

SAS Technical News

Summer 2008

*For Higher
Customer Satisfaction,
We Bridge
the SAS System
Between
Customer's World.*

CONTENTS

- 1 特集1 SAS 9.2のご紹介(プラットフォーム編)
- 7 SAS Academic News
- 12 Q&A
- 14 SASトレーニングのお知らせ
- 16 新刊マニュアルのお知らせ
- 16 最新リリース情報
- 16 SAS Technical News送付についてのご案内

特集1

SAS 9.2のご紹介 (プラットフォーム編)



1. はじめに

SAS Systemの最新リリースとなるSAS 9.2の提供が開始されます。SAS Technical Newsでは、特集記事として今回から3つのフェーズにわたってSAS 9.2をご紹介いたします。1回目の今回は、SAS 9.2 Enterprise Intelligence Platformを中心にSAS 9.2リリースの概要についてご紹介します。

なお、本文にはリリース前の情報が含まれており、プロダクトのスクリーンショットなどは開発段階のものを使用しています。実際リリースされるものと異なる場合がありますので、あらかじめご了承ください。

2. SAS 9.2とは?

現在提供しているSAS 9.1.3を拡張したこの新しいリリースは、これまでにユーザーの皆様から寄せられたさまざまなご意見を取り込んだものとなっています。たとえば、プロシジャなどの追加によりさらに高度な分析を可能にしたり、グラフなどのビジュアル面を向上させたり、クライアントツールのUIをより使いやすいものに変更するなど、より多くのお客様に満足いただけるシステムを目指して開発されました。また、インストールと設定も以前より簡単になりました。

2.1 Business Intelligenceとは?

まずここで簡単にSAS 9からの重要なコンセプトのひとつである、Business Intelligence(以下、BI)のおさらいをします。BIという言葉は、元ガートナーグループのアナリスト、ハワード・ドレスナー氏が1989年に使ったことが最初とされており、たとえば次のように定義されています。

ビジネス上の価値を引き出すことを目標として、企業内に蓄えたさまざまなデータを収集して分析し、その結果を表示・配布する仕組みの総称。ERPパッケージやCRMソフトなどからもたらされるデータの分析を専門家に依存せず、経営者や社員が必要な情報を自在に分析し、経営計画や企業戦略などに活用することを目指している。

出典：日経コンピュータ IT Pro Keyword

現在、現場にあふれる業務データを経営判断に活かすため、多くの企業がBIの導入に取り組んでいます。

2.2 SAS Enterprise Intelligence Platformについて



SAS 9から提供している、SAS Enterprise Intelligence Platformは、SASが提供する統合型BIプラットフォームです。散在するデータの統合、加工、蓄積から分析、活用、予測まで、BIの定義に含まれるすべての要素をSASのテクノロジーで提供します。このプラットフォームはBase SAS、SAS/STATなど従来の対話モードで利用されるプロダクト群「SAS Foundation」の基礎部分と、SASコードを直接記述しなくてもその機能を簡単に使用することができるさまざまなクライアントツール群で構成されています。従来、SASは難しいというイメージがありましたが、ここで提供しているクライアントツール群は、さまざまなニーズを持つユーザーがスキルセットにあわせて直接活用できるよう、工夫されています。たとえば、データ統合を担当する方には、データの抽出、加工、ロード(ETL)をGUIで実行するツールを用意し、またエンドユーザー向けとして、Office製品に追加されるSASのアドインメニューを利用したレポート作成を可能にし、さらにエグゼクティブクラスなどレポートを参照する方向けには、簡単な操作で必要なレポートを参照できるWebアプリケーションを用意しています。これにより、SASに関する知識の有無にかかわらず、『必要な人が必要なときに必要な形で情報を活用すること』を可能にします。

従来、BIはある特定の部門の部門サーバーのみで活用されることが多いものでしたが、SAS Enterprise Intelligence Platformの導入でBIは社内全体で使用できるインフラとなり、上記のようにさまざまなバックグラウンドのユーザーが活用できるようになります。そこで、すべてのユーザーがスムーズに使用するための環境整備とセキュリティ設定などのシステム管理は、安定した運用をするうえで非常に重要になります。そのため、インテリジェンスプラットフォームを支える管理者とプラットフォームのユーザーのはっきりした役割分担が必要です。

SAS Enterprise Intelligence Platformでは、煩雑になりがちなシステム管理をSASの環境設定やユーザー定義、ライブラリ定義などをメタデータとして一括管理することで、運用面でもサポートしています。

SAS Intelligence Platformについては、弊社Webサイトを参照してください。
<http://www.sas.com/japan/technologies/architecture/index.html>

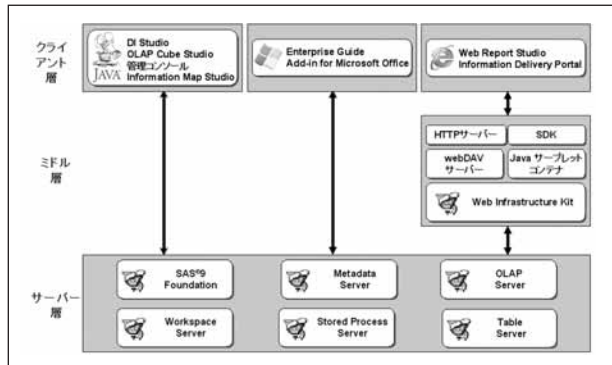
SAS 9の概要などについては、下記のドキュメントを参照してください。
 『SAS Technical News Summer 2004』

特集「SAS 9.1のご紹介」

http://www.sas.com/japan/periodicals/technews/pdf/04_sum.pdf

2.3 システム構成図

SAS Enterprise Intelligence Platformの代表的なシステム構成は、このようになっています。



SAS Enterprise Intelligence Platformは、さまざまなジョブを実行するサーバー層、Webアプリケーションを実行するためのコンポーネントで構成されるミドル層、そしてクライアント層の3つの層でシステムを構成しています。システムの流れを簡単にまとめると、次のようになります。ユーザーは、まずクライアントからサーバー層のMetadata Serverに接続し、「自分が該当のアプリケーションを実行可能なユーザーか」、「どのサーバーを使用するか」などの定義情報(メタデータ)を入手します。そのメタデータをもとに、サーバー層を構成する各サーバー(Workspace Server、OLAP Serverなど)に処理がリクエストされ、実行結果がクライアントに返されます。

SAS Intelligence Platformの構成、各サーバー機能などに関しては下記のドキュメントにも記載しておりますので、参照してください。

『SAS Technical News Summer 2004』

特集「SAS 9.1のご紹介」

http://www.sas.com/japan/periodicals/technews/pdf/04_sum.pdf

『SAS Technical News Spring 2005』

特集「Open Metadata Architecture入門」

http://www.sas.com/japan/periodicals/technews/pdf/05_spr.pdf

3. ユーザーの業務別活用例(ユーザーマッピング)

この章では、SAS Enterprise Intelligence Platformに含まれる代表的なプロダクト(クライアントツール)をどのようなユーザーの業務に対応しているかという観点からご紹介します。まずは、下記の図を参照してください。

システム管理者	データ管理者	分析担当者	分析レポート担当	一般ユーザー	エグゼクティブ
アクセス権の設定などシステム全体を管理するユーザー	ETLプロセスの開発などデータ管理を担当するユーザー	データマイニングなど複雑な分析を行なうユーザー	データを分析し、さまざまな種類のレポートを作成するユーザー	主にExcelなどでレポートを作成するユーザー	レポートを参照し、意思決定に活用するユーザー
管理コンソール	Data Integration Studio	Enterprise Miner SAS/STAT SAS/ETS など	Add-in for Microsoft Office		
		Enterprise Guide	Web Report Studio		
Information Delivery Portal					

ユーザーマッピング例

3.1 システム管理者 SAS管理コンソール



SAS 9では基本的にすべての情報は「メタデータ」として管理します。メタデータには、SASの環境設定、サーバー情報、ユーザー定義、ライブラリ定義などあらゆる情報が格納されていますが、これらを一括して管理するプロダクトがSAS管理コンソールです。システム全体をひとつのツールで管理でき、管理者側の負担を軽減します。SAS管理コンソールおよびメタデータに関しては、下記のドキュメントもあわせて参照してください。

『SAS Technical News Spring 2005』
 特集「Open Metadata Architecture入門」
http://www.sas.com/japan/periodicals/technews/pdf/05_spr.pdf

『SAS Technical News Spring 2005』
 特集「Open Metadata Architecture運用入門」
http://www.sas.com/japan/periodicals/technews/pdf/05_sum.pdf

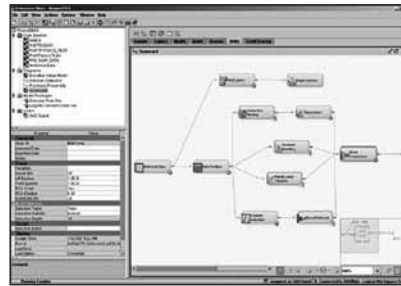
3.2 SASを使ってETLプロセスを実行したいユーザー SAS Data Integration Studio



データの抽出、クエリなどのデータ変換や、SAS/ACCESSがサポートしている他社製データベースへの接続などETLプロセスの開発がGUIベースで実行できるプロダクトです。SAS Data Integration Studioには、一般的に良く使われる変換プロセスがテンプレートとして用意されており、複雑なフローをすばやく作成することができます。バックエンドでは、SAS Foundationに含まれるコンポーネントが実行されており、SASコードが生成されています(生成されるSASコードを確認したり、編集したりすることも可能です)。ここで開発されたETLプロセスは、すべてSAS Metadata Serverが管理するメタデータリポジトリに保存されます。このプロダクトの詳細およびSAS 9のETL関連プロダクトについては、下記のドキュメントを参照してください。

『SAS Technical News Autumn 2005』
 特集「SAS ETL Serverのご紹介」
http://www.sas.com/japan/periodicals/technews/pdf/05_aut.pdf
 基本的な機能などは変わりませんが、製品名など変更された点があります。

3.3 SASの高度な分析機能を活用したいユーザー SAS Enterprise Miner



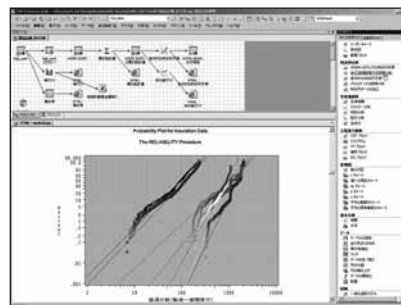
蓄積された膨大な量のデータをSASの処理能力により、高速かつGUIベースでサンプリング、データ探索、加工、モデリング、評価まで一連のデータマイニングプロセスを実施するプロダクトです。クラスター分析や回帰分析、決定木、ニューラルネットワークといった定番の分析手法からDMニューラル、二段階モデルなど幅広いモデルをサポートしています。分析フローや分析結果を共有化することができます。データマイニングは、通信会社、銀行、証券会社、クレジットカード会社、および保険会社では、不正行為の検出、マーケティング・キャンペーンの最適化、リスク管理などに活用され、医療業界では、処置や検査、投薬の有効性を予測するために役立てられています。また、小売業者では、クーポンやイベントの効果を評価し、さまざまな購買者に最適なキャンペーンを提案するために利用されています。

SAS/STAT、SAS/ETSなど



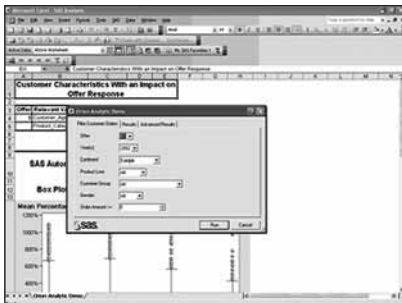
SAS Foundationに含まれるSAS/STAT、SAS/ETSなど、長年にわたってSASが提供してきた統計解析関連プロダクトを活用していただくことにより、さまざまな種類の分析を実行できます。これらのプロダクトを使用する場合は、主として対話モードでコードを実行していただきます。

3.4 SASのさまざまな分析機能を活用したいユーザー SAS Enterprise Guide



さまざまなデータソースにアクセスし、ウィザード形式のタスクを活用して、さまざまな種類のデータ加工や分析を実行し、定型レポートからアドホックレポートまで柔軟に作成できるWindowsクライアントアプリケーションです。OLAPキューブにアクセスし、データの探索をすることも可能です。SASの機能をノンプログラミングで使用できますが、コードを記述して独自の処理を実行することもできます。そのため、SASの利用経験のある方から利用経験のない方まで、幅広いユーザーに活用していただけます。実行した分析プロセスは、プロセスフローダイアグラムを使用して管理できます。また、このアプリケーションで実行したプロセスをストアプロセスとして他のプロダクト(SAS Add-in for Microsoft Officeなど)から再利用したり、分析結果やレポートをポータルに配信したりすることが可能です。さらに、バッチで実行していた既存のコードをタスク化するなど、これまでの資産を有効活用できます。

3.5 SASを日々のレポート作成に手軽に活用したいユーザー SAS Add-in for Microsoft Office



SASの機能をMicrosoft Officeアプリケーションのアドインとして使用できるプロダクトです。使い慣れたMicrosoft Excelから、SASのデータにアクセスしてExcelの制限を越える大容量データを読み込み、あらかじめ分析担当者が用意したストアプロセスによるSASを活用した分析やアドインに含まれるウィザードを使用したグラフ作成、クエリーの実行などが可能です。SASを今まで使用したことがないユーザーでも、普段Officeアプリケーションを使用しているときと同じ感覚で、SASの機能を活用したレポートを作成できます。

3.6 SASのBIを活用したレポートを参照したいユーザー SAS Web Report Studio



分析結果をもとにビジネスの意思決定をされる方々向けに用意された、簡単な操作で必要なレポートを参照・生成できるWebアプリケーションです。すでに担当者によって作成されたレポートを参照するだけでなく、自らクエリーやレポートを実行できることで、必要な情報をより希望に近い形式で、タイムリーに参照することが可能になります。またWebアプリケーションですので、使用するクライアントPCにブラウザがインストールされていれば、ほかのソフトウェアのインストールなどは必要ありません。

SAS Information Delivery Portal



SAS Web Report Studio、SAS Enterprise Guideなどで作成したレポートや、ストアプロセス、その他のアプリケーションなどを表示、管理できるWebポータルです。ユーザーが必要な情報で、なおかつ参照などの権限が与えられているものに対してWebから簡単にアクセスできます。

4. SAS 9.2 から新たにサポートする環境

この章では、SAS 9.2から新たにサポートする環境について、ご紹介します。

4.1 オペレーティングシステム

新たにサポートされるオペレーティングシステムには、次のものがあります。

サーバー:

UNIX・・・Solaris 10、HP-UX 11iv2、AIX 5.3、Linux 2.6 Kernel distribution
Microsoft Windows(x64版OSにも対応)・・・
Microsoft Windows Server 2003

クライアント(x64版にも対応):

Microsoft Windows XP Professional、Microsoft Windows Vista (Enterprise、BusinessおよびUltimate Edition)

4.2 Web Application Server

Web Application Serverとしては、新たにJBoss Application Server 4.2.0(Tomcatが含まれます)、IBM WebSphere Application Server 6.1、BEA WebLogic Server 9.2をサポートします。

4.3 ブラウザ

SAS 9.2からは、Internet Explorer 7がサポートされます。また、Firefox 2.0もサポートされます。

4.4 Java環境

これまで、Java 1.4.xをサポートしていましたが、SAS 9.2からJava 5をサポートします。

5. SAS 9.2の拡張点

この章では、各プロダクトのSAS 9.2からの主な拡張点についてご紹介します。今回は、SAS Enterprise Intelligence Platformに含まれるクライアントアプリケーションを中心にご紹介します。SAS Foundationプロダクト(Base SAS、SAS/STAT、SAS/GRAPHなど)の拡張点については、次回の特集(SAS Foundation編)でご紹介する予定です。

5.1 ETLプロセス(SAS Data Integration Studio 4.2)の拡張点 マッピングおよびプロパゲーションの拡張



今回から次の点が拡張されます。

- ・自動的にマッピングやプロパゲーションが可能
- ・カラムに対するプロパゲーションの変更をサポート
- ・プロパゲーションの実行方向を設定可能
- ・データ変換もしくはカラムレベルでマッピング/プロパゲーションをブロック

マッピングは、あらかじめ用意したルールに基づいて自動的に実行することができます。たとえば、「数値を文字値(もしくはその反対)として扱う」、「最初の6文字が同じであれば、自動的にマッピングする」などのルールを設定することができます。

#	Column	Column	Column Description	Expression	Type
1	Region	Region			Character
2	Subsidiary	Subsidiary_WestCoast			Character
3	Stores	Stores	Number of Stores		Character
4	Sales	Total S	Total Sales		Numeric
5	Inventory	Total I	Total Inventory		Numeric
6	Returns	Total R	Total Returns		Numeric

上のスクリーンショットは、このルールに基づいた自動マッピングのものです。ここでは、「最初の6文字が同じであれば、自動的にマッピングする」というルールが設定されているので、SubsidiaryとSubsidiary_WestCoastのカラムがマッピングされています。ルールは、テキスト形式の設定ファイルを用意しており、そこで設定します。また、自動的にマッピングやプロパゲーションをするか、プロパゲーションの方向を最初からもしくは最後からのどちらかにするかなどは、プリファレンスでまとめて設定することができます。

メタデータレポートنگ



定義されているジョブやライブラリに含まれるテーブルなどに関するメタデータの詳細レポートをHTML形式で表示できるようになります。これは、管理者の方にとって便利な機能です。

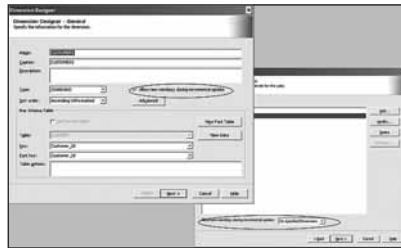
たとえば、テーブルのレポートでは、テーブル名、作成日、変更日、テーブルの作成者、使用されているスキーマなどの情報が確認できます。また、ジョブのレポートでは、各ジョブの作成日、変更日、ソーステーブル、ターゲットテーブル、ジョブの作成者・変更者などを確認できます。

プロセスフローの拡張

作成したプロセスを実行すると、現在どのプロセスを実行していて、実行結果がどうであったかフロー上にわかりやすく表示します。

5.2 OLAPキューブの拡張点

キューブの増分更新(Incremental Cube Update)が可能に



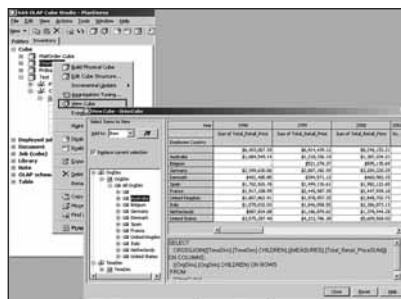
これまでお問い合わせの多かった既存のOLAPキューブの更新が、今回のリリースから可能になります。更新方法として下記の2つの方法が提供されます。

- ① 既存キューブにデータを追加する
- ② 既存キューブをもとに別名のキューブを作成し、その後、置き換える

①の方法は、既存のキューブに新しいデータを増分として追加するもので、キューブのファイル名などは変更されません。キューブの構造が保持されるので、更新作業中でもユーザーはキューブを参照したり、探索したりすることができます。

②の方法は、段階的な作業になります。最終的にキューブを置き換えるまでは、ユーザーが既存キューブを参照したり、探索したりすることができます。別名のキューブを作成するのでキューブの名前を変更したいというニーズにも対応できます。

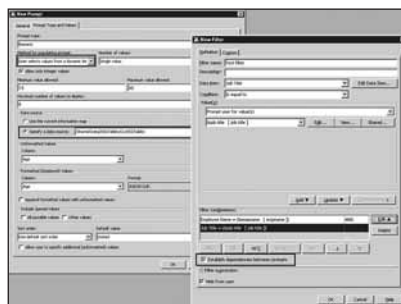
キューブの内容をより詳しく確認



SAS 9.1.3ではキューブの内容を確認する場合、キューブのプロパティから構造だけ確認することができました。SAS 9.2からは、View Cubeという機能が追加され、選択したキューブの構造だけではなく、含まれるデータの内容も参照できるようになります。これにより、登録されているキューブのより詳細な内容を確認できます。

5.3 インフォメーションマップの拡張点

リレーショナルデータに対するダイナミックプロンプトとカスケードプロンプト



ダイナミックプロンプトとカスケードプロンプトの追加で、インフォメーションマップで利用できるプロンプトがより便利になりました。ダイナミックプロンプトは、データから動的に選択できる値を取得してリスト表示するプロンプトです。ユーザーはプロンプトに自分で値を入力することなく、表示されるリストから選択することが可能になります。もう一方のカスケードプロンプトは、プロンプトに依存関係を持たせるもので、複数のプロンプトを組み合わせて使用できます。カスケードプロンプトでは、はじめのプロンプトが絞り込んだ結果を次のプロンプトに引き継いでいきます。たとえば、売り上げデータのレポートに使用するために地域と国と都市の3種類のプロンプトに依存性を持たせたカスケードプロンプトを作成したとします。1つめの地域を選択するプロンプトで地域に「アジア」を選択すると、そのデータが2番目の国を抽出するプロンプトに引き継がれ、さらに、そこで「日本」を選択すると、3番目の都市を選択するプロンプトではここまでで絞り込まれた「日本」のデータが使用されます。つまり、OLAPキューブでデータをドリルダウンするように、データを絞り込んでいくことを可能にするプロンプトです。

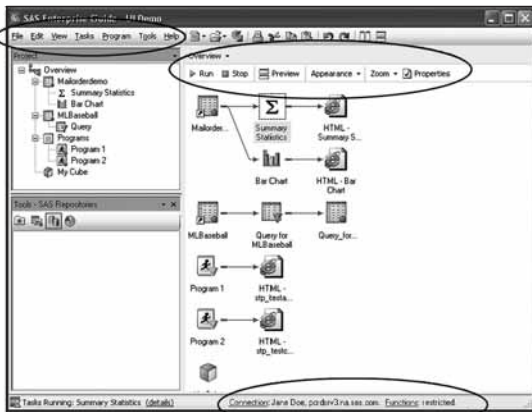
インフォメーションマップのハイパーリンクサポート



Web Report Studio向けにインフォメーションマップでハイパーリンクをサポートできるようになります。これにより、Web Report Studioで作成したレポートのテーブルで、ここでハイパーリンクに設定したテキストやイメージを追加することができます。たとえば、製品一覧のテーブルに、関連するWebサイトへのリンクやURLから取得した製品イメージの画像を表示することが可能になります。

5.4 レポートの拡張点

[デスクトップレポート作成環境(SAS Enterprise Guide 4.2)での拡張点]
ユーザーインターフェイスの拡張



最上部にあるメニューの数を減らし、実際に分析のフローなどを表示するウィンドウにその作業に対応するメニューをコンテキストツールバーとして表示するようにしたことで、ユーザーインターフェイスがよりシンプルでわかりやすいものに変更されました。また、以前からユーザーによって使用できる機能を制限したいというご要望がありましたが、今回から管理者が設定したユーザーのロールにより、使用できる機能を制限することができるようになりました。上の図のように、ログインしているユーザー名と

使用できる機能についての情報が右側のウィンドウの一番下の部分に表示されます。そのリンクをクリックすると、どの機能が使用可で、どの機能が使用不可になっているか確認できます。

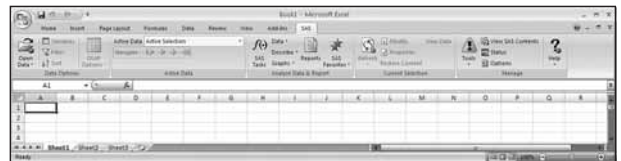
クエリービルダーの拡張



計算式を含むカラムの設定では、これまで、ユーザーが式を手入力していましたが、ウィザードが使用できるようになり、より設定しやすくなりました(上の図がその例です)。また、集計を含むカラムの式を自動的に検出して、表示することができ、具体的に確認できるようになりました。さらに、フィルタ間の関連付けをクエリービルダーのフィルタデータタブで直接変更できるようになりました。

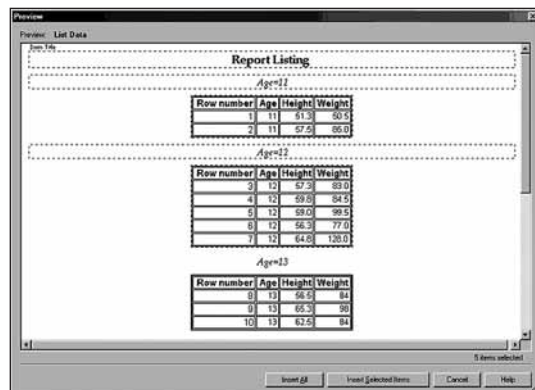
[Microsoft Office環境との統合(SAS Add-in for Microsoft Office 4.2)での拡張点]

Microsoft Office 2003とMicrosoft Office 2007両方をサポート



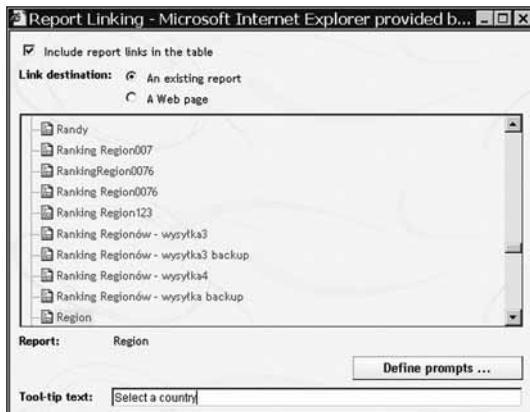
Microsoft Office 2003とMicrosoft Office 2007の両方をサポートします。また、Microsoft Office 2007のリボンをサポートしています。

出力のプレビューとレポートに挿入するアイテムの選択



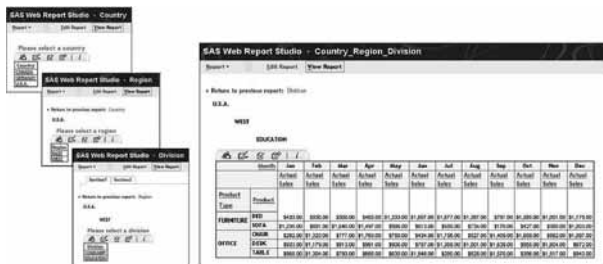
ストアドプロセスとして用意されているプロセスを実行した際、もしくはSASタスク(グラフ作成など)を実行した際に、Excelのワークシートなどに挿入する前に別ウィンドウで結果をプレビューできるようになります。さらに、プレビュー後、出力された結果をすべて作成中のレポートに挿入するか、選択したものだけ挿入するか選択できます。これにより、必要なテーブルやグラフだけをレポートに取り込むことができます。

5.5 Webからのレポート参照の拡張点(SAS Web Report Studio 4.2) レポートのリンクの拡張



作成するレポートに同一レポート内の他のページや他のレポートへのリンクを追加できるようになります。また、レポートから直接ストアドプロセスにリンクすることも可能になります。この拡張により、Web Report Studioで作成するレポートがより使いやすくなります。

プロンプトとフィルタの拡張



SAS Web Report Studioでも「ダイナミックプロンプト」および「カスケードプロンプト」が使用できます。また、大文字小文字を区別しないフィルタやプロンプトの実行などが可能になります。

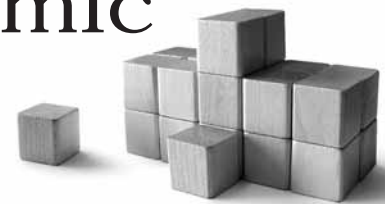
インフォメーションマップの拡張点の説明と重複しますが、「カスケードプロンプト」では、複数のプロンプトが依存性を持つため、はじめのプロンプトが絞り込んだ結果を次のプロンプトに引き継ぐことができます。たとえば、上の図のように、国と地域と部門の3種類のプロンプトに依存性を持たせたカスケードプロンプトでは、1つ目の国を選択するプロンプトで選択された「アメリカ」だけのデータが2番目の地域を選択するプロンプトに引き継がれて使用されるというように、データを絞り込んでいくことができます。

また、ダイナミックプロンプトでは、プロンプトで抽出できる値をデータから直接動的に取得して表示するので、ユーザーが直接値を入力する必要がなく、便利です。

6. 終わりに

今回の特集では、SAS 9.2のプラットフォームのご紹介をしました。他にも多くの拡張点がありますが、この特集でSAS 9.2に興味を持っていたいただければ幸いです。次回の特集では、SAS 9.2 Foundationについてご紹介する予定です。

SAS Academic News



はじめに

SAS Academic Newsでは、教育機関でSASをうまく活用するための情報共有等を目的とし、本誌を通じてドキュメントには無い、生のお客様の声や、よりSASを快適にご利用いただく方法、またSASの認定制度のご紹介をいたします。

事例紹介

春号に引き続き、同志社大学様にてSASを授業で活用するまでの準備等、前号よりもう一步踏み込んだ内容をご紹介します。

宿久 洋

同志社大学文化情報部/文化情報研究科
教授 博士(工学)

SAS: 前回のSAS Technical News Spring号2008で、同志社大学文化情報学部/文化情報研究科の教育理念、国際的に活躍できる人材の育成と学生が学ぶ学術範囲について伺いました。その中で3回生の、特に高度な統計解析を学ぶ学生、統計科学を専門とする大学院に進学する学生が主にSASを使用すると聞きしました。平成19年度から引き続き平成20年度もこの4月から勉強会・ゼミ・授業等が始まりましたが宿久様が学生にSASを使わせるに至ったこれまでの経緯や理由をお聞かせ頂けますか。

宿久: 私のゼミに所属する学生の多くは、将来的には統計学を生かせる職業につくことを希望しており、その中の半数が進学を希望しているという状況です。研究者としてはもちろん職業人としても理論だけではなく実践力を身に付けることは必須であり、本学部ではすべての学生にプログラミング教育、統計解析ソフトの教育を行っていますが、私の研究室ではそれに加えて、SASの修得を必須の条件としています。これは、現在の社会におけるシェアと学生の希望職種から考えた選択です。なお、他の統計解析ソフト(SPSSなど)の知識も講義を通して習得させるようにしています。

SAS: SASが使用されている人数(規模感)はどのくらいでしょうか。

宿久: 講義については、情報メディア棟(1人1台のパソコン、3 in 2 outの教材提示システム)で40名を対象に講義をしています。教材と教員用パソコンの画面を同時に見せながらの演習です。

ゼミでの勉強会については、ゼミ室(1人1台のパソコン)で1名が講師役で事前に予習をし、その内容を全員(10名)で演習しています。教材はBase Program 1です。

SAS: そのなかで現状及び今後解析で使用する検定手法はどんなものでしょうか。

宿久: これは、多種多様です。全員が既に学んでいる手法は各種検定、簡単な分散分析、各種多変量解析です。

ゼミでは、今後データ解析コンペティションへの参加を企画しており、データマイニング法を中心として取り組む予定です。その準備として、講義では取り扱えなかった、決定木、アソシエーション分析などをEMを利用しながら習得していく予定です。

SAS: それにしても、昨年初めてSAS講座を開講した際にはかなりの準備期間が必要だったのではないのでしょうか。具体的にどのくらいかかりましたか。そして教員の悩みの種は教材かと思いますが、どうされましたか。本年度も引き続きSASを取り入れた理由をいただけますか。

宿久: 1年くらいです。色々な解説書を集めたり、アメリカで行われたSASの教員向けセミナーにでたりしました。幸い、同志社大学はSASのサイトライセンスを所持しており、改めての購入やソフトのインストール作業は不要でした。教材や講義内容についての検討のみを行えばよかったのでそんなに負担はありませんでした。

SAS: 準備期間のなかで何か宿久様が特別に行ったことはございますか。

宿久: これは、上記にも書いたように、米国開催のTraining for Educatorに出たことでしょうか。実際の講義の内容はその段階では決めていませんでしたが、できればデータマイニングを取り入れようと思っていましたので、データマイニングのコースを取りました。参加者は全世界から集まっていたのですが、MBAの大学院の担当教授が多かったですね。参加後は、いくつか送ってもらったAcademic Trainers Kitsの資料の内容を確認し、講義内容を検討したくらいですね。

SAS: 授業の内容について教えていただけますか。

宿久: 結局、Academic Trainers Kits(英語版)をそのまま利用した講義を行いました。SASの認定プログラムの受験を意識していたのであえて英語の教材としました。内容はBASE Program Iです。統計学の理論については、数多くの講義を配置していますので、ちょっとした復習程度で主に使い方中心の講義となりました。

SAS: 授業以外に学生への宿題または、学生自信がSASを習得するに当たりどのくらいの時間が必要と考えますか。

宿久: これは、完全に人に依存すると思います。統計の知識はひと通りあり、英語にも不自由しない、かつ、簡単なプログラムの経験があるという学生であれば基本的な部分は数週間で使えるようになるでしょうか。実際、数名の学生はこのレベルでした。一般にはこれらの条件のどれも仮定できませんのでプログラムベースのSASを使えるようになるには相当の時間を取るのではないのでしょうか?その意味ではEGはありがたいですね。どうしてもプログラムがだめな学生はEGでやれば、他のGUIベースの統計ソフトと同様に利用できるようになるでしょう。

SAS: 宿久研究室では、授業教材のほかにもどのようなものを参考文献としていますか。それはすぐに探せるものでしょうか。

宿久: 上記のAcademic Trainer Kitsの中からいくつか選択して勉強会で使っています。SASの日本語の本はあるにはあるのですが、少し古いですね。新たな執筆者が書いてくれるとありがたいのですが、この点はRの勢いに負けている気がします。

SAS: 学生からのSASについての質問は授業の内容と同様に、宿久様がすべて解答されているのでしょうか。

宿久: 私はヘビーユーザではないので、通り一遍のことしか答えられません。結構、自分たちで話しながらやっているようです。

SAS: 授業の進め方についてお聞きします。他大学様では、「SASの使用方法を教えるのに時間がかかり、肝心の授業の中身が進まない」などの話を時々耳にします。宿久研究室での授業の成功の秘訣を教えてください。また、他大学様へこんな点を注意すべきといったことがあれば教えてください。

宿久: これは、先ほども書いたように、他のところでは参考にならないかもしれません。現在、SASを利用した講義を受けているのは3回生で、そこまでにExcel、SPSSでの演習、10コマ以上におよぶ統計関係の講義を受けた学生です。私の場合はSASに特化して講義をすればよく、しかも他のソフトとの違いを説明しながらという形で行うものでしたので、ご指摘のような困難さはありませんでした。知識習得を目的とした講義の援用ソフトとしては、SASは敷居が高いですね。

SAS: 運用で困った点などございましたでしょうか。また他ソフトと比べ不便な点などあれば教えてください。

宿久: 運用上は特に感じていません。これは、本学の事務サイドのサポートのおかげもあると思います。各研究室でSASを利用する際も丁寧なインストールガイドを準備してくれていますし、インストール上のトラブルも事務で対応してくれています。少なくとも運用上では、SAS社と直接やりとりすることなくほとんどの問題が解決しています。

SAS: 上記を踏まえSASへの改善リクエストなどございますか。

宿久: EMでテキストマイニングができるようになるとありがたいです。

SAS: SASを授業に取り入れてよかった点はございましたでしょうか。

宿久: これは、今後の結果をみて判断したいと思っています。

SAS: 2008/4~2008/6までの授業(ゼミ等)で学生様の反応や感想はいかがでしたか。

宿久: 講義の方はまだ手探り状態ですね。少なくとも、他の統計解析ソフトとは違うという感覚は持ったようです。ゼミの方は私の手をほとんど煩わせることなく、学生主体で進めてくれます。近々、認定試験を受けるとのことですので結果を楽しみにしています。

SAS: ありがとうございます。引き続き後期の授業状況を次号のSAS Technical News Autumnで伺います。

SAS認定試験 SASグローバル認定プログラム (SAS Global Certification Program)の 傾向と対策

SASグローバル認定プログラムは、SAS製品について深い知識をもつユーザーを認定するための制度で、SASが使用できることを証明できる認定試験です。

今回はシリーズの第一弾として、SAS認定試験を実際に受講・合格された経験談を、住商情報システム株式会社の村本様よりいただきました。皆様の認定試験に関する何らかの動機付けや参考になれば幸いです。

SAS認定試験合格について



住商情報システム株式会社
産業システム事業部 統計解析ソリューション部
SAS Certified Advanced Programmer
村本 和昭

私が現在所属する部署では、医薬品開発における統計解析・データマネジメント・市販後調査・EDCソリューションをメインにトータルなITサービスをお客様に提供しております。

その中で私は、製薬企業向けに臨床試験の統計解析業務支援を行っております。具体的には解析仕様書の作成、SASプログラミングによるデータ加工、解析図表の出力、バリデーション等を担当し、日々の業務をこなしながら、お客様の業務標準化や効率化を提案させていただいております。

SAS認定試験については2004年12月にSAS Base Programming Examに合格し、そして2005年2月にSAS Advanced Programming Examに合格しました。

当時SASプログラマーとして3年が経とうとした頃、業務にも慣れ今度は後輩を育成する立場となって、自身の技術を確認する目的とステップアップを目指しSAS認定試験にチャレンジしました。また部署としてもSAS資格取得を推奨しており、同僚の合格が刺激にもなり自分もやってみようという意識が芽生えました。

SAS認定試験の受験にあたり、勉強期間はそれぞれ約1~2ヶ月くらいかけました。業務内容(とくにデータ加工や図表作成の技術)が出題範囲と近いこともあり、今まで経験した技術を見直し、日常の業務ではあまり利用しない処理やステートメント等を中心に勉強しました。試験に向けての準備として、まず出題範囲からSASのマニュアル・システムヘルプをチェックし対策ノートを独自に作りました。さらに各処理の動きを体で覚えるために、自分でサンプルデータとプログラム集を作成し、目的に応じたステートメントやオプションのパターンを可能な限り試してみました。これらの作業は通常業務後、準備を行っていました。またすでに合格している先輩や同僚から出題の傾向を聞くことも貴重な情報となりました。

実際の試験について出題問題及び回答の標記はすべて英文であり回答は四択なので、英語が苦手な方は抵抗があるかもしれませんが(私もそうでした)。ただ英文を選択する問題だけではなく、ステートメントやプログラムの実行結果について、四択から選択する問題もあり、後者の問題数のほうが多い為それらを確実に回答していくことが、今回の合格に結びついたと考えています。回答内容の確認も、焦らず余った時間に何度も見直しを心がけました。なお合否は受験したその場で確認できます。

認定試験準備のなかで、試験勉強の過程において、通常業務へも活用できるような処理やステートメント等の機能を見出したことは新たな喜びとなりました。Base試験ではDataステップにおけるループ処理とその応用。Advanced試験ではSQLプロシジャ(Dictionaryテーブルも含む)がそうでした。

参考までに、DataステップとSQLプロシジャを例に出題傾向を簡単にまとめてみました。

< Dataステップの処理 >

SASの内部構造を理解し、Dataステップにおける各オプザベーションの動きを意識したプログラミングができるようになったのは、自分にとっては大きな収穫でした。Dataステップ内のDOループに加算処理やOUTPUTステートメント等を組み合わせることによって、処理の可能性が広がりました。Base試験にはDataステップに関する多くの問題が出題されました。

< SQLプロシジャ >

SQLプロシジャは、たとえばDataステップやSORTプロシジャを複数組み合わせていた処理を1つのプロシジャで表すことができます。またSQLには様々なデータ結合処理やマクロ変数の作成も可能です。それらをDataステップで再現することがSQLプロシジャを理解するうえで重要であり、試験対策としても有効でした。SAS Ver.8のシステムヘルプがよくまとまっていたので勉強に利用しました。

例)データセット[A]のオプザベーション数をマクロ変数[&X]へ代入する処理

・Dataステップによる記述

```
data _null_;
  set A end=eof nobs=nobs;
  if eof then call symput('X', put(nobs, best.));
run;
```

・SQLプロシジャによる記述

```
proc sql noprint;
  select count(*) into :X from A;
quit;
```

これらの技術は、試験対策はもとより、業務において効率的なプログラミングを行う上で非常に役に立ちました。そしてマクロ技術もあわせて汎用的な標準プログラムの開発も可能となりました。

合格から数年が経ち、現在は後輩にSAS認定試験を受験するよう勧めています。できればSAS Base試験を合格した直後にAdvanced試験を受験することが飛躍への近道であると、当時の自分の姿と重ね合わせて指導しております。業務と直結する資格があるということはとても励みになり、私自身、この経験を活かして業務の幅を広げると共に、さらに上位の試験への挑戦を考えています。

2008年4月現在、弊社には30名以上のSAS認定試験合格者が在籍しております。会社としてもSAS認定試験合格を奨励しており、合格者には一時金が支給されます。また、SAS社から認定を取得することで、お客様の信頼はもとより、業務の拡大に貢献しています。今後も社内にてSAS認定試験合格者を増やし、お客様へより質の高いサービスの提供を目指していきます。

以上

コラム「SAS四方山話」

コラムは、シリーズで東京医科歯科大学の大橋渉様にSASを使用する上で実体験に基づいたお話をいただきます。大橋様がSASを初めて触ったところの体験や、現在の医学界においてのデータ解析ソフトについてなどをシリーズでお伝えいただきます。

第2回 「SASを難しく見せるものとは？」

大橋 渉

東京医科歯科大学 歯学総合研究科大学院 生命情報学
情報医学センター 臨床・生物統計グループ 特任助教

1) SASの存在感

春号の話と若干重複しますが、私の職場において某クリックだけでできる(?)統計ソフトウェアの講座を開催しようとしたところ、定員一杯の60名以上の申し込みがありました。それに対し、後日SAS講座を開催しようとしたところ、何と申し込みは5名程度だったのです。その時は、「SASの知名度はそんなに低いのか? 医業ではスタンダードだろ?」などと考えてしまい、SASユーザーとしては非常に悲しい思いをしてしまったことも覚えています。

さて、筆者は現在東京医科歯科大学において統計相談の窓口業務を担当させて頂いているのですが、その相談の多くは論文作成のための統計処理方法に関するものです。雑誌によりまちまちではありますが、近年は特に統計処理が適切になされていない論文は受理されない傾向が益々顕著です。さらに、投稿規程の国際的基準とされる「生医学雑誌への投稿のための統一規定」にも、「論文には統計処理に用いたソフトウェアを明記すること」がキッチリと記載されています。また、これも雑誌の種類によるのですが、簡易表計算ソフトやそれらのアドオンソフトなどによる統計処理結果が「信頼性に欠ける」として却下されてしまう事例もあるようです。近年のこのような動向も手伝ってでしょうか、年々相談の件数はもちろんですが、相談の内容も以前よりマニアックなものに(?)変化してきていることが実感できます。

このようなお話をすると、「医学研究者の皆様もようやく統計処理について真面目に考えてくれるようになったものか」とも思えそうですが、同時に「統計処理をどうすればよいのでしょうか?」という相談以上に「統計処理を行ってください」という相談(要求?)も増加傾向にあります。統計処理に対するジャッジが厳しくなっている一方で、研究者自らが統計処理を行えないような状況になってしまっていることが予想されますが、やはり完全に人任せにしてしまうのは考え物です。確かに統計処理について真面目に考えてくれるのはいいのですが、「重要であるからこそ、全て統計の専門家にお任せしたい」というのはいかがなものでしょうか? 生医学雑誌への投稿のための統一規定をはじめとした各種投稿規程のチェックリストには、「統計の専門家の監修を受けましたか?」という項目はありますが、「統計の専門家にやってもらいましたが?」とはありません。たとえ生物統計学の専門家であっても、臨床研究を専門とする医学の研究者より先その研究領域に詳しいはずがありませんから、勿論筆者の勉強不足という指摘もありますが、あくまで一般論として、やはり臨床の研究者自らが統計処理を行えることが理想と言えるでしょう。あくまで個人的意見ですが、臨床の研究者が実験ごとに毎回統計家に対しデータの内容や実験の背景等々について延々と説明する時間を考えた場合、どうしても臨床の研究者自らが統計処理を学習する方が有意義であるようにも思えるのです。臨床医の皆様を中心とした、日々多忙な研究者の皆様には酷な話かも知れませんが、説明を繰り返す時間に統計処理を学習できてしまうかも...と考えるのでした。

冒頭の状況によれば、「やはりプログラミングよりクリックが好まれるのは仕方の無いところなのか」と思ってしまうのは当然なのですが、上記の「統計処理を行ってください」とやって来る研究者の皆様...いえ、「どのようにすればよいのでしょうか?」とやって来られる相談者の皆様も何故か、「やってくれるならばSASで」との要求があるのです。相談の内容から判断して、

特にSASによる解析結果しか認められないような研究機関やジャーナルのための統計処理ではないにしても、相談者の皆様も「同じやるなら、結果に説得力があった方がいいですよ」と、何故かSASによる結果に関する情報はしっかりと持っていたりするので。正直なところ、ここに来るまではSASの知名度は低いものと思っていたのですが、それは筆者の思い違い、実際にはSASの説得力を臨床の研究者の皆様はしっかりとご存知だったのでした。冒頭の事例と同様に、プログラムタイプのフリーのソフトウェアであるRの講習会を開催しようとしたときにも、やはり希望者は5名程度でした。どうやら、「プログラミングを覚えるまではちょっと...。どうしても自身で言うならクリックだけで出来るものが...」、以前の言葉を用いるならば「生物統計はやりたいけれど、やっぱりSASが難しくねえ」というような状況なのでしょう。

2) SASのイメージ

Rはともかくとして、筆者は少なくともSASに関しては「難しそう」「プログラミングが嫌」等々の話はしょっちゅう耳にします。その「難しそう」の中には、おそらく統計学に対する「難しそう」も多分に含んでいるのでしょうか? 先の方の事例も考慮すれば、やはり「プログラミングに対する「難しそう」というイメージも多分に含んでいるのです。確かにクリックだけよりは若干(?)難しいのかも知れませんが、果たしてSASのプログラムは一般的に思われているほど難解で複雑怪奇なものなのでしょうか? データステップといくつかのプロシージャの組み合わせのみで一部機能を使いこなすことも十分に可能ですし、少なくとも生物統計学を学ぶことを断念しなければならないほどの(?)難易度であることはないでしょう。実際に筆者も、クリックタイプのソフトウェアで統計処理をやり始めたユーザーがその後プログラミングに魅せられてしまう事例を結構目にしましたし、さらに多くのユーザーがその使い勝手の良さに目覚めてしまった(?)事例も目にしています。たとえば筆者が学部4年生にSAS実習を行ったときには、90分でExcelデータの読み込み方、180分でそのデータの簡単な抽出・加工方法などが理解できたほどでした。授業後のアンケートでも「思ったより簡単だった」と考えている学生も割近くを占めていることから、やはり筆者としては「SASは必要以上に難しいと考えられている」と思えてならないのです。では、その理由とは一体...?

3) プログラミングだけが理由なのか?

やはりプログラミングであること以外にも理由はあると思いますが、そのひとつとして筆者は「統計ソフトウェアの領域特化」の仕組みによる部分があると考えます。SASは特に他のソフトウェアと比較して医業や金融といった特定領域における利用が多いために、本来はその特定領域の理論や作法が難しいものが、何故かSASが難しいと混同されてしまっていることにあると思います。

もう一点、非常に穿った見方ではありますが敢えて誤解を恐れずに書きます。SASユーザーの一部が(多分ほんの一部であると思うのですが)、「いかにSASユーザーの特権的であるかのように思い、SASを他のソフトウェアと比較して特別なものであるように見せている」「SASを他のソフトウェアと比較して高尚なものであるかのように思い込み過ぎている」部分も少なからず感じられます。早い話が、理由はともかくとしてユーザー数が増加するのを好ましく思わないユーザーも少なからず存在しているようです。たとえばユーザー数の少ないソフトウェアは、当然売れる数も限られていますから教本の種類も少ないまま、高いままになってしまいます。そうなりますと数少ないユーザーの中で、確実に売れそうな層向けの教本しか流通しなくなりますから「初心者向け」「入門編」というカテゴリーは当然品薄になり益々敷居が高く「見えてしまう」ようになる...(以下繰り返し)のスパイラルが完成してしまいます。LinuxやRの事例もそうですが、ソフトウェアを育てるのは他ならぬユーザー...なのではないでしょうか?

以上

カスタマーリレーション

社会学におけるSASの利用について、昨今の国内事例では、SASを使用している社会学という分野での使用が少ないのが実情ですが、海外においては活発に利用されています。今回、新企画として労働政策研究所・研修機構の平田周一様より、社会科学におけるSASの利用についてご紹介させていただきます。

社会学データ分析におけるSASの使用

労働政策研究・研修機構
平田 周一

私は、社会学を専門として学び、現在は労働政策関係の研究を行う独立行政法人の研究員として働いている。また、兼任講師として大学で統計理論、統計解析法の講義を行っている。この度、SASインスティテュートジャパンから社会科学におけるSASの使用例について何か書けないかと提案された。いつもTechnical Newsの記事に感心したりしている自分は決してSASのヘビー・ユーザーではないと思っていたし、困ったなとも思ったのだが、それを承知で自分なりに統計解析はなぜSASなのか、またSASをどう使っているのかについて3回に分けて書いてみたい。

- 背景 -

SASを使用してデータ分析を行うようになってから、およそ20年が経つ。他の統計ソフトウェアも使用しているが、大概は、最終的にSASを使用して報告書などをまとめている。なぜ、SASを使っているかにはいくつかの理由があるが、SASを使うようになったのかについての理由は簡単である。職場にSASが導入されたからだ。大学院で勉強していた時は他の統計ソフトを使用していた。大学院を卒業して、今の職場に入った当時、職場に用意されていた統計ソフトは、コンピュータ会社が独自に作成した無償のソフトだった。このソフトは、そこそこの能力を持っていたが、高度な多変量解析を行うには力不足が否めなかった。しばらく経って、SASが導入された。それまで、職場の統計ソフトに不満を持っていた私は飛びつくようにSASを使うようになった。

当時は、まだ大型コンピュータでSASを動かしていた時代だった。SASを起動すると大型コンピュータの端末のディスプレイ画面に、今のSASでいうコマンドエディターとログ画面が上下半分ずつ表示され、コマンドエディターにプログラムを書いてサブミットするというものだった。最近の統計ソフトの多くは、ブルダグメニューを使って大部分の処理を行えるものが多い。私は、古い人間なので、学生時代に統計ソフトを使い始めたころからプログラム(もどき)を書くことをしていたので、当時のSASの使い方には抵抗はなかった。大型コンピュータでSASを動かしていた時と、現在、PC上でSASを動かしている時とで大きな差はない(もちろん、SASにもアナリスト等のブルダグメニュー機能はあるが、デフォルトでついていないわけではない)。このSASコードというプログラムを書くということが、最近、SASを敬遠する人が多い要因の一つとなっている。しかし、私にとっては、このプログラムを書くことが、SASを使い続ける要因の一つでもある。

- なぜSASなのか -

プログラムを書くことについての利点については、あとで述べてとして、SASを使い続けている他の理由を幾つか挙げたててみたい。

1. 多機能である

SASには統計手法をはじめ多くの機能が用意されていることについては異論を唱える人は少ないだろう。しかも、私見だが、新しい統計手法に対する反応が早いと思う。後で詳細を述べるが、社会学でEvent History Analysisと呼ばれる統計手法がある(経済学などではDuration Analysis等と呼ばれる)。この手法に関するプロシジャーをいち早く取り入れたのはSASだと思う。私がSASを使い始めた1980年代後半、Event History Analysisについて日本ではほとんど知られていなかったが、SASには、この手法に関するプロシジャーがすでに用意されていた。最近では、ミクロな調査データとマクロな統計変数の関係を分析できるMultilevel Analysisという手法も注目されているが、これについてもSASはいち早く導入している。新しい統計手法に対するSASの対応の早さを示す例ではないが、Event History AnalysisとMultilevel Analysisを用いて経時的データの分析を行う方法を示す"Applied Longitudinal Analysis (J. D. Singer & J. D. Willet 2003 Oxford Univ. Press)"というテキストに書かれている内容を、実際の様々な統計ソフトを用いた例がUCLAのホームページの中にあるが、

すべての章について例が上げられているのはSASだけである。これなどもSASの機能の多様さを示す例だろう。

2. 海外の研究者の多くがSASを使用している

SASを使っていて重宝していることは、海外の研究者の間でSASが使用されていることだ。私は、個人的に海外の研究者と交流し、ときには共同研究も行っているが、海外の研究者の多くがSASを使っているの、いわば、SASを言語の一つとして活用している。このようなことがあった。アメリカの大学の研究者と共同論文を書くことになった。お互い、ある会合で一度会っただけで、それぞれどのような研究分野に関心を持っているか以外、お互いの詳しいことは何も知らない。個別に、データの分析を行い、結果を電子メールでやり取りして初めてわかったのだが、お互いにSASを使っていたのだ。以降、SASのコードやアウトプットを直接相手に送りあうことで円滑にコミュニケーションをとることができた。また、アメリカの官庁統計データの多くは生データを公開しているが、多くの場合、SAS形式のデータやSASのインプット・フォーマットを提供してくれる。これも、アメリカでSASが多く使用されている証拠だろう。

3. SASに関する文献が多い

海外の研究者の間でSASの使用者が多いということに関連しているが、SASに関する文献が多い。日本では、SASの参考書の数は決して多くないので、意外に思われる方もあると思うが、アマゾンで、「洋書」及び「SAS」という言葉がタイトルに含まれているものを検索すると膨大なタイトルが出てくる。また、統計学のテキストなどでも、付録としてSASのコードが掲載してあるものがたくさんある。邦訳されているものとしては、A. Agrestiの『カテゴリーカルデータ解析入門(サイエンス社)』では、すべての章の中で論じられている例について再現するためのSASコードが掲載されている。データについても、Agrestiで検索し、彼のホームページを見つければそこから本に掲載されているSASコードとともにダウンロードできる。このような例は珍しいことではない。

このように、海外でSASの使用者が多いのは、先にふれたように、SASが非常に多機能であるということが大きな要因となっていると思われるが、もうひとつ、SASがコードを書いて操作することを基本として買っていることも要因となっているのではないだろうか。ブルダグメニューで操作するのは、特に初心者にとっては、確かに魅力的だろう。しかし、ブルダグメニューに頼りすぎるとどのような操作が行われているかを忘れてしまうと応用力を養えない。その点、SASはコードを書いて指示を与えることを基本としているので、自分が今何をしているのかを知ることができるし、どこをどう変えれば何ができるのかを考えることができる。だからこそ、SASコードを交換し合うことによって知識を交換することが可能になるのであり、多くのテキストにSASコードが掲載されているのだろう。

SAS以外の統計ソフトにおいても、シンタックスやコードと呼ばれる簡易プログラムを書くことができる。しかし、SASの利点はSASコードを書いて操作できる範囲が非常に広い点にあると思う。前にふれたEvent History Analysisに関して第一人者であるシカゴ大学の山口一男教授のEvent History Analysis(1991 Sage Publications)では、分析例をSASやSPSSを用いて具体的に示してあるが、その中で、Event History Analysisを行う際に必要なデータの操作について述べたのに続いて、「このような操作はSASを用いれば簡単に書くことができる(同書p.24筆者訳)」としている。Event History AnalysisとSASについては回を改めて詳しく述べる予定だが、これも、山口教授が日常的にSASを用いているからこそであり、アメリカでSAS使用者の数が多しことを示す例だろう。

また、日本で出版されている「SASでらくらく統計学(得津一郎・高橋秀世著 有斐閣ブックス)」というテキストの中で、中心極限定理について、標本の大きさを変えると標本分布がある一定の値を平均値とする正規分布に近づくことを実際にシミュレートする例がSASコードを用いて示されている。他の統計ソフトを使ってこのような例があるのは、私は見たことがない。同書で示されているプログラム自体は、他のプログラミング言語、例えば、Visual Basicなどを使って書くことは可能だろう。しかし、他のプログラミング言語と比べると、SASコードの方が簡単だ。その上、シミュレートした結果をSASのグラフ機能を用いて簡単にグラフ表示させることができる。グラフ表示まで他のプログラミング言語を用いて行うのは大変な作業になるだろう。SASコードは他の統計パッケージと比べると非常に多機能で、他のプログラミング言語と比べると簡単に書けるという特徴をもっているのだ。

ここまで、抽象的にSASの持つ利点について述べてきた。今回は、Event History Analysisに話題を限定し、SASの持つ利点について具体的に述べる。

Q&A



区切り文字を用いた文字列の連結
文字列変数に含まれる数値でソート
変数グループ毎に転置

Excel上の連続していないセルにDDEでデータを出力する方法
Enterprise Guideでのユーザー定義出力形式の利用
Windows XP HomeへのSASのインストール
Enterprise Guide 4.1のMicrosoft Windows Vistaでのサポートについて
日本語を含めた任意の日時出力形式
回歸直線式の表示

Q

複数の文字変数の値をカンマ区切りで連結し、一つの変数に格納することはできますか？

A

バージョン9より追加された、CATX関数を使用すると区切り文字を指定して、文字列の連結を行うことができます。

例

```
DATA _NULL_;
  dlm=",";
  char1="山田";
  char2="太郎";
  char3="15";
  char4="A";

  results=CATX(dlm, OF char1-char4);
  PUT results;
RUN;
```

出力例

```
山田,太郎,15,A
```

CATX関数の詳細については、TechNews 2004, Autumn号特集記事に掲載されていますので、あわせてご確認ください。

http://www.sas.com/japan/periodicals/technews/pdf/04_aut.pdf

Q

下のようなデータがあり、文字列変数の中の数字順でソートを行いたいのですが、このままだとA-1 A-10 A-12 A-2 A-3 A-4 となり、数字順にソートされません。数字順にソートさせるにはどうしたら良いでしょうか？

```
id age
A-1 10
A-2 15
A-3 9
A-5 5
A-8 6
A-9 3
A-10 12
A-12 7
B-3 9
B-5 3
```

```
B-8 4
B-9 6
B-11 7
B-15 10
```

A

ソートする前に、id変数の中の数字の先頭に0を追加する必要があります。0を追加するには次の処理を必要があります。

1. 変数を「-」文字で分割します。
2. 数値文字列の先頭に0を付加します。
3. 前後の空白を取り除き、分割した文字列を連結します。

1.の処理にはSCAN関数、2.の処理にはZw.dフォーマット、3.の処理にはCATX関数を使用します。以下のサンプルをご確認ください。

例

```
DATA test;
  INPUT id $ age $;
CARDS;
A-1 10
A-2 15
A-3 9
A-5 5
A-8 6
A-9 3
A-10 12
A-12 7
B-3 9
B-5 3
B-8 4
B-9 6
B-11 7
B-15 10
;
RUN;
DATA test2;
  FORMAT id3 age;
  SET test;

  /* 先頭1文字を抽出 */
  ID1 = SCAN(id,1,'-');

  /* 数値文字列を抽出 */
  ID2 = PUT(INPUT(SCAN(id,2,'-'),best.),Z2.);

  /* 数値の先頭に0を付けて連結 */
  ID3 = CATX('-',id1,id2);
RUN;

PROC SORT DATA=test2(KEEP = id3 age);
  BY id3;
RUN;
```

SCAN関数、Zw.dフォーマット、CATXの詳細については、弊社FAQサイト、およびTechNewsに掲載されていますので、あわせてご確認ください。

SCAN関数: <http://www.sas.com/japan/service/technical/faq/list/body/ba154.html>

Zw.dフォーマット: <http://www.sas.com/japan/service/technical/faq/list/body/ba245.html>

CATX関数: http://www.sas.com/japan/periodicals/technews/pdf/04_aut.pdf

Q 次のデータセットを変数グループ毎に転置するにはどうしたら良いでしょうか？

[データセット] -----> [転置後]

NO	NAME	NO	dn1	dn2	dn3
1	aaaaaa	1	aaaaaa		
2	aaaaaa	2	aaaaaa	bbbbbb	
2	bbbbbb	3	cccccc	dddddd	eeeeee
3	cccccc	4	ffffff		
3	dddddd				
3	eeeeee				
4	ffffff				

A TRANSPOSEプロシジャのBYステートメントにて変数名を指定すれば変数グループ毎に転置が実行されます。

例

```
DATA a;
LENGTH no 8 name $ 6;
INPUT no name $;
CARDS;
1 AAAAAA
2 AAAAAA
2 BBBBBB
3 CCCCCC
3 DDDDDD
3 EEEEEE
4 FFFFFF
;
RUN;

PROC SORT DATA=work.a;
BY no;
RUN;

PROC TRANSPOSE DATA=work.a OUT=work.b (DROP=_name_) PREFIX=dn;
BY no;
VAR name;
RUN;
```

TRANSPOSEプロシジャのBYステートメントについては、オンラインドキュメントの以下の記載をご確認ください。

<http://support.sas.com/onlinedoc/913/getDoc/ja/proc.hlp/a000063664.htm>

Q DDEを利用してエクセルにデータをアウトプットする際に、セルを飛ばしながら出力する方法はあるでしょうか。例えば、エクセルのある行の1列～10列までデータを出力する場合に、5つのデータを偶数列のみに出力するようなケースを想定しています。

A 変数と変数の間にタブを出力することで、ご希望の処理が実現できます。

```
FILENAME xl DDE "EXCEL| [BOOK1] SHEET1!C1:C10";
DATA _NULL_;
SET sashelp.class;
t='09'x;
FILE xl DSD DLM=t;
PUT t name t age t sex t weight t height;
RUN;
```

Q Enterprise Guideで、任意のライブラリに作成・保存した出力形式が、ユーザー定義の出力形式として一覧に表示されず、利用できません。出力するにはどうしたら良いでしょうか。

A 任意のライブラリに作成した出力形式を利用するには、FMTSEARCHオプションでそのライブラリを指定する必要があります。

まず、一度Enterprise Guideを閉じ、クライアント側のファイル、
C:\Documents and Settings\userid_xxxx%\Application Data\SAS\Enterprise Guide\4\EGAuto.sas
を編集し、FMTSEARCHオプションを追加して、指定したライブラリの出力形式を利用可能にします。

例)ライブラリ名が"myLibrary"の場合

"EGAuto.sas"ファイルを編集し、以下のような記述を追加してファイルを上書き保存します(記述場所は任意)。

OPTIONS FMTSEARCH =(WORK LIBRARY myLibrary);

必要となるライブラリは、ご利用の環境によります。

Q Windows XP Home editionのマシンにSASのインストール/ソフトウェア更新を行おうとしたところ、うまくできませんでした。

A Windows XP Home editionは、稼動対象OSではありません。システム必要条件を確認のうえ、稼動対象のOSをご利用願います。

この件に関するSASノートを提供しております。

<http://support.sas.com/kb/9/585.html>

また、システム条件のドキュメントは以下でご確認いただけます。

http://www.sas.com/japan/service/documentation/installcenter/the_sas_system/9.1.3_TS1M3/win/win/sreq.pdf

<http://support.sas.com/kb/21/079.html>

Q Enterprise Guide 4.1は、Microsoft Windows Vistaをサポートしていますか？

A インストールメディアのケースに記述されているRev番号が、37以上のCD-ROMに含まれておりますEnterprise Guide 4.1は、Windows Vistaをサポートいたします。

Rev番号は、SAS 9.1.3インストールメディアのケース(黒いバインダーの前面または青いボックス側面)に記述されていますので、インストール作業前に必ず確認します。お手持ちのインストールメディアのRev番号が37未満である場合は、弊社担当営業までご連絡ください。

なお、インストールの際は、Rev番号が37未満のときと同様に、「System Requirements Wizard for SAS Enterprise Guide 4.1」および「SAS Enterprise Guide 4.1」と記載されているCD-ROMを使用して、作業を行なっていただく必要がございますので、ご注意ください。

Q 日時データを日本語を含めた形式で出力するにはどうしたら良いでしょうか。

A NLDATMを使用すると、日本語を含む任意の書式で日時を表した文字列を生成することができます。

構文 NLDATM(datetime,descriptor)

datetime SAS日時値

descriptor 任意の文字列を含んだ日付と時間の出力形式を指定する変数または式

詳細はオンラインヘルプ、またはオンラインドキュメントをご参照ください。

<http://support.sas.com/onlinedoc/913/docMainpage.jsp>

下記に例を示します。

```
DATA sample;
  var1 = NLDTM(DATETIME(), '%Y年%M月%D日%p%I時%M分%S秒');
  var2 = NLDTM(DATETIME(), '現在の時刻は%Y年%M月%D日%p%I時
%M分%S秒です。');
RUN;
```

Q GPLOTプロシジャにて散布図を描く際、SYMBOLステートメントにてI=Rのオプション指定を行い、単回帰直線をグラフ上に描いています。さらに、PLOTステートメントにて、REGEQNオプションを追記することによって、モデル式が左下に表示されているのですが、この位置を左上などに変更できないでしょうか。

A モデル式は、REGEQNオプションを指定することによって、グラフ上に表示させることはできますが、その形式、および位置などを指定することに対応しているオプションはありません。ただし、ANNOTATE機能を用い、単回帰直線のモデル式をテキストとして位置を指定し、表示させることはできます。この場合、最初にREGプロシジャを実行し、パラメータ推定値を求めた後、ANNOTATEデータセットを作成することが必要となります。詳細に関しては、以下のサンプルプログラムをご参照ください。

```
/* 回帰直線のパラメータ推定値をデータセットに出力 */
ODS OUTPUT PARAMETERESTIMATES=est;
PROC REG DATA=sashelp.class;
  MODEL weight=height;
RUN;
QUIT;

/* ANNOTATEデータセットの作成 */
DATA anno;
  LENGTH text $50;
  SET est;
  RETAIN XSYS YSYS '1' POSITION '6' WHEN 'A' SIZE 1.5
  COLOR 'red' STYLE 'simplex' text;
  /*表示する回帰式をテキストとして保持*/
  IF VARIABLE='Intercept' THEN DO;
    FUNCTION='LABEL'; X=2; Y=90;
    text="Regression Equation:"; OUTPUT;
    text=TRIM(DEPENDENT) || " = " || TRIM(LEFT(PUT(ESTIMATE,10.4)));
  END;
  /*傾きの推定値が正である場合には、'+ 'を式に追加*/
  ELSE DO;
    FUNCTION='LABEL'; X=2; Y=85;
    IF ESTIMATE < 0 THEN
      text=TRIM(text) ||
      TRIM(LEFT(PUT(ESTIMATE,10.4))) || " * " || VARIABLE;
  ELSE
    text=TRIM(text) ||
    " + " || TRIM(LEFT(PUT(ESTIMATE,10.4))) || " * " || VARIABLE;
  OUTPUT;
  END;
RUN;

/* ANNOTATE機能を用いてグラフを作成 */
PROC GPLOT DATA=sashelp.class;
  PLOT weight*height / ANNOTATE=anno;
  SYMBOL I=R V=dot C=blue;
RUN;
QUIT;
```

SAS Training

SASトレーニングのお知らせ

特別トレーニングコースの開催のご案内

「SASを利用したData Cleaning Techniques」コース
(2日間)

日程: 2008年8月25日(月)~26日(火) 10:00 ~ 17:00 (東京会場)

価格: 147,000円(税込) / チケットのお取り扱いはありません。

受講対象: データの中の不正な値を見つける様々なテクニックを身に付けたい方

前提知識: 「SASプログラミング」を受講済みか、同程度の知識のある方
学習内容: 本コースでは、ローデータあるいはSASデータセット中に

含まれるエラーを見つけ出す方法について学びます。ここで学ぶ方法には、DATAステップを使ったプログラムと「FREQ」「MEANS」「UNIVARIATE」などのSASプロシジャを活用して、例えばGENDER(性別)やRACE(人種)といったカテゴリ変数における無効値や、無効な数値を、ユーザー定義による基準あるいは自動外れ値検出に基づいて検出する方法などの他、複数のファイルからサブジェクトIDを見つけ出すプログラムや、重複レコードを検出するプログラムなどが含まれます。

コンテンツ: Ron Cody, Ed D. の「Cody's Data Cleaning Techniques Using SAS, Second Edition」を利用して、データ管理/分析の担当者が生データを得た時に実施しておいた方がよい基本的なチェック方法を順に解説・実習します。

- ・カテゴリ情報が正しく入力されているか
- ・数値情報が正しく入力されているか
- ・欠測値が存在するか
- ・実データのチェック
- ・複数のデータがある時のチェック
- ・ダブルエントリのVerification
- ・SQL言語を利用した方法
- ・Data setのvalidation

テキスト: 本コースは、下記英語版の書籍『Cody's Data Cleaning Techniques Using SAS, Second Edition』を使用し、日本語で説明を行います。また、テキストの内容を簡単にまとめた補助資料を用意します。

【注意】受講申込者は、トレーニング申込とは別に事前にご購入ください。(上記書籍を既にお持ちの方は除きます。)購入方法は、下記『書籍購入方法』をご参照ください。

書籍購入方法: 『Cody's Data Cleaning Techniques Using SAS, Second Edition』

注文番号: 61703/価格: 7,665円(本体価格: 7,300円)

別途発送手数料が1,050円かかります。

下記URLの注文用紙に必要事項をご記入のうえ、03-3533-3781までファックスでご注文ください。

<http://www.sas.com/japan/manual/order.pdf>

注文用紙のお客様通信欄に『8/25-26トレーニング受講』とご記入ください。

「SASによる臨床試験データの統計解析【前編】」コース（2日間）

日程: 2008年8月28日(木)~29日(金) 10:00 ~ 17:00 (東京会場)
 価格: 147,000円(税込) / チケットのお取り扱いはありません。
 受講対象: SASシステムによる臨床試験統計解析に従事している方(上級者)
 前提知識: 下記の4コースを受講済みか、同程度の知識のある方

- ・「医薬向け分散分析」
- ・「医薬向けカテゴリカルデータ解析1」
- ・「医薬向けカテゴリカルデータ解析2」
- ・「医薬向けSASによる生存時間解析」

学習内容: 上記テキストは、新医薬品の臨床試験において頻繁に使用される統計解析について、その理論的側面よりもSASによる実際の解析手順、結果の解釈を主眼として書かれています。

【前編にて学習: 2008年8月28日-29日開催】

1章では、予後因子を調整した解析(連続型、離散型、time-to-event型のエンドポイント)と治療 共変量の交互作用の検討について学習します。2章では、閉検定手順、再抽出に基づく多重比較と多重エンドポイントの解析について学習します。

【後編にて学習: 2008年11月以降予定】

3章では、安全性・診断データにおける基準範囲の推定と比較について学習します。4章では、群逐次デザインにおける解析(繰返し有意差検定、エラー消費関数、確率的打ち切り検定)を種々のSASマクロプログラムを用いて学習します。

【特別編: リクエストに応じ開催予定】

第5章では、欠測データに対する高度な解析手法(尤度に基づく解析、多重補完法など)について学習します。

コンテンツ: 【前編にて学習: 2008年8月28日-29日開催】

1. Analysis of Stratified Data
2. Multiple Comparisons and Multiple Endpoints

【後編にて学習: 2008年11月以降予定】

3. Analysis of Safety and Diagnostic Data
4. Interim Data Monitoring

【特別編: リクエストに応じ開催予定】

5. Analysis of Incomplete Data

テキスト: 本コースは、下記英語版の書籍『Analysis of Clinical Trials Using SAS: A Practical Guide』(SAS Press)を使用し、日本語で説明を行います。また、テキストの内容を簡単にまとめた補助資料を用意します。

【注意】受講申込者は、トレーニング申込とは別に事前にご購入ください。(上記書籍を既にお持ちの方は除きます。)購入方法は、下記『書籍購入方法』をご参照ください。

書籍購入方法: 『Analysis of Clinical Trials Using SAS: A Practical Guide』(SAS Press)

注文番号: 59390 / 価格: 13,440円(本体価格: 12,800円)

別途発送手数料が1,050円かかります。

下記URLの注文用紙に必要事項をご記入のうえ、

03-3533-3781までファックスでご注文ください。

<http://www.sas.com/japan/manual/order.pdf>

注文用紙のお客様通信欄に『8/28-29トレーニング受講』とご記入ください。

新規トレーニングコースの開催のご案内

「SAS Web Report Studioカスタマイズと管理」コース（1日間）

日程: 2008年9月16日(火) 10:00 ~ 17:00 (東京会場)

価格: 57,750円(税込) / チケット捺印数 1

受講対象: SAS Web Report Studioをカスタマイズしたい管理者やビジネス分析を行う方

前提知識: SAS Web Report StudioやSAS管理コンソールの使用経験が必要です。

学習内容: ・SAS Web Report Studioの権限を管理します。

・SAS Web Report Studioの表示をカスタマイズします。

・SAS Web Report Studioのスケジュールや配信の機能を管理します。

コンテンツ: SAS Web Report Studioの探索

・SAS Web Report Studioの操作

・SAS Intelligence Platformの定義

・SAS Intelligence Platformの複数階層

SAS Web Report Studioの権限

・配信権限 ・利用可能な権限の表示 ・SAS管理コン

ソールでの権限設定 ・SAS Web Report Studioの権

限の利用 ・権限の割り当て ・ユーザー権限の設定例

・権限割り当ての復元

SAS Web Report Studioのカスタマイズ

・イメージやスタイルシートフォルダの作成 ・イメージや

スタイルフォルダの新規フォルダへのコピー ・イメージの

プロパティ修正 ・イメージファイルのインポート

・セッションタイムアウト間隔の変更

「SAS Information Delivery Portalカスタマイズと管理」コース（1日間）

日程: 2008年9月17日(水) 10:00 ~ 17:00 (東京会場)

価格: 57,750円(税込) / チケット捺印数 1

受講対象: SAS Information Delivery Portalをカスタマイズしたい管理者やビジネス分析を行う方

前提知識: SAS Intelligence Platformの使用経験が必要です。また、「SAS ビジネス・インテリジェンス/データ・インテグレーション・アプリケーション概要」コースを受講済みか、同程度の知識のある方

学習内容: ・新しいポータルテーマを作成します。

・ポータルのグループコンテンツを作成・管理します。

コンテンツ: SAS Information Delivery Portalの探索

・SAS Information Delivery Portalの操作

・コースシナリオ ・ページ操作

SAS Information Delivery Portalの管理

・ポータル・ユーザー権限の操作

・ポータル・コンテンツのセキュリティ

ポータルの表示のカスタマイズ

・カスタマイズテーマの開発

・デフォルトのプリファレンスの変更

SAS Information Delivery Portalの編集(自己学習)

・Public Kioskの管理

・セッションタイムアウト間隔の変更

・ログ出力情報と場所の編集

SAS Institute Japan株式会社では、今後も多岐にわたったトレーニングコースを追加していく予定です。コース内容・日程等の詳細は、順次弊社Webサイトに公開しますので、以下のURLをご参照ください。

<http://www.sas.com/japan/training/>

その他、トレーニングに関する情報については、上記のURLをご参照いただくか、下記トレーニング担当までお問い合わせください。

トレーニングに関するお問い合わせ先

● TEL 03-3533-3835

● FAX 03-3533-3781

● E-mail JPNTraining@sas.com

New Publications

新刊マニュアルのお知らせ

「Cody's Data Cleaning Techniques Using SAS, Second Edition」
http://support.sas.com/publishing/bbu/companion_site/61703.html
データクリーニングのためのプログラムやマクロについて記載された本です。
こちらより内容の一部をご覧いただけます。

<http://support.sas.com/publishing/pubcat/chaps/61703.pdf>

注文番号: 61703

価格: 7,665円(税込)

SASマニュアル申込用紙および最新のPublication Catalog(マニュアル案内パンフレット)は弊社ホームページにて公開しておりますので、併せてご利用ください。
<http://www.sas.com/japan/manual/>

マニュアルに関するお問い合わせ先

- TEL 03-3533-3835
- FAX 03-3533-3781
- E-mail JPNBooksale@sas.com

Latest Releases

最新リリース情報

PCプラットフォーム

Windows版	SAS 9.1.3	9.1 TS1M3
64-bit Windows (Itanium)版	SAS 9.1.3	9.1 TS1M3

UNIXプラットフォーム

Tru64版	SAS 9.1.3	9.1 TS1M3
SunOS/Solaris版	SAS 9.1.3	9.1 TS1M3
HP-UX版	SAS 9.1.3	9.1 TS1M3
HP-UX(Itanium)版	SAS 9.1.3	9.1 TS1M3
AIX版	SAS 9.1.3	9.1 TS1M3
Linux(Intel)版	SAS 9.1.3	9.1 TS1M3
ABI+版	SAS 6.11	TS040

ミニコンコンピュータプラットフォーム

OpenVMS AXP版	SAS 6.12	TS020
OpenVMS VAX版	SAS 6.08	TS407

メインフレームプラットフォーム

IBM版(OS/390, z/OS)	SAS 9.1.3	9.1 TS1M3
富士通版(F4, MSP)	SAS 6.09E	TS470
日立版(VOS3)	SAS 6.09E	TS470
CMS版	SAS 6.08	TS410

Information

SAS Technical News 送付についてのご案内

SAS Technical Newsは次の方を対象にお送りしています。

- ・ SASコンサルタントとしてご登録の方
- ・ SAS Technical Newsの購読をお申し込みいただいている方

今後SAS Technical News購読が不要の方、配信先の変更等をご希望の方は、下記URLよりお手続きください。

配信停止

<http://www.sas.com/japan/corporate/material.html>

配信先変更手続き

http://www.sas.com/japan/sas_privacy.html#inquiry

SAS Technical News Summer 2008

発行
SAS Institute Japan株式会社

テクニカルニュースに関するお問い合わせ先

テクニカルサポートグループ

TEL:03-3533-3877

FAX:03-3533-3781



SAS Institute Japan株式会社

www.sas.com/japan/

東京本社
〒104-0054
東京都中央区勝どき1-13-1
イヌイビル・カチドキ
Tel 03 (3533) 6921
Fax 03 (3533) 6927

大阪支店
〒530-0004
大阪市北区堂島浜1-4-16
アクア堂島西館 12F
Tel 06 (6345) 5700
Fax 06 (6345) 5655