

最速さいそくで

収益しゅうえきにつなげる

完全自動かんぜんじどうの

データ分析たかふん

The Automated Data Analysis for Fast Profits

高橋威知郎たか かし い ち ろう

Ichiro Takahashi

トップデータサイエンティストが教える、ワンランク上の分析手法!

はじめに

「なんか、ビッグデータのなこと、やってよ！」

「AI(人工知能)で、なんとかしといて！」

「よし、デジタルトランスフォーメーションだ！」

会社のエライ人から、このようなことをいわれたビジネスパーソンは、たまたまのものではない。何をすればいいのかわからない。

多くの人は、データサイエンスや機械学習、AI(人工知能)やデジタル革命など、近しい分野の書籍を何冊か買い込み理解しようと努力する。確かに書籍を買って読めば、なんとなく見えてくることもある。しかし、何をすればいいのかまではわからないのではないだろうか。

わかっているのは、自分がデータを使って何かしら成果を出さなければならないこと。企業であれば、それは「収益をあげること」であるケースが多い。しかも、単に収益をあげるだけでなく、データを活用することで何かしら違いを見せなければならない。今までよりも、より速く、より確実に、より効率的に、よりスマートに収益につなげる必要がある。

本書は、あなたが「最速で収益につなげる完全自動のデータ分析」に向けて動き出すための、地に足の着いた実践的な知識を共有する。

では、そもそも「データを使い最速で収益につなげる」ためには、一体何が必要になるのだろうか。

それは、主に以下の2つである。

- ① データ分析の民主化
- ② 定量分析の完全自動化

数は力である。データを活用し成果を出すには、より多くの人が、当たり前のように、気軽に、データ分析ができた方がいい。少なくとも、データ分析の効用を知っていた方がいい。それが、「データ分析の民主化」である。

しかし、ただデータ分析をただけでは、ビジネスで成果をあげるのは難しい。データ分析の「活用の質」が問われる。データ分析の「活用の質」は、定量分析の結果を「どう解釈し、整理し、次につなげるか」にある。データ分析と聞くと定量分析のイメージが強いが、本当に重要なのは、次につなげるために、できるだけ解釈や整理、洞察といった定性的なブレインワークに時間を割けるようにすることだ。そのために「定量分析の完全自動化」を目指す。

したがって、「最速で収益につなげる完全自動のデータ分析」とは、「データ分析の民主化」と「定量分析の完全自動化」を目指した、質の高いデータ分析の活用のことである。

千里の道も一歩から

ビッグデータの要素技術やその周辺技術は目覚ましく進展した。しかし、どんなにデータがあっても、どんなにマシンが高機能化しても、どんなに分析技術が発展しても、組織内でデータが活かされるかどうかとは別問題である。何事もそうだが、目新しい技術を取り入れれば一足飛びに売上が伸びるほど、この世界は甘くはない。

そもそも、手元にあるリトルデータすら上手く活用できないのに、膨大な量のビッグデータを活用することなど、夢のまた夢だ。

リトルデータすらビジネスで活かせない組織に、RやPython、SAS、SPSSなどといった分析ツールや、BI(ビジネス・インテリジェンス)ツール、ML(機械学習)プラットフォームなどを導入しても、ただデータを溜め、それをこねくり回すだけで終わってしまう可能性が高い。それでは、収益が拡大したり、業務が効率化したりすることもなく、実りの小さなIT投資になってしまうだろう。ストレートにいうなら、お金の無駄である。実際はお金が無駄になるどころか、新しいツールやシステムなどの導入コストだけでなく、それらと付き合う作業工数が増え、逆に企業利益を悪化させつづける要因となるかもしれない。

では、一体何が足りないのだろうか。

答えは明確である。データをビジネスに活かすための「勘所」を持った人財が

足りないのだ。勘所を養うためには、データをビジネスに活かした「経験」が必要である。とはいえ、それほどゴイ経験が必要なわけではない。ちょっとした経験で十分である。

では、どうやってその経験を積みばいいのだろうか。

まずは「誰でも使える分析ツール」を使い、手元にあるリトルデータを、ビジネスに活かすことから始めるべきである。

誰もが使える分析ツール Excel

多くの企業が使っている表計算ソフトがある。Excelである。これほど浸透しているソフトウェアも珍しい。単に表計算のためだけに使うのではなく、企業や人によって様々な用途で使われている。

例えば、申請書や帳票、レイアウト図など。さらに、Excel VBAを使いこなせば、その可能性は飛躍的に広がる。

ビジネスパーソンにとって、Excelは非常に身近で、それほど扱いが難しいわけでもなく、工夫次第で色々使える便利なツールである。

そして、このExcelは分析ツールとしても十分に使える。もちろん限界もある。SASやSPSSなどの商用の高機能な分析ツールに比べれば、まだまだ機能は足りない。

だが、ここで思い描いてほしい。

もし、この「誰でも使える分析ツールであるExcel」が、SASやSPSSなどの商用の高機能な分析ツールなみの機能を備えたら、どんなに素晴らしいことだろうか。

「Excel でデータ分析している ≒ データ分析素人」ではない

私は大学卒業後約20年間、データ分析一筋で生きてきた。紆余曲折はあったものの、常に「実践的」なデータ分析を心がけてきた。最初は安全保障系のデー

タ分析を、その後は事業戦略やマーケティング、製品開発、営業系、Web系などのデータ分析を、最近ではIoT(モノのインターネット)系のセンサーやヘルスケアデバイスなどのデータ分析を行ってきた。

データ分析に長年携わる中で、私はExcelの圧倒的な存在感と拡散パワーを目の当たりにしてきた。

多くのデータ分析従事者は、「Excelでデータ分析している≒データ分析素人」と考えがちである。私もかつてはそのように思っていた。

データ分析従事者たるもの、常日頃から最新の分析技術動向に目を光らせ、高度なデータ分析を実現するため、RやPython、SAS、SPSSなどの分析ツールをエレガントに使いこなし、データを切り出すためにSQLを縦横無尽に扱い、分析ツールに実装されていない手法は自ら学術論文を読み込みC言語などでコーディングする。そのようなものだと、ある時期まで私も思い込んでいた。しかし、それは大きな間違いであった。

どんなにすごい分析手法を使ったとしても、活用されないデータ分析は存在しないに等しい。まったくの無価値であり、単なるデータ分析従事者の趣味で終わってしまう。

では、どうすればいいのか。その答えがExcelである。

例えば、データ分析を得意とする人が好むRやPythonなどの分析ツールを前提に、データ活用の文化を組織内に広めようとしても、なかなか広まらない。ところが、Excelを前提にした途端、この文化が急速に広まるがあった。

直接RやPythonで分析した方が、お手軽で早く分析できる。しかし、多少面倒でもExcelで分析し、活用できるように環境を整えた方が、データ活用の輪が広まる。輪が広まると、多くの人が当たり前のようにデータを活用しはじめる。

この事実に対して、当時の私はどうしても納得することができなかった。「たかがExcel……」という気持ちが私の心のどこかにあったからかもしれない。深く反省すべき点である。自分の趣味を他人に押しつけても意味はない。重要なことは、データが組織の中で躍動し、ビジネスに貢献することである。であれば、Excelの圧倒的な存在感と拡散パワーを使わない手はない。

Excel を最強の分析ツールにする

最近、Rというフリーのデータ分析ソフトウェアが注目を浴びている。それにはいくつか理由がある。

- ① 無料である
- ② 参考書籍がたくさんある（自習できる）
- ③ 最新の分析手法が使える

そのため、データ分析の初学者から研究機関の研究者、データ分析従事者まで、色々な分野の人が世界中で利用している。

例えば、こういうのはどうだろう。

圧倒的な存在感と拡散パワーを持つExcelに、Rの持つ分析ツールとしてのパワーを取り込み、Excelを最強の分析ツールにする。そして、ビッグデータ活用の第一歩として、手元にあるリトルデータを分析し、ビジネス成果をあげる。

その間データ活用の「勘」が養われ人が育ち、さらに組織はビッグデータ駆動型の組織へと変貌する。

事実、このRはExcelをユーザインターフェースにして使うことができるのだ。RExcel(参考文献：Thomas Baier and Erich Neuwirth (2007), Excel :: COM :: R, Computational Statistics 22/1, pp. 91-108)である。RExcelのアドインをインストールさえすれば、誰でもExcelからRを使える。

そして、このRExcelを進化させたのがautstatのscExcel(<http://www.autstat.com/>)である。このscExcelを使うことで、ExcelからRの優れた分析機能にアクセスする。scExcelによって、Excelのデータ分析機能が大幅に強化される。

要は、まずは誰でも使える分析ツールExcelを中心に使い、手元にあるデータをビジネスに活かすことから始めればよいということだ。

では、そのためには何をどうすればいいのか？ その答えが本書である。

本書であなたが得られること

第1章の「基礎編」では、現場でデータ活用を実現するための考え方について述べている。現場では管理のためのPDCA (Plan-Do-Check-Act、計画-実行-評価-改善) サイクルよりも、現場が動くためのOODA (Observe-Orient-Decide-Act、観察-方向づけ-決定-実行) ループを使った方が、データ活用が上手くいきやすい。

第2章の「指標 (KPI など) 設計編」では、データ活用する上で重要な指標 (KPI など) の設計の考え方やつくり方を述べている。指標を制する者がデータ分析を制する。データ活用を考えたら、まず指標 (KPI など) 設計をしよう。

第3章から第5章にかけては「Excel 定量分析」について述べている。

第3章では、現状把握のためのExcel 定量分析について、主に「異常検知」や「要因分析」などを取り扱う。

第4章では、将来予測のためのExcel 定量分析について、主に「予測モデル」の構築方法と使い方などを取り扱う。

第5章では、決断のための定量分析について、主に「AHP」(Analytic Hierarchy Process、階層的意意思決定法) を使った意思決定法について取り扱う。

第6章では、ビジネス成果を確実にするための「コツ」について述べている。1つはビジネス成果の「金額換算」。もう1つは定量分析の「完全自動化」である。

すべてのデータ分析の成果を「金額換算」するクセをつけることで、ビジネス成果への感度が高まり、データ分析のビジネス貢献が加速する。

また、そもそも一口にデータ分析といっても、できるだけ時間をかけるべきところと、そうでないところがある。

前述した通り、定量分析は「完全自動化」を目指し、ブレインワークの時間を確保するべきである。そうすることで、データ分析の「活用の質」を向上させ、ビジネスへの貢献が加速する。

また、特別付録として、Excelを最強の分析ツールにするための設定マニュアルがダウンロードできるページを巻末に設けている。解説しているのは「Excel 分析ツール」と「Excel ソルバー」「scExcel(旧REExcel)」の3つの設定方法である。

分析するためのツールは、Excel以外でも構わない。しかし、多くのビジネスパーソンにとって、Excel以上に身近で愛されているツールは、おそらくない。

このような非常に身近なツールを使い、データがビジネスで活かされる意義は大きい。これこそ「データ分析の民主化」である。

再度繰り返す。本書の目指す「最速で収益につなげる完全自動のデータ分析」とは、Excelを上手く活用することで「データ分析の民主化」と「定量分析の完全自動化」を目指すことである。

データ分析の民主化を目指すことで、データ分析人財は幾らか増えるだろうし、完全自動化を目指すことで、部分的であっても自動化が進み、データ分析のスピードと質が高まる。少なくとも、何かしらのビジネス成果を手に行けるはずである。やって損はない。

本書を通じてデータ分析人財が育ち、ビジネスの現場で成果を出せるデータ分析が、日本社会にもっと浸透することを願っている。

また、本書内で使用したデータなどは、セールスアナリティクスのサイト(<https://www.salesanalytics.co.jp/books/>)で公開している。ぜひダウンロードしていただき、手を動かしながら定量分析をマスターし、その先にある定量分析の完全自動化を目指して創意工夫を重ねていただければ幸いである。

はじめに	2
千里の道も一歩から	3
誰もが使える分析ツールExcel	4
「Excelでデータ分析している≒データ分析素人」ではない	4
Excelを最強の分析ツールにする	6
本書であなたが得られること	7

第1章 基礎編

① なぜ、データ分析が実務で活かされないのか	14
ビッグデータどころか、	
リトルデータさえ活用できない悲しい現実	14
リトルデータで、データ活用の勘所をつかむのが早道	16
Excelレベルの分析をものにしたとき、	
データ活用の世界は一変する	19
② データ活用の現場では、PDCAサイクルを捨てよ	22
PDCAサイクルは、現場には向かない	22
PDCAサイクルが現場の足かせになるとき	24
現場を躍動させるOODAループ	26
③ OODA×Excelで現場レベルのデータ活用が実現する	29
PDCAサイクルとOODAループの関係	29
OODAループとデータ分析	30
O (Observe、観察)とデータ分析	30
O (Orient、方向づけ)とデータ分析	33
D (Decide、決定)とデータ分析	35
OODAループはスキルではなく意識の問題である	39

第2章 指標 (KPIなど) 設計編

① 指標を制する者がデータ分析を制する	42
なぜ、指標がデータ分析の「キー」になるのか?	42
成果の出るビジネス指標の5つの条件SMART	46
どの指標を最重視するかで、ビジネス成果は変化する	49
② 指標は、管理会計の指標を中心に 3種類の指標で組み立てよ	51
管理会計とは何か	51
あなたのデータ分析は、 管理会計の「どの部分」を変化させるのか	52
3種類の指標で細分化し、アクションにつなげる	58
③ 指標とOODAループ	61
OODAループでモニタリングする指標は絞る	61
モニタリング指標は目的変数と呼ばれ、 その要因は説明変数と呼ばれる	63
現場を無視した指標に未来はない	65

第3章 今あるデータで「現状把握」するためのExcel定量分析

① O (Observe、観察) の分析フロー	68
② 問題のありそうな指標を検知する定量分析「異常検知」	82
「異常検知」とは?	82
すぐできるExcel定量分析—「異常検知」編	87
ステップ1 (時系列モデルの構築)	87
ステップ2 (外れ値スコアの算出)	123
ちょっと高度なExcel定量分析—「異常検知」編	130
scExcel (旧REExcel) による残差のカーネル密度推定と 混合正規分布	130
scExcelによるAR (自己回帰) モデルの自動構築	142

3	問題の原因を探る定量分析「要因分析」	152
	「要因分析」とは？	152
	すぐできるExcel定量分析－「要因分析」編	161
	ステップ1(目的変数の特定)	161
	ステップ2(統計モデルの構築)	162
	ステップ3(要因特定)	168
	ちょっと高度なExcel定量分析－「要因分析」編	170
	scExcel(旧RExcel)による回帰型の統計モデルの構築	170
	とても使い勝手のいい正規化回帰モデル	183
4	「現状把握」のまとめ	195
	「現状把握」とは何だったのか？	195
	定量分析の「完全自動化」に向けて－「現状把握」編	196
	「現状把握」の次にすべきこと	197

第4章 進むべき未来の「方向性」を示すためのExcel定量分析

1	0(Orient、方向づけ)の分析フロー	200
2	今後の見通しをつける定量分析「将来予測」	209
	「将来予測」(予測モデルの構築と予測)とは？	209
	すぐできるExcel定量分析－「将来予測」編	213
	ステップ1(説明変数の選択)	213
	ステップ2(予測モデルの構築)	246
	ステップ3(将来の予測)	260
	ちょっと高度なExcel定量分析－「将来予測」編	270
	Ridge回帰による予測モデル構築	270
	ExcelVBAでRを操作し自動化する	277
3	「方向づけ」のまとめ	283
	「方向づけ」とは何だったのか？	283
	定量分析の「完全自動化」に向けて－「方向づけ」編	284
	「方向づけ」の次にすべきこと	285

第5章 次に何をすべきかを「決定」するためのExcel定量分析

① D (Decide、決定) の分析フロー	288
② 意思決定を支援する定量分析「評価決定」	291
AHPによる「評価決定」とは?	291
すぐできるExcel定量分析ー「評価決定」編	292
ステップ1 (評価基準のウェイトの算定)	293
ステップ2 (代替案のスコアの算定)	296
ステップ3 (対策案のレコメンド)	298
ちょっと高度なExcel定量分析ー「評価決定」編	299
③ 「決定」のまとめ	307
「決定」とは何だったのか?	307
定量分析の「完全自動化」に向けてー「決定」編	308
「決定」の次にすべきこと	309

第6章 ビジネス成果を産む「実行」の実現に向けてすべきこと

① Outputの先のOutcomeまでこだわれ	312
② Outcome (ビジネス成果) を「金額換算」する	314
4つの金額を見積もる	314
「業務プロセス」を描き「工数コスト」を見積もる	318
③ できる限り「自動化」し、ブレインワークを増やせ	320
時間のかけ方を工夫する	320
定量分析は「完全自動化」を目指そう	322

おわりに	325
------	-----