

2024年5月10日 全10頁

データ分析を経営にどう活かすか？

「ラストワンマイル」を乗り越えるために

データアナリティクス部 コンサルタント 江藤 俊太郎

[要約]

- 企業におけるデータ分析は、経営課題を可視化し、意思決定につなげるために行う。データ分析の手法がデータサイエンスであり、データ分析は、データドリブン経営、DXにつながる最初の一步でもある。
- 時に、データ分析チームとマネジメント層（経営陣や意思決定権者）の間には「ラストワンマイル」（あと一步の距離）があるといわれる。そこで、その距離を乗り越える対策を提案する。
- 「ラストワンマイル」を乗り越えて、データ分析を企業文化に定着させる必要がある。

1. データ分析とは？

本稿では、企業におけるデータ分析¹を、データを用いて企業の経営課題を可視化し、経営判断、意思決定、業務上の施策につなげるための分析と定義する。そして、データサイエンスはデータ分析の主たる手法と考える。

データサイエンスとは、多種大量のデータから有用な情報を得るために、採用される科学的な方法²を指す。数学、統計学、人工知能（AI）、コンピュータ・エンジニアリングの分野の原則と実践を組み合わせた学際的なアプローチ³とされる。

データドリブン経営（DDM）とは、データ分析により得られた知見を活かし、戦略的な意思決定につなげる経営のことを指す。

DXとは、「世界最先端デジタル国家創造宣言・官民データ活用推進基本計画」（2020年7月17日閣議決定）によれば次の通りである。

¹ データ分析は、英語では data analytics, data analysis, data assay と表記されることがある。それぞれ微妙にニュアンスに違いがあるようだ。

² 出所：Cambridge English Dictionary

³ 出所：<https://aws.amazon.com/jp/what-is/data-science/>（2024年4月22日閲覧）

「企業が外部エコシステム（顧客、市場）の劇的な変化に対応しつつ、内部エコシステム（組織、文化、従業員）の変革を牽引しながら、第3のプラットフォーム（クラウド、モビリティ、ビッグデータ/アナリティクス、ソーシャル技術）を利用して、新しい製品やサービス、新しいビジネスモデルを通して、ネットとリアルの両面での顧客エクスペリエンスの変革を図ることで価値を創出し、競争上の優位性を確立すること」。

DXはデジタル技術を用いた企業全体の仕組みの改革であり、データドリブン経営はDXを支える重要な要素である。以上により、データ分析の手法がデータサイエンスであり、データ分析の目的がデータドリブン経営、そしてデータドリブン経営の目指す姿がDXである、とまとめられる。

以上はデータ分析と関連用語の定義である。実際の事例はどうだろうか。以下は、20年近い歴史のある大阪ガスのデータ分析専門部署、ビジネスアナリシスセンターが行ったデータ分析の事例である。幅広い経営・業務課題の分析を行ってきたことがわかる。

（図表1）データ分析の事例（大阪ガス）

・ガスの需要予測	・市場価格リスク評価
・価格弾力性分析	・信用リスク評価
・社員健康診断データ分析	・パソコン操作ログ分析
・時刻別電力需要予測	・プラント運転最適化
・電力価格予測	・機器や設備の故障予知
・顧客ターゲティング分析	・機器や設備の故障診断
・シフトスケジューリング	・エリアマーケティング分析
・配船計画の最適化	・ガス機器の環境性評価
・タンク運用の最適化	・車両配置の最適化
・顧客の声のテキストマイニング	・顧客接点分析

出所：河本薫「最強のデータ分析組織 なぜ大阪ガスは成功したのか」日経BP社（2017年11月）

さて、マネジメント層（経営陣や意思決定者）とは、多くの局面で、経営・業務課題を解決すべく、何らかの決断を迫られる人たちである。彼らにとって、データ分析の結果を活用したいというニーズは強いものがある。

図表 2 は、総務省（2020）が、製造業、情報通信業、エネルギー・インフラ、商業・流通業、サービス業に属する大企業・中小企業 2003 社にアンケートを採ったものである。各企業が、分野別に今後のデータ活用予定について回答している。

「経営企画・組織改革」に限っていえば、63.3%（=A+B）の企業がデータを未活用であるが、うち 43.6%（=B/(A+B)）が今後活用したいと考えている。既に活用している 36.8%（=C+D）と併せると、64.3%（=B+C+D）の企業が（引き続き・これから）データを活用する予定があるという結果が出ている。

一方で「生産・製造」、「物流・在庫管理」「保守・メンテナンス・サポート」では比較的データ活用が進んでおらず、活用していきたい企業も全体（2003 社）の約半分程度の状況である。

（図表 2）今後のデータ活用予定

			経営企画・ 組織改革	製品・サービスの 企画、開発	マーケティング
			n=2003	n=2003	n=2003
未活用	今後もデータを活用する予定はない	A	35.7%	37.8%	37.5%
	今後はデータを活用していきたい	B	27.6%	29.0%	30.3%
活用中	今までと同程度にデータを活用していきたい	C	9.7%	8.6%	8.1%
	今まで以上にデータを活用していきたい	D	27.1%	24.6%	24.1%
全体		=A+B+C+D	100.0%	100.0%	100.0%
未活用		=A+B	63.3%	66.8%	67.8%
	うち活用していきたいの割合	=B/(A+B)	43.6%	43.4%	44.7%
全体	うち活用していきたいの割合	=B+C+D	64.3%	62.2%	62.5%
			生産・製造	物流・在庫管理	保守・ メンテナンス・ サポート
			n=2003	n=2003	n=2003
未活用	今後もデータを活用する予定はない	A	50.7%	51.4%	46.2%
	今後はデータを活用していきたい	B	31.1%	33.8%	36.0%
活用中	今までと同程度にデータを活用していきたい	C	4.8%	4.4%	6.3%
	今まで以上にデータを活用していきたい	D	13.4%	10.4%	11.5%
全体		=A+B+C+D	100.0%	100.0%	100.0%
未活用		=A+B	81.8%	85.2%	82.2%
	うち活用していきたいの割合	=B/(A+B)	38.0%	39.7%	43.8%
全体	うち活用していきたいの割合	=B+C+D	49.3%	48.6%	53.8%

出所：総務省「デジタルデータの経済的価値の計測と活用の現状に関する調査研究」（2020）より、
大和総研作成

大和総研コンサルティング本部が扱う「人事データ分析」は、「経営企画・組織改革」に分類されるだろうが、さまざまな引き合いが来ている。属人的な経験と勘から脱却して、データに基づく意思決定を行いたいという流れはこれからも強くなっていくだろう。

2. データ分析とマネジメント

(1) ラストワンマイル問題

データ分析の重要性について、マネジメント層も当然理解しているといえるだろう。しかし、データ分析を社内外のチームに依頼すると、満足のいく結果が得られないこともある。時に、データ分析チームとマネジメント層の間にはラストワンマイル（あと一步の距離）があるといわれる⁴。例えば、Kaggle（2017）の調査では17.7%のデータサイエンティストが、データサイエンスを用いた分析結果が意思決定に使われなかったと回答している⁵。マネジメント層からみて、ラストワンマイルは、例えば以下のようなものである。

ア) 「本当にそうなのか」

データ分析の成果物には「可能性がある」「示唆される」「と思われる」「確率が高い」などの言葉が並ぶ。それらは、学術的にみると正確な表現である。しかし、貴重な経営資源を用いて決断しないといけない立場であるマネジメント層にとって、データ分析チームの結論はどこか腰が引けているような印象を受ける。したがって、分析結果は意思決定の材料にまで至らない、ということが起こりうる。

さらに、データ分析の成果物は、例えば経営企画部門の優秀なスタッフをもってしても電卓（と表計算ソフト）だけでは検証できないことがありうる。したがって、データ分析の結果が正しいかどうか、分析者以外の誰にもわからない場合が多い。マネジメント層は「本当にそうなのか」「誰が保証してくれるのか」といいたくなるのである。

イ) 「結局どれくらいになるのか」

マネジメント層は数字を背負っている。意思決定が売上高、コスト、その他 KGI、KPI にどれくらい影響するかを常に意識している。例えば、「人事データ分析」のうち退職者分析なら退職率がどれだけ下がるのか、が最重要関心事項だ。つまりマネジメント層は、説明変数⁶が変われば目的変数⁷についてどれだけのインパクトがあるか、について鋭敏な問題意識を持っている。そしてこの分析によって決定された施策にどのような効果があるのかがわからなければ、分析そのものの費用対効果もつかめない。

しかし、高度なモデルになればなるほど、説明変数の変化に対してどれくらい（目的変数が）

⁴ Scott Berinato. “Data Science and the Art of Persuasion”. Harvard Business Review. January-February 2019. <https://hbr.org/2019/01/data-science-and-the-art-of-persuasion>, (2024年4月22日閲覧)

⁵ <https://businessoverbroadway.com/2018/03/18/top-10-challenges-to-practicing-data-science-at-work/> (2024年4月22日閲覧) にまとまっている。

⁶ 事象を説明する変数。退職者分析の場合、業務のやりがい、残業時間など。

⁷ 事象の目的となる変数。退職者分析の場合、退職確率のこと。

変化するかが、なかなか読みづらくなる。前稿「退職確率を分析する」

(https://www.dir.co.jp/report/consulting/pbo_analysis/20221020_023331.html) を執筆したところ、周囲のコンサルタントからは、主に2通りの反応が返ってきた

- A. ノモグラム等、各説明変数が結果に与える影響がわかりやすい。
- B. 予測精度が低い。モデル構築にあまり現実的ではない(=強い)仮定が必要となる。モデルに盛り込める説明変数の数に制約がある。

Aは主にマネジメント層よりの立場でものを考える人たちである。しかしデータ分析に詳しい立場からは、Bのような指摘が多かった。

しばしば説明可能性と予測精度はトレードオフの関係にある、といわれる。実際「退職確率を分析する」で使ったのは伝統的な数理統計モデルであったので、強い仮定が必要な反面、各変数が退職確率に与えるインパクトについて説明が容易であった。それは、そのようなモデルを採用し、現実を捨象したからである。一般に、複雑で精度が高いモデルになればなるほど、直感的な理解が難しい傾向にある。従業員のやりがいが増えたら、退職確率が何%下がるのか、単純には説明できなくなる。マネジメント層は「結局(影響は)どれくらいになるんだ」といいたくなるのだ。

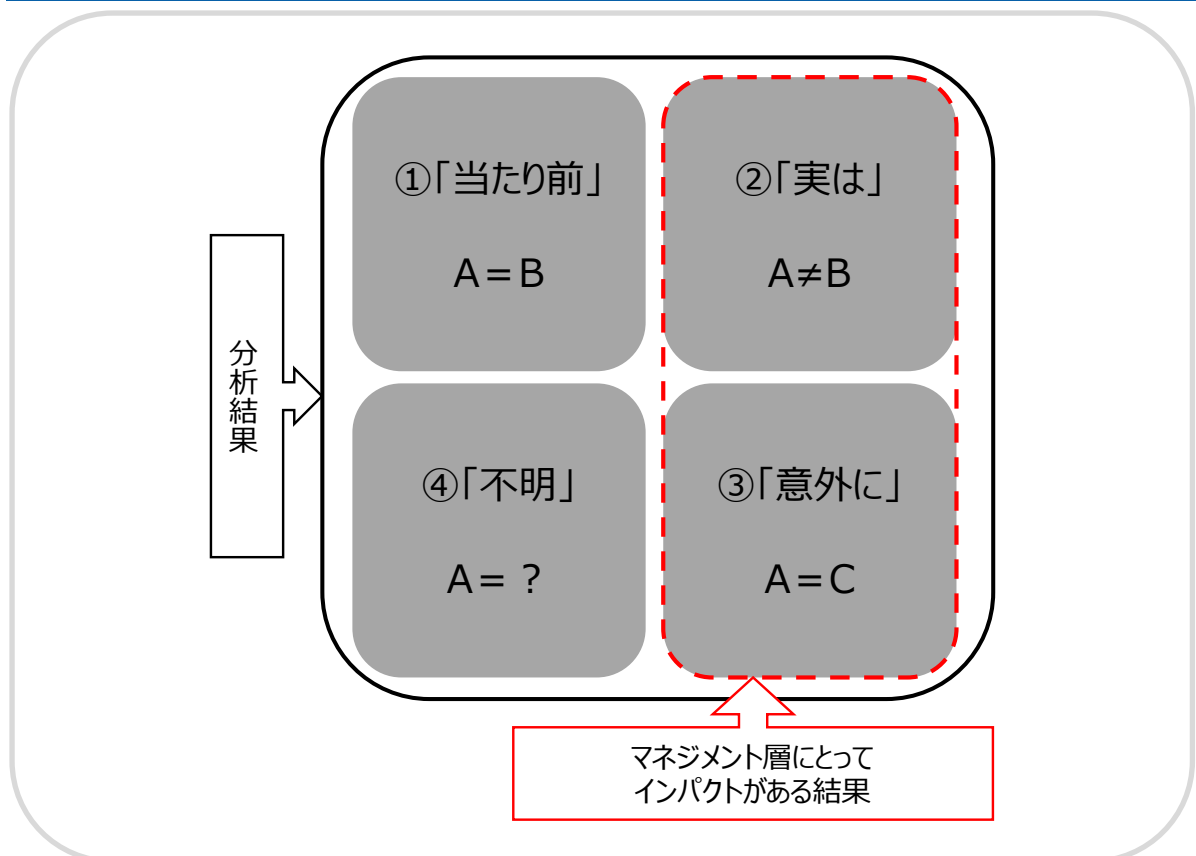
ウ)「当たり前ではないか」

20年程前、「おむつとビール」⁸の例がはやったけれども、データ分析の成果を扱った記事・書籍などを読むと「AとBが結びついていてと思っていたが、意外なことにAはCに結びついていて」という結論が好まれる傾向にあるようだ。確かに、データ分析を行う上でそういう隠れた相関関係がみつき、そこに何らかの因果関係が推定されることは起こりうる。データ分析の結論を以下の4つにまとめてみた。(図表3)

- ①「当たり前」
- ②「実は」間違っている
- ③「意外に」全く予想のつかなかったことに関連性がある
- ④「不明」判断がつかない

⁸ 週末の夜、スーパーマーケットで「おむつとビール」を同時に買う顧客が多かったので、その背景となる事情を調べ、売り場を近接させたところ、両者の売上が上がったというエピソードのこと。確かに古い事例であるが、根拠のない話ではないようだ。<https://canworksmart.com/diapers-beer-retail-predictive-analytics/> (2024年4月22日閲覧)

(図表 3) 分析結果とインパクト



出所：大和総研

マネジメント層にとっては③に最大のインパクトがある。②も、従来の施策の転換を迫られるので、かなりインパクトがある。しかし、結果的な成果物は①、④に終始してしまうということが、十分に起こりうる。マネジメント層にとっては、コストをかけたのに、「当たり前じゃないか」という感想を抱いてしまうのだろう。

(2) ラストワンマイルの処方箋

ア) 「本当にそうなのか」への対策

データ分析に、確実性を求めるのはそもそも難しい。不確実性を乗り越えていくのはそもそもマネジメント層の役割である。とはいえ、経営判断にデータ分析の結果を意思決定に採用することがためられる場合、対策は、まず「徐々に」取り入れることである。施策はマイナスの影響の出にくいもの、コストがかからないものから始めて、データ分析のPDCAサイクルを回していき、やがて大きな動きにつなげていく（ただし、それは施策の話であって、分析するデータそのものの規模は最初から大きい方が望ましい）。

また、コストはかかるが、検証する体制を構築することは不可能ではない。社内にデータサイエンティストがいれば、社外の分析結果を検証させる。今、多くの企業でインハウスのデータサイエンティストを採用・育成している。図表4によれば、データを活用していると答えた企業の4割（全体の回答社数2003社を母数とすると3割）がデータ分析を行う専門部署を置いていると回答している。育成と検証業務を兼ねた有効な活用法として提案したい。

(図表4) データの分析体制

	全体	大企業	中小企業
	n=1426	n=805	n=621
各事業部門のデータ分析が専門ではない人	47.3%	44.7%	50.6%
各事業部門のデータ分析専門の担当者	42.4%	54.9%	26.2%
データ分析を行う専門部署の担当者	41.9%	51.7%	29.3%
外部に委託	12.1%	14.8%	8.5%
アライアンスやコンソーシアムなど他社等を交えた共同分析	5.0%	7.6%	1.6%

出所：総務省「デジタルデータの経済的価値の計測と活用の現状に関する調査研究」（2020）より、大和総研作成

社内にデータサイエンティストがいなければ、複数の企業、独立性が担保できれば単一企業の2つのチームに任せるのも一つのやり方である。データ分析の性質上、全く同じ結果が出ることはないが、適正なプロセスで行われていることを確認できれば良い。

イ) 「結局どれくらいになるのか」への対策

マネジメント層にリニア（一次関数的）な因果関係⁹だけではないロジックで分析していることを理解してもらう必要があるだろう。データ分析チーム側の課題としては、従来以上に可視化の手法を洗練させ、よりマネジメント層の理解を促進することである。可視化にはいろいろな手法が用いられる。Python では主に、matplotlib、seaborn といったライブラリを使って可視化することができる。適切なみせ方を工夫する必要があるだろう。

（図表 5）可視化のさまざまな手法

縦棒・横棒グラフ	ヒートマップ
円グラフ	ウォーターフォールチャート
ドーナツグラフ	ツリーマップ
折れ線グラフ	サンバーストチャート
各種 3D グラフ	レーダーチャート
ボックスプロット	地図
散布図（行列）	ワードクラウド
バブルチャート	インフォグラフィック
ジョイントプロット	等
並行座標プロット	

出所：小久保奈都弥「データ分析のための Python データビジュアライゼーション入門」翔泳社（2020年8月）より、大和総研作成

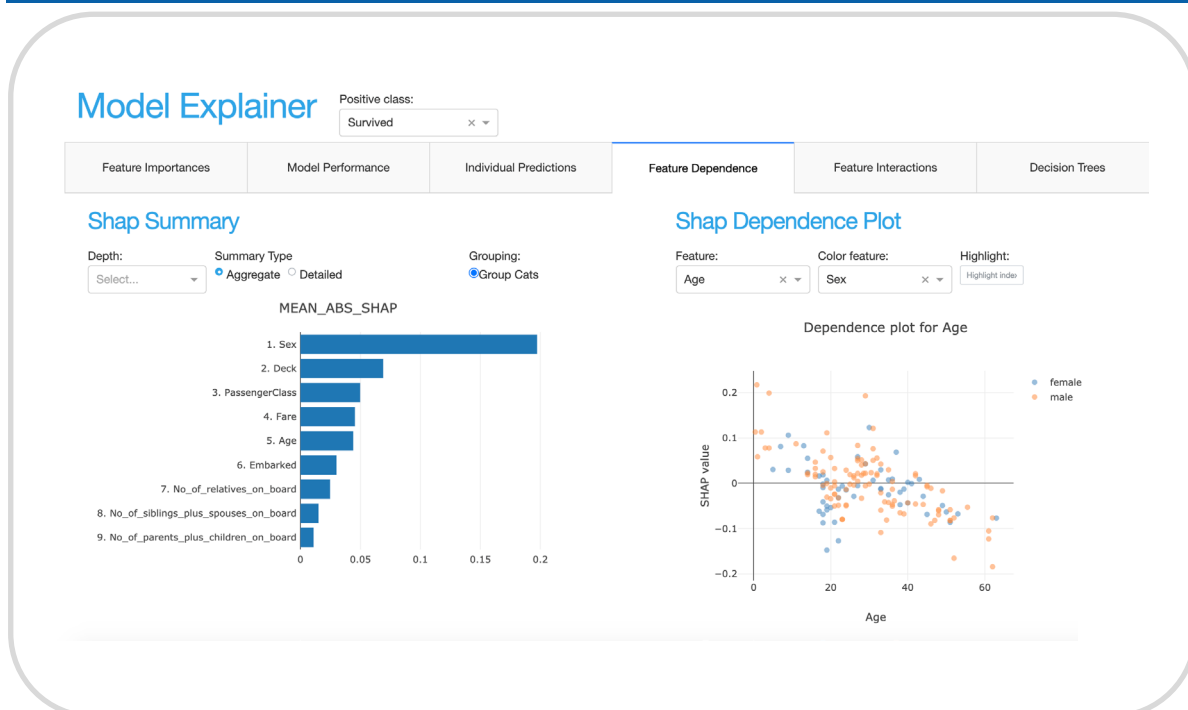
Python のパッケージ「Explainer Dashboard」なども便利である。図表 6 は、タイタニックの生存者分析をまとめたダッシュボードである。各変数が個人の生存確率の予測値にどれだけ影響しているか判断することができる。これは、技術的には、SHAP 値¹⁰と呼ばれる方法を用いている。

SHAP 値もその一例だが、最近は機械学習・AI のブラックボックス化を回避すべく、「説明可能な AI (eXplainable AI, XAI)」の研究が進められている。一方で、XAI はまだまだ開発途上の技術であり、そもそも、より複雑なモデルにおいて、はたして説明可能性が成立するのか議論があることは付記しておきたい。

⁹ リニアな因果関係とは、例えば、「 $y=ax+b$ 」といった関係である。また、そもそもデータサイエンスの世界では、因果関係を証明することはきわめて難しいことの一つとされている。

¹⁰ SHAP (SHapley Additive exPlanations) 値とは、協力ゲーム理論におけるシャープレイ値のコンセプトを機械学習モデルに応用し、予測値に対する説明変数の寄与度を計算する手法。

(図表 6) Explainer Dashboard



出所 : <https://explainerdashboard.readthedocs.io/en/latest/dashboards.html> (2024年4月22日閲覧)

ウ) 「当たり前ではないか」への対策

データの束をポンと渡して、これで何かみつけよ、といった場合、例えばそれほどの業務知識がないデータ分析チームの場合まず何が「当たり前」なのかわからない。これで結果が「当たり前」ばかりであったとしても、当然といえよう。そこで、マネジメント層がどういう仮説を検証してもらいたいかは、データ分析チームに事前に明らかにしておく必要がある。分析前の企業ヒアリング・相互ディスカッションなどを十分に行うことが理想だ。そうすれば、データ分析チームは「当たり前」を確認する過程で「意外な」仮説を発見することができるかもしれない。

さらにいえば、当たりの結論にも意味があり、マネジメント層はよりアクセルを踏むことができる。「実は(間違っている)」の場合、現在の施策の再検討が望まれる。また「不明」の場合、データが不足しているのか、より適切な分析手法が必要なのか、さらなる検証が必要ということであり、決して無駄足ではない。

3. データ分析の最終目標は何か

まとめると以下のようになる。

データ分析チームは、マネジメント層が決断する立場にあること、数値の結果を常に求められていること、インパクトのある結果を求めていること、を理解しなければならない。また、可視化を洗練し、正確さとわかりやすさのトレードオフを常に意識して、どちらかに偏らずに分析を行うことが肝要である。

マネジメント層は、データ分析で確言はできないこと、できれば検証できる体制づくりを整えること、リニアな因果関係のみで説明できない事象があること、仮説をしっかりと提示すること、意外性ばかりを求めないこと、を意識する必要がある。

そのためには、両者のさらなるスキルアップとコミュニケーションの促進が重要なことはいうまでもない。このようにして、「ラストワンマイル」を乗り越えてデータ分析を企業文化に定着させる必要がある。

企業におけるデータ分析は「分析のための分析」ではない。目標は、直面する経営課題を可視化し、解決することである。そして、前述のDXの定義より、DXの一部としてのデータ分析の役割は、顧客体験を更新し、競争優位性を獲得するプロセスに貢献することにある。

例えば、こういう例が考えられよう。「品質の維持は大前提として、(顧客の)待ち時間こそが顧客の体験価値を大きく毀損し、競争力を喪失する契機となる」という仮説のもとに、全面的にシステムと人員計画を組みなおすと決定したあるDX企業があるとする。このDXを進めるのに、最初の仮説の立案から施策の検証に至るまで、データ分析は不可欠だ。

もちろん、最初からDXの一部としてデータ分析をやる選択肢ばかりではない。まずは直面する経営課題をコツコツ可視化し、分析していくことも順当である。

データサイエンスは企業に大きな未来を拓く。データ分析はその最初の一步である。

—以上—

参考文献

- 高橋威知郎、白石卓也、清水景絵「データサイエンティストの秘密ノート」SBクリエイティブ（2016年11月）
- 尾花山和哉、株式会社ホクソエム「データ分析失敗事例集」共立出版（2023年8月）
- 森下光之助「機械学習を解釈する技術」技術評論社（2021年8月）