



Base SASを使った便利なTIPS&テクニック ～ DATAステップ **関数** 編 ～

SAS Institute Japan株式会社
カスタマーサービス本部
石井 宏司

Agenda

- 意外に知らないDATAステップ関数：
 - ✓ 6.12 ~ 8.2での拡張機能を振り返って
 - ✓ 正規表現による文字列操作
 - ✓ 関数のネストに関するTips

DATAステップ関数の増強(6.12)

- SCL関数の多くをサポート
 - ✓ SASファイルや外部ファイルの情報取得が容易に。

- %SYSFUNC、%QSYSFUNC
 - ✓ 処理スタック、配列、マクロ関連などの一部を除く全てのDATAステップ関数をマクロから利用可能

SAS 6.12新関数

■ データセット関数

- ✓ EXIST ライブラリのメンバの存在を確認する。
- ✓ OPEN/CLOSE データセットを開く/閉じる。
- ✓ ATTRC(N) データセットの文字(数値)属性の値を返す。
- ✓ FETCH/FETCHOBS オブザベーションをデータセットから読む。
- ✓ REWIND データセットの先頭にポインタを置く。
- ✓ NOTE/POINT/ DROPNOTE ノートマーカを記録/位置付け/削除する。
- ✓ CUROBS 現在のオブザベーションの番号を返す。
- ✓ DSNAME IDに対応するデータセット名を返す。

SAS 6.12新関数

■ 変数関数

- ✓ GETVARC/GETVARN データセットの文字(数値)変数値を返す。
- ✓ VARFMT/VARINFMT 変数に割当てた出力(入力)形式を返す。
- ✓ VARLABEL データセット変数に割当てたラベルを返す。
- ✓ VARLEN データセット変数長を返す。
- ✓ VARNAME データセット変数名を返す。
- ✓ VARNUM データセット変数の数を返す。
- ✓ VARTYPE データセット変数のデータタイプを返す。

SAS 6.12新関数

■ 外部ファイル関数(1)

- ✓ FEXIST/FILEEXIST 外部ファイルの存在を確認する。
- ✓ FOPEN/FCLOSE 外部ファイル(ディレクトリ)を開く/閉じる。
- ✓ FDELETE 外部ファイルを削除する。
- ✓ FREAD/FWRITE FDBと外部ファイルのデータ移動。
- ✓ FGET/FPUT 変数とFDBのデータ移動。
- ✓ FNOTE/FPOINT レコードポインタを記録/位置付ける
- ✓ FREWIND ポインタをファイル先頭に配置する。
- ✓ FAPPEND 現在のレコードをファイル末尾に追加する。

SAS 6.12新関数

■ 外部ファイル関数(2)

- ✓ FILENAME ファイル参照名を割当て/解除する。
- ✓ FILEREF ファイル参照名の割当を確認する。
- ✓ FINFO 外部ファイル情報を返す。
- ✓ FOPTNUM/FOPTNAME 外部ファイル情報の項目数/項目名を返す。
- ✓ FRLEN 現在のレコードサイズを返す。
- ✓ FCOL/FPOS FDB内でのカラム位置を返す/設定する。
- ✓ FSEP FGET 関数のトークン デリミタを設定する。

SAS 6.12新関数

■ 外部ファイル関数(3)

- | | |
|--------------------|--------------------------|
| ✓ DOPEN/CLOSE | ディレクトリを開く/閉じる。 |
| ✓ DREAD | ディレクトリメンバの名前を返す。 |
| ✓ MOPEN | ディレクトリ ID とメンバ名でファイルを開く。 |
| ✓ DINFO | ディレクトリについての情報を返す。 |
| ✓ DNUM | ディレクトリ内のメンバ数を返す。 |
| ✓ DOPTNUM/DOPTNAME | ディレクトリ情報の項目数/項目名を返す。 |
| ✓ PATHNAME | ファイル/ライブラリ参照名から物理名を返す。 |

SAS 6.12新関数

■ ライブラリ/カタログ関数

- ✓ CEXIST SAS カタログ(又はエントリ)の存在を確認する。
- ✓ LIBNAME ライブラリ参照名を割当てる/解除する。
- ✓ LIBREF ライブラリ参照名の割当を確認する。

SAS 6.12新関数

■ 特殊関数

- ✓ GETOPTION SAS システム オプションの値を返す。
- ✓ SPEDIS 2 つのワードの非対象なスペルの隔りを返す。
- ✓ SYSGET 指定したホスト環境変数の値を返す。
- ✓ SYSMSG エラー/警告メッセージを返す。
- ✓ SYSRC システム エラー番号、又は終了状態を返す。
- ✓ SYSTEM SAS セッション中にOSコマンドを発行する。

そしてSAS 8

- 他プログラミング言語などで標準的に装備されている機能が追加された。
 - ✓ SCLの拡張とともに他言語ユーザーからの敷居を低く。
- 変数情報がより詳細に取得可能になった。
 - ✓ データ構造に依存しにくいコーディングが容易に。
- グローバルソフトウェアとしての条件を強化
 - ✓ Web環境、ナショナルランゲージサポートの強化。
- 地味ながらもSASware Ballot の結果を反映。
 - ✓ 皆様のSASでありつづけるために。

ちょっと寄り道 SASware Ballot

米国SAS本社が例年行っている調査で、SASのシステムに追加してほしい機能にユーザーが投票できる機会でもあります。

世界中のすべてのSASユーザーが調査に参加できます。

SASware Ballot調査は年末にWebを通して実施されます。

この SASware Ballot Survey ウェブページ が利用可能な期間中は、米国 SAS 社サイトのトップページ、Service&Supportのページ、Technical Supportのページなどからダイレクトにリンクが張られ、年明けに調査が締め切られた後、その結果がSUGIで発表されます。ユーザー様からの提案は、SAS社スタッフへのご提案の他、

Phone: +1-919-677-8008 または

Email: suggest@sas.com でも受け付けています。

URL <http://support.sas.com/techsup/news/sasware.html>

SAS 8 新関数

数学関数

- COMB n 個の母集団から r 個を抽出す組合せ数を返します
- CONSTANT 数学や計算機で用いられる定数を返します
- DEVIANCE デビアンスを返します
- FACT 階乗値を返します
- PERM n 個の母集団から r 個を抽出す場合の順列の数を返します

SAS/STATからの移行

- PROBBNRM 二変量正規分布の有意確率値を返します
- PROBMCMC 多重比較のパーセント点/有意確率値を返します

変数情報関連

- VNAME(X)/VLABEL(X) 変数名/ラベルを返します。
- VTYPE(X)/VLENGTH(X) 変数の型/長さを返します。
- VFORMAT(X)/VINFORMAT(X)
変数に割り当てられている出力/入力形式を返します。
- VFORMATN(X)/VINFORMATN(X)
変数に割り当てられている出力・入力形式の名称部のみを返します。
- VFORMATD(X)/VINFORMATD(X)
変数に割り当てられている出力/入力形式の少数桁数部を返します。
- VFORMATW(X)/VINFORMATW(X)
変数に割り当てられている出力/入力形式の総桁数部を返します。
- VARRAY(X) 配列であるかどうかを返します。
- VINARRAY(X) 配列要素であるかどうかを返します。

SAS 8 新関数

日付関連

- DATDIF 2つの日付間の差を日単位で返します。
- YRDIF 2つの日付間の差を年単位で返します。
- JULDATE7 ユリウス暦を7桁の整数として返します。

特殊関数

- IORCMMSG _IORC_自動特殊変数に関するメッセージを返します。
- MISSING 引数が欠損かどうかを返します。
- SLEEP 指定時間、DATAステップの実行を中断します。

64ビットサポート

- ADDRLONG / PEEK(C)LONG / POKELONG / PTRLONGADD
64ビット環境でのメモリアドレスを操作します。

SAS 8 新関数

WEB技術関連

- HTMLENCODE/HTMLDECODE

‘<’や’&’などのHTML上で特別な意味を持つ特殊文字のエンコード/デコード

- URLENCODE/URLDECODE

‘=’や’:’などのURLエスケープ文字列のエンコード/デコード

NLSサポート

- EUROCURR

ヨーロッパ通貨同士を変換し、値を返します。

- NLDATE / NLDATM

任意の日付 / 日時要素文字列を取得します。

SAS 8 新関数

会計関数

- CONVX/CONVXP

列挙キャッシュフロー/定期キャッシュフローストリーム のコンベクシティを返します。

- DUR/DURP

列挙キャッシュフロー/定期キャッシュフローストリーム の修正期間を返します。

- PVP

定期キャッシュフローストリームの現在価値を返します。

- YIELDP

定期キャッシュフローストリームの満期利回りを返します。

SAS 8 新関数

DBCS関数 (SBCS用に対応するものが存在しないもの)

(SO/SI 対応)

- KCOMPARE 文字列の比較結果を返します。
- KSTRCAT 2つ以上の文字列を連結します。
- KTRUNCATE 指定した長さに文字列を切り捨てます。

(その他)

- KCOUNT 文字列内の2バイト文字の数を返します。
- KCVT 他の2バイトコードに変換します。
- KUPDATE / KUPDATEB ... (左辺SUBSTRの代替として)
文字値の内容を挿入、削除、置き換えます。

SAS 8 で拡張された関数

- PUT 出力のアラインメント調整が可能になった。
- SCAN 右側からの単語走査が可能になった。

・・・仕様変更

- QUOTE 末尾ブランクを取り除かなくなった。
TRIM、TRIMN関数の併用が必要な場合あり。

正規表現関数およびルーチン

- **正規表現**(Regular Expression) - 正則表現、有理表現とも。
『特定の文字列ではなく、文字列の一部を一般化して表現するための手法。もともとはコンピュータ言語理論の分野において、字句(変数名や予約語、その他の識別子)を一般化して定義するために考案された表現手法。通常のコンピュータ利用では、ドキュメントからの文字列検索時などに、検索したい文字列すべてを指定するのではなく、文字列の任意の一部を置き換え可能な状態で検索する場合などに用いる。』(アスキーデジタル用語辞典より)

正規表現関数およびルーチン

rx=**RXPARSE**(*pattern-expression*)

- パターンを解析して値を返します。

position=**RXMATCH**(*rx*, *string*)

- パターンが一致するサブストリングの先頭を見つけ、値を返します。

CALL RXFREE(*rx*)

- ほかの正規表現 (RX) 関数やCALLルーチンによって割り当てられたメモリを開放します。

正規表現関数およびルーチン

- **CALL RXSUBSTR**(*rx*, *string*, *position*<, *length*<, *score*>>)
 - パターンが一致する位置、長さ、およびサブstringのスコアを検索します。
- **CALL RXCHANGE**(*rx*, *times*, *old-string*<, *new-string*>)
 - パターンが一致する1つ以上のサブstringを変更します。

RX関数における正規表現

- 2種類の構文
 - ✓ 検索構文 RXMATCH、RXSUBSTR 用
 - ✓ 置換構文 RXCHANGE 専用
- 記述可能な要素
 - ✓ 文字列
 - ✓ 文字クラス
 - 組込み
 - ユーザー定義
 - ✓ タグ、スコア、NOTやOR条件などの制御文字

RX関数における正規表現

■ 組込み文字クラス

- ✓ ? 一文字のワイルドカード
- ✓ \$A / \$L / \$U アルファベット / 小文字 / 大文字
- ✓ \$D 数字
- ✓ \$W 空白、タブ、キャリッジリターンなど
- ✓ \$C、\$I SAS名として有効な文字 / 先頭文字

■ 組込みパターン(不定長)

- ✓ : 任意長 のワイルドカード
- ✓ \$N SAS名
- ✓ \$F 数値
- ✓ \$Q クォートされた文字列
- ✓ \$(n)、\$[n]、\${n} ネストされた括弧の任意レベル部分

RX関数における正規表現

■ オプションなど

- ✓ \$P / \$S 語頭 / 語尾 条件
- ✓ ^ / ~ NOT条件
- ✓ | OR 条件
- ✓ @n パターンマッチング開始位置
- ✓ * / + パターンのくりかえし
- ✓ #=n 任意スコア値の割当
- ✓ < > タグ: 置換時再利用の為の部位指定

文字列操作関数との比較

- INDEX
- INDEXC
- INDEXW
- VERIFY
- TRANSLATE
- TRANWRD

- SCAN

関数のネスト

文字列に関する情報を数値として得られる関数など

- LENGTH
- INDEX
- (比較演算)

戻り値を REPEAT、SUBSTR、SCAN等、数値を引数にとる関数の入力として使える

Ex.

```
SEG = SCAN('未婚女性 未婚男性 既婚女性 既婚男性',  
           1 + (SEX='M') + (MARRIAGE='Y')*2 );  
PARM = SCAN(SUBSTR(STMT, INDEX(COMMAND, 'PARM=')+5), 1);  
MASK = SUBSTR(REPEAT('*', LENGTH(PASSWORD)), 2);
```


関数のネスト

- マクロ化して簡易に

概念的に一つの関数のように扱いたいものを間違いを減らし使い易くする効果。

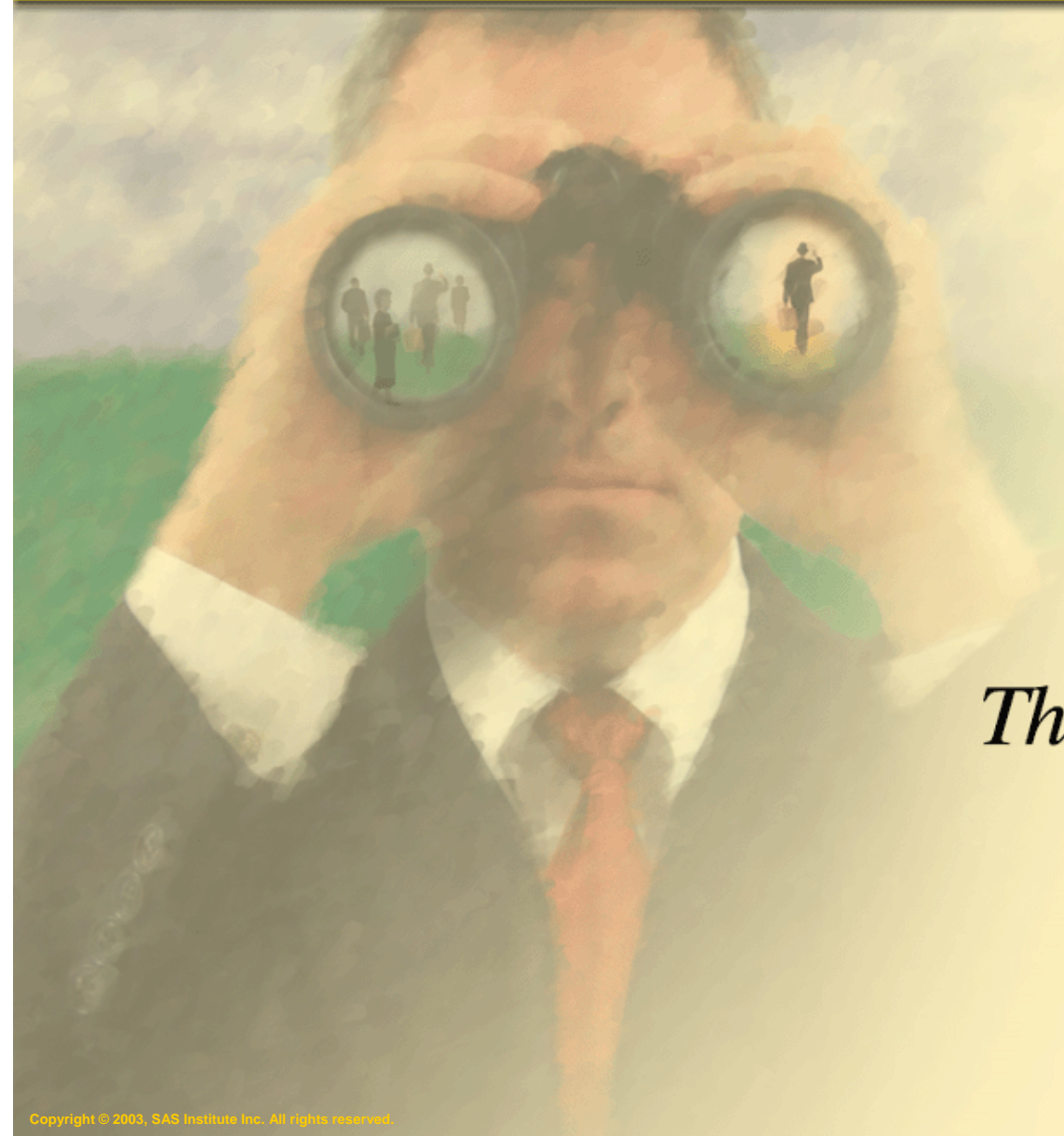
サンプル:

%MEDIAN、%FACT

%KANAS2D、%KANAD2S

本ご紹介したサンプルプログラム

- <http://www.sas.com/japan/usergroups/wg/archive/index.html#kansai030523>
- More...
Support.sas.com のFAQページ
http://support.sas.com/techsup/faq/data_step/logifunc.html



The Power to Know®