STN
SAS Technical News
For Higher Customer Satisfaction, We Bridge the SAS System Between Customer's World.
SUMMER 2012

特集 …… 01
SAS® Enterprise Guide® 5.1
アナリティクスを
もっと簡単に、迅速に。

SAS Academic News …… 09
- 随想「マーケティングとデータ解析」
- コラム「SAS／JMPとの歩み」

Q&A …… 15
SASトレーニングのお知らせ …… 18
最新リリース情報 …… 20

*a little bit, but...*

NEW

# SAS® Enterprise Guide® 5.1

## アナリティクスをもっと簡単に、迅速に。

2012年2月にSAS® Enterprise Guide® 5.1がリリースされました。SAS Enterprise Guideとはマウス操作によりSASを簡易に扱うことを可能にするクライアントソフトウェアで、最新バージョン5.1は、SAS 9.3に対応しています。SAS Enterprise Guide 5.1では多くの新機能が追加され、既存機能も拡張・改善されています。本特集では次の4章に分けてSAS Enterprise Guide 5.1をご紹介します。

1. SAS Enterprise Guideとは　2. SAS Enterprise Guide 5.1の新機能　3. SAS Enterprise Guide 5.1ご利用開始までの流れ
4. SAS Enterprise Guide 5.1 の学習リソース

# 1 SAS Enterprise Guide とは

## 1.1 SAS Enterprise Guideの製品概要

SAS Enterprise Guide（以後、EGと略）とは、マウスによるメニュー操作を主体としたSASの利用を可能とするクライアントソフトウェアです。SASプログラミングの知識を持たない方でも、データの加工・集計・分析・グラフ作成など従来の SAS と同様の処理を行うことができます。

また、EGでは入門ガイド、ヘルプ、操作メニューを日本語でご利用いただけます。

## 1.2 SAS Enterprise Guideを使うには

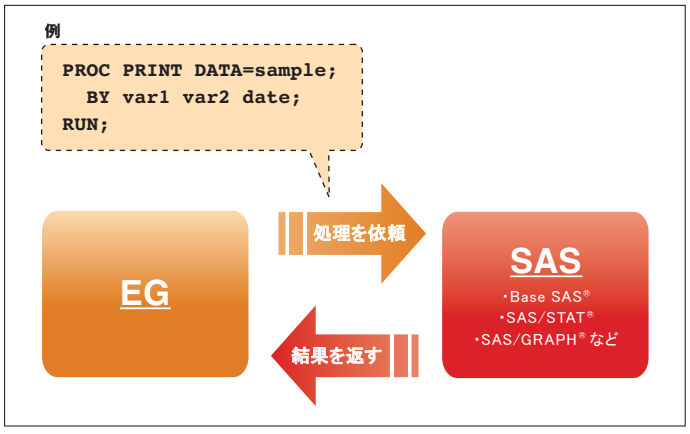
EGはPC版SAS 9の基本パッケージ（例：Analytics Pro）で追加費用なくご利用いただけます。SAS のインストールメディアよりインストール可能です。サーバー版SASの場合は追加のご契約が必要になるため、弊社営業担当者までご連絡ください。

## 1.3 SAS Enterprise Guideの仕組み

EGが動作する仕組みについて簡単にご説明します。

EGは従来の SAS と連携して動作します。ユーザーがEGで操作を行うと、EGは自動的に SASプログラムコードを生成し、その処理を従来のSASへ依頼します。依頼を受けた SAS は処理を行い、結果をEGへ返します。その処理結果は EG 上でレポートやグラフという形で表示されます。

### EGとSASの関係

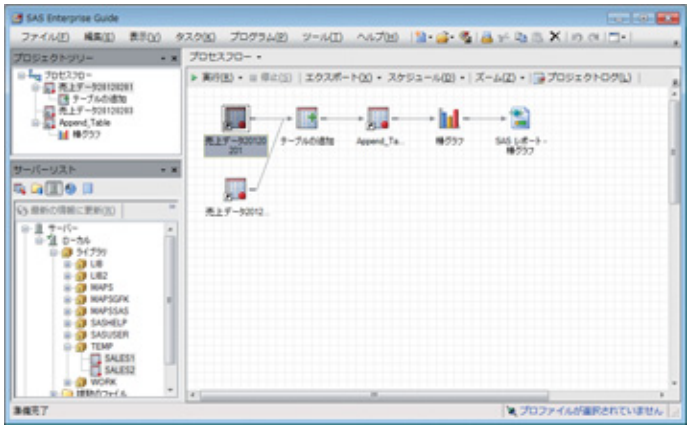



内部的には上述の連携が行われますが、ユーザーでは特にこのような技術を意識することなく利用できます。

## 1.4 SAS Enterprise Guideの基本画面

EGでは次のようなアローダイアグラム（プロセスフローと呼びます）の上に処理を配置することで分析・データ加工アプリケーションを作成します。

### EGの基本画面



一連の処理は左から右、上から下へ順番に実行されます。データの内容によって処理を分岐することや、ユーザーの入力によって処理が変化する動的アプリケーションも作成できます。

プロセスフローの利点は、処理の流れが可視化されることで他のメンバーが作成した分析・データ加工アプリケーションも比較的容易に読み取ることができるという点です。例えば上のスクリーンショット（EGの基本画面）では2つの売上データを結合し、売上データの棒グラフを作成していることが読み取れます。

作成した分析・データ加工アプリケーションは「プロジェクト」というファイル形式（.egp）で保存可能です。保存したプロジェクトはスケジュール実行したり、他のメンバーと共有したりすることができます。

## 1.5 SAS Enterprise Guideの分析・データ加工機能

EGによる分析・データ加工は主に「タスク」というメニュー画面から行います。タスクはメニューバーより選択可能で、用途別に多数のタスクが用意されています。本章ではその一例をご紹介します。

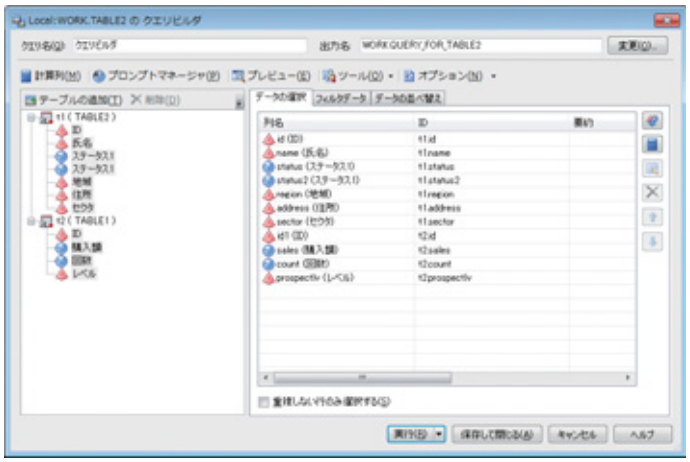
### ● クエリビルダ（データ加工）

クエリビルダとは多様な機能を持つデータ加工専用タスクで、列の並べ替え・列の要約・条件フィルタ・計算列の作成・テーブル結合などに対応しています。クエリビルダを起動するには、メニューバーより「タスク」⇒「データ」⇒「クエリビルダ」を選択します。

### クエリビルダの起動

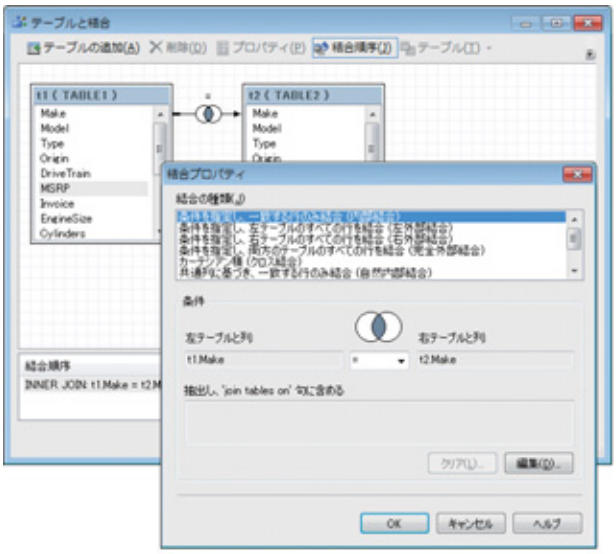
クエリビルダの基本画面は次のようになります。

### クエリビルダの基本画面



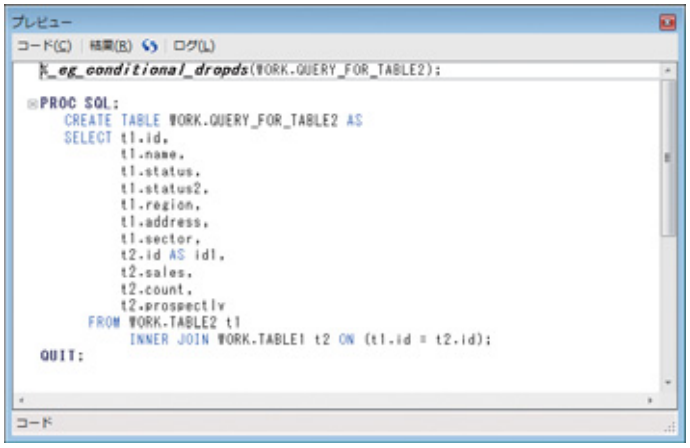
例えば、クエリビルダを使って2つのテーブルを横に結合する処理を行う場合、上記の画面で「テーブルの追加」ボタンを押し対象テーブルを選択した後、結果テーブルに含める列をドラッグ&ドロップで右のペインに追加します。最後に、実行ボタンを押すことで結合が行われます。
従来のSASではDATAステップ、もしくはSQLプロシジャを用いたプログラムコードを記述する必要がありましたが、クエリビルダを使えば上述の操作のみでテーブル結合を行うことができます。
なお、より複雑なテーブル結合を行いたい場合は「テーブルと結合」ボタンを押します。次のようなウィンドウにて詳細な結合条件を指定できます。

### 結合条件の指定



なお、クエリビルダを実行すると内部的に次のようなコードが生成・実行されます。こちらはクエリビルダのプレビューボタンを押すことで閲覧、編集可能です。

### クエリビルダによって生成されるコード



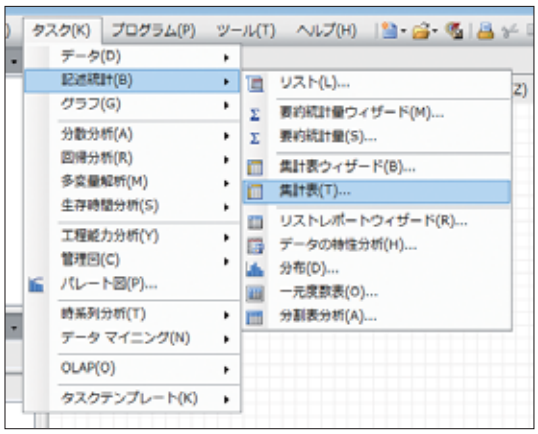
```
%_eg_conditional_dropds(WORK.QUERY_FOR_TABLE2);

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_TABLE2 AS
   SELECT t1.id,
          t1.name,
          t1.status,
          t1.status2,
          t1.region,
          t1.address,
          t1.sector,
          t2.id AS id1,
          t2.sales,
          t2.count,
          t2.prospectlv
      FROM WORK.TABLE2 t1
           INNER JOIN WORK.TABLE1 t2 ON (t1.id = t2.id);
QUIT;
```

### ● 集計表タスク（クロス集計表）

EGの記述統計メニューでは、さまざまな表を作成するためのタスクが用意されています。例えば、集計表というタスクではマウス操作で高度なクロス集計表を作成できます。集計表タスクを起動するには、メニューバーより「タスク」⇒「記述統計」⇒「集計表」を選択します。

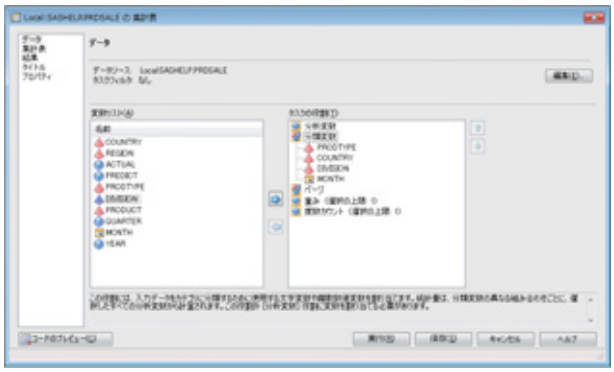
*how EG to operate!*

### 集計表タスクの起動




メニューバー集計表タスクではまず対象となるテーブルを指定し、メニュー画面で分析変数や分類変数などの表で用いる変数を指定します。

### 集計表タスクの基本画面

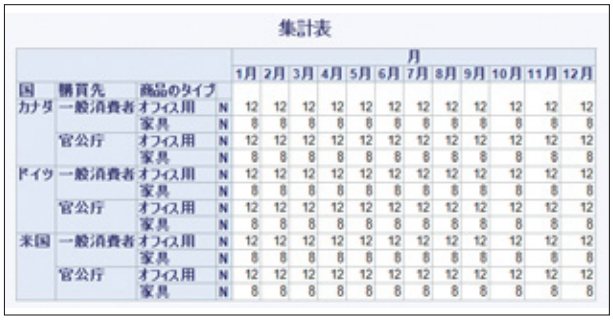


次に、マウスによるドラッグ&ドロップで表を作成します。右のプレビューペインより表の形を整形可能です。また、統計量を追加することも可能です。

### 集計表の整形

実行ボタンでタスクを実行すると、次のようにクロス集計表が作成・表示されます。ここで、結果ウィンドウのエクスポートボタンを選択することで表をHTMLやPDFとして出力することも可能です。

### 集計表タスクの結果

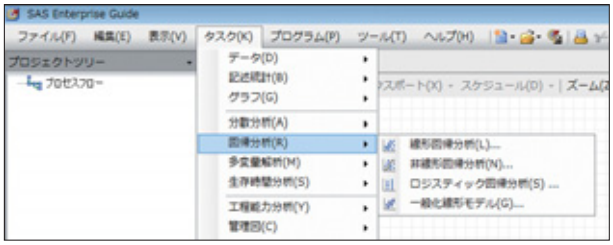



### ● 線形回帰分析タスク（線形回帰分析）

EGでは多数の統計解析タスクが用意されています。これらのタスクを実行すると、内部的にはSAS/STATなどの統計プロシジャを含んだプログラムコードが生成されます。従来のSASであれば、統計プロシジャの構文を学習する必要がありますが、EGのタスクを使えば比較的簡単にさまざまな統計解析を行うことができます。

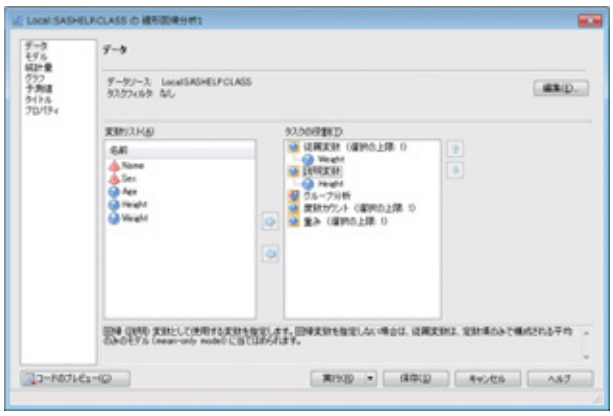
例えば、あるデータに対して線形回帰分析を行う場合、まずメニューバーより「タスク」⇒「回帰分析」⇒「線形回帰分析」を選択します。

### 線形回帰分析タスクの起動




次のようなタスク画面より、データ項目より従属変数と説明変数を選択します。

### 線形回帰分析タスクの基本画面

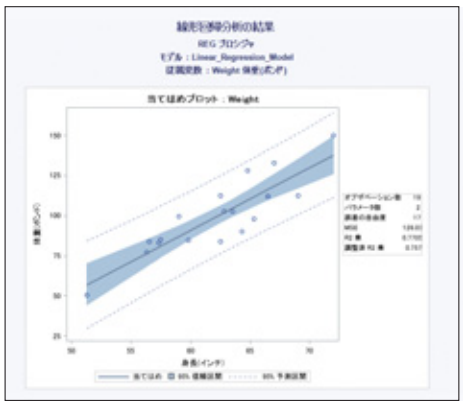


必要に応じて、結果に含めるプロットの種類、モデル、統計量などを選択します。

### グラフ、統計量のオプション画面

実行ボタンを押すと、線形回帰分析の結果とグラフが作成されます。

### 線形回帰分析の結果



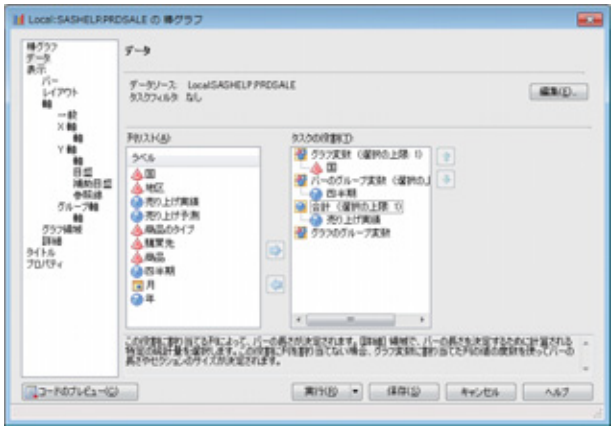

### ● 棒グラフタスク（棒グラフ）

EGでは棒グラフ、円グラフ、等高線グラフ、ヒストグラム、曲面プロットなど多彩な種類のグラフを作成することができます。例えば棒グラフタスクを作成する場合、メニューバーより「タスク」⇒「グラフ」⇒「棒グラフ」を選択します。入力データを選択した後、チャートの種類を選択します。

### 棒グラフ 形状の選択

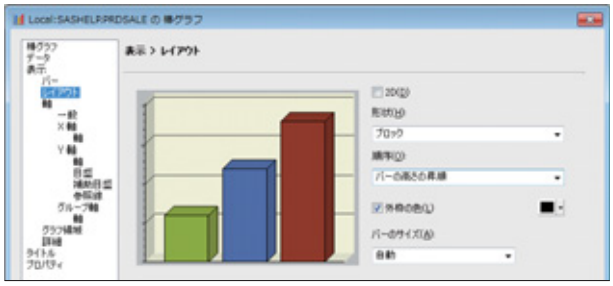
次に、変数に対して役割を割り当てます。役割を割り当てるには、変数を選択し、タスクの役割ボックス内の役割名の1つまでドラッグします。なお、各役割の意味は下の説明文ボックスに表示されます。

### グラフ変数の選択



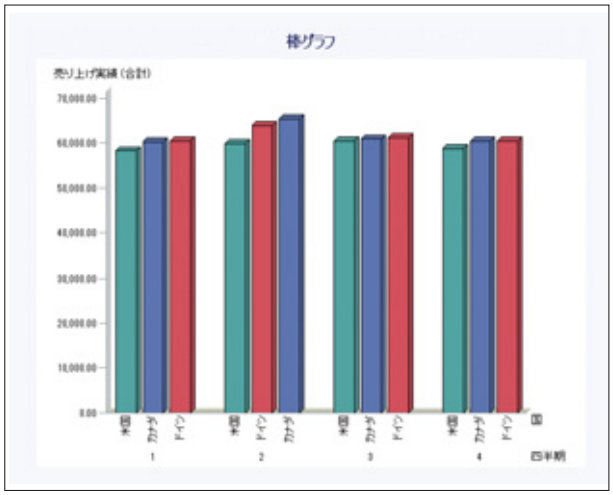
必要に応じて目盛・軸・バー・参照線などレイアウトのオプションを選択します。

### バーの表示形式を選択



実行ボタンを押すと、次のようにグラフが作成・表示されます。

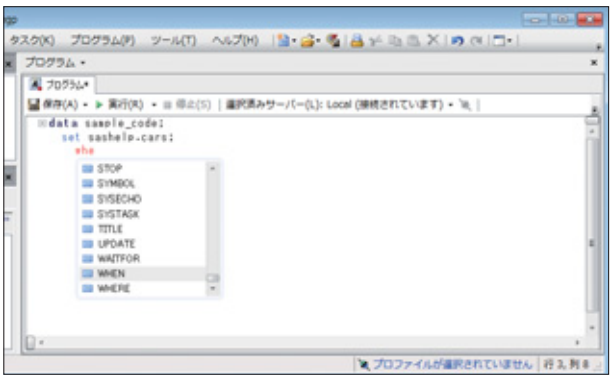
### 棒グラフタスクの結果



### 1.6 SAS Enterprise Guideによるプログラム開発

EGはプログラムの開発環境としても利用できます。従来のSASと同じように拡張エディタを使ったプログラムの入力、実行が可能です。特にEG 4.3以降ではSASの構文やデータセットなどのメンバー名が自動補完される機能（オートコンプリート機能）が実装されています。作成したSASプログラムはプロセスフロー上に配置することができます。このため、データ加工はEGのタスクで、統計分析は既存の分析プログラムで行うという使い方も可能です。また、EGが自動生成したプログラムコードを編集して再利用するという高度な使い方も可能です。

**EGの拡張エディタとオートコンプリート機能**



## 2 SAS Enterprise Guide 5.1の新機能

EG 5.1では多くの機能が追加・拡張されましたが、基本的にはEG4と同じメニュー、同じ操作でご利用いただけます。本章では、新機能の中から5つご紹介します。

- 64ビット版EGの追加
- データエクスプローラツール
- クエリのタスクテンプレート
- 拡張エディタの機能拡張
- ODS統計グラフの表示

### 2.1 64ビット版SAS Enterprise Guideの追加

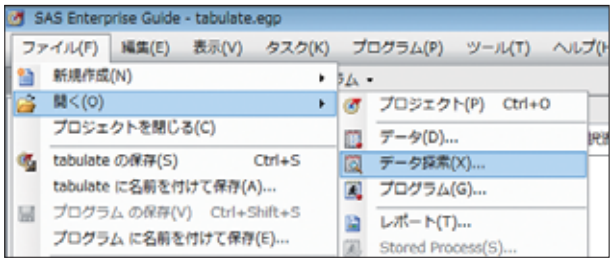
EG 5.1は、32ビットと64ビットのどちらのアプリケーションとしても使用できるようになりました。32ビットMicrosoft Windows OSを使用している場合は、32ビットバージョンのSAS EGを使用する必要があります。64ビットWindows OSを使用している場合は、32ビットと64ビットのどちらのバージョンのEGも実行できます。

64ビット版EGではより多くのメモリを活用できるため、32ビット版EGよりも大きいファイルのインポート/エクスポートが可能になりました。

### 2.2 データエクスプローラツール

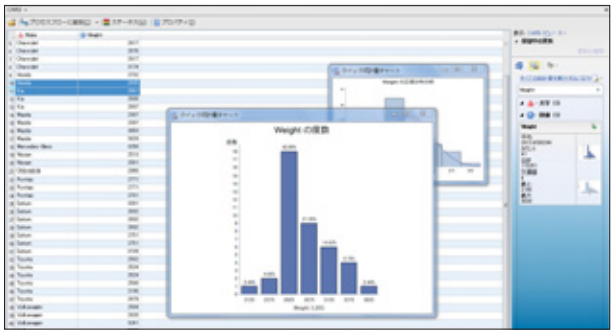
EG 5.1より、データエクスプローラツールが使えるようになりました。データエクスプローラツールを起動するには、メニューバーより「ファイル」⇒「開く」⇒「データ探索」を選択します。

**エクスプローラツールの起動**




データエクスプローラツールを使うと、データの各列の基本的なチャートや統計量を容易に表示できます。また、データの取捨選択、並べ替え、フィルタの設定などを手軽に行えます。

**データエクスプローラツールの操作画面**



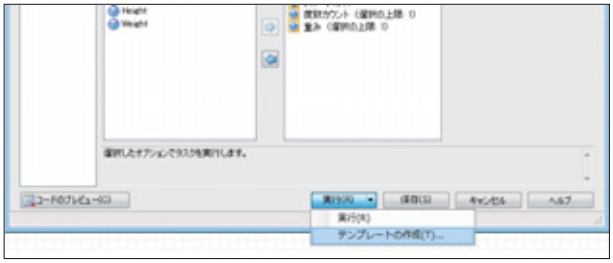
データエクスプローラを使うことでデータの特性や傾向を把握できます。

### 2.3 クエリのタスクテンプレート

タスクテンプレート機能を使用して、クエリビルダで作成したクエリを保存することができるようになりました。

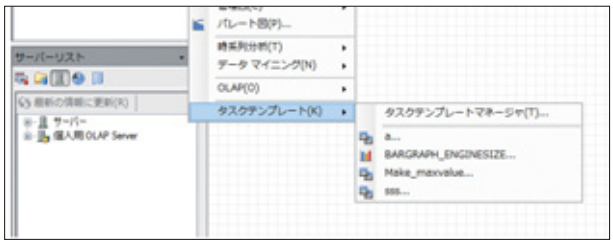
タスクの実行ボタンの隣の矢印を選択し「テンプレートの作成」を選択することで、クエリビルダで設定したオプションやパラメータ値をテンプレートマネージャに保存することができます。

**クエリの保存**




保存したタスクテンプレートはタスクテンプレートマネージャより呼び出して再利用できます。これによりタスクを再作成する手間を省けます。また、一度作成したテンプレートは以後のセッションにも引き継がれるため、異なるプロジェクトでも使用することができます。

### タスクテンプレートの呼び出し

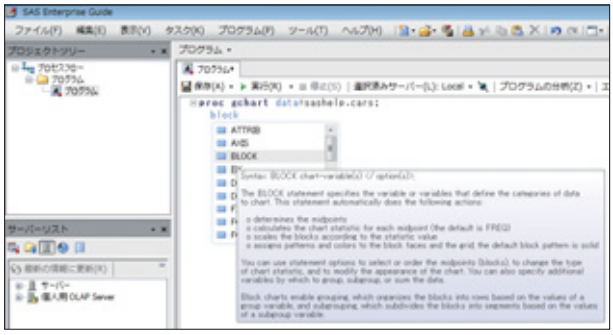



テンプレートマネージャのインポート・エクスポート機能を用いて他のユーザーとタスクテンプレートを共有することもできます。タスク単位での共有が可能となることで、プロジェクトの開発効率を上げることが期待できます。

## 2.4 拡張エディタの機能拡張

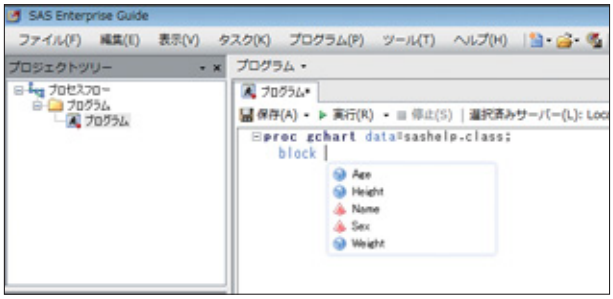
EG 5.1では拡張エディタの構文ヘルプがポップアップで表示されます。調べたい構文の上にマウスをかざすと、構文ヘルプが表示されます。EGでプログラム開発を行う場合、関数やシステムオプションなど構文に関する不明点は拡張エディタ上で速やかに解決できます。

### 拡張エディタの構文ヘルプ機能



さらに拡張エディタのオートコンプリート機能が拡張されました。以前のバージョンでは補完の対象はSASテーブル名のみでしたが、EG 5.1では列名まで補完されます。列名の指定が必要な構文を入力すると、列名の一覧が表示されます。

### 列名の自動補完

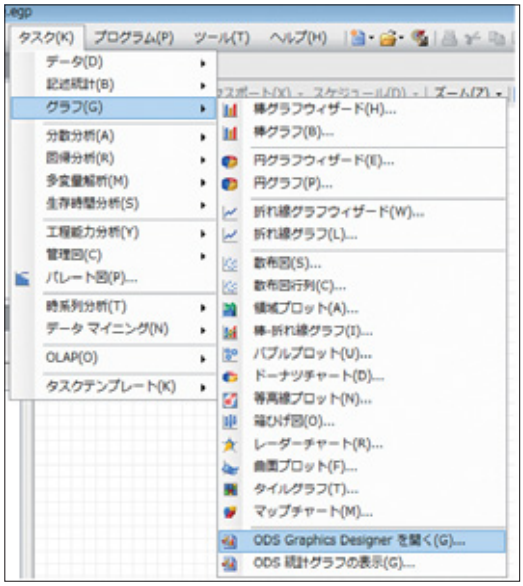



## 2.5 ODS統計グラフの表示

「ODS統計グラフの表示」タスクを使用することで、ODS Graphics Designerでカスタマイズしたグラフテンプレートを EGで利用することができます。

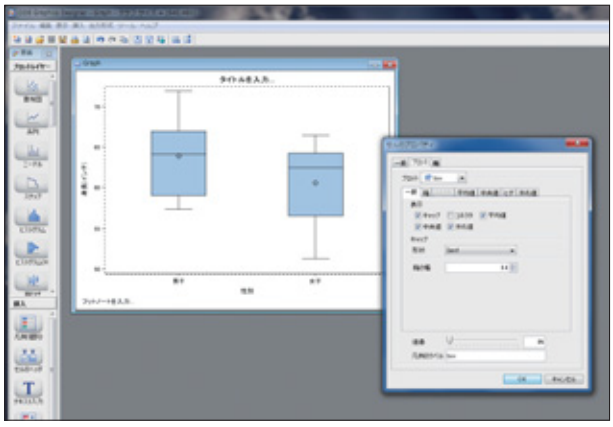
ODS Graphics Designerとは、ODS統計グラフをマウスによる操作で柔軟にカスタマイズすることを可能とする機能です。EG 5.1から起動する場合、「タスク」⇒「グラフ」⇒「ODS Graphics Designerを開く」を選択します。

### ODS Graphics Designerの起動




プログラムを使ってODS統計グラフのレイアウトを変更する場合、TEMPLATEプロシジャやGraphic Template Languageの知識が必要でしたが、ODS Graphics Designerでは次のようなメニュー画面より軸や凡例などの部位を簡単にカスタマイズできます。

### ODS Graphics Designerの操作画面



カスタマイズされたレイアウトはグラフのテンプレートとしてファイル（.sdg）に保存可能です。

EG 5.1の「タスク」⇒「グラフ」⇒「ODS統計グラフの表示」を選択することで、上述のファイルを呼び出し、テンプレートをベースとしたODS統計グラフを作成することができます。

# 3 SAS Enterprise Guide 5.1 ご利用開始までの流れ

まだEG 5.1をご利用されていない方を対象に、EG 5.1導入の流れをご説明します。

## 3.1 システム必要条件の確認

まず、オペレーティングシステムがEG 5.1に対応していることを確認する必要があります。EG 5.1は下記のオペレーティングシステムでサポートされます。

> Microsoft Windows XP Professional (Service Pack 3を適用)
> Microsoft Windows Vista (Service Pack 1を適用)
> Microsoft Windows Server 2003 (Service Pack 2を適用)
> Microsoft Windows Server 2008
> Microsoft Windows Server 2008 R2
> Microsoft Windows 7

また、ご利用いただいているSASのバージョンがEG 5.1に対応していることを確認する必要があります。EG 5.1はSAS 9.3 TS1M1以降でサポートされます。ただし、リモートホストで動作するSASへ接続する場合に限り、SAS 9.2 TS2M0以降でサポートされます。
より詳しいシステム必要条件は、次のページよりご確認いただけます。
http://www.sas.com/japan/service/documentation/installcenter/eguide/5.1/sreq.pdf

## 3.2 SAS 9.3 TS1M1の入手

EG 5.1は、PC版SAS 9.3 TS1M1の基本パッケージのインストールメディアに含まれます。SAS 9.3 TS1M1より前のバージョンをご利用のお客様はバージョンアップを行う必要があります。バージョンアップに関するお問い合わせは弊社テクニカルサポートまでご連絡ください。

## 3.3 SAS およびEG 5.1のインストール

EG 5.1をインストールするには、SAS 9.3 TS1M1 のインストールDVDよりインストールウィザードを起動し、プロダクトの選択画面で「SAS Enterprise Guide」を選択します。

## 3.4 Hot Fixの適用

弊社が提供するHot Fixには既知の問題に対する修正や機能の改善が含まれます。EGに対するHot Fixは次のページにて公開しておりますが、適宜適用いただくことを推奨致します。

**32bit版 EG 5.1のHot Fix**
http://ftp.sas.com/techsup/download/hotfix/HF2/G35_win.html

**64bit版 EG 5.1のHot Fix**
http://ftp.sas.com/techsup/download/hotfix/HF2/G35_wx6.html

# 4 SAS Enterprise Guide 5.1 学習リソース

本章ではEG 5.1を学習・活用するためのリソースをご紹介いたします。

## 4.1 入門ガイド

EG 5.1の入門ガイドはオンラインで公開しております。入門ガイドではEGの基本操作を、スクリーンショットを交えて日本語でご説明しています。EGを初めて操作される方は、入門ガイドをご覧いただくことをお勧めします。
http://support.sas.com/documentation/onlinedoc/guide/tut51/ja/

## 4.2 ヘルプドキュメント

EG 5.1のメニューバーより「ヘルプ」⇒「SAS Enterprise Guideヘルプ」を選択することでヘルプドキュメントが起動します。ヘルプドキュメントはEG 5.1の機能や操作方法について日本語で説明しています。

**ヘルプドキュメントの起動**

**ヘルプドキュメントの画面**

## 4.3 トレーニングコース

弊社にて有償トレーニングコースをご提供しています。EGの入門から実践的なコースまでご用意していますので、ご活用ください。
http://www.sas.com/japan/training/schedule.html

# 5 おわりに

本特集をご覧いただくことでSAS Enterprise Guide 5.1の魅力をご理解いただければ幸いです。

# SAS Academic News

SAS アカデミック・ニュース

本号より、多摩大学大学院経営情報学研究科客員教授 朝野熙彦（あさのひろひこ）先生による新連載がスタートします。朝野先生は、日本のマーケティング分野におけるデータ解析の第一人者であり、マーケティング・サイエンスの実践的な研究者として高名な先生です。
先生のマーケティングに関する豊富な経験と知識を、ユーモアを交え連載いただきます。
第2回、成蹊大学新村先生のコラム「SAS／JMPとの歩み」では、ビジネスとしてSASを使用して本格的な計算サービス活動の開始です。新村先生の統計に対する情熱とプロフェッショナルであるがゆえの現実と理想が語られています。

随想
「マーケティングとデータ解析」

コラム
「SAS／JMPとの歩み」

## 随想 「マーケティングとデータ解析」

朝野 熙彦
多摩大学大学院客員教授

### 第1回 行き当たりばったり半生記

#### 1. 正直に自己紹介

このたび光栄にもSTNへの連載を仰せつかりました朝野です。大学ではマーケティング・サイエンスとマーケティング・リサーチを講義しております。真面目一辺倒のSASテクニカルニュースの中で、リラクゼーションのための唯一のコーナーを担当したいと思います。読者の皆様にお楽しみいただければ幸いです。まず今回は私の行き当たりばったりの半生記から始めて、連載をスタートさせてください。

本誌前号のSPRING 2012を読みますと、私の尊敬する新村先生が社会人第一歩の時代を書いておられました。さすがに一流の研究者だけあって20代でいきなりライフワークに出会い、その当初から研究上の問題意識を明確にされ、20代にして独自の提案を発表し始めたとのことです。自分自身と比べるとどこにも共通点がなく恥ずかしい限りですが、ただ一つ「たまたまの偶然」が人生を決めた、という点だけは私も共通していました。

私は大学時代、将来何をしたいという意欲もなく、そもそも実社会に出るのが嫌でゴロゴロと過ごしていました。スキーとギターだけはしっかりやりましたが、生産的なこと、自分の将来に役立つことは一切しないという牧歌的な気質の学生でした。三つ子の魂百までも、この性格は今でも変わっていません。

さて、就活もしないまま4年の夏休みに入ったころのことです。あるマーケティング・リサーチの会社が求人しているという情報を、客員教授の先生が知らせてくれました。その時期になっても志望業種が決まっていなかったというだけの理由で、私が受験することになりました。それがマーケティングの世界に私が入るきっかけになったのです。大学ではマーケティングなど習ったことがありませんでしたし、調査もしたことがないので、マーケティング・リサーチという仕事は面白そうには思えなかったものです。

入社試験には英文和訳の問題が出ましたが、英文の中にconsumerという単語が出てきました。前後の脈絡からしてコンソメでは意味が通じないなあ、と困ったことを覚えています。試験は出来なかったのに合格でした。後から合格の理由を聞きますと「人手不足なので誰でもよかった」という話でした。入社試験はただのセレモニーだったのですね。

#### 2. 仕事との幸せな出会い

さて、たまたまの偶然で入ってしまったマーケティング・リサーチ会社でしたが、そこでは1年目から営業の仕事をさせられました。企業を回って歩いて、何かマーケティングでお困りのことはありませんか、調査でお手伝いしましょうか、と聞いて回る仕事です。調査企画といえば聞こえはよいのですが、要するに御用聞きですね。もちろんマーケティングで困っていない会社などあるはずがなく、いろいろとご相談を受けました。いろいろといっても一言でいえば「自分が担当するビジネスを成功させて会社が儲けるにはどうしたらよいか」につきます。依頼先の企業は、どうしたらよいかが分からないから相談しているのですから相談されるのは前例のない難題ばかりです。会社に戻って社長に案件を報告しますと、それは自分で解決してよ、という返事です。指示もなければアドバイスもありません。「社長でできるくらいなら社員は雇わない」という誠にもっともな理屈です。これが私にはとても嬉しかったのです。

前例のない新しいマーケティング課題に応えるハウツーマニュアルなど存在しません。だから自分の頭で考えて解決するしかありません。それが規則と有職故実の嫌いな自分にぴったりだったのです。私が就職したのは、老舗の調査会社でしたが、それでも規模的には間違いなく中小企業でしたので、ごく少人数の社員がそれぞれ個人タクシーのように独立して仕事をしていました。調査という仕事は調査企画にはじまり調査票の作成、調査の実施、データの集計、そしてレポート作成とお客様への報告、という一連のプロセスを経るのですが、小さい会社なものですから、担当者が全ての仕事を一人でやることになります。今日、企業らしくなった一部のリサーチ会社では仕事を業務別に分担していると聞きます。その方が経営的には効率がよいのでしょうが、一人の人間が一つの仕事を一貫して行うのは、個人にとって貴重な経験になります。最近、生産管理でも単一工程単能工から複数工程多能工への移行が試みられているそうです。

そういうわけで、学生時代には何の関心もなかったマーケティング・リサーチでしたが、仕事をしてみると実に面白くエキサイティングでした。面白いということは易しいとか楽だという意味ではありません。

その逆で悩ましい課題ばかりで、時には顧客のビジネスを助けてあげられなかった失敗もありました。勝ち戦のマーケティングだとマスコミで喧伝されがちですが、その舞台裏には負け戦が山のように隠れています。その負け戦のいくらかでも勝ち戦にもっていくか、せめて傷を少なくするかに腐心した日々でした。私は幸いなことに在職中にたくさんの難しい仕事に恵まれました。テーマパーク、住宅、家電品、飲食料品、日用雑貨品、ペット用品などいろいろな商品を手がけました。企業秘密のためにあまりオープンにはできませんが、悩みどころの一端なりとも今後紹介させていただきます。

## 3. 会社員時代のアフターファイブ

次は会社員時代の時間外の話です。調査企画や集計分析の仕事を通して、自分にはマーケティングの知識とデータ解析の能力がないことは明確に自覚できました。学生時代に勉強をしなかったのですから当たり前です。その当時勤務先には社員教育の体制がありませんでしたので、細々と独学で勉強をすることにしました。

例えば大学入試センターを会場として昭和61年から月例会が開かれた多変量解析研究会がありました。これは多変量解析の権威である柳井先生を主催者とする全くの私的な研究会でした。岩崎先生（現、日本統計学会理事長）、岡太先生（現、日本行動計量学会理事長）などそうそうたる先生方が研究発表をされていました。内容は難しくて私にはちんぷんかんぷんでしたが、夜間に開かれた研究会でしたので、会社員の私でも参加することができました。

また当時東大の片平先生から、温泉旅行に誘われたこともあります。確かに温泉は好きなので、遠慮もせず伊豆にお供しました。同好の士だという人達が10人くらい集まって、飲んだり騒いだり、という結構な集まりでした。温泉旅行はその後何年も続きました。そのうち、その温泉旅行は実は某学会の研究部会の活動であったことに気づくのです。メンバーはマーケティング分野で有名な先生方でした。飲んでいたのは研究上の思索を深めるためであり、騒いでいたのは互いの研究について真剣に議論を交わしていたのであり、温泉につかっていたのは将来の研究構想を練っていたのだ、という説明もできなくはありません。しかし当時の私にはただ飲んで騒いでいただけとしか思えませんでした。もちろんそれで文句はありませんが。この温泉旅行がマーケティング・サイエンスの世界に私が入るきっかけになったのです。マーケティング・サイエンスには分かりきった知識などほとんどなく、一方で未解決の問題なら山のようにある刺激的な分野だと感じた次第です。

## 4. イーカゲン男の人生訓

その後、研究らしい研究もしていないのにある大学に呼んでもらうことになりました。会社員としての楽しい仕事と高給優遇の身分を捨てるのは残念でしたが、好奇心が勝って転職してしまいました。大学に移ってみると教員の仕事は傍目で想像していたよりも忙しく、雑用もたくさんあって驚きました。しかし自分が選んだ道ですので、愚痴を言うべきではありません。愚痴ではありませんが、怠け者で無計画な自分自身が得た教訓があります。

### ①幸も不幸も受け止め方次第

大学生は就職先として有名な大企業を志望する人が多いようです。たぶんその傾向は昔も今も変わらないでしょう。私自身は全く有名でない中小企業に就職したのですが、それが不幸だとか不運だとは感じませんでした。当時の就職先には、部下を懇切丁寧に指導してくれる上司もいませんでした。しかし、それがかえって幸いでした。自分でやらなければ何も始まらない、という社会人としてもっとも大切なことを学んだからです。

### ②勉強はいつスタートしても遅くない

これは、学びに終わりはないという意味です。もちろん、大学時代からしっかり勉強していて悪いことはありません。しかし社会に出てから初めて勉強の必要性に気づいて勉強を始めるのも素晴らしいことではないでしょうか。近年では働きながら学べる社会人大学院がいくつも出来て、そこでは経営学やマーケティングの高度な実学を学ぶことができます。私もビジネススクールの一教員ですので志ある諸姉諸兄の入学をお待ちしています。

### ③行き当たりばったりで構わない

自分自身のことを振り返ると、他人に人生を計画的に生きろ、などとはとうてい言えません。常に筋書き通りに進むことはよいことなのでしょうか。逆に想定外の事態が起きたときの対応力が心配になります。予定調和からはずれてしまったときにこそ新鮮な発見が生まれるでしょうし、リスクマネジメントも必要になります。人生など紆余曲折が当たり前だと思います。

## 5. マーケティング・データの特徴

マーケティングで扱うデータには次のような特徴があって、伝統的な統計的手法がストレートには使えません。次回以降の予告編の意味で、3つだけ問題点をあげておきます。

### ①要因がコントロールできない

Yが結果でXが要因、Zが余分な変数だとします。フィッシャーの実験計画法では、Y=f(X)というモデルにおいて実験者がXを割り付けることができ、Zは一定にしたり無作為化するなどして、その影響を除去することができました。それはフィッシャーが麦の品種や肥料の種類をXとして割り付けを行い、農作物の収穫量Yの変動を観測したからです。したがって要因Xには誤差は含まれません。ずばり実験計画法と分散分析が活躍できる研究課題でした。

一方マーケティングではYをブランド選択だとか製品の購買量といった市場反応にとります。もし仮にYの要因Xが何であるかが分かったとしても、それを企業がコントロールできるのでしょうか。消費者の性別や収入などを企業の都合がいいように変えることはできません。また交絡する余分変数を除去することもなかなか困難です。消費者行動の原因と結果の法則性はどうやって明らかにすればよいのでしょうか。

### ②観測変数は名義尺度か順序尺度

これもマーケティングらしい特徴です。コミュニケーション活動やプロモーション活動の狙いは何でしょうか。それは自社の商品を、消費者の知名集合や選択集合に入れてもらうことであり、さらには他社と比べて選好順序を高めてもらうことです。

前者は名義尺度で後者は順序尺度です。あなたが飲む可能性があるビールは何ですか？　Aビール、Bビール、Cビール。それではあなたの好きな順はどうなりますか？　一番好き⇒Bビール、二番目は⇒Aビール、三番目は⇒Cビール。

これらのデータには尺度の単位が存在しないので、データに対して四則演算が適用できません。分散も積率相関も計算できません。計量的なデータを前提にして作られたデータ解析法は利用できない、という問題が出てきます。

③文脈依存性

観測値が独立で同一の確率分布に従わないようなデータが多いのもマーケティング・データの特徴です。例えば、グループ・インタビューの発言を考えますと、AさんBさんCさんの発言は、お互いの発言内容が依存関係にあって独立ではありません。個々の発言を抜き出して単独に解釈することはできませんし、座談会の第1グループと第2グループの発言をプールして集計することも意味がありません。このような文脈依存性の高いデータをどうやって統計解析すればよいのか、難しい問題です。

冒頭に述べましたように、私は行き当たりばったりの人間ですので、今後の執筆予定もあてにはなりません。今念頭においているのは、次のような内容です。

- 統計学の独習記
- 因子分析と主成分分析を区別しよう
- 主成分への回帰とは何か
- 特異値分解の面白さ
- SAS/IMLについて
- SASの魅力はどこにあるか
- コンジョイント分析の応用
- マーケティングにおけるデータ解析の悩み
- プレゼンテーションのしかた
- クリエーターとリサーチャーのバトル
- ネット上の統計情報ソース
- 分析データはどこにあるか
- グラフを使ったトリック
- データを用いた我田引水

**略歴：**

千葉大文理学部卒業後市場調査会社に就職、埼玉大大学院修了、千葉大・筑波大講師、専修大・都立大・首都大教授を経て多摩大学大学院客員教授。学習院マネジメントスクール講師、日本マーケティング・サイエンス学会論文誌編集委員、日本行動計量学会理事。

主な著書に『アンケート調査入門』東京図書（編著）、『最新マーケティング・サイエンスの基礎』講談社、『Rによるマーケティング・シミュレーション』同友館、『入門共分散構造分析の実際』講談社、『魅力工学の実践』海文堂出版、『入門多変量解析の実際第2版』講談社、『新製品開発』朝倉書店（朝野熙彦・山中正彦著）などがある。

# コラム「SAS／JMPとの歩み」

新村　秀一
成蹊大学 経済学部教授 理学博士

## 第2回 SASの計算サービスとSASによる研究成果

前号のつづきです。

### 5. 大阪府立成人病センターとの個人研究（1975～1984）

大学卒業の1971年から4年間、NECから出向して行った私の社会人最初の判別分析の研究（2章参照）は、大阪府立成人病センター野村医師の開発した「フィードバックのかかった枝分かれ法」にかないませんでした（2章参照）。枝分かれ論理は、その後の人生で決定木分析やAIとも深く関係してきます。

1975年にSCS社へ戻り大阪の社長直轄の企画部門に配属となりました。1976年の翌年から正式にSASを用いた計算サービスで初めて自立して、東京で仕事をすることになりました。東京に移っても、1984年までの10年間は、大阪府立成人病センターの先生方との稼ぎのない個人研究と統計の独学の苦しい模索の時代でした。当時研究職という職種がなかったので、私一人が仕事の制約のない気楽な身分としてエンジョイしていると同僚や後輩から見られていたようです。

#### 5-1 野村先生との研究

野村先生からは、急性期心筋梗塞の予後予測の研究テーマを与えられ、重回帰分析と判別分析を使った幾つかの論文を出しました[12-13]。そして、1978年に初めてトロントで開かれた医療情報の国際会議で発表しました[14]。稼ぎのない自分をよく国際会議に出してくれたものと思います。実はこの国際会議の座長が、レーダー信号のノイズから敵機を検出するために開発されたROC曲線（Receiver Operating Characteristic Curve、受信者動作特性曲線）を計量診断学に取り入れた医者のLusted博士です。野村先生らが翻訳し、献本を受けていました[15]。発表前の打ち合わせで博士から「君は何分発表したいか」と聞かれ、舞い上がって15分と答えるところ50分といってしまったようです。同じセッションで東大医用電子の開原成允先生[16]も発表されていましたが、先生から小声で間違いを指摘していただきました。国際会議初デビューの懐かしい思い出です。

#### 5-2 疫学部との共同研究

成人病センターの疫学部には、がん検診などの大量の調査データが整備されており、これらのデータを使い疫学部の先生方と共同研究をしました。駆け出しの私には恵まれすぎた環境です。しかし、このことがその後自分でデータを集める習慣が身に付かなかった原因でもあります。

共同研究は沢山あります。その中の代表的な研究として、SASの豊饒な判別手法をROC曲線で比較評価したものがあります。[17-18]は、胃X線像の検査データを各種の判別手法で分析した結果と医師診断の

結果を比較したものです。[19]は、乳がんに適用したものです。

ROC曲線を医学診断に持ち込んだのは医師でもあるLusted博士で、彼の本を野村先生が翻訳したものを献本されていました。各種の判別分析の判別成績の比較評価に利用したのは、恐らく私が世界で最初だと思います。判別成績は、Fisherの2群が同じ正規分布と考えた場合、2群のケース数が異なってもプールした分散共分散を用いて計算されます。ケース数に比例して考える場合、事前確率の指定で変更する必要があります。さらに医学診断等の統計ユーザーの要求を取り入れてリスクを導入することもあります。少なくとも3つの判別成績が異なります。最近でも、研究テーマがなくなってきたためか大学院生や若手研究者にFisherの線形判別関数(LDF)で得られた判別結果より、判別境界を動かすことでよりよい判別結果を探す研究が発表されています。私の研究[17-19]は、そのような研究は無駄であり、判別得点の範囲を20段階あるいは10段階で分割し、それを判別境界にしてROC曲線を描いて評価すればよいというものです。また、LDF、2次判別関数、ロジスティック回帰、FUNCATといったSASの判別手法と数量化Ⅱ類や医師診断などの異なった判別結果の比較評価をROC曲線で行えば、異なった判別手法の評価も行えます。

今日ROC曲線はJMPのロジスティック回帰に取り入れられています。日本で判別分析の評価にいち早く私がとりいれたのは少し自慢です。また、スペクトル診断やROC曲線による判別結果の評価の論文がありますが、私の論文が引用されていないのはさびしい限りです。

## 5-3 清水先生との共同研究

清水先生との共同研究[20-21]は、大気汚染研究に発表したものが、米国環境保護庁(EPA)のデータベースに翻訳され登録されたのも、単に統計処理をお手伝いしただけですがうれしい限りです。

# 6. SASで統計の計算サービスを開始(1977年~)

SASで統計の受託計算サービスを開始して、自分の統計の勉強に役立つような高度な統計手法を使った仕事が舞い込むことを期待しました。ところが最初のころは、FREQプロシジャによる単純集計や多重クロス集計の仕事ばかりです。

### (1) クロス集計の思い出

初期のユーザーとして、某スーパーに段ボール数箱の多重クロス集計の出力を納品しました。1頁幾らという契約であり、また集計の出力頁が多かったため営業担当と2人で軽トラックに帳票を積んで納品に行くと、先方の担当者が1頁1頁検品し始めました。そして、多重クロス集計表の数を数え始めました。複数頁に出力されたものを1頁で精算したいということです。やりあっても埒が明かないので、請求書を書き直し提出することで合意して帰社した記憶は今も鮮明に残っています。それまでそのような単価の交渉でシビアな経験をしたことがなかったので、流通業の厳しさを身をもって実感しました。

### (2) 思わぬ伏兵

また別のユーザーから大量のクロス集計の処理が持ち込まれました。私は、SAS社との販売に関する交渉のため渡米する予定が入っていたため、この仕事の全般を後輩に任せました。後輩が行う作業の中で一部の変数を書き換える必要があり、手作業でIBMの80カラムのカード修正が数千枚におよび、その作成作業が大変であったと後日当人からこぼされたことがありました。

今日、SASのエディターやSASのマクロ機能で簡単に処理できますが、当時はこの作業が思わぬ伏兵でした。数年後に彼は家業に転身しましたが、彼の同期生から半分冗談に「私が無理な仕事をさせたからやめることになった」とたびたび言われました。

### (3) 調査のプロの神業に驚く

また、当時コンピュータ設備を持たない数社の調査会社の研究員がコンピュータを処理時間精算で借りに来ていました。多くの場合、18時以降の夜間割引制度を利用して、24時ごろまでの夜間処理です。もっとも、夜間の方が依頼してから出力されるまでのスループットがよかったこともあります。

ある女性研究員が、3重か4重クロス集計表をチェックしているので、「結果が分かるのですか?」と聞きました。「これぐらいのクロス集計だと、頭の中で構成できます」といわれて驚きました。

### (4) 某大手就職情報会社との継続作業

継続的に一番大きな仕事は某大手就職情報会社です。最初は情報部門からの仕事が多かったのですが、その内、企画部門や人事測定部門が主流になりました。

当時、同社は、就職する学生用に採用企業の分厚い情報誌を作成し、学生に配っていました。その中の記事に学生のアンケートなどの分析レポートを作成するために、SASを使った仕事が複数の女性担当者から引っ切りなしに来ます。

何人かの担当者が、先方の作成したアンケート調査表を事前に見ないで、アンケート票を受け取りパンチ入力してIBMカードを作成し、分析して返すことを数日のサイクルでこなしていたのです。その中で女子学生のアンケート集計の分析で、担当者からどうすればよいかと相談を受けました。質問項目の一つが「あなたの尊敬する女性の名前を自由記述してください」というものです。しばらく考えて、費用がかかるがそのままパンチし、「PROC FREQ;」で単純集計し、同じ女性をまとめたリストを作り、それを再度SASで一つのコードに変換してまとめ、無事難局を乗り切りました。例えばサッチャー首相は、5~6個のバリエーションがあったように記憶しています。

これ以外の仕事も十分な打ち合わせなしで、担当者が同社の女性担当者と主として電話でやり取りしてルーチンワークをこなしていました。稀に土日に私の自宅に処理が間違っているというクレームの連絡が入ったこともあり、東京女子大や横浜国大出といった才媛からのクレームも、今から振り返れば楽しんでいたように思います。

同社は、私が在職中にSASを導入しましたが、SASのプログラム作成の仕事で長く取引が継続しました。

### (5) 大学教育の拾いもの

大学にうつつ統計教育を行って驚いたことは、多くの教員はクロス集計あるいは分割表に対する深い知識やその重要性を認めていないことです。確かに、研究のテーマとして魅力的ではないかもしれませんが、学生が社会に出て一番需要が多いのではないかと思います。統計教育は、自分の専門を教えるのではなく、統計学の手法全体を体系的に社会で必要とするものを

教えるべきだと強く思います。

また、2*2の分割表の場合、Fisherの直接確率で独立性の検定のp値が簡単に計算できます。私は、成蹊大学の1年生の統計入門で、推測統計学の仕組みの説明に用いています。p値の計算が簡単に説明できるのはこれしかありません。しかし、実際の仕組みを教えれば推測統計学は文科系の学生にとってもなじみやすくなります。文献[6]では、Fisherの直接確率、ローンの計算、GoogleのPageRank技術を文科系の学生にもわかりやすく解説していますので、定型外郵便物で切手を張った封筒を大学まで送っていただければ、金額の範囲内で[6]以外の各種資料を含めて送ります。

## 7. 心に残るSASを使った統計の仕事

### (1) 気象予報システムに採用

重要顧客の一つは、東京気象協会、関西気象協会、東北気象協会といった気象協会グループの仕事です。特に東北気象協会の研究者とは相性がよく、各観測所のランダムな雨量の計測値から2重フーリエ変換を使って格子点の雨量を推計するプログラムをFortranで組んだり、SASのロジスティック回帰で降雨確率を予測したりしました。これはその後、「気象庁の予報システムに組み込まれましたよ」という彼の弾んだ電話での報告を受け、大きな達成感と喜びを感じました。

ロジスティック回帰の確率表現が、医学診断に続いて日本の公的な社会システムに取り入れられた先駆けと考えられます。

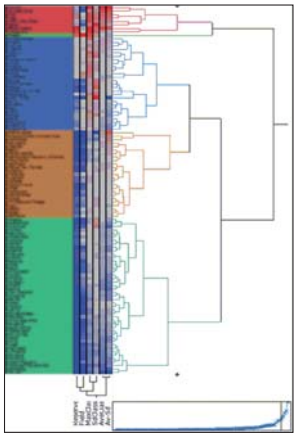
### (2) 某大手金属鉱山

大手金属鉱山会社がアフリカの多分ニジェールかどこかでウラン鉱の探索をやっていました。昔は金鉱山とか石油などの鉱脈を見つけるのは、山師(データマイニングのマイニング)といわれ家庭を顧みず僻地に分け入り調査をする大変な作業だったようです。しかし、その当時は衛星で地上からの電磁波などを検出するリモートセンシングが実用化されてきていました。既存のウラン鉱等のデータを標準パターンとして似た地区を、K-means法で見つけ、その後で確度の高い現地調査するということです。

実は2010年度の社会人の大学院生に東大理学部の修士を卒業し、あるエンジニアリング会社の技術者の方が私の統計の学生でした。わずか2単位の授業で20頁以上の統計レポートを出すことを課していましたが、この方は38頁の見事なレポートを提出してくれました。その中で、次のような主要な産油国の油田のクラスター分析をJMPのWard法でしています。私は授業では、最近隣法、最遠隣法、重心法、群平均法を説明した後、Ward法は説明がめんどうなので行っていません。またグラフはデフォルト主義なのですが、彼は私がウイーンで監修した解説書[24]やHelpで調べて分かりやすいグラフを多用したレポートを作成しました。授業で教えた以上の内容で、ただ脱帽です。今年から博士課程に進学することになりました。

**各クラスターを構成する主要な産油国は次のとおりである。**

| | |
|---|---|
| クラスター1(赤) | 中東巨大産油国 |
| クラスター2(緑) | 東欧産油国 |
| クラスター3(青) | 伝統的大産油国 |
| クラスター4(黄土) | 新興産油国 |
| クラスター5(青緑) | その他の中小産油国 |



クラスター分析結果(デンドログラム)

### (3) 企業評価

判別分析の応用として、ニューヨーク州立大学のAltman准教授による優良企業と倒産企業の判別の研究があります。単に財務データを説明変数とした2群判別の研究ですが、この分野の先駆的な例として有名なものです。同氏は、現在はコンサルタントとして企業評価で活躍しており、一つのことをテーマにして一生の糧を得ることに才能のない私には、うらやましい限りです。

日本では、ある商社系のSCSと同業の情報システム企業が倒産判別システムを作っていました。私はこれまで親会社の商社とは縁がありませんでしたが、業務部門から取引先の倒産判別をしたいということで手伝いました。100社程度の実際の取引先で、倒産した企業と優良企業のデータで2群判別を行いましたがあまり芳しい成果が出ませんでした。説明変数間に多重共線性があり判別係数の符号が常識と反対になっていることを説明したのですが、担当者に納得してもらえませんでした。あとで、日経新聞のCASMAのさきがけシステムの移行を手伝った際、親会社の限られた取引先であること、一人の担当者による調査研究であったことも上手くいかなかった理由と考えるようになりました。

日経新聞社がコンピューターシステムを米国のある企業からIBMへ移行するのに伴い、統計ソフトもSASに変えたいということで移行の手助けをしました。優良企業と不良企業の財務データをもらいSASへの移植方法の説明を兼ね、判別分析の分析レポートを作成しました。分析結果の説明を行いましたが、その後の会議は参加できませんでした。会議の内容を聞くと、企業担当の経済記者を交えて判別分析の結果を再評価するということです。それから暫くして、日経新聞に優良企業100社の紹介記事とそれらの企業の広告が載っていて、統計分析の結果に付加価値をつける商売のうまさにため息が出ました。

## 8. 大学人の間で悪名が広がる

### 8-1 丹後先生との論争

丹後先生は、医療統計に大きな足跡を残した東京工業大学出身の統計家です。丹後先生が東京都の医療機関の研究員であった時代、独自の統計ソフトの開発を行い「医用電子と生体工学[22]」にその成果を発表しました。1980年代は、丹後

先生に限らず心ある多くの統計家が独自の統計ソフトを開発し発表されていました。例えば九州大学の浅野先生を中心とするNISANシステムは、開発などに参加されている研究者も多く最大規模でした。また、小林龍一立教大学元教授や芳賀敏郎東京理科大学元教授のように個人で開発されている方も多かったようです。これらの統計ソフトウェア開発の目的は、自己の統計の知識をソフトウェアで実現したいという欲求であり、外国製の高価な統計ソフトに代わって安い価格で提供したいという欲求などがその開発の動機でした。

しかし私の考えは違っていました。統計に代表される理数系のソフトは、言語に影響されにくい世界的な商品であるべきです。最初は、IBMや日本の汎用機メーカーは統計や数理計画法ソフトを自社開発し、自社の高額なハードウェア・ユーザーに差別化して開発販売していました。例えば、一時日本のPC市場を5割以上押さえていたNECが、自社の理数系のパッケージを販売していましたが、その数理系パッケージソフトウェアをユーザー数が多いPCでも稼働させることは、その当時考えていないようでした。そこで、メーカーの機種に制約を受けない独立企業のパッケージがメーカーのパッケージを駆逐するのは、歴史的必然と考えていました。

また個人の研究家がパッケージを開発するには幾つかの弊害があると感じていたました。パッケージは最初のコンセプトが重要だと思います。特に、SASのようにデータからプロシジャを独立させれば、新規の手法の登録（プロシジャを置き換えるだけで分析手法を変更できる）が容易になります。次に、大きな投資が伴うマニュアルなどの準備や変更・追加などの整備が定期的に必要となります。最後に、多くのユーザーを獲得することでソフトウェアの使用頻度が高まるため、バグの発見や新規機能のユーザーの要望が吸収できるためです。ここがソフト文化と書籍文化の異なる点であると思います。書籍は個人の力量が重要ですが、パッケージは創始者の構想力と、ユーザーによる要望を取り入れた継続的な発展が重要であると思います。そして、新規にパッケージを開発する場合、世界中の既存の同製品を調べ、それを超えていなければ開発すべきでない、というのが私の考えです。

以前にVAX版SASの展示をしているとき、恐らく大学の先生と思われる方から、「自分で作らなくて米国の物を販売していて恥ずかしくないですか」と詰問されたことがあります。その問いに対し、私は「理数系のソフトは世界でNo.1になるかトップグループに入らなければ意味がありません。そこがハード製品との大きな違いです。機能がある水準を満たしていれば、ハードは価格が安ければセグメント化された顧客に売れますが、ソフトウェアは無償にしても使われないという傾向があります。また、既存の製品に後れを取った開発を行うことは、能力の無駄使いです」と答えました。

浅野長一郎先生から折に触れ、「君のおかげで、企業にNISANの説明に行っても、SASの話を持ち出されて往生した」とたびたび言われました。しかし、このグループの研究者からうとまれず一番親しくさせてもらったのは、私にとって非常に幸運であったと感じています。私の大学への転籍を強く働きかけてくれたのは、浅野門下生の渡辺美智子東洋大学教授と山口和範立教大学教授です。また博士号の取得を進めていただいたのは、垂水共之岡山大学教授、田中豊岡山大学元教授ですが、パソコン統計解析ハンドブック（共立出版）の著者です。統計ソフトの普及という面では、まったく異なる意見を持っていました。

惜しむらくは、数理計画法のORソフトで、日本における1980年代のような統計ソフト普及の熱気を感じる時期が一度もなかったことが、日本におけるORの低迷を象徴しているのではないかと思っています。

## 8-2 NLINとMATRIXプロシジャ

SALSは、東大の研究者が開発した非線形回帰ソフトです。[23]のSALSに関する論文の中にSASのNLINプロシジャに関する比較がありました。私はその比較に関して、NLINプロシジャに関して誤解があると思い、MATRIXプロシジャ（SAS/IMLの前身）でNLINプロシジャの機能を再現し、収束域に関する質問論文を情報処理学会に投稿しました。

しばらくして、鈴木筑波大学教授（現副学長）が笑って、「大学人の中には、SASのことを批判的にいうと必ずあなたが登場してくるので、SAS坊やという人がいますよ」と忠告されたことがあります。私はそれほど好戦的な人間でもなく、単に向学心に燃え、違うものは違うのではないかといい、また疑問があればそれを議論したかっただけなのです。

**文献：**


[6] 新村秀一（2009b）. 数学でできる問題解決法、成蹊大学一般研究報告、42/4、1-52.

[12] 野村裕、宮川定吉、若杉茂俊、戸山靖一、新村秀一（1976）. 臨床時系列データの解析、医用電子と生体工学、14(3)、207-213.

[13] 野村裕、新村秀一（1979）. 急性心筋梗塞の予後、医学のあゆみ、110-13、871-878.

[14] Y.Nomura & S.Shinmura (1978). Computer assisted prognosis of acute myocardial infarction, MEDINFO 77, Shires / Wolf, editors IFIP 517-521, North Holland Publishing Company.

[15] L.B.ラステッド（野村裕／中村正彦訳）.（1976）.臨床診断への新しい道 意思決定の理論と実際.コロナ社.

[16] 川村昇、新村秀一、開原成允（1979）. 医学データ解析のための統計プログラムパッケージ. 医学のあゆみ、109-11、620-626.

[17] 新村秀一、鈴木隆一郎、中西克己（1981）. 胃X線像の各種判別分析、オペレーションズ・リサーチ、26-1、51-60.

[18] S.Shinmura, T.Suzuki, H.Koyama & K.Nakanishi (1983). Standardization of medical data analysis using various discriminant methods on a theme of breast diseases, MEDINFO 83, J.H. Van Bemmel, M.J, Ball and O. Wigertz editors, 349-352, North-Holland Publishing Company

[19] 新村秀一、鈴木隆一郎、中西克己（1983）. 各種判別手法を用いた医療データ解析の標準化 — マンモグラフィによる乳癌の診断 — 医療情報学、3-2、38-50.

[20] 清水忠彦、常俊義三、河野洋、新村秀一（1975）. 光化学スモッグによる自覚症状の分類、大気汚染研究、9-4、734-741.

[21] T.Shimizu, Y.Tsunetoshi, H.Kono&S.Shinmura (1975). Classification of subjective symptoms of junior high school students effected by photochemical air pollution.Traslated for NERC-Library,EPA,from the original japanese by LEO KANCER Associates, P.O.Box 5187 Redwood City, California 94063, Nov,1975(TR 76-213)

[22] 新村秀一（1980）. 丹後論文に対する質問. 医用電子と生体工学、18-6、455-456.

[23] 新村秀一（1984）. 中川、小柳「非線形最小二乗法のソフトウェア」についての討論 — SASの評価について、情報処理、25-7、697-707.

[24] J.Sall他（2004）. JMPを用いた統計およびデータ分析入門. SASジャパン.

# Q&A

- 集落抽出法(クラスターサンプリング)
- PUT関数にフォーマット名を変数で渡したい
- IN演算子にて整数を範囲指定する方法
- Windows版SASにて拡張エディタを新規に表示する方法
- SASプログラムの複数行のコメントアウト
- BODYTITLE指定時、タイトルを複数ページに表示させる
- AMOでの簡易分析について
- エクスプローラからのダブルクリックでExcelが起動しない場合に関して

**Q** クラスター(cluster)ごとに多くのオブザベーションを含むデータセットがあります。データセット全体に対するランダムサンプリングではなく、クラスターレベルでのランダムサンプリングはできますか。

**A** 集落抽出法、クラスターサンプリングと言及される手法に関してですが、SAS® 9.2 TS2M3以降では、SURVEYSELECTプロシジャのSAMPLINGUNIT、CLUSTERステートメントにて、クラスターを示す変数を指定します。

**例**

```
                    /* SAMPLINGUNIT ステートメントを指定 */
PROC SURVEYSELECT DATA=indata OUT=outdata
                  METHOD=SRS SAMPSIZE=3 SEED=123;
  SAMPLINGUNIT cluster;
RUN;
```

SAS 9.2 TS2M3より以前のリリースでは、クラスターの水準リストに基づく無作為抽出を行い、入力データと結合することで、行うことができます。

**例**

```
              /* グループ変数 CLUSTER の水準リストを作成 */
PROC FREQ DATA=indata NOPRINT;
  TABLES cluster / OUT=temp1(KEEP=cluster);
RUN;
          /* グループ変数の水準に対し、無作為抽出を実行 */
PROC SURVEYSELECT DATA=temp1 OUT=temp2
                  METHOD=SRS SAMPSIZE=3 SEED=5678;
RUN;
             /* 抽出した水準のみ、オブザベーションを保持 */
DATA outdata;
  MERGE indata temp2(IN=a);
  BY cluster;
  IF a;
RUN;
```

この点に関しては、以下のUsage Noteにて記載があります。
http://support.sas.com/kb/24/555.html

**Q** PUT 関数に指定するフォーマット名を文字変数で渡したいのですが、エラーになります。何か方法はありますか。

**A** フォーマット名を文字変数で渡す場合、PUT関数の代わりにPUTC関数(文字フォーマットの場合)、またはPUTN関数(数値フォーマットの場合)を使用する必要があります。

**例**

```
DATA _NULL_;
  LENGTH pn $10  pc $10;
  INPUT fn :$32. fc :$32.;
  n=TODAY();
  c='AbC';
  pn=PUTN(n,fn);     /* フォーマット名を文字変数 FN で指定 */
  pc=PUTC(c,fc);     /* フォーマット名を文字変数 FC で指定 */
  PUT pn= pc=;
CARDS;
yymmdds10. $upcase3.
yymmddd10. $quote5.
;
```

**Q** ある変数の値が、1から10および50のデータを取得したい場合、プログラムを簡略化する方法はありますか。

**A** SAS® 9.1.3 SP3以降では、IN演算子にて整数値の範囲を指定する場合、:(コロン)を使用することができます。

**例**

```
DATA test;
  DO id=0 TO 100;
    value=INT(RANUNI(123)*1000);
    OUTPUT;
  END;
RUN;
DATA test2;
  SET test;
  WHERE id IN(1:10 50);
RUN;
```

**Q** Windows版SASにて「表示」メニューより拡張エディタを新規に表示することができますが、これらのメニューを介さずに表示する方法はありますでしょうか。

**A** SASのコマンドボックスに「WPGM」とコマンドを入力することで可能です。WPGMコマンドについてはSASオンラインドキュメンテーションをご確認ください。

**WPGM Command: Windows**
http://support.sas.com/documentation/cdl/en/hostwin/63047/HTML/default/viewer.htm#n0d3iwsy8e3pixn1ksvepabkrruq.htm

**Q** SASプログラムを複数行コメントアウトすることはできますか。

**A** コメントアウトしたいプログラムを範囲選択し、CTRLキーおよび/キーでコメントアウトすることが可能です。また、コメントアウトされたプログラムを範囲選択し、CTRLキーおよびSHIFTキーおよび/キーでコメントアウトを解除できます。

**Q** ODS RTFのBODYTITLEオプションを使った場合、複数ページに渡る表のタイトルが2ページ目以降に出力されません。

**例**

```
                    /* 2ページ目以降にタイトルが出力されない */
ODS RTF FILE="c:¥temp¥test.rtf" BODYTITLE;
TITLE1 "タイトル";
PROC PRINT DATA=sashelp.cars;
RUN;
ODS RTF CLOSE;
```

2ページ目以降にもタイトルを出力する方法はありますか。

**A** SAS®9.2の新機能であるODS TAGSETS.RTFを使うことで実現可能です。ODS TAGSETS.RTFではタイトルやフッターを繰り返し出力するUNIFORMオプションが実装されています。

**例**

```
                    /* 2ページ目以降もタイトルが出力される */
ODS TAGSETS.RTF FILE="c:¥temp¥test2.rtf" UNIFORM;
TITLE1 "タイトル";
PROC PRINT DATA=sashelp.cars;
RUN;
ODS TAGSETS.RTF CLOSE;
```

なお、ODS TAGSETS.RTFではデフォルトでタイトルがRTFのBODY領域に出力されるためBODYTITLEオプションは指定不要です。

**Q** SAS® Add-In for Microsoft Officeを使用して統計量を算出し、業務で必要なレポートを作成しています。
その際に、任意のデータの平均値、合計数を参考にしたい場合でも要約統計量タスクなどを使用して適切に設定しないと、それらの値を取得することができません。もっと手軽にそれらの統計量を取得する便利な機能はありますか。

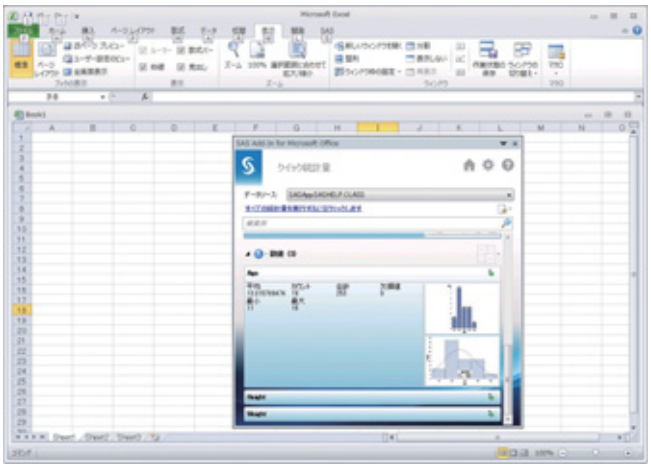
**A** SAS Add-In for Microsoft Office 4.305以前のリリースでは、各種タスクなどを使用しいくつかのステップを経て、期待する結果を得ることができました。
SAS Add-In for Microsoft Office 5.1から各種処理に簡単にアクセスでき、詳細な設定をしなくても基本的な統計量を算出できる「クイックスタート」機能が追加されました。
この「クイックスタート」機能は、下記の機能から構成されます。

1. SAS Add-In for Microsoft Officeで行えるタスクをグラフなどを用いて視覚的に説明し、どのタスクでどのようなアウトプットが得られるか確認してタスクを選べる「タスクギャラリ」

2. 基本統計量を算出する「クイック統計量」

3. 選択したデータに基づき簡易的にグラフ作成する「自動作成チャート」

例えば、任意の数値データにおいて、平均、合計、最小値などの基本的な統計量を参照したい場合、「クイック統計量」機能を選択します。入力データを選択することで、下記の図のように基本統計量とその数値データに基づいた簡易的なグラフを確認することが可能です。また、Excelのスプレッドシート上に出力する必要はありません。



**Q** Excelを起動してからExcelのBookを開くことはできるのですが、エクスプローラからのダブルクリックでExcelが起動しなくなってしまいました。AMOをインストールしてから発生するようになったのかどうかは分かりません。Excelのファイルをダブルクリックで開けるようにするにはどうしたらよいでしょうか。開こうとした際に表示されるポップアップメッセージは次の通りです。

Windows7の場合:

> プログラムにコマンドを送信しているときに、エラーが発生しました。

**A** ご質問の状況には、いくつか想定できるものがございます。以下の各事項についてIT部門のご担当者様とご相談ください。

**1. Officeのファイル検証機能により、Excelの起動時に時間がかかっている**

ご利用のExcelが2003であり、開こうとしているファイルがネットワークドライブである場合に該当する可能性があります。
AMOの有効・無効には関連はありませんが、Microsoftの修正モジュールの適用によっては発生する場合があります。
最新のMicrosoftのモジュールにて修正されているかと思われますのでMicrosoftのファイル検証機能に関するFix(KB2553065)についてご検討ください。

**2. インターネットに接続できないPCでの証明書の確認による遅延**

AMOをインストールしてから発生するのであれば、この事象が考えられます。
PCがインターネットに接続できないなど外部サイトへの接続が制限される環境においてInternet Explorer(IE)のセキュリティ設定で証明書確認が有効である場合、証明書の確認に時間がかかることで起動エラーとなる場合があります。
Officeではこの IEの設定を継承しているため、外部サイトへの接続が制限されている環境の場合、IEの設定を解除することをご相談ください。
弊社では次のNotesにてこの情報を公開しております。
**Problem Note 43085: Microsoft Office applications might require a long time to load if the SAS® Add-In for Microsoft Office is enabled**
**http://support.sas.com/kb/43/085.html**

上記ではIEの詳細設定から[セキュリティ]の項目にある[発行元証明書の取り消しを確認する]のオプションをオフにしていただくことで、起動時間の短縮が期待できます。
Office2003の場合の関連情報は次のMicrosoftのアドレスにて記載があります。
**Office の証明書取り消しに関する処理**
**http://www.microsoft.com/japan/office/ork/2003/seven/ch23/SecA03.htm**

**3. ExcelでのDDEに対するオプションにより動作できない**

これはAMOの有効・無効にかかわらず発生する問題となり、オプションの変更を意図せず行ってしまった場合が考えられます。
このオプションを無効にすることについてご検討願います。
Excelで[Dynamic Data Exchange(DDE)を使用する他のアプリケーションを無視する]の設定が有効になっている場合、エクスプローラからの起動(dde利用)に失敗します。
弊社では次のNotesにてDDE利用に関する設定についての情報を公開しています。
**Usage Note 7017: "Physical file does not exist" error messages can appear when using DDE**
**http://support.sas.com/kb/7/017.html**

関連するMicrosoftの情報には次の情報があります。
ご利用のOS種別ごとに若干エラーメッセージが異なるようです。
**Excel エラー:プログラムにコマンドを送信しているときに、エラーが発生しました**
**http://support.microsoft.com/kb/211494/ja**

**4. Excelのセキュリティ設定により動作が妨げられている**

こちらもAMOの有効・無効には関係はありません。
Excelのセキュリティ設定の変更で影響を受ける可能性があります。
Excelのセキュリティセンターでの「信頼できる場所」にExcelのBookが存在する場所を追加することをご担当者様へご相談ください。
関連するMicrosoftの情報には次の情報があります。
**Excel 2010、Excel 2007またはExcel 2003でレジストリ ポリシー設定によりブロックされている種類のファイルを開くと、エラー メッセージが表示される**
**http://support.microsoft.com/kb/922848/ja**

**Office 2010の信頼できる場所の設定を計画する**
**http://technet.microsoft.com/ja- jp/library/cc179039.aspx**

# SAS Training

## SAS トレーニングのお知らせ

### SAS® Enterprise Guide®研修（中・上級コース）のご案内

前回の「トレーニングのお知らせ」に引き続き、SAS Enterprise Guide（以下、EG）の研修コースをご案内します。

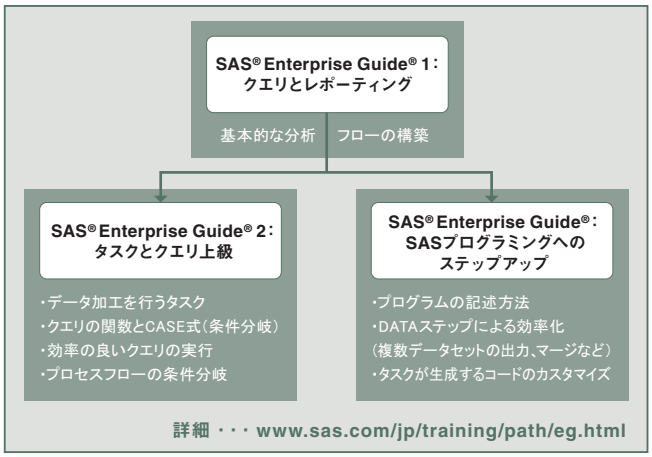



図1：研修フローと主な学習内容

前回紹介した「SAS® Enterprise Guide® 1：クエリとレポーティング」コースを受講いただくと、EGを使用して基本的な分析フローを構築できるようになります。
このフローは単純化すると以下になります。
①データ加工：「クエリ」を使用して分析に必要なデータセットを作成
②分析：「タスク」を使用してレポートを作成

今回ご紹介する2コースは基本的なフローが構築できるようになった方がさらなる機能追加や処理効率向上を図った際、つまずきそうなポイントを押さえているため、EGの上級者へと効率よくステップアップできる、おすすめの研修です。後半の［トレーニング・コース開催のご案内］では開催日程も記していますので参考にしてください。

#### ■SAS® Enterprise Guide® 2：タスクとクエリ上級

このコースはデータ加工に焦点を当てて、クエリやタスクにおける上級なトピックを解説します。
通常データ加工は「クエリ」を用いて行いますが、クエリでは一部実現できない加工もあります。例えば「転置」と呼ばれるデータセット中の行と列を入れ替えるような変換です。このコースではそのような加工を行う「データ加工用のタスク」を紹介します（図2）。

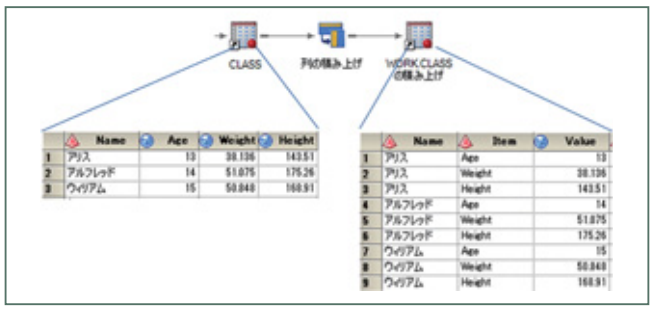



図2：「列の積み上げ」タスクによるデータ加工の例

また、クエリでさまざまな関数やCASE式（条件分岐）を駆使すると、複雑なデータ加工でも1回のクエリで実現できるようになります。実行するクエリの回数はデータ入出力の回数へとつながりますので、処理時間の短縮といった効果が期待できます。

#### ■SAS Enterprise Guide：SASプログラミングへのステップアップ

このコースを受講すると、EGの背後で生成されるSASのコードを把握して、さらなる処理速度の向上や機能追加ができるようになります。タスクの背後では、その設定に応じたSASのプロシジャを実行するプログラムが生成されます。ただし、プログラミング上では提供されているが、タスク上のオプションとしては提供していない設定項目もあります。
このコースでは、タスクのウインドウから［コードのプレビュー］→［コードの挿入］とする方法など、マウス操作のみの設定では足りない部分を補足する方法を習得できます。この方法を使用すると、図3の身長列のようなデータ値に従って背景色を切り替えるようなことが実現できます。

| 行番号 | 名前 | 年齢 | 身長（センチ） |
|---|---|---|---|
| 1 | アリス | 13 | 143.510 |
| 2 | アルフレッド | 14 | 175.260 |
| 3 | ウィリアム | 15 | 168.910 |
| 4 | キャロル | 14 | 159.512 |
| 5 | ジェームズ | 12 | 145.542 |

図3：コードの挿入によるデータ値の背景色変更

また、クエリの背後ではSQLというコードが作成されます。一つのSQLからは一つのデータセットしか出力できません。このためたくさんのデータセットを出力したい場合は、複数のクエリを作成する必要があります。
大量のデータを処理して、複数のデータセットに振り分けていくような場合は何回もデータセットを読み込みなおす必要が出てきます。DATAステップのプログラムを記述すれば、1回の読み込みで複数ファイルの出力が行えます。この場合は読込回数が減ることにより、パフォーマンスの向上も期待できます。
プログラミングの習得はEGの操作と比べるとハードルが高めですが、これを乗り越えるとより機能性・効率性の高いフローが構築できるようになります。

### 定期トレーニング・コースの開催のご案内

#### ●「SAS Enterprise Guide2：タスクとクエリ上級」コース（2日間）

［日　　程］
2012年9月20日（木）～21日（金）10:00 ～ 17:00（東京会場）
［価　　格］
115,500円（税込）／チケット捺印数2
［受講対象］
高度なSAS Enterprise Guideのテクニックを習得したいSAS Enterprise Guideの使用経験者

[コンテンツ]

■イントロダクション
・コース・ロジスティクス
・コースの概要
■データメニューの利用
・データタスク イントロダクション
・テーブルの追加タスク
・列の分割タスク
・列の積み上げタスク
・転置タスク（自己学習）
・無作為抽出タスク
・データの並べ替えタスク
■クエリビルダによる関数の利用
・関数イントロダクション
・値の加工
・データタイプの変換
■クエリ出力のオプション
・クエリ出力のオプション
■プロンプトと条件処理
・複数の値を持つプロンプト
・範囲のプロンプト
・プロジェクト・ステップの条件処理
■CASEロジック
・値の再コード化
・置換の条件
・CASE式の記述
・ネストしたCASE式の記述（自己学習）
■グループ化とフィルタ
・グループ化と要約データ
・要約データのフィルタリング
・詳細データと要約データを含める
・グループ内での要約データのフィルタリング

[Web Page]

**www.sas.com/jp/training/course/eg2.html**

## ●「SAS Enterprise Guide：SASプログラミングへのステップアップ」コース（2日間）

[日　　程]

2012年8月16日（木）～17日（金）10:00 ～ 17:00（東京会場）

[価　　格]

115,500円（税込）／チケット捺印数2

[受講対象]

SAS Enterprise Guideのユーザーで、SASプログラムのコーディングや、SAS Enterprise Guideによって生成されるSASコードの修正について学習し、製品の利用価値を高めたい方

[コンテンツ]

■なぜSASプログラミングを学習するのか？
・コース概要
・SAS Enterprise Guideの基礎の復習
・SASプログラミングを理解することの利点
・SAS Programmingの基本
■SASプログラミングの基本
・SASプログラム イントロダクション
・SASプログラミングの概念
・構文エラーの診断と修正
・SASライブラリへのアクセス
・リレーショナルデータベースへのアクセス（自己学習）
■SASデータセットの読み込み
・DATAステップとフィルタとクエリタスクの比較
・DATAステップの基本
・DATAステップの処理
・条件付き処理
・その他の演算子と出力形式（自己学習）
■生データファイルの読み込み
・区切り記号付きデータの読み込み
・データエラーの調査
・PDVでの値の保持（自己学習）
■SASデータセットの組み合わせ
・DATAステップを使用したデータのマージ
・データの連結
・データの追加（自己学習）
■タスクによって生成されるコードの拡張
・生成されたコードの理解
・生成されたコードとの対話
・ヘルプの利用（自己学習）
■出力コンテンツの制御
・タイトルやフットノートの拡張
・コードによる実行時プロンプトの利用
・条件によるハイライト
■出力の種類の制御
・アウトプット・デリバリー・システム（Output Delivery System）
・HTML出力のカスタマイズ
・RTF出力のカスタマイズ
・PDF出力のカスタマイズ

[Web Page]

**www.sas.com/jp/training/course/eg_prog_9_2.html**

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。

**www.sas.com/jp/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。

**トレーニング担当**
TEL: 03-6434-3690　　FAX: 03-6434-3691
E-mail: JPNTraining@sas.com

# Latest Releases
## 最新リリース情報

### PCプラットフォーム

| | |
|---|---|
| Windows版(32-bit/64-bit) | SAS 9.1.3 / 9.2 / 9.3 |
| 64-bit Windows(Itanium)版 | SAS 9.1.3 / 9.2 |

### 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 / 9.2 / 9.3 |

### UNIXプラットフォーム

| | |
|---|---|
| SunOS/Solaris版 | SAS 9.1.3 / 9.2 / 9.3 |
| HP-UX版 | SAS 9.1.3 / 9.2 / 9.3 |
| HP-UX(Itanium)版 | SAS 9.1.3 / 9.2 / 9.3 |
| AIX版 | SAS 9.1.3 / 9.2 / 9.3 |
| Linux(Intel) 版 | SAS 9.1.3 / 9.2 / 9.3 |

### SAS Technical News入手

SAS Technical Newsは、右記のURLから入手できます。 http://www.sas.com/jp/periodicals/technews/index.html

発行:SAS Institute Japan株式会社

STN SAS Technical News SUMMER 2012

■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:03-6434-3680 FAX:03-6434-3681

SAS Institute Japan株式会社

本社
〒106-6111
東京都港区六本木6-10-1
六本木ヒルズ森タワー 11F
Tel 03(6434)3000
Fax 03(6434)3001

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

www.sas.com/jp

このカタログに記載された内容は改良のため、予告なく仕様・性能を変更する場合があります。あらかじめご了承ください。SASロゴ、The Power to Knowは米国SAS Institute Inc.の登録商標です。その他記載のブランド、商品名は、一般の各社の登録商標です。
Copyright©2012, SAS Institute Inc.All rights reserved.