



データサイエンス

No.25(2022)

和歌山大学図書館 パスファインダー

ラーニング・アドバイザーがおすすめる…

データサイエンス

はじめに

近年、DX(デジタルトランスフォーメーション)や AI、IoT という言葉を耳にすることが多くなったのではないのでしょうか。これらの普及にはデータサイエンティストという職業の方々が深く関わっています。このパスファインダーは、データサイエンティストを目指したい方や興味がある方向けに作成しました。データサイエンティストに求められる知識である「統計学」と「プログラミング」の基礎について学ぶための本を主に取り上げていきます。

発行：2022.01

和歌山大学図書館

作成：ラーニング・アドバイザー
(システム工学研究科)

1. はじめに/イントロダクション

1-1. 「データサイエンス」とは？

★統計学、情報工学など、様々な領域の手法を用い、有意義なデータを引き出すための研究分野のことです。その分野を専門とする職業のことを「データサイエンティスト」といい、企業のDX推進や様々な分野でのAI/IoTの導入が進む現代社会で需要が高まっている職業の一つとされています。



1-2. 学習するにあたってのポイント

★基本的にデータサイエンティストに必要な能力として、一般的に以下の3つが挙げられています。

- ①プログラミング等のITスキル
- ②統計学や機械学習の知識
- ③マーケティング・ビジネスに関する知識・スキル

特に①や②の能力はデータサイエンティストの基本的なスキルになるので、しっかりと学習しましょう。ただ実際、本やインターネットで勉強しようとする、内容が難しく挫折してしまうことが多いです。そのため、このパスファインダーでは、初学者向けに統計学とPythonプログラミングの基礎を学習できる本を紹介しています。

また、オススメの勉強手順は、以下の通りです。

- ①データサイエンスの導入本を読んで、興味を深める
- ②統計学とプログラミングの基礎を学ぶ
- ③実際にデータを触ってみる

そして、勉強手順③を行う際に以下の2点を意識してください。

- ①課題設定と目的の明確化(手持ちのデータを解析して、どんな情報が欲しいのかを決める)
- ②仮説をしっかりと考えること

これを意識せずにデータ分析をすると何をやっているのかわからなくなり、泥沼にはまってしまう。そのため、課題設定・目的・仮説の三つについてしっかりと考えてからデータ分析をしましょう。

※勉強手順③を行う際、kaggle という競技プログラミングサイトを使用することをオススメします。(<https://www.kaggle.com/>)

初心者向けのコースなどもあり、何より多くの種類のデータを提供しているのですのでオススメです。競技者にはランクがあり、上位ランカーになると就職活動でも役立ちます。

1-3. 一般向けに書かれた資料・読み物

- ◆ シン・ニホン：AI×データ時代における日本の再生と人材育成/ [安宅和人]著,
[ニュースピックス], 2020年

(配架場所：[2F・開架図書] 請求記号：[332.107||AK] 書誌ID[1000469166])

★AI やデータサイエンスの領域がなぜ重要なのか、そして強い日本になるためにどのようにこの領域を活用していくべきなのかが述べられています。専門的なお話はあまり登場せず、分かりやすく平易な言葉で理解しやすい内容に仕上がっています。AI・データサイエンスまわりの教養を身に付けるのにピッタリの本です



- ◆ 文系のためのデータサイエンスがわかる本/[高橋威知郎]著, [綜合法令出版], 2019年

(配架場所：[2F・開架図書] 請求記号：[336.17||TI] 書誌ID [1000460388])

★「文系のための」というタイトルですが、文系理系問わずためになる一冊だと思います。作者の実務経験をもとに、データ分析をする上での心得えを分かりやすい文章や、イラストなども用いて書かれているので、初学者にオススメの一冊です。



2. 学習用資料

2-1. 統計学

- ◆ 統計学の図鑑（まなびのずかん） / [涌井良幸], [涌井貞美]著, [技術評論社], 2015 年

（配架場所：[2F・開架図書] 請求記号：[417||WY] 書誌 ID：[1000473904]）

★統計学の入門書です。

図やイラストが多く、わかりやすいので、
初めて統計学を学ぶ方でも理解できます。



- ◆ Excel によるやさしい統計解析：分析手法の使い分けと統計モデリングの基礎 / [荒川俊也]著, [オーム社], 2020 年

（配架場所：[2F・開架図書] 請求記号：[417||AT] 書誌 ID：[1000467701]）

★統計学を一通り学習したら、実際に手を動かして
実践してみましょう。いきなり Python とか R などの
プログラミングから始めると挫折しやすいので、
まずは Excel から始める方がいいです。

（使うツールの違いだけで、やることは一緒です。）

この書籍では、実例をもとに web サイトで配布されて
いるデータを用いて、統計学を実際に手を動かしながら
学ぶといった内容になっています。



★ その他の学習コンテンツ

- ◆ 統計 Web <https://bellcurve.jp/statistics/>

- ◆ ヨビのり確率・統計

<https://www.youtube.com/watch?reload=9&v=Bj8fkq533Dc>

2-2. プログラミング[Python]

- ◆ **東京大学のデータサイエンティスト育成講座：Python で手を動かして学ぶデータ分析 / [塚本邦尊], [山田典一],[大澤文孝]著, [マイナビ出版], 2019 年**

(配架場所：[2F・開架図書] 請求記号：[007.609||TK] 書誌 ID：1000457693)

★Python の文法の説明から始まり、numpy や Pandas などのライブラリーの紹介など詳しく解説されています。難しく書かれていないので、初心者でも理解できると思います。また、後半部では簡単な機械学習の実装の仕方まで載っているので一通り網羅的に学習したい方にもオススメです。



- ◆ **Python1 年生：体験してわかる!会話でまなべる!プログラミングのしくみ/ [森巧尚] 著, [翔泳社], 2017 年**

(配架場所：[2F・開架図書] 請求記号：[007.64||MY] 書誌 ID：[1000438064])

★Python の解説本の中では一番易しい本だと思います。イラストを多く交えながら解説しているので、非常にわかりやすく、初心者の方でも挫折することはないと思います。オススメの一冊です。



★ その他の学習コンテンツ

★ Python の超基本的な部分をたった 1 時間で学べます【プログラミング入門講座】

<https://www.youtube.com/watch?v=HyU3XL2F9GE&list=PLavQwENTsEBUH7XZHGfKjTQqDovNYPmB&index=15>

◆ 中学生でもわかる Python 入門シリーズ

https://www.youtube.com/watch?v=rPCmv_W4Ng&list=PLCZyyif9kAwX7rIXKrJTIPQs9Nk7TBAHO&index=1

3. より深く学習したい方向けに

3-1. 専門書

- ◆ 44 の例題で学ぶ統計的検定と推定の解き方 / [上田拓治]著, [オーム社], 2009 年
(配架場所 : [2F・開架図書] 請求記号 : [417.6||UT] 書誌 ID : [1000473905])
- ◆ 欠測データ処理: R による単一代入法と多重代入法 (統計学 One Point) / [高橋 将
宜][渡辺美智子]著, [共立出版], 2017 年
(配架場所 : [2F・開架図書] 請求記号 : [417||TM] 書誌 ID : [1000473906])
- ◆ Python 機械学習プログラミング : 達人データサイエンティストによる理論と実践 /
[Sebastian Raschka][Vahid Mirjalili]著,[東京:インプレス],2020 年
(配架場所 : [2F・開架図書] 請求記号 : [007.13||RS] 書誌 ID : [1000466534])

★本文中で紹介されている図書の探し方★

和歌山大学図書館トップページの OPAC 検索窓に、
10 桁の書誌 ID を入力して検索してください。

<https://www.wakayama-u.ac.jp/lib/> 和歌山大学図書館

