

「モダン Excel」は筆者の造語のようだ。世間一般に使われている言葉ではない。

モダン Excel といっても、従来の Excel と別の何かではありません。マイクロソフトの関係者によれば、Excel 2010 に「PowerQuery (以下、パワークエリ)」と「Power Pivot for Excel (以下、パワーピボット)」というアドインが追加されたとき、それ以前の Excel と区別する意味で、「**モダン Excel**」という愛称が使われるようになったようです。

本書で解説するような経営ダッシュボードを作れるようになったのは、Excel に“**104 万行の壁**”が実質的に存在しなくなったことと大きな関係があります。このようにいうと「Excel の最終行は 104 万 8576 行のままだ」と思うかもしれません (図 3)。しかし「モダン Excel」を使えば、この壁を軽々と超えることができます。

パワークエリは、データを分析できるようにするための「ETL ツール」だとよくいわれます。

ETL は「Extract (抽出する)」「Transform (変換する)」「Load (読み込む)」の頭文字です。

Extract (抽出する)	Excel、Access、CSV、Web、リレーショナルデータベース、JSONドキュメント、Azure など、さまざまなデータソースからデータを抽出する
Transform (変換する)	データの突合、クレンジング、追加、削除などにより、データを分析・集計・利用できるように変換する
Load (読み込む)	変換したデータを「テーブル」「ピボットテーブル」「ピボットグラフ」に読み込めるほか「接続の作成のみ」として詳細データをパワークエリに格納したままとすることもできる (52 ページ図 36 参照)

図 1 ETL は、「抽出する」「変換する」「読み込む」を表す英単語の頭文字。これらの作業を通じて、分析に必要なデータを準備する

<失敗例>

空白列があると、2つの表と認識してしまう。

C 列が空白なので、ソートすると、左右のデータが合わなくなる。

	A	B	C	D	E	F
1	伝票日付	伝票番号	得意先コード	売上金額	区分	
2	2020/1/31	223043	80110200	1,000	G1	
3	2020/2/28	223044	80110201	2,400	G2	
4	2020/3/31	223045	80110202	2,400	G3	
5	2020/4/30	223046	80110203	2,400	G4	
6	2020/5/31	223047	80110204	1,000	G5	
7	2020/5/31	223048	80110205	1,000	G6	
8						
9						
10						

	A	B	C	D	E	F
1	伝票日付	伝票番号	得意先コード	売上金額	区分	
2	2020/5/31	223047	80110200	1,000	G1	
3	2020/5/31	223048	80110201	2,400	G2	
4	2020/4/30	223046	80110202	2,400	G3	
5	2020/3/31	223045	80110203	2,400	G4	
6	2020/2/28	223044	80110204	1,000	G5	
7	2020/1/31	223043	80110205	1,000	G6	
8						
9						
10						

図2 上図の表(=範囲)を、A列の「伝票日付」セルを選択した状態で「降順」に並べ替えると、下図のようになり、「伝票日付」と「伝票番号」に対し、「得意先コード」や「売上金額」などが不整合となってしまった

表を「テーブル」に指定すれば、ソートしても左右データの整合は維持される。

図3 ベタ打ちの表を「テーブル」に変えるには、対象範囲を選択して(1)、「挿入」タブにある「テーブル」ボタンをクリック(2,3)。開く画面で「先頭行を…」にチェックを入れて(4)、「OK」を押す(5)

「1列1項目」「1行1件」の“きれいな”データにする——「データクレンジング」といいます。

ここでは代表的なデータクレンジング手法を3つ取り上げます。

- ①改行文字を除去する
- ②余計な空白文字を取り除く
- ②結合されたセルを正規化する

(セル内で)「Att」+「Enter」キーで改行すると、改行位置には目に見えない“改行文字”が入力されてしまうからです。パワークエリでこのC3セルのデータを取り込むと、

「販売場所」ではなく「販売#(lf)場所」として認識されてしまいます。「#(lf)」が改行文字のコードです。これが存在すると、(データベースの)データ処理上の思わぬトラブルが発生しかねません。よって、この改行文字は除去する必要があります。

~~~~~

## LF 【Line Feed】 ラインフィード

LF とは、ASCII 文字コード体系に定められた制御コードの一つで、文字の入出力位置の改行（次行の同じ位置への移動）を意味するもの。

~~~~~

先日 Twitter を眺めていると「俺の 4 時間を、返せ！」という衝撃ツイートを目にしました。何事かと読み進めると、VLOOKUP 関数でデータ同士の突き合わせ（データ突合）をしても、全然うまくいかず 4 時間浪費してしまった、というのです。時間浪費の原因は非常に単純な話で、**文字列後ろの半角スペースだった**そうです。こうした“目に見えないスペース問題”は、実務でよくあることです。

上 記①～③のデータクレンジングを従来の Excel の機能で行おうとすると、工数がかかり、とても面倒です。しかし、モダン Excel のパワークエリなら、クリック操作だけであつという間に可能です。具体的には、

- ① 「クリーン」機能で改行文字を除去する
- ② 「トリミング」機能で空白文字を取り除く
- ③ 「フィル」機能で連続データを入力する

という操作になります。

横持ちデータはデータベースにならない。

	A	B	C	D	E
1	販売実績表				横持ちデータ
2	支店	商品A	商品B	商品C	総計
3	東京	27,500	23,000	34,900	85,400
4	名古屋	28,500	24,000	35,900	88,400
5	大阪	30,722	26,222	38,122	95,066
6	総計	86,722	73,222	108,922	268,866
7					
8					
9					
10					
11					
12					
13					

図28 行の項目と列の項目を交差させて見るクロス集計表。こうした形式の表は「横持ちデータ」とも呼ばれる

図29 図28と同じ内容を「縦持ちデータ」に置き換えた例。「1列1項目」になっているため、コンピューターで処理しやすくなる

	A	B	C	D	E
1	販売実績表				縦持ちデータ
2	支店	属性	値		
3	東京	商品A	27,500		
4	東京	商品B	23,000		
5	東京	商品C	34,900		
6	名古屋	商品A	28,500		
7	名古屋	商品B	24,000		
8	名古屋	商品C	35,900		
9	大阪	商品A	30,722		
10	大阪	商品B	26,222		
11	大阪	商品C	38,122		
12					
13					

同じ種類の値が1つの列にまとめられている(1列1項目)

パワーピボットを使うメリットをまとめると、

- ①ファイルサイズを削減でき、処理スピードを上げられる
- ②いわゆる“104万行の壁”を乗り越えてビッグデータ解析が可能になる
- ③ 後述する DAX（データ分析式）で複雑な計算も簡単にできる



図1 データ分析は大まかに3つのプロセスに分けられる。取得したデータを構造化することを「データモデリング」と呼ぶ

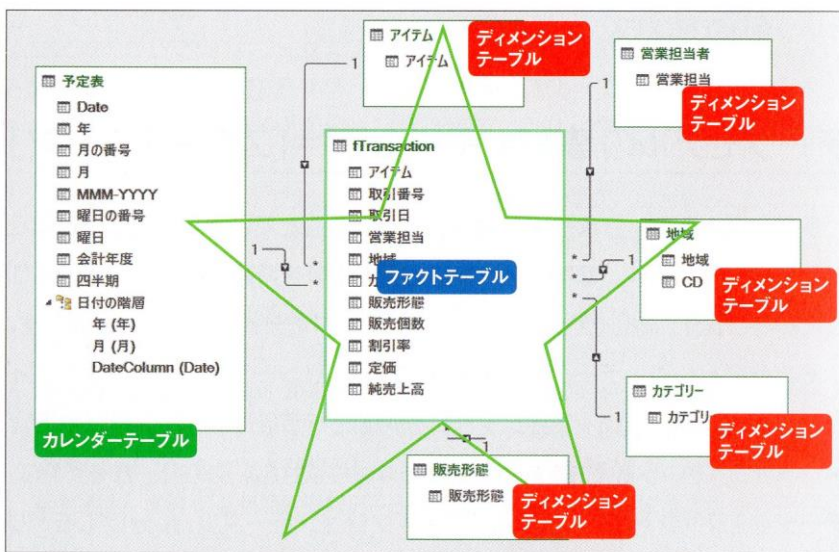


図2 データ分析の前提となる星形の表「スタースキーマ」(本書の事例)

「DAX (Data Analysis eXpressions)」

DAX は、パワーピボットのデータモデルに関連付けられた数式言語で、既存のデータから新しい情報を作成するのに役立つ“データ分析式”です。パワーピボットを使いこなすには、DAX の理解が不可欠になります。

^ キャレット

そもそも、**モダン Excel** は「列」で考えます。これは従来の Excel が「セル」で考えるのと大きな違いです。モダン Excel 初心者は、まずこの違いを理角牢する必要があります。

DAX というデータ分析式の中でも、**タイムインテリジェンス関数**はモダン Excel で頻繁に用いられる特徴的な関数の一つです。年、四半期、月、日などの期間=タイムを使用して、データを操作=インテリジェンスし、前年比や累計などを計算する際に便利な関数です。

例えば、「SAMEPER 旧 DLASTYEAR」という関数を用いると前年同月比などを簡単に計算できます。

②は、**プログラムの可読性**という、Excel 関数式や VBA のコードにも共通することです。上級者の中には複雑で高度な数式やプログラムを一気に記述する方がいますが、何をやっているのかひと目で分からないような数式やコードは、ほかの人が手を加えたり引き継いだりできなくなり、業務に差し障ることもあるので、やめてほしいと思います。これは内部統制評価における IT 監査での経験談でもあり、老婆心ながら申し上げます。

DAX を学習すると目にすることが多くなる用語の一つに「**スカラー値**」があります (図 A)。これは Excel という「セル」のイメージです。リストやレコードや配列のことを「**非スカラー値**」と呼びます。

また、「**ブール値**」「**ブール式**」という言葉も目にする機会が多く旬ります。この「**ブール**」とは「**真偽**」のことです。DAX では、真 (True) = 「1」、偽 (False) = 「0」と計算します。

経営ダッシュボード

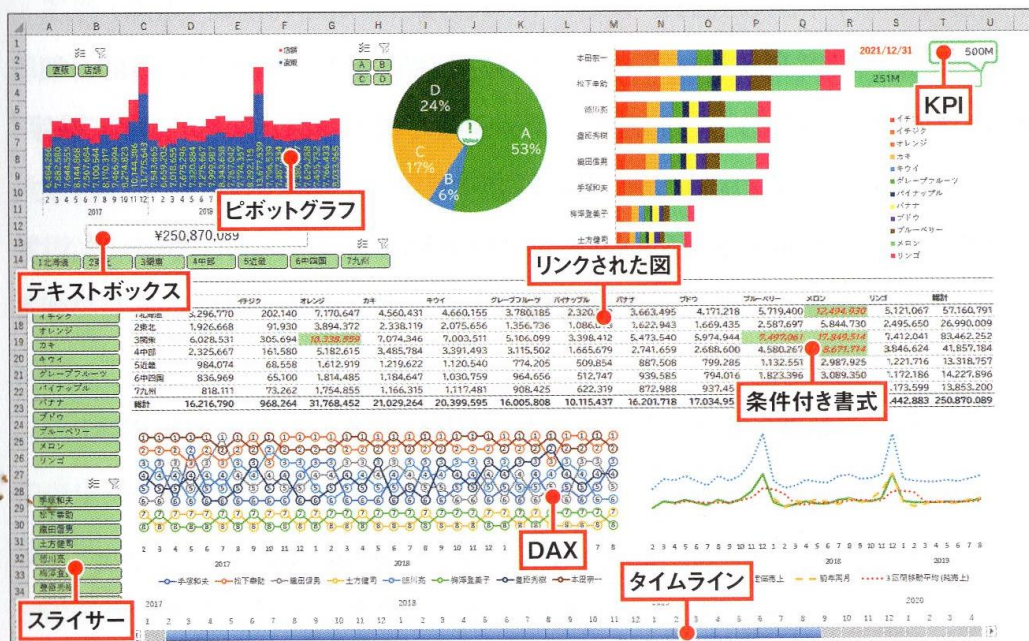


図 1 本書で作成する「経営ダッシュボード」の例。図中に示した機能を用いている

第1章で紹介したように、「パワークエリ」を使えば、「ETL（抽出・変換・読み込み）」によってデータクレンジングが可能となります。「フォルター接続」をすれば、ファイルの追加・削除によるデータ更新も自動化できて便利です。さらた第2章で解説したように、「パワーピボット」を使えば、関連するテーブル同士を簡単に「リレーションシップ」でつなぐことができ、新たな視点でデータを見ることができるようになります。そして、「DAX（Data Analysis eXpressions）」というデータ分析式を使うことで、目的に沿ったデータ分析が自在となり、しかも、クリックするだけで即座にデータの更新が可能となります。

人間が瞬時に把握できる KPI はせいぜい 20～25 個。

...

そこで、経営に必要な数字だけを 1 カ所にまとめて、しかも最大でも **20 個程度の数字にとどめた経営ダッシュボード**という**経営管理の考え方が**登場します。経営実態を分かりやすく把握できることから、次第に浸透し、現在に至っています。

ある 1 人の人物がモダン Excel を用いてシステムを構築し、**その人だけが計算やデータ処理の流れを知っている状態**というのは、**内部統制的に問題がある**ことは容易に想像できるでしょう。仮にその人物が不正な行為をしても、誰も気付くことができません。不正がなくても、その人が退職などでいなくなり、その人だけが知っている自動化の仕組みで計算ミスや不具合が起きれば、誰も検証・修正できないというリスクにさらされることになります。

こした野良ロボットがはびこらないようにするには、どうすればよいでしょうか。内部統制的に考えれば、その退治の**秘訣は 2 つ**です。

- ① 本格稼働する前に組織による「承認」を経る
- ② 第三者が検証できるように「コメント」機能を活用する

2種類のコメント方法	
シングルラインコメント	コメントの始まりに「//」(半角スラッシュ2個)を入力し、1行ずつコメントを入れる
マルチラインコメント	コメントの始まりに「/*」(半角スラッシュと半角アスタリスク)、終わりに「*/」を入力し、複数行にコメントを入れる

図7 モダンExcelに共通するコメントの入力方法には2種類ある。なお、パワーピボットでは「--」(半角ハイフン2個)でシングルラインコメントを入力できる

テキストボックスの作成時、

ドラッグすると枠線付きのテキストボックスになりますが、クリックすると

枠線なしの透明なテキストボックスになります。

ヒトの目線は「Z」を書くように、左上→右上→左下→右下の順に動くといわれ、これを「**Zの法則**」と呼びます

パワークエリで ETL（抽出・変換・読み込み）を行い、1列1項目のきれいなデータを「フォルダー接続」により取り込んで、パワーピボットでデータ相互のリレーションシップを取り、適切なデータモデルとし、DAXでメジャーを書き、ピボットテーブルとピボットグラフで可視化の土台を作る。あとは蓄積したデータを決められたフォルダーに投げ入れる都度、「すべて更新」ボタンをクリックすれば、「経営ダッシュボード」を常に最新の状態に保つことができ、「ファクトベースでのデータ分析」が可能となる――。

これが、「**モダン Excel**」の真骨頂です。