
Windows	Simplified Chinese	euc-cn or ms-936
Windows	Traditional Chinese	big5 or ms-950

From the documentation, it looks like this should have worked

```
data eg;  
set in.eg(encoding='ms-950');  
run;
```

```
NOTE: Format VISF was not found or could not be loaded.  
NOTE: Format ECGRSLF was not found or could not be loaded.  
NOTE: Format YNF was not found or could not be loaded.  
ERROR: Some character data was lost during transcoding in the dataset  
ERROR: Some character data was lost during transcoding in the dataset  
ERROR: Some character data was lost during transcoding in the dataset  
NOTE: The data step has been abnormally terminated.  
NOTE: The SAS System stopped processing this step because of errors.
```

# More Encoding Options

support.sas.com | KNOWLEDGE BASE | SUPPORT | LEARNING CENTER | COMMUNITY

Product Documentation > SAS 9.2 Documentation

## SAS(R) 9.2 National Language Support (NLS): Reference Guide

PDF | Purchase

Contents | **About**

- What's New in SAS 9.2 National Language Support (NLS) Reference Guide
- NLS Concepts
- SAS Language Elements for NLS
  - Data
    - Data Set Options for NLS
      - Data Set Options for NLS by Category
      - ENCODING= Data Set Option**
      - OUTREP= Data Set Option
    - Formats for NLS
    - Functions for NLS
    - Informats for NLS

Previous Page | Next Page

### ENCODING= Data Set Option

Overrides the encoding to use for reading or writing a SAS data set.

**Valid in:** DATA step and PROC steps

**Category:** Data Set Control

#### Syntax Description

Syntax	Syntax Description
ANY	specifies that no transcoding occurs.
<b>ASCIANY</b>	specifies that no transcoding occurs when the mixed encodings are ASCII encodings.
EBCDICANY	specifies that no transcoding occurs when the mixed encodings are EBCDIC encodings.
<i>encoding-value</i>	specifies an encoding value. For details, see <a href="#">Encoding for NLS</a> .

After much trial  
and error ...

#### ASCIANY

specifies that no transcoding occurs when the mixed encodings are ASCII encodings.

#### EBCDICANY

specifies that no transcoding occurs when the mixed encodings are EBCDIC encodings.

#### *encoding-value*

specifies an encoding value. For details, see [Encoding for NLS](#).

# Success at Last!

```
1      options nofmterr;
2      libname in 'C:\test_drug\china';
NOTE: Libref IN was successfully assigned as follows:
      Engine:          V9
      Physical Name: C:\test_drug\china
3
4      data eg;
5      set in.eg(encoding='asciiany');
6      run;

NOTE: Format VISF was not found or could not be loaded.
NOTE: Format ECGRSLF was not found or could not be loaded.
NOTE: Format YNF was not found or could not be loaded.
NOTE: There were 1509 observations read from the data set IN.EG.
NOTE: The data set WORK.EG has 1509 observations and 27 variables.
```

## ASCIANY

Transcoding normally occurs when SAS detects that the session encoding and data set encoding are different. ASCIANY enables you to create a data set that SAS will not transcode if the SAS session that accesses the data set has a session that encoding value of ASCII. If you transfer the data set to a machine that uses EBCDIC encoding, transcoding occurs.

# New Methodologies

# Game Changers

---

- In the CRO business, clients now routinely ask for the SAS programs that produce the analyses
  - How do you provide easily understandable programs without giving away intellectual property?
  - For any company, the brains of the reporting and analysis systems are the macros
  - By delivering the expanded macro code, you remove the key concerns from an intellectual property standpoint while preserving the ability to reproduce the deliverables



# Providing Programs to FDA

---

- Similar problems arise for pharma companies when the FDA requests SAS programs
  - Programs may also be required as part of ADaM standards
- Most pharmaceutical companies and CROs have big macro systems for dataset creation and reporting.
  - The probability of these macros running on the FDA or client's systems is close to zero.
  - These systems represent a substantial investment in intellectual property.

# Solution Through Google Search

---

- How do you provide the SAS code without it being impossible to follow, and without giving away vital methodology?
- I found the basis for the answer through a Google search
  - Search: CAPTURING SAS PROGRAMS MPRINT
  - Result: “Capturing SAS Macro Code into an Executable SAS Program”, David J. Jemiolo, DataCeutics, Inc., Pottstown, PA, NESUG, 2002.

# Jemiolo's Method

---

- The method has two steps:
  - 1. Run the macro using `OPTIONS MPRINT MFILE`. This writes the expanded code from the macro to a file with all macro code and macro variables resolved.
  - 2. Clean up the code with post-processor. This includes addition of standard header block, basic commenting, indenting, white space, or any other standards to be enforced

# In a Nutshell ...

---

- Jemiolo (2002) states
  - “Although MPRINT and MFILE are very useful in capturing SAS Macro code, they still lack the ability to directly capture non-macro SAS code. The “trick” ... is to enclose the MPRINT, MFILE option, and %INCLUDE statements within a dummy macro called %\_RUNIT\_. When %\_RUNIT\_ is called, not only will the decoded SAS Macro calls be captured, but any other non-macro code that may exist in the original SAS program, will also be retained. All captured code is written to a text file...”

# Jemiolo: Basic Code

---

```
ods listing close;
%macro _runit_;

    filename incode "&dir.&del.&infile..sas"
    lrecl=2048;
    filename mprint "_temp_"    lrecl=2048;

    options mfile mprint;

    %include incode;

    options nomprint;

%mend _runit_;

%_runit_;
ods listing;
```

# Extensions to Production System

- This simple code fragment is the basis of a SAS program that writes SAS programs for transfer to clients and the FDA

```
/* -----  
Program: CO_A03125_erythema.sas  
Author: Mike Todd  
Project: Intendis  
Protocol: A03125  
Purpose: Erythema tables: women taking oral contraceptives  
Libnames, formats, RTF templates: see start.sas  
Date: 02MAY08 21:23  
----- */  
ods listing close;  
title; footnote;  
/* --- start of code for "subgroup_select". --- */  
  
PROC SQL;  
CREATE TABLE WORK.subgroup_select AS SELECT SUBGROUPS.PROTOCOL,  
SUBGROUPS.RANDID,  
SUBGROUPS.USUBJID,  
SUBGROUPS.POP,  
SUBGROUPS.POPX,  
SUBGROUPS.SEX,  
SUBGROUPS.SEXX,  
SUBGROUPS.RACEH,  
SUBGROUPS.RACEHX,  
SUBGROUPS.ETHCAT,  
SUBGROUPS.AGE,  
SUBGROUPS.AGECAT,  
SUBGROUPS.RXGRP,  
SUBGROUPS.RXGRPX,  
SUBGROUPS.DURROSA,  
SUBGROUPS.PREV_ROSA,  
SUBGROUPS.SUBGROUP_CO AS subgroup  
FROM A03125.SUBGROUPS AS SUBGROUPS  
WHERE SUBGROUPS.POPX = "Per-Protocol" AND SUBGROUPS.RXGRPX = "Aza Gel"  
AND SUBGROUPS.SUBGROUP_CO NOT IS NULL ;  
QUIT;  
/* --- End of code for "subgroup_select". --- */
```

My paper 'Using SAS Enterprise Guide to Generate Standalone Programs' covers this in detail.

# Summary

---

- Looking for clear solutions to grab and use
- Solutions based on PROCs or functions are preferable to macros
  - Fewer validation issues
- Leveraging others experience can save a lot of development time and expense
- Cautions
  - Not every solution is good
  - Even a good solution may not mesh with your organization