



統計解析にスプレッドシートを使用することのリスク

内容

- 1 概要
- 1 スプレッドシートに人気がある理由
- 7 スプレッドシートの代替手段
- 10 IBM SPSS Statistics を利用した場合の学習曲線
- 11 結論
- 11 IBM ビジネス・アナリティクスについて

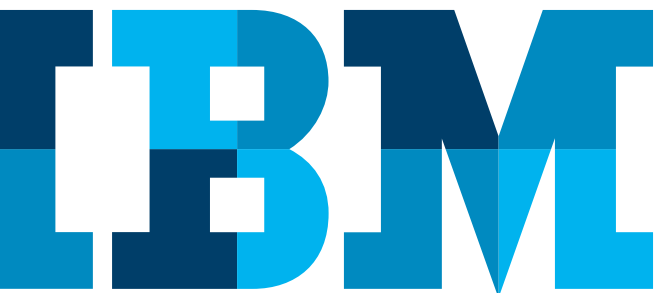
概要

スプレッドシートは、統計解析をはじめデータ分析に幅広く使用されています。確かにスプレッドシートは、有用なツールといえます。しかしながら、役に立つのはある特定の分野に限られると言えます。そもそもスプレッドシートでの実行が意図されていないタスクに使用したり、スプレッドシートの機能の限界に近いタスクや限界を超えているタスクに使用する場合、スプレッドシートは多少リスクを伴う可能性があります。この資料では、スプレッドシートを使用して統計解析を行う（またはその計画を立てる）場合に考慮すべきいくつかのポイントをご紹介します。また、より適切な方法となる代替手段についても説明します。

スプレッドシートに人気がある理由

スプレッドシートは、その使いやすさゆえに、計算の実行に適した魅力的な手段です。ほとんどの方が使い方を知っているか、または知っているつもりになっています。さらに、スプレッドシート・プログラムは、標準的なデスクトップ用コンピューターにプリインストールされているため、既に使用可能な状態になっています。

スプレッドシートは、特定のジョブにとっては素晴らしい発明であり、優れたツールです。しかし、大抵の場合、スプレッドシートはその能力を超えたタスクを実行するように求められます。それはまるで、英語の表現でいうところの「金づちしか持っていない人は、すべての問題を釘と見なしがちだ」という古いことわざのようです。しかし、問題によっては、ドライバー、接着剤、またはベルトのバックルの方が適切に対処できる場合もあります。



さらに、スプレッドシートは使いやすいという考え方は、ある意味、錯覚とも言えます。スプレッドシートから答えを得るのは常に簡単ですが、正しい答えを得るとなると、必ずしも簡単というわけではありません。

しかしながら、何か他のもの（よく知らないテクノロジーやツール）を使用するという決定も、必ずしも容易なわけではありません。代替手段を考慮する際に、次の2つの質問が思い浮かびます。このツールはどれほど役立ちますか？ また、使い方を習得するのはどの程度難しいでしょうか？

最初の質問の答えは、データ分析の規模および複雑さに左右されます。標準的なスプレッドシートは、処理できるレコード数に制限があるため、ジョブの規模が大きい場合には、スプレッドシート以外のツールの方が役に立つ場合があります。

スプレッドシートは統計解析に有用ですが、実行するよう意図されていないタスクに使用する場合、いくつかの制限があります。

複雑さに関しては、データの表面的なレビューだけが必要な場合は、スプレッドシートは適切なツールと言えます。しかし、データに貴重な情報が含まれている可能性があるが、すぐに明らかにならない場合、または詳細な分析を実行したり、隠れたパターンを見つける必要がある場合、スプレッドシートには必要な機能が備わっていません。

考慮すべき他の要因は、必要な分析の精度です。大規模なデータ・セットで作業したり、複雑な計算を実行する場合（あるいはその両方の場合）、スプレッドシートの結果は信頼できない可能性があります。絶対的な精度が必要な場合、スプレッドシートは十分とは言えません。その代わりに、もっと信頼できる別の精密なツールを検討する必要があります。

最後に、単に限られた量の履歴データを分析するだけの場合は、スプレッドシートで間に合います。しかし、信頼できる予測をしたり、傾向を引き出す場合、特に大規模なデータ・セットが関係する場合には、もっと適切なツールがあります。

さて、統計の計算を行うのにスプレッドシートの代わりとなる手段を検討するため、2番目の質問である「使い方を習得するのはどの程度難しいですか？」の答えを考えてみましょう。

しかし、その前に、数値計算以外のタスクにもスプレッドシートは使用される、という点には留意しておいたほうがよさそうです。例えば、スプレッドシートをデータベースであるかのように使用して、リストの作成や管理が行われることがよくあります。ここでも、規模や複雑さの原則が当てはまります。つまり、ある程度の限度を超えたら、データの構造化、データ整合性の保守、監査証跡の作成などの規則が組み込まれた適切なデータベースを使用する方がはるかに妥当です。

スプレッドシートについて覚えておくべき2つの事柄

スプレッドシートで作業する際には、2つの重要な事項を念頭に置く必要があります。それは、スプレッドシートは作成するのが複雑であるという点と、スプレッドシートはエラーを起こしやすいという点です。

スプレッドシートは実際はコンピューター・プログラムである

スプレッドシートのレイアウトを設計する場合、コンピューター・プログラムを作成していることとなります。Microsoft Excel などのスプレッドシート・プログラムは、「非手続き型プログラミング言語」というものを使用します。Visual Basic で Excel 用の手続き型プログラムを作成することも可能ですが、セルに数式を入力するという日常の作業は、非手続き型プログラミング言語で行うものです。

スプレッドシートを作成することは、コンピューター・プログラミングと同じくらい複雑である可能性があります。

通常、プログラミング言語といえば、BASIC、C、Java、FORTRAN などが挙げられます。これらはすべて「手続き型言語」で、それぞれに対象の言語のプログラム向けに作成された、首尾一貫した方法論があります。というのは、プログラムが正常に動作する上で、この規則に厳密に従うことが極めて重要であることが、ここ数年で明らかになっているからです。そのように厳密に従っていても、複雑なプログラムに正しい数値を生成させるには、膨大な量のテストおよびデバッグ作業が必要です。

非手続き型プログラミングは、意思決定、複雑さ、およびミスの可能性の多さについては、手続き型プログラムと変わりありません (ただし、最も単純な手続き型プログラムはこれに該当しません)。

手続き型コンピューター・プログラムは、標準的なソフトウェア開発方法論を使用して、二重三重にチェックされます。対照的に、スプレッドシートは、企業の経営にとって極めて重要であるとしても、通常は一人の人が作成したものです。詳しくチェックされたり、テストされることはほとんどありません。また、たいていの場合、ほとんどまたはまったく検証なしで実運用に投入されます。例えば、収益予測や将来の投資計画など、経営における重要な意思決定でさえ、スプレッドシートが生成する数値に基づいて行われます。

スプレッドシートはエラーを起こしやすい

スプレッドシートのエラーの頻度に関して、多くの調査が行われています。この調査によると、全シートの 90% には最低 1 つのエラーが含まれているようです。調査は、基幹業務のスプレッドシートの目視検査によって実施されました。そのため、他の多くのエラーは検出されなかった可能性があります。また、多くの場合、エラーの修正作業が新たなエラーを引き起こすことも判明しました。

調査によると、全スプレッドシートの 90% には最低 1 つのエラーが含まれていることが明らかになりました。

世界で実際に生じた、損失を伴うスプレッドシート・エラーの例には、次のようなものがあります。

- 「...データ入力エラーで 11 万 8,387 ドルを損失」¹
- 「...1,100 万ドルもの巨額な退職金のエラーの原因は、欠陥のあるスプレッドシートだった」²
- 「...企業がスプレッドシートにより 3,000 万ドル相当のミスを犯した」³
- 「...6 億 4,400 万ドルの誤記載:「急造」プロセスに潜んでいたスプレッドシート・エラー」⁴

スプレッドシート・エラーのタイプ

スプレッドシート・エラーは、主に 3 つのタイプに分けることができます。

「最も親切な」タイプは、いわゆる関数エラーです。このエラーが発生すると、スプレッドシートは動作を停止するので、見つけるのが最も簡単です。誤った数値は表示されず、エラー・メッセージが表示されるか、何も行われません。

次に、外れ値エラーがあります。このエラーが含まれる場合、スプレッドシートは動作しているように見えますが、数値は正しくありません。多くの場合、このタイプのエラーは、ある程度の結果を想定していたものの、その結果が予測値と一致しないという事実に気付いたユーザーによって発見されます。

スプレッドシート・エラーのタイプには、関数エラー、外れ値エラー、ステルス・エラーがあり、重大度は低いものから高いものまでさまざまです。

最も深刻なエラーは、ステルス・エラーと呼ばれるものです。このエラーは、不正確な結果を作成しますが、誰もそれが間違いであると気付きません。検査をパスし、真実として受け入れられます。ステルス・エラーは、正しい結果がどのようなものか誰にもわからない (統計計算の場合に起こりやすい) か、または数値と予測値との差がごくわずかなので妥当と思われることが原因で起こります。

スプレッドシート・エラーがどれほど面倒な結果をもたらしてきたかについては、多くの事例があります。その 1 つは、2006 年 1 月にカリフォルニア州ネバダシティで起きた事柄です。このときには、500 万ドルの赤字予算が見つかりました。予算スプレッドシートは、以前に使用していたものと同じでしたが、新年度のデータを入力したときに、数式が気付かずに上書きされてしまいました。幸いなことに、これは外れ値エラーで、市議員がすぐに気付きました。とは言え、財務担当重役は丸 1 日、修正作業にあたりました (その間、他のエラーもたくさん見つかりました)。

別の事例では、2003 年以降、ある大学で数人の学生の成績の平均にエラーが見つかりました。その数値はまったく意味をなさないものでした。計算を手作業で行った後、調査官は成績を修正しました。そして、スプレッドシートの式のエラーがカット・アンド・ペースト操作によって生じたものであることを突き止めました。この操作の際に、絶対セル・アドレスと相対セル・アドレスの相違を考慮に入れていなかったのです。また、スプレッドシートは主任がチェックしていましたが、入念に調査したのは最初の行だけで、たまたまその行だけが正しかったのです。

スプレッドシート・エラーの原因

スプレッドシートのユーザーは、どの係数がエラーの原因となるのか理解している必要があります。残念なことに、ここにリストできないほど多くのエラーがあります。主なものは次のとおりです。

- **ロジックに含まれるミス:** 誤った関数を呼び出した、加算の代わりに減算した、数式を作成する際に括弧を省略したなど、単純なミスです。このようなエラーは、スプレッドシート内のセルの暗黙の関係によって生じる可能性もあります。
- **誤ってコピーされた数式:** 別の場所から式を読み取って入力すると、多くの場合、カット・アンド・ペーストの実行時にエラーが発生します。既存の式を新しい場所にコピーすると、通常は参照元のセルが変更されるため、正確を期すために変更の結果を確認することが重要となります。
- **誤って上書きされた数式:** 式が含まれるセルは数値のように見えますが、ユーザーが見ているのはすべて式の結果です。そのため、既に数式が含まれているセルに数値を誤って挿入すると、その数式が上書きされ、セルの内容が定数に変わります。他の数式がこのセルの結果に左右される場合、このエラーは深刻な事態を招くおそれがあります。
- **組み込み関数の誤用:** 誤った関数が使用される可能性があります。例えば、AVERAGEA を使用すると、テキストおよび間違えた入力が、AVERAGE ではなくゼロと評価され、エラーが無視されます。残念なことに、このタイプのミスは非常に犯しやすいミスです。

- **省略された係数:** 単に何かを除外するのは、非常に簡単です。それは、式やデータであることもあれば、その両方である場合もあります。このタイプのエラーは、既に完成しているスプレッドシートに新規データが追加される場合に、かなりの頻度で発生します。すべてのデータが入力されていなかったり、新規セルのいくつかが、関係するすべての式に含まれていなかったりするケースがその例です。
- **データ入力エラー:** 運が良ければ、データ入力エラーは外れ値エラーになりますが、常にそうなるとは限りません。例えば、3.5 の代わりに 3.5 と入力した場合、スプレッドシートはこれを数値ではなくストリングと見なします。その結果、そのセルを参照するすべての数式で、値ゼロが使用されます。別のミスは、3/5 と入力した場合です。これは日付と見なされ、計算では非常に大きい数値になります。

他にも多くの可能性があります。例えば、数字と式の両方を含む列をソートする場合、数字に加えて式をソートすると、計算エラーになる可能性があります。

スプレッドシート・エラーは多くの理由で発生します。例えば、論理のミス、正しくない数式、誤って上書きされた数式、および組み込み関数の誤用などがあります。

最後に、スプレッドシート・ソフトウェアの信頼性の問題があります。スプレッドシートの開発者は、ソフトウェアのパッチや修正を頻繁に発行します。しかし、2008年に Gregg Keizer は、⁵ Excel でエラーを修正するためのパッチが、計算で新たなエラーを引き起こすことが判明したと報告しています。実際に以前より多くのエラーが発生しました。(ソフトウェアは5年前に発行されたものですが、問題は検出されておらず、検出されたとしても、修正が加えられていませんでした。)パッチの問題だけではありません。いくつかの調査によると、複雑な数学的プロシージャーや大規模データ・セットを対象とした場合、スプレッドシートは、たとえ正しくコーディングされていたとしても、十分に正確なツールとは言い難いことが分かっています。

スプレッドシートは幅広く使用されており、統計の教育においてさえ使用されているため、多くの記事で、Excel の統計プロシージャーのエラーが詳しく解説されてきました。また、かなりの数の Web サイトで、高度な分析に関するスプレッドシートの欠陥が強調されています。(本資料の末尾にある参考文献をご覧ください。)これらの調査結果を要約すると、「Excel の本格的な分析により、スプレッドシートでは結果の精度をある程度犠牲にしなければならない場合があるため、大規模なデータ・セットや複雑なデータ・セットには適していないことが示される」ということとなります。

スプレッドシートの使用に関連するその他の問題

スプレッドシートのユーザーは、特殊タイプのデータの処理、予測、およびデータ管理になると、制約を受けることがあります。

特殊タイプのデータへの対応

多くのタイプの調査に共通するいくつかのデータ・タイプでは、特殊な対応が求められます。

スプレッドシートを使用した調査を記録する場合、欠損値またはカテゴリー・データを正確に表現するのは特に難しい可能性があります。

頻繁に発生する問題の1つは、欠損データ値を処理する方法です。スプレッドシートで作業する場合、欠損値の処理に注意する必要があります。そうしたデータにゼロを割り当てる場合、結果として一定範囲の値の平均などをゆがめてしまうことがあります。欠損値を示すためにセルにストリングを入力する場合、ストリングを無視する式もあれば、それをゼロと評価する式もあります。場合によってはゼロは有効な値ですが、欠損値を示すには別の方法が必要です。しかし、そのようにする場合、データの入力時に欠損値の指定方式を一貫して使用するよう注意するだけでなく、今後スプレッドシートを変更する際に、これまで使用していたデータ規則を無効にしないように、用いる手法を丁寧に文書化する必要があります。また、上記の手法のどれも、欠損値を確実に代入するための方法として承認されているわけではないことに注目してください。

カテゴリー・データを処理する際に、別の特殊な状況が生じます (多くの場合、調査結果において直面します)。例えば、調査上の質問で 1、2、3、4 という 4 つの値を「はい」、「いいえ」、「不明」、「回答できない」を表すように割り当てるとします。スプレッドシートを使用してこのタイプのデータを保管する場合、値とその意味を文書化するために特別の努力をする必要があります。そして、データが正しく入力され (各回答に割り当てられた正しい値を使用)、意味のある方法で処理されるようにする必要があります。そうしないと、スプレッドシートの開発者が配置転換になったら、意味がなくなってしまいます。

将来の予測

一般的に、スプレッドシートは過去の事象に関する情報や関係を取り出すために使用されます。しかし、組織は将来起こる可能性がある事象に関心を持つようになってきました。Excel の最新版には、FORECAST、TREND、GROWTH などの専用関数があります。これらは、既存のデータに基づいて新規値を予測するためのもので、一定の範囲のプラグイン・プログラムも使用できます。しかし、問題となるのは、それらの機能の信頼性と精度です。どの場合でも、まじめな数学者が結果の妥当性を検査するために使用するようなテストを一切実施していません。

データ管理

スプレッドシートは、セル・レベルが処理の中心対象なので、データ管理の点でいくつかの難しい問題があります。概念的に単純な変更 (開始時間の変更、新規メンバーの追加、数式の変更など) であっても、数十、場合によっては数百もの他の変更を伴うことがあります。

一部のスプレッドシート・アプリケーションには、将来の傾向および結果を予測する関数が含まれているとしても、これらのメソッドは多くの場合、信頼できず、不正確です。

1 つの単純な変更であっても、セル、行、列の挿入または削除、一定の範囲のセルにまたがる数式の編集またはコピー、スプレッドシート全体の再構成が必要になることがあります。これらの操作は時間がかかるだけでなく、実際にさらに多くのエラーを誘発する可能性があります。

スプレッドシートは、完成した後であっても必ずと言ってよいほど、新規データを追加する必要があります。しかし、新しい数字をどのように収容する必要がありますか? 1 つの方法は、式を拡張してすべての新規データを組み込むように、スプレッドシートを設定することです。これに伴う問題は、いくつかの式が不注意に拡張されて、含めてはならないデータを組み込む可能性があるということです。他方、シートが自動的に拡張するように設定されていない場合、含めなければならないデータが除外されてしまう可能性があります。いずれの場合も、変更を慎重にチェックしない限り、スプレッドシートが正しい結果を生成できない可能性があります。

スプレッドシートの代替手段

これまで、この資料では、スプレッドシートが統計解析には不適切であるか、少なくとも面倒な場合があるという状況を概説してきました。ただし、スプレッドシートには価値がないと言っているわけではありません。少数の変数に対して単純なテストを実行する作業の場合、スプレッドシートはどのツールにも劣らないほど適切なツールです。

とは言え、既に述べたように、スプレッドシート・プログラムは汎用のソフトウェアです。プラグインの有無にかかわらず、分析ツールの範囲には限界があり、また、スプレッドシート・プログラムのアルゴリズムは統計解析専用のソフトウェア・プログラムほど厳格に設計もテストも行われていません。

IBM SPSS Statistics は、プログラミングを行わなくても、堅固で徹底的な統計情報分析を行う機能を提供します。

大工は、10 本程度の木材を切断する場合にはのこぎりを使用し、飾り棚を作成する場合には専門の道具に持ち替え、建築用の大量の木材を扱う場合は電力工具を使用します。それと同じように、頑健で徹底的な分析を行いたいと願うなら、そのジョブ専用で作成されたツールを使用する必要があります。その 1 つが IBM SPSS Statistics です。

SPSS Statistics は、1968 年以降、開発およびテストが継続的に行われています。その期間中、多くの形式の統計解析がこのソフトウェアに組み込まれてきました。また、式を実行するアルゴリズムが、学界、研究室、および事実上あらゆる種類の企業の開発者とユーザーの双方によってテストされました。そのため、ユーザーは、ソフトウェアの徹底したテストが実施され、その結果は信頼できることが判明しているという確信を持つことができます。

プログラミングを何も行わなくても、ユーザーは非常に広範囲に及ぶ統計解析を実行できます。さらに、ユーザーの理解力が向上すれば、既にソフトウェアに組み込まれている高度なメソッドをすぐに応用することができます。

当然のことですが、SPSS Statistics は、スプレッドシートには行えなかった方法で統計計算を処理するように最適化されています。実際に、ソフトウェアは、データ入力から、意思決定者向けのレポートの作成に至るまで、あらゆるポイントの統計作業に合わせて最適化されています。

SPSS Statistics によるデータ入力

SPSS Statistics の場合、データ入力プロセスは、使用するデータ・タイプを定義することから始まります。これらの定義はかなり詳細にわたります。例えば、すべてのデータ・タイプには長い名前と短い名前があります。(最適な名前は、表およびグラフの注釈付けに使用される名前です。) さらに、単純な例を示すために入力できるデータのタイプ (数値またはテキスト) を指定できます。この時点で、エラー・チェックの第 1 レベルが実施されます。データは、定義されたタイプの特性に適合する必要があります。そうしないと、受け入れられません。データ・タイプもその他のレイアウト特性も、誤って修正することはできません。データ間の関係も変更できません。データ入力は単なるデータ入力に過ぎず、プログラミングと混ざり合うことは決してありません。⁶

SPSS Statistics で提供されるデータ検証およびエラー・チェックのメカニズムは、非常に包括的です。文書のメタデータ全体または選択した変数値の個別比較により、2 つのデータ・ファイルまたは 2 つのデータ・セットを比較して、それらの間の相違を特定できます。自動プロシージャが、調和しないように見える値を見つけます。この方法で、ほとんどのタイプミスに対処できます。とは言え、値が範囲内にあるが、入力された他の数値と比較したときに何らかの異常がある場合、SPSS Statistics はそれを発見し、それについて尋ねます。

SPSS Statistics に組み込まれたデータ検証およびエラー・チェックのメカニズムは、入力データが有効で正しいものであることを確認するのに役立ちます。

データ分析の準備: SPSS Statistics の手法

上述のとおり、多くの場合、分析に使用できるデータは不完全であることがよくあります。例えば、調査では、回答し損ねたり、質問に回答することを選ばなかった人がいるかもしれません。既に指摘したとおり、スプレッドシート内の不完全なデータを処理すると、多くの問題が引き起こされます。調査員は、SPSS Statistics を使用して、使用可能なデータを調査し、欠落している項目の値を計算します（「代入」というプロセス）。6 つの診断レポートの 1 つを使用してデータを調査し、欠落しているデータのパターンを明らかにすることができます。

あるいは、要約統計を推測して、欠損値を代入することもできます。その場合、自動化プロシージャを使用して、データの特徴に基づいて最も適切な代入メソッドを選択します。さらに、すべてのデータが存在しているかのように分析を実行できます。この場合、非常に現実的で数学的に有効であるという意味で、データが存在していると言えます。

分析用のデータを準備するその他のデータ手順には、データの分散を考察し、外れ値を調べ、データを編成または「ビンニング」することが含まれます。これらの手順により、使用するアルゴリズム（単純ベイズやロジット・モデルなど）が効果的に機能します。SPSS Statistics が実行するこれらのデータ準備手順は、スプレッドシート・プログラムでは実行するように設計されていません。

SPSS Statistics を使用した統計解析

SPSS Statistics が分析モードに切り替わり、出力の生成に必要なアクションを実行するとき、データは変更されません。データはプロセスへの入力としてのみ使用されます。また、プロットおよびグラフの素晴らしい配列を含む、多様な形式で使用可能な出力は、別のウィンドウに表示されます。

SPSS Statistics によって分析されたデータは、プロットおよびグラフを印象的に並べるなど、多様な形式に出力できます。

さらに、どのタイプの分析を実行する場合でも、変更しなくてもさまざまなデータ・セットで何度も保存および実行できる構文形式で、自動的にプログラムが作成されます。（ただし、必要に応じて変更は可能です。）

さらに、SPSS Statistics には、上級者であれば、Programmability Extension を使用して新規プロシージャおよび機能を実装できるという利点もあります。この拡張機能により、R 統計プログラミング言語、Python、.NET、または Java を使い慣れたユーザーは、新規のアルゴリズムまたは関数を製品に直接組み込むことができます。上級者は、自分で作成した新機能向けのネイティブ GUI を作成し、プログラマー以外のユーザーがそれを使用できるようにすることもできます。そうすれば、それらのユーザーは迅速かつ効率的に自力で分析を実行できます。

SPSS Statistics でベストな結果を表示

SPSS Statistics ユーザーはグラフ・ビルダーを使用して、散布図マトリックス (SPLOM)、ヒストグラム、および人口ピラミッドなどの一般的に使用されるグラフをより簡単に作成できます。統計的工程管理図を扱うユーザーは 1 次および 2 次の管理図の両方においてルール・チェックを要求できます。これによって精度が高まり、プロセスが正常に作動しているかどうかを理解しやすくなります。

SPSS Statistics のグラフィック表示システムにより、作成と編集の両方の段階で制御が可能となり、実動設定でのワークロードを軽減するのに役立ちます。

スプレッドシートと異なり、分析結果は、グラフボード・テンプレート選択機能で利用可能なマップ・テンプレートを使用して地理的に表示することができます。コロプレス地図 (カラー・マップ)、ミニ・グラフ付きマップや、オーバーレイ・マップを含む、多様なタイプの視覚化が作成できるので、ユーザーがより効果的に計画、予測およびターゲット設定をすることができるようになります。SPSS Statistics にはいくつかのマップ・ファイルが同梱されています。あるいは、マップ変換ユーティリティを使用して、既存のマップ・シェープファイルをグラフボード・テンプレート選択機能で使用できるように変換することができます。

SPSS Statistics の出力は、適切な条件付き書式での手順のダイアログから直接表を作成することにより、すぐに表示に使用できるようにすることができます。また、出力文書への一般的な編集を自動化することもできます。

グラフや表は、Windows、Mac や Linux のデスクトップ環境、iPod、iPhone や iPad、Android フォンやタブレット (バージョン 2.1 以上)、および Windows 8 機器を含む、幅広い機器でいつでもどこでも表示可能です。

IBM SPSS Statistics を使用した将来の考察

スプレッドシートは多くの場合、予測の実行、つまり履歴データや不確定の入力に基づいて将来のイベントを推定するために使用されます。例えば、一般的なビジネス・アプリケーションは、前年度の結果に基づいて次の 2 四半期の収益を予測します。これらの計算はスプレッドシートを使用して行うこともできますが、ビジネスにおける季節性などの要因を明らかにしたり、what-if 分析つまり他の複数の変数に基づいてシナリオを作成することは、SPSS Statistics などの数学的に頑強なソフトウェアでのみ実行可能です。

スプレッドシートはある種の予測の実行には使用できますが、複雑な変数または状況を考慮に入れるには、SPSS Statistics などのツールが必要です。

IBM SPSS Statistics の場合の学習曲線

この資料の初めの部分で、新規ソフトウェアの使用を考慮する際に、よく思い浮かぶ質問を提起しました。それは「習得するのはどの程度難しいですか?」というものです。

SPSS Statistics の場合、答えは「まったく難しくありません」。スプレッドシートと同様に、WYSIWIG インターフェースがあるため、すべて見えるようにはっきりと配列されており、使い慣れたメニューやツールバーのレイアウトを使用して、その機能にアクセスできます。プログラムの統計機能は論理的にグループ分けされています。そのため、あるものを選択すると、関連するオプションがポップアップ・ウィンドウに表示され、必要なオプションを選択して「OK」または「実行」ボタンをクリックすると、計算が実行されます。

SPSS Statistics は習得が容易なツールで、組織が複雑なデータの分析をスプレッドシートから数学的に頑強な分析に移行するのを支援します。

さらに、SPSS Statistics には、非常に包括的なチュートリアル、極めて詳細なヘルプ・ファイル、およびケース・スタディーが組み込まれ、ビジネスおよび調査といった状況での統計解析の使用例が詳しく説明されています。これらによって、ユーザーは、統計の初心者レベルから熟練したアナリストのレベルにまで素早く到達することができます。当然ながら、企業は、オンデマンドの Web ベースの研修を含め、いくつかの研修オプションを提供しています。さらに、あらゆるタイプの背景のアナリストの間で長く使用されてきたという理由で、サード・パーティー製の付加的な学習リソースも入手できます。この中には、他のユーザーからのヒントを掲載するオンライン・ディスカッション・ボード、教育的書籍やビデオ、テキストおよびワークブックなどが含まれます。

結論

この資料を作成するにあたり、著者は現時点で提起する価値があるいくつかの点を発見しました。第一に、スプレッドシートは一般に認識されているよりもはるかに広範囲にわたって使用されています。しかも、多くの場合、他のソリューションを探し出すことなく採用されています。第二に、スプレッドシートの使用で発生するエラー率は、他の計算形式で発生する許容エラー率よりも高い可能性があります。第三に、スプレッドシートはさまざまなタイプの問題に対処するように求められますが、その中にはプログラムの機能にまったく適合していないものがあります。

データ・セットは固有であり、スプレッドシートの使用方法もまた固有です。

スプレッドシートがご自分のニーズに十分適しているか、それとも SPSS Statistics などの専門化されたツールの方が役に立つかを理解するには、通常必要とする分析タスクを実行して、各プログラムがデータをどのように処理するかをご自身で確かめるのが最善です。

SPSS Statistics を体験するのは簡単です。IBM にお問い合わせいただくか、次の Web サイトで無料のソフトウェア評価版をダウンロードしていただくことができます。ibm.com/software/analytics/spss/products/statistics データが既にスプレッドシートに含まれている場合、SPSS Statistics はそのデータを容易にインポートできます。データがインポートされたら、使用できる分析のタイプを評価し、現在のケースまたはいくつかの状況で、汎用のスプレッドシート・プログラムの代わりに統計解析専用のツールを使用する利点があるかどうかを確認できます。

IBM ビジネス・アナリティクスについて

IBM ビジネス・アナリティクス・ソフトウェアは、企業がよりスマートにビジネス活動を行い、ライバル企業をしのぐ業績を上げられるように支援する、データを重視した洞察を提供します。この包括的なポートフォリオには、ビジネス・インテリジェンス、予測分析と意思決定管理、パフォーマンス管理、およびリスク管理のソリューションが含まれています。

ビジネス・アナリティクス・ソリューションにより、企業は、顧客分析などさまざまな分野において業績に重大な影響を及ぼす可能性がある傾向やパターンを特定し、視覚化することができます。企業は、シナリオの比較、潜在的な脅威や機会の予測、経営資源に関する計画、予算および予測の改善、リスクと期待収益との対比、ならびに規制上の要件への対応を実現します。企業は、分析を広く利用可能にすることにより、戦術的および戦略的意思決定を調整してビジネス目標を達成することができます。ibm.com/software/jp/analytics



日本アイ・ビー・エム株式会社
〒103-8510
東京都中央区日本橋箱崎町19-21

IBM のホーム・ページはこちらからご覧になれます。

ibm.com

IBM、IBM ロゴ、ibm.com および SPSS は、世界の多くの国で登録された International Business Machines Corporation の商標です。他の製品名およびサービス名等は、それぞれ IBM または各社の商標である場合があります。現時点での IBM の商標リストについては、次の Web サイトをご覧ください。ibm.com/legal/copytrade.shtml

Microsoft、Windows、Windows NT および Windows ロゴは、Microsoft Corporation の米国およびその他の国における商標です。

Java およびすべてのJava 関連の商標およびロゴは Oracleやその関連会社の米国およびその他の国における商標または登録商標です。

本資料は最初の発行日の時点で得られるものであり、予告なしに変更される場合があります。すべての製品が、IBM が営業を行っているすべての国において利用可能なものではありません。

これらの資料は、特定物として現存するままの状態を提供され、商品性の保証、特定目的適合性の保証および法律上の瑕疵担保責任を含むすべての明示もしくは黙示の保証責任なしで提供されます。IBM 製品は、IBM 所定の契約書の条項に基づき保証されます。

¹ <http://archive.columbiatribune.com/2006/feb/20060222news009.asp>

² http://articles.marketwatch.com/2005-11-09/news/30780581_1_eastman-kodak-robert-brust-kodak-spokesman-gerard-meuchner

³ <http://www.abc.net.au/news/newsitems/200506/s1394937.htm>

⁴ <http://www.gao.gov/atext/d04754t.txt>

⁵ Gregg Keizer, "Microsoft fixes Excel math mistake," Computerworld (March 2008).

⁶ データ入力には、IBM SPSS Data Collection Data Entry などの特殊用途の製品や、企業の調査研究ソフトウェア製品ラインの他の製品が推奨されます。これらはデータ入力時のエラーをチェックする目的で設計されています。

© Copyright IBM Corporation 2013



Please Recycle