

DX時代にマネージャーが知っておくべき データ活用と必要な人材とは



目次

「経験と勘」からデータドリブンな意思決定へ	2
データドリブンな意思決定で ROI は 13 倍 生産性は 5~6%向上する	2
マネジメント人材にこそ データを理解し伝えるスキルが必須	3
データ分析をするうえで重要な 4 つのポジション	4
普通のビジネスマンが当たり前 データ活用できる「市民データサイエンス」	4
市民データサイエンス実現に必要な 4 つのポジション	5
4 つのポジションには 適性がある	5
データ活用課題に対するデータビークルのアプローチ	7
データサイエンティストは データ加工に時間を使っている	7
専門スキルがない人でも 直感的に利用できるデータ分析ツール	8
診断的分析と予測的分析を可能にする 拡張アナリティクスツール	10
データ分析からアクションをどう導き出すか	12
データ活用で経営改善や 人事採用をスムーズに	12
データ分析からのアクションまでの 4 つの考え方	13
データ活用のサイクルを 回すための人材育成	13

「経験と勘」からデータドリブンな意思決定へ



データドリブンな意思決定で ROI は 13 倍 生産性は 5~6% 向上する

新型コロナウイルス感染症の影響を受け、2020 年の 1 年で社会は大きな変化を遂げました。

内閣府が運営する V-RESAS (<https://v-resas.go.jp/>) (新型コロナウイルス感染症が地域経済に与える影響を可視化した地域経済分析システム) では、さまざまな企業のデータを集めて公開しています。それを参照してみると、例えばイベントのあり方が大きく変化しているということがわかります。

2019 年のいわゆるビフォーコロナの時期にはオフラインイベントが盛んに行われていましたが、現在ではオンラインに切り替わってきていることがデータから読み取ることができるのです。

これまでオフラインのイベントを開催することがマーケティングの勝ちパターンだった会社は、オンラインに移行しなければなりません。しかし、移行するにしてもどんな媒体で行えばいいのか、どこにフォーカスすればいいのか、従来の「経験と勘」では導き出せないような市場の変化が起きています。

そこで重要になるのが「データをもとにした意思決定」です。

実際にデータ分析をすることによってどんなメリットが得られるのか、経済学の専門家たちがさまざまな事例を集めて実証検証を行っています。それによると、**データを分析した結果に基づいて判断する企業は、平均して ROI が 13 倍になるという結果が得られています。**これは例えば、1,000 万円投資をすると、1 億 3,000 万円のリターンがあるということを意味します。また、別の

角度から分析によれば、**データドリブンな意思決定をする会社はそうでない会社と比べて生産性が 5~6%ほど高いという研究結果も出ています。**

例えば 100 億円の事業を 1 年で 200 億円に成長させたいという場合、データを活用するだけでは難しいかもしれません。しかし、100 億円の事業を数千万円から 1 億円グロスさせるようなアイデアをデータから見つけることは可能であるということです。

データから見出した実行しやすいアイデアを選び、少しずつ試していけば、100 億円から 105 億円に売上を伸ばしたり、100 億円かかっていたコストを 95 億円に削減したりといったことができます。

マネジメント人材にこそ データを理解し伝えるスキルが必須

今から 10 年以上前、これからはビッグデータの時代だと盛んに言われました。データ分析をするために、高額なハードウェアやソフトウェアを購入した企業もあるでしょう。ここ 10 年ほどデータ活用を課題とし、データサイエンスに投資した企業も多かったのではないのでしょうか。

社内にデータ活用できる人材がいない場合、コンサルティングファームなど社外のデータサイエンティストに外注をする必要があります。その結果、現場とのコミュニケーションがいまぐれなかったり、社内にデータ分析のノウハウが蓄積しなかったりといったことから、社内でデータ活用人材の採用や育成をする動きも出てきました。

そうして内製化を進めても、今度はデータ活用

人材として実務経験を積んだ社員が転職してしまったり、すぐ辞めてしまったりして、投資がむだになったという話もありました。

近年、我々のところに寄せられる相談も、DX 人材やデータサイエンス人材をどうすればよいかといったものが中心になってきています。

世の中が DX 人材を求めるようになった転換点として挙げられるのが、Google のチーフエコノミストを務めた経済学者ハル・ヴァリアン氏の発言です。

彼は 2009 年に、「今後 10 年間で一番セクシーな仕事は、統計学の専門家である」と発言しています。この発言がきっかけとなって、2010 年代の 10 年間は統計学や機械学習の知見を持った人材の争奪戦が繰り広げられてきたわけですが、その次の 10 年間、2020 年代はいったいどんな時代になるのでしょうか？

そのヒントも、ハル・ヴァリアン氏のインタビューの中に隠れています。「**統計学の専門家はセクシーだ**」と言った直後、彼は「**統計家の仕事は、データを理解して価値を引き出す仕事の、ほんの一部にすぎない**」ということを述べています。

データから洞察を得て、何が大事なのかを理解し、それを伝えるというスキルはとても重要です。データサイエンティストのみならず、マネジメントに携わる人たちこそそうしたスキルを身につけるべきなのです。

このことから私たちは、データサイエンス人材に任せておけばいいのではなくマネジャー自身が直接データにアクセスして理解し、使えるようにならなければいけないというのが、2020 年代のトレンドになると考えています。

データ分析をするうえで重要な 4 つのポジション



普通のビジネスマンが当たり前 データ活用できる「市民データサイエンス」

IT 分野を中心に、投資に関する調査や助言を行うグローバル企業「ガートナー」は、ハイプ・サイクルという理論モデルを提唱しています。ハイプ・サイクルとは、「さまざまな技術はすべて最初は黎明期で、誰も注目していなかったところから『過度な期待』のピークを迎え、その後一旦幻滅期に入るが、必要とされる技術は啓蒙活動が続けていくことで安定期を迎える」という一連のサイクルをいいます。

ハイプ・サイクルでは、AI 技術はいま幻滅期に入っていると整理されていますが、その中で「市民データサイエンス」という言葉が注目されるといわれています。市民データサイエンスとは、データ分析の専門家でない普通のビジネスマン

が、当たり前のようにデータサイエンスと技術を活用してより意思決定の精度を上げ、正しい判断につなげたり、生産性を上げたりすることをいいます。ガートナーでは 2020 年から 2023 年にかけて、そういった技術が本格的に普及していくと予想しています。

これまで多くの企業では、データを高速に処理するためにハードウェアやクラウドコンピューティングに投資してきました。その次に、データ分析ができるような人材の採用や育成に投資してきました。しかし、ここで終わりではありません。なぜなら、分析結果のレポートを見ているだけでは価値は生まれないからです。

そのチャンスをどういうやり方で取っていくのか。あるいは、そういったコストをどう削減していくのか。会社の中で具体的な施策を考えて、そ

れを実行に移そうという意思決定が必要となります。その後、実際に現場で手を動かした結果をさらにデータ分析することで、再投資をするか、施策を取りやめるのか判断することができます。

このように、データ・分析・意思決定・現場というサイクルを回していくことがデータ活用の理想ですが、データ分析に投資をしても、分析結果をもとに新しいアクションを取りましようと思意思決定する人がいなければ、価値にはつながりません。

市民データサイエンス実現に重要な

4つのポジション

市民データサイエンスの実現には、次の4つのポジションが重要です。実際このサイクルがうまく回って意思決定できている組織は、この4つのポジションについて誰がどの役割をするか、ちゃんと機能しています。

4つのポジションの1つ目が、「ボス」です。数字と理屈で意思決定のリーダーシップを取る人のことです。データ分析から導き出されたロジックはレポートを読めば誰でも理解できます。その先の意思決定の部分でリーダーシップをとる人のことです。

2つ目。分析結果から仮説を考え、どんなアクションを採れば機能しうるかを考える、業務の「エキスパート」が必要です。最低限のデータサイエンスのリテラシーを持った上で、周りのスタッフとコミュニケーションが取れるビジネスのエキスパートがいるかどうかはプロジェクトの成功にとって非常に重要です。現場の事情やオペレーション、つまり、ヒト・モノ・お金がどう動いているかを具体的にイメージできる人であれば、分析結果の意味を理解することができます。

とでしょう。

3つ目はデータマネージャーです。社内のデータやITシステムについて土地勘がある人。データが蓄積していく仕組みやオペレーションの流れをイメージできる人がいれば、データの意味やデータの取り方を相談することができます。

4つ目が分析担当者です。分析担当者には、ITとハードワークが苦にならない人が向いているでしょう。ここでいうハードワークというのは、試行錯誤とも言い換えられます。データマネージャーやエキスパート、ボスといったポジションの人たちと分析結果をディスカッションして、違う見方で分析をやり直すといった試行錯誤ができれば、数学がそれほど得意ではなかったり、プログラムが書けなくても役割を果たすことができます。

4つのポジションには 適性がある

この4つのポジションごとにどんな人が向いているのか、具体的に紹介しましょう。

「ボス」や「エキスパート」の仕事は、社内で行い組むべき課題を整理して、分析結果から打ち手につなげることです。これらは本来、管理職や経営者の仕事ですが、コンサルティングファームなどが代わりを務めることがあります。かつてボストンコンサルティンググループにいた現早稲田大学教授の内田和成先生に、コンサルタントには何が重要か聞きました。

彼が答えたのは、次の3つです。**1つ目は、頭脳労働であることから基本的に地頭がよいほうが望ましいということ。2つ目は、頭がよくても人に好かれる人でなければ伸びないということ。3**

つ目として、知的タフネスであること。新しいことを試したり、周りと議論をしたりするとき、そこに興奮して楽しめるような人が向いているのだそうです。

これはボスやエキスパートといった人たちも同様でしょう。そして、当然自分たちの業界や業種の業務知識が求められます。さらには、会社内外にそれぞれ人的ネットワークがあり、調整ができるかどうか。こういったスキルを持つことが、いわゆる「仕事のできる大人像」になっていきます。「ボス」や「エキスパート」にはリスクテイキングをして周りを巻き込んでいくリーダーシップがある方が向いています。

「データマネジャー」に向いているのは、大規模なデータを現実的な速度で扱える人です。例えば、来月の企画を考えているときに、月末にならなければ揃わないデータを参考にするとなると、ボトルネックになってしまいます。また、既に何らかの言語でコードが書けるかどうかも大事なポイントです。Pythonだけでなく、C 言語や Java など命令型の言語や SQL にもある程度詳しいことが望ましいでしょう。さらに、通信や計算環境など、インフラ周りの設定ができるとよいでしょう。大規模なデータであれば分散環境を作っていく必要があるので、Hadoop などの経験が役に立つこともあります。

最後に「データ分析者」の役割は、業務課題に「何がどれだけ関係するか」を明らかにすることです。この役割を果たすにあたっては、業務のオペ

レーションや不測の事態に対する理解が大切です。数学やプログラミングの優先度はそこまで高くありません。数字を見て最低限何ができればよいかというと、どんな分析手法を用いればいいか、分析の際どんな注意が必要か、意義と限界を理解しておくことです。そして、論文や書籍で先行研究や先行事例を取り込んで、何かおもしろい分析ができないか、アンテナを張っておくことがポイントです。

最近では上記の 4 つのポジションに加えて、機械学習のエンジニアが必要とされているという話もあります。マシンラーニングのエンジニアであれば、Python プログラミングの知見はほぼ必須でしょう。それに加えて、音声や画像といったバイナリデータを扱う経験も求められます。また、論文を読む習慣が大切ですし、数学的な素養も必要とされます。機械学習についての新しい手法やアイデア、解決策が投稿される arXiv というサイトを毎日チェックすることも役に立ちます。

上記に加えて、計算環境を自前で整えるリテラシーも必要です。演算を自分たちの手前のハードでやるのか、それとも Azure や AWS といったクラウドサーバでやるのかで、精度が上がったり下がったりします。そういったチューニングを重ねる忍耐力や、新しい便利なツールへのアンテナを張ることが、専門的な機械学習エンジニアの適性として重要です。

データ活用課題に対するデータビークルのアプローチ



データサイエンティストは データ加工に時間を使っている

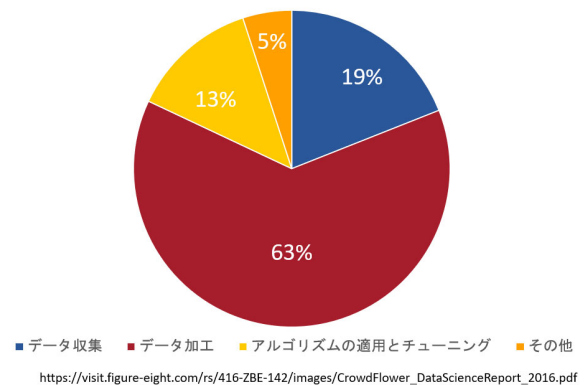
データ・分析・意思決定・現場というサイクルを回すことはとても大切です。DX 組織を考えていくうえで、この 4 つの中の足りない部分を補完することが重要になってきます。しかし、残念なことに、データ活用意識がないデータマネジャーがいた場合、セキュリティが完璧で業務が円滑に回っていても、データ分析や活用といった話には至りません。

あるいは、データ分析をする側があまりに業務を知らなさすぎるというパターンもあります。すると、数字上の分析結果は正しくても、業務上機能しないという問題が起こります。また、意思決定する人が責任を取りたくない場合や、新しいことに挑戦するのが嫌いなタイプであればそこで止まってしまいますし、現場の方々が数字に弱すぎて、データに基づく施策を取りたくて

も心に響かないということもあります。そうした中で、データマネジャーという仕事は非常に大事なポジションです。

ここで、データサイエンティストが何に時間を使っているのかを見てみましょう。この円グラフはアメリカのクラウドソーシング会社がデータサイエンティストの方に「何に一番時間を使っているか」を尋ねた結果です。

データサイエンティストは何に一番時間を使っているのか



グラフを見ると、過半数がデータ加工に時間を費やしていることが分かります。このアンケートからも読み取れるように、アルゴリズムに集中できているデータサイエンティストは意外と少数です。そこで我々データビークルでは、データ活用を阻む課題を解決するひとつの手段として、データ加工をサポートするプロダクト『dataDiver』を作りました。

dataDiver の特徴は以下の3点です。

1点目ですが、分析用のデータ準備は大きな手間であり、センス頼みという状況です。データマネージャーはさまざまなデータ活用の機会を経て経験を積むことでスキルを磨くことができますが、体系的な教科書がないためにどうしてもセンス頼み、経験頼みになってしまいます。そこでdataDiverは、蓄積している業務用のデータから、分析用データに加工して生成する過程を自動化することで、センスや経験に頼らないデータ加工を可能にしました。

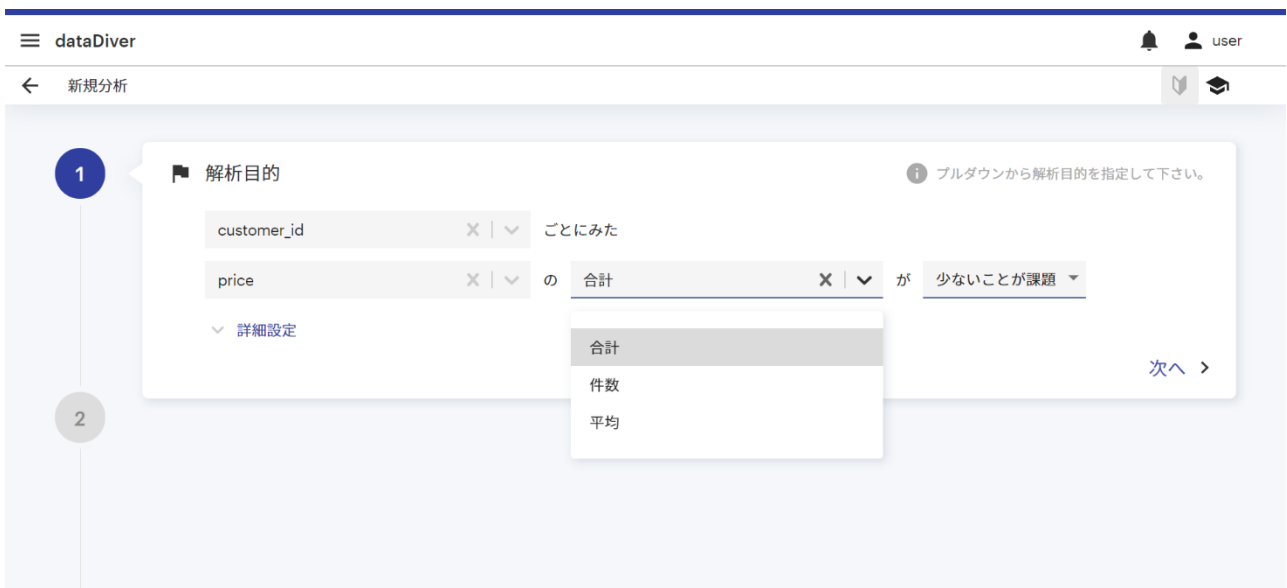
2点目として、変数や分析手法の選択は非常に貴重な専門的スキルで、これができる人材を採用するのは多くの会社にとって難しいことです。

そこで、dataDiverはこの部分も自動化し、誰でもわかりやすい分析モデルを探索できるようにしました。膨大な数の変数をすべて使うのではなく、人間がわかりやすい範囲に絞ったうえで、一番精度の高い予測モデルを作ります。

3点目として、そこまで自動化したところで、意思決定をする方が必ず統計学に詳しいとは限りません。そこで、分析結果は日本語の文章やグラフを自動的に生成して、できるだけわかりやすく表示するようにしました。そうすることで、意思決定をする人、あるいは現場の人に触っていただきやすくするツールがdataDiverなのです。

専門スキルがない人でも 直感的に利用できるデータ分析ツール

dataDiverの画面は非常にシンプルです。画面の左上を見ていただくと、解析目的の下に「customer_id」と入力されていますね。データベースに入っているデータは正規化されているという表現がされていますが、dataDiverではデータがどの粒度であったとしても、このIDをうまく使って集計し直すようにできています。



The screenshot shows the 'dataDiver' web application interface. At the top, there's a navigation bar with a hamburger menu, the text 'dataDiver', a bell icon, and a user profile icon labeled 'user'. Below this is a sub-header '新規分析' (New Analysis) with a back arrow and a book icon. The main content area is a light blue box with a vertical sidebar on the left containing two numbered steps: '1' (selected) and '2'. Step 1 is titled '解析目的' (Analysis Purpose). It contains a form with two input fields: 'customer_id' and 'price', each with a dropdown arrow and a close button. To the right of these fields is a label 'ごとにみた' (Viewed by). Below the 'price' field is a dropdown menu with options: '合計' (Total), '件数' (Count), and '平均' (Average). To the right of this menu is a label 'の' (of). Further right is another dropdown menu with options: '合計' (Total), '件数' (Count), and '平均' (Average). To the right of this menu is a label 'が' (that). Finally, there's a dropdown menu with options: '少ないことが課題' (Low count is a problem) and '多いことが課題' (High count is a problem). A blue button labeled '次へ >' (Next) is at the bottom right. A small information icon and text 'プルダウンから解析目的を指定して下さい。' (Please specify the analysis purpose from the dropdown menu.) are at the top right of the form area.

従来は、この「どんな形で集計し直せばいいか」という部分がセンス頼みでした。dataDiver では、蓄積したデータの中から、顧客ごとに良い悪いを見つけたり、商品ごとに良い悪いを見つけたり、スタッフごと、あるいは事業所ごとに良い悪いを見つけたりと…いったように、さまざまな切り口から集計することができます。

その下、2行目には「price」と書かれていますが、この行ではお客様のどこを見て良い悪いを判断するかを指定しています。お客様の購入金額の合計なのか、件数なのか、平均なのか、お客様ごとのテーブルからデータをどう集計するかが自動的に提案されてきます。

そして、最終的に購入金額が多いことが課題なのか、少ないことが課題なのかを選びます。これらたった 4 つを選んでいただくだけで、数十秒後に次のような分析結果が表示されます。

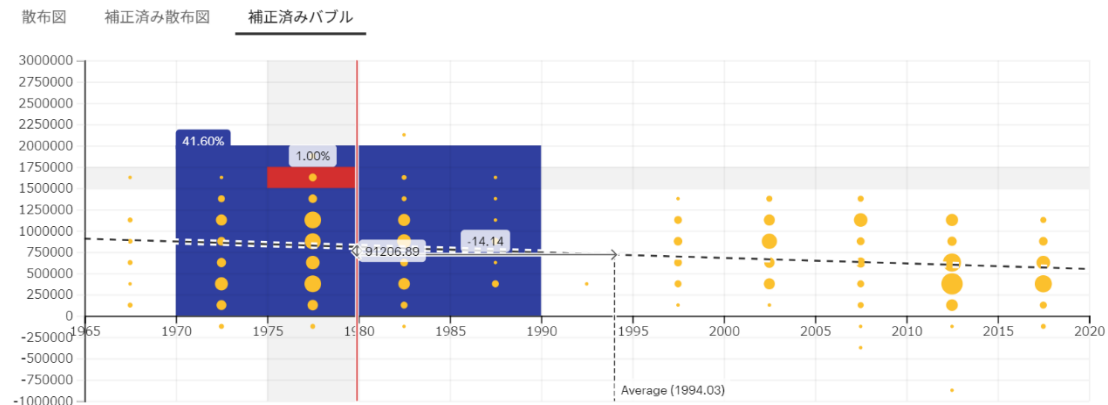
こうした分析結果を見ることで、ソフトウェアを買ってくれた人が優良顧客であるということが分かってくるかもしれません。あるいは、顧客マスタの中からお取引先様の業種が重要だということが分かるかもしれません。

分析結果では、ポジティブなトピックが青字で、ネガティブなトピックが赤字で表示されます。そして「クリアさ」を見ていただくと、それが統計的に誤差の範囲と言われるものなのか、それとも再現性が高く信頼できる結果なのかを判断することもできます。

次に、「詳細」という棒グラフのボタンを繰り返しクリックすると、横軸に説明変数、縦軸に課題として設定した購買金額の合計が表示されます（写真は次ページ）。これを見ると、「90 年代に創業した会社がお取引様にいない」ということが分かります。



顧客企業設立年の値 1 増える毎に 購買金額 の合計が6450.83低い傾向にあります



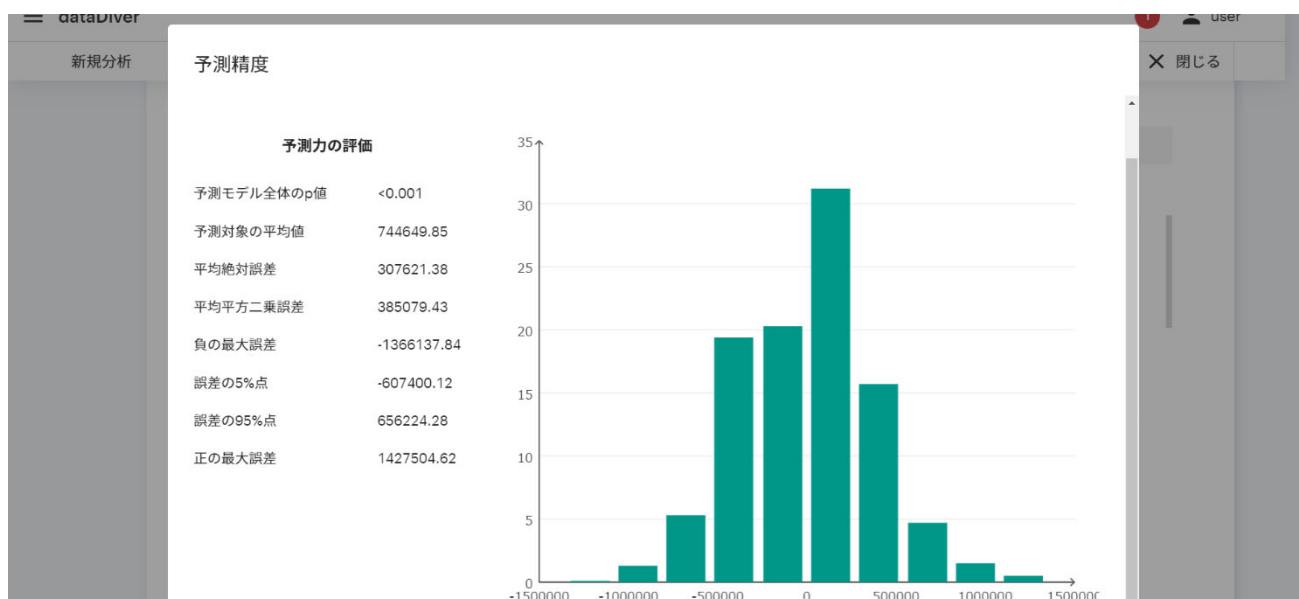
ほかにも「1980 年代に創業した会社の平均的な客単価がどれだけ高いか」が分かったり、「1970 年代から 90 年代の 20 年間で創業した会社が何%を占めているのか」といったことが、自動的に可視化されます。

予測精度についてもワンクリックで確認できます。たとえば「この条件だと平均して購入金額がこれぐらいになる」という予測に対して、どれぐらいぶれるのか、バッファをどれぐらい見ればいいのかなどについて、シミュレーションも簡単に行うことができます。

診断的分析と予測的分析を可能にする 拡張アナリティクスツール

前述したガートナーは、データの扱い方を 3 つに分類しています。

1 つは、過去に何が起きたのかを正確に把握する記述的分析です。みなさんが BI ツールで行う分析は、この記述的分析です。記述的分析手法からもう一段階踏み込んだ分析として、診断的分析と予測的分析の 2 つが挙げられます。



診断的分析は、過去に何が起きたかという数字だけを見るのではなく、なぜこうなったのか原因を見つけ、対策を取るもの。予測的分析とは、このままの状況が続いた場合にどれくらいの数値に落ち着きそうかを予測するものです。

弊社のクライアント様には引き続き BI ツールを使って記述的分析をしていただきながら、診断的分析や予測的分析を dataDiver で自動化していただいています。

実は、データビークルを創業した 2014 年当時、dataDiver が一体何を分析するツールなのか、分類するジャンルがありませんでした。それが近年、ガートナーで拡張アナリティクスという言葉が使われるようになりました。

拡張アナリティクスとは、データ準備と洞察の生成、そして洞察の可視化を自動化するツールです。ガートナーでは拡張アナリティクスツールがあれば多くの状況でデータサイエンティストの関与が不要となると予測しています。

業務知識のある人や意思決定する人自らがマーケティングや営業、人事、あるいはオペレーション改善について分析をし、より高度なことについてはデータサイエンティストにお願いするといった流れが、2020 年代のトレンドになっていくと考えられています。

データ分析からアクションをどう導き出すか



データ活用で経営改善や 人事採用をスムーズに

オラクルや MySQL といったビジネスデータ系の AI 構築自動化ソリューションの中で、dataDiver は国産クラウドサービス No.1 を獲得しました。dataDiver は実際にどんな場面で活用されているのでしょうか。

例えば、**隠れた優良顧客を見つけたいという場合**、EC サイトの購買履歴や POS データ、MA 接触歴、商品マスタといったデータを使用します。これらの ID を dataDiver の中で連結して分析することができるのです。

ところで、何をもって優良顧客というのでしょうか。たくさん購入してくれる人とそうでない人では、何が違うのでしょうか。**お客様の属性や**

購買パターンから、こういう商品を買っている人は優良顧客であるといったことが見えてくれば、新しいキャンペーン企画や営業活動、あるいは商品企画といった場面でさまざまな施策を考えることができます。

また、dataDiver を活用することで需要予測を行うことも可能です。例えば、「商品が売れました」「商品を出荷しました」といったデータを持つ会社はたくさんあるでしょう。営業活動の一環としてどんなイベントを開催したか、どんな CM を打ったかといったプロモーション履歴と商品の特性を掛け合わせることで、その商品がいつ、どれだけ売れるのかを予測することができます。生産や仕入れを適正化し、コストや売上を正確に把握することができれば、経営状態をより良い方向に導くこともできるでしょう。

人事採用に活かすこともできます。採用時の評価やテスト成績、あるいはどんな資格を持ち、どんなキャリアを歩んできたか、また、ストレステストを含めた心理検査の結果などを掛け合わせることで、ハイパフォーマーの特徴を分析します。営業の数字を伸ばしている人や 360 度評価で優秀だと評価されている人たちがどの時点でハイパフォーマーとなったのか、逆に、離職してしまう人にはどんな特徴があるのか。そういった分析結果や予測を採用プロセスに反映することで、全体のパフォーマンスが上げられる可能性があります。

データ分析からのアクションまでの

4 つの考え方

データ分析からのアクションの考え方には、大きく 4 つあります。

1 つ目は「変える」。ある特徴を持った人が優良顧客やハイパフォーマーになるといった場合、その特徴を持たせるために変化を促す施策を打ちます。優良顧客の特徴を持たせるためにお客様に広告を打ち、今まである商品を買っていなかった人にも買わせることで優良顧客になってもらうのです。

2 つ目は「狙う」。優良顧客が男性であったり女性であったりするように、顧客の属性が変えられない場合には、その属性の人たちが集まる場所でイベントを行ったり、よく見ている媒体に広告を出したりして、効率的にリーチするよう施策を打ちます。

3 つ目が、「大丈夫にする」。ある特徴を持った人は優良顧客になりにくいという事実があったとき、その特徴を持っても大丈夫にするサービスやツールを提供するという意味です。例え

ば、「男性はあまり買わない」という商品があった場合、男性特有の買いにくい理由があるのかもしれませんが。そういった人たちでも大丈夫なツールや、大丈夫にするサービスを提供しようというのが 3 つ目のアクションです。

4 つ目は「最適化する」というアクションです。

データ分析の結果から満足いく精度の予測モデルが完成したら、その予測値に応じて最適化するリソースは何か考えましょう。例えば、仕入れや生産量、人員数など、適切に配分することができれば無駄なコストを抑えて経営状態を改善することができます。

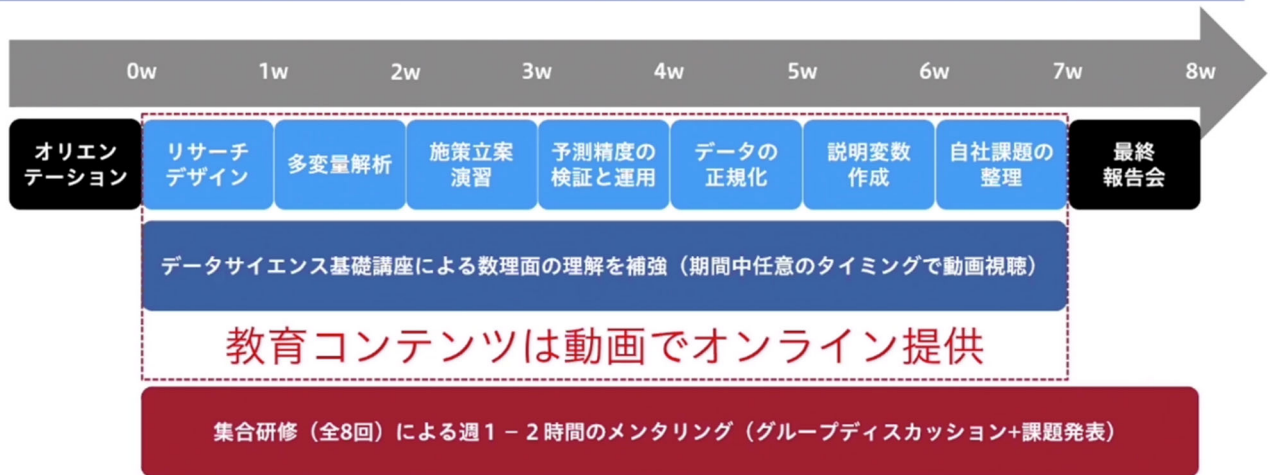
ここまで、意思決定にフォーカスをしてアクションの考え方を紹介しましたが、いずれにしても「データ・分析・意思決定・現場」のサイクルを回していただくことがデータ活用の理想です。このサイクルを一周回すことができれば、さまざまな分野に応用することは難しいことはありません。

データ活用のサイクルを 回すための人材育成

データビークルでは、こうしたデータ活用のサイクルを回すための人材育成メニューを新たに開発しました。

このプログラムは、プログラミングの過程を省力化し、分析の設計や施策の立案といったメタスキルにフォーカスしているのが特徴です。演習ではいきなり自社の業務課題に取り組むのではなく、親しみやすい問題設定とリアルなデモデータを用意して、8 週間をかけて仮想的な OJT に取り組んでいただきます。

データ活用人材研修プログラム概要

 DataVehicle


集合研修は、ディスカッション・発表を中心に実施

想定研修時間各週約5時間×8回=合計40時間

※実際はオンラインコンテンツの視聴時間、課題実施時間で前後します。

冒頭でもお話したように、「経験と勘」が通じにくくなった時代だからこそ、データドリブンな意思決定が価値を持ってくるのではないのでしょうか。それを実現するために、市民データサイエンス型組織を作っていきましょう。

（了）