

# 文系新人が目指すDXの担い手「データサイエンティスト」

SBIホールディングス株式会社 社長室ビッグデータ担当  
アシスタントマネジャー 田中紀帆氏

## はじめに

コロナ禍真っ最中で、卒業旅行すら行けずに文系私立大学を卒業した私は、2021年4月に証券・銀行・保険をコアとする総合金融グループであるSBIグループの持株会社、SBIホールディングス株式会社に新卒入社しました。それから約1年半が経ち、配属先である社長室ビッグデータ担当で、AIを活用した地域金融機関のDX（デジタルトランスフォーメーション）の担い手としてデータサイエンティストとなりました。

本稿では、文系で田舎の高校を卒業し、大学でデータサイエンスを専門的に学んできたわけでもなければ、理数系の科目が得意でもなかった私がどうやってデータサイエンティストになったのかについて紹介させていただきます。

## 1. DXの第一歩としての人材教育

データやデジタルを活用してビジネスを変革させ、新たなサービスやビジネス・モデルを構築することで、お客様に新しい便利な価値を提供し、企業自身の組織・文化・社員も変化し続けるのがDXと考えます。あらゆる産業でDXが必要と叫ばれている一方、IT化やデジタル化に向け、会社の中にデータやデジタル技術を取り込んでいくことさえ実践するのは大変難しいものです。とりわけ地方でDXを実践し、新しいビジネス・モデルを起こしていくには大きなハードルがあるように思えます。そこで大きな強みになるのはAIを使いこなせる人材です。

AIを使いこなせる人材は、一般的にデータサイエンティストと呼ばれています。さらに、①プログラミング・ITスキルと②統計知識に加え、③データを用いて意思決定できるビジネス知識も必要とされており、一般的には非常にハードルが高いと思われています。それでも世界中で多くの人がデータサイエンティストになろうと日々学んでおり、日本でも多くの社会人や学生が学んでいます。そして、データサイエンティストが活躍する企業は、どの地域にあっても日本全国または世

界で競争できる可能性を手にします。

しかし、DXを実現するために、データサイエンティストを確保することが企業にとって強い力になるとしても、どのように育成すればよいのかは悩ましい課題です。そこで、文系大学を卒業後、社会人1年目にデータサイエンティストを目指して学んだ私の経験をもとに、DXに必要な人材教育のステップをお伝えできればと思います。

### データサイエンティストに必要なスキル

- ①プログラミング・ITスキル
- ②統計知識
- ③データを用いて意思決定できるビジネス知識

## 2. データサイエンティストへのステップ

私は、大学時代から社会人1年目を通じて3つのステップで学んできました。まずは①データに興味を持つことと、次に②工夫して手軽にプログラミングを学ぶこと、そして③現場の生の声や感覚を丁寧に聴くことです。

本当に一人前のデータサイエンティストになるためには、さらに経験を積む必要がありますが、AIをビジネスに活用するという第一歩においては、このステップで始められると考えています。

### データサイエンティストへの3つのステップ

- ①データに興味を持つ
- ②工夫して手軽にプログラミングを学ぶ
- ③現場の生の声や感覚を丁寧に聴く

## 3. データに興味を持つ

最初のステップとしては、非常にシンプルで簡単ではありますが、とても重要なポイントである①データに興味を持つ、ということです。

大学では商学部でマーケティングを専攻していましたが、所属していたゼミと消費者行動論の講義において、顧客分析からターゲティングを実施

する機会があり、そこで初めてデータサイエンスという領域を知りました。講義の中で顧客ターゲットのためのデータ分析を実施しましたが、当時はプログラミングスキルが全く無い上に、マーケティングの知識も学習の途上であったため、何をどのように分析すればよいのかよくわかりませんでした。

一方、マーケティングを学習していく中で、インターネットやAIの普及によって、ビジネスにデータの活用が必要不可欠なものになっていると確信しました。そこからデータに強く関心を持つようになり、データサイエンティストが活躍する弊社への入社が決まりました。

繰り返しになりますが、まずはデータに興味を持つことが大事です。どのようなきっかけであっても構いませんが、日常的に意識することで自然と目につくようになり、データがあると何か意味があるのではと考えられるようになるはず。そこに楽しみを見出すことが出来れば、最初のステップはクリアです。

#### 4. 工夫して手軽にプログラミングを学ぶ

新卒で入社してから、データサイエンティストとして本格的に勉強を開始し、まずはプログラミングの習得に注力しました。最近では、特別な教室に通ったり、経験者が身近にいて教えてもらったりすることが無くても、オンライン学習によって、かなり多くの事が学べる環境が整っています。

私は、paiza、SIGNATE等のプログラミング学習用サイトの他、総務省による統計力向上サイト、データサイエンス・スクール等を活用していきました。オンラインで学習できるため、日本全国どこに住んでいても何時でも繰り返し学習することが可能です。

プログラミングの学習においては、このような手軽に学べる機会を積極的に利用することで、上司や先輩が忙しく質問し難い時でも自分で学習を進めることができ、ゲーム感覚で同期と切磋琢磨しながらプログラミングスキルの向上を目指しています。

このように、プログラミングにおいても学習方法は普及しつつあり、データサイエンティストへの間口もさらに広がっていくものと考えています。

なお、「日本データマネジメント・コンソーシアム (JDMC)」のYouTubeチャンネルにて、「Z世代の新入社員データサイエンティスト」として、若手社員が工夫してプログラミングを学ぶ姿勢についてインタビューを受ける機会をいただきました (下写真)。

#### 5. 現場の生の声や感覚を丁寧に聴く

入社して半年ほどの学習期間を経てから、実際にお客様を担当して、データサイエンティストの実務を開始しました。

まず実感したのは、分析対象の領域における専門知識が不足していることでした。データを扱う



引用：[https://www.youtube.com/watch?v=pLLxpwbBX\\_k](https://www.youtube.com/watch?v=pLLxpwbBX_k)  
 左上がJDMCの司会の方、右上が先輩の鍋倉さん、右下が同期の谷岡さん、左下が筆者田中

プログラミングスキルがあっても、分析対象のことを知らなければ役に立ちません。そこで、現場の人の生の声や感覚的な部分をたくさん聴くことが重要と解りました。

例えば、データ分析を進める際、実際に現場で業務に携わっている担当者に、どのように業務をしているのか、どのようにデータが作り出されているのかについて詳しくヒアリングします。現場に携わっている方にとっては当たり前の業務であり、感覚的にやっていることかもしれませんが、データを扱う上では、そこをしっかりと理解することが欠かせません。現場の日常に潜む材料の中から、気づきや、改善に向けた事実の発見につながるのです。

また、相手方と密にコミュニケーションを取ること、データへの理解や業務知識の習得だけではない、プラスの副産物を生み出します。業務上発生したデータや、業務の流れを説明することで、これまで現場で行われていた事が言語化され、整理されることで、現場の方の業務・データの理解向上につながります。データの分析や活用への提案を繰り返す中で、現場から改善の要望やアイデアが出てくることも度々あります。

このような取り組みを繰り返す事により、データサイエンティストと現場の信頼関係が構築され、実際の現場での改善施策にデータ分析の結果やAIが活用されるようになります。分析やAIも現場で使われなければ意味がありません。ビジネスの意思決定につなげるためにも現場とのコミュニケーションは非常に大事だということを、実際の業務を通じて強く実感しました。

## 6. AutoMLが大いに助けになる

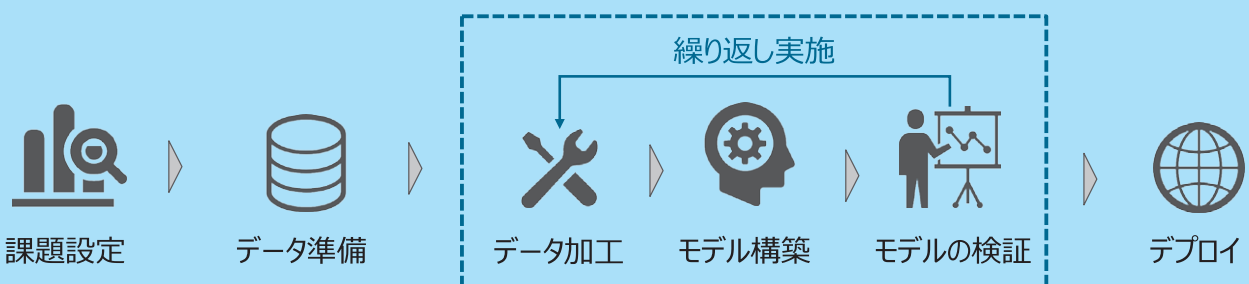
データサイエンティストの仕事をする中で、まだまだ①ITスキルと②統計知識が不足していると感じる場面が多々あります。そこで大いに助けになるのがAutoML (Automated Machine Learning=自動化された機械学習) です。データサイエンティストのAI構築業務の中でも高いスキルが要求される"データ加工"、"モデル構築"、"モデルの検証"のプロセスを自動化するAutoMLを活用することで、AIを構築するハードルが大きく下がります。

しかも、AutoMLのサービスは有料のものから無料のものまで数多く提供されており、多くの人が手軽に活用できるようになっています。

代表的な AutoML

	名称等	特徴
無償版	auto-sklearn	AutoML OSS の老舗
	AutoGluon (Amazon)	最短 3 行で高精度な AutoML を実現する OSS
	H2O (H2O.ai)	GUI のみで操作可能な Java 実装の AutoML プラットフォーム
	PyCaret	コードを大幅に削減可能な既存ライブラリのラッパー
	AutoKeras	Keras ベースのディープラーニング AutoML OSS
	Model Search (Google)	Google 発のディープラーニング AutoML OSS
有償版	DataRobot	大企業を中心とする 100 社以上の企業で導入済
	Prediction One (Sony)	直感的に操作しやすい GUI (Graphical User Interface)
	MatrixFlow	シンプルな機能性で、正確な売上予測、不良品判別、社員の能力分析等を実現する
	dotData	データ収集・加工からモデル設計、可視化、運用までのプロセスを数ヶ月から数日に短縮できる

## データサイエンティストの業務フロー



AutoMLによって自動化



## 7. 市民データサイエンティスト

AutoMLのようなデータサイエンスに関するツールやソリューションは、多様なものが提供されています。これらを用いることにより、高度に専門的なITスキルや数学の知識が無くても、簡単にデータ分析やAI構築ができるようになっていきます。データサイエンティストのように専門的な作業は出来ないが、ツールを使って現場の知識やノウハウを用いて、データ分析・AI活用を行う人材を市民データサイエンティストと呼びます。DXが多くの企業・産業で拡大していく中では、市民データサイエンティストの必要性も高まっています。

データサイエンティストを会社に極少数だけ育成したとしても、現場でデータを活用できる市民データサイエンティストをたくさん育成することによって、DXを推進できる体制は整います。データサイエンティストとして一人前になるには3年以上かかると言われていますが、市民データサイエンティストは約1年で十分なスキル習得が可能と言われています。

そのため、データサイエンティストの育成を進めるとともに、オンライン学習や現場でのデータ・AIの活用により、市民データサイエンティストも育成する教育体制を整備することで、DX実践の素地が多くの会社で構築できると考えています。

### さいごに

新型コロナウイルス拡大の影響によって、消費者の行動や需給のあり方が大きく変容している今、現状を分析し変化を捉えることが、ビジネスの意思決定において必要不可欠なものになっています。企業が自らを変化させ続けDXを実現していくためには、データを活用する、活用できる人材

を育成する、環境を整備する必要性がますます高まっていると考えられます。

私がコロナ禍で新卒入社し、これまでと大きく異なる環境での社会人生活をスタートさせて、1年半が過ぎました。そして、データサイエンティストとしてのスキル習得・経験を積む中で、私が学生時代に感じていたよりも、データ・AI活用の門戸が非常に広く開いている事は明らかでした。また、地域に関わらず、どこにいてもオンラインでデータ活用の教育が受けられるようになっている現状は、地方のDX促進につながっていくと実感しています。

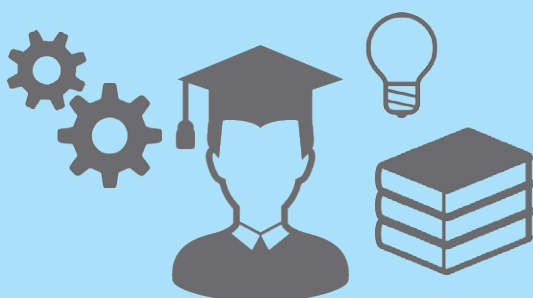
私が学んできたデータサイエンティストとしてのステップは、未知の領域ではありましたが、コロナ禍による変化とDXが求められる環境に後押しされ、多くの知見や経験を得やすいものでした。今回、私が学んできたステップを共有する機会をいただいたという事は大変貴重でありたいことであると感じています。これからも、データ・AIを活用しDXに取り組む方に向けて、私の経験から伝える事ができるように、AIを活用した地域金融機関のDXの担い手として、データサイエンティストの業務に日々取り組んでまいります。



筆者 田中紀帆

### データサイエンティスト

統計・機械学習に関する知識をもつ人材



### 市民データサイエンティスト

統計・機械学習以外の領域の業務知識をもつ、業務としてデータ分析を実施する人材

