

2024年2月22日
データサイエンティスト集会 in VRC

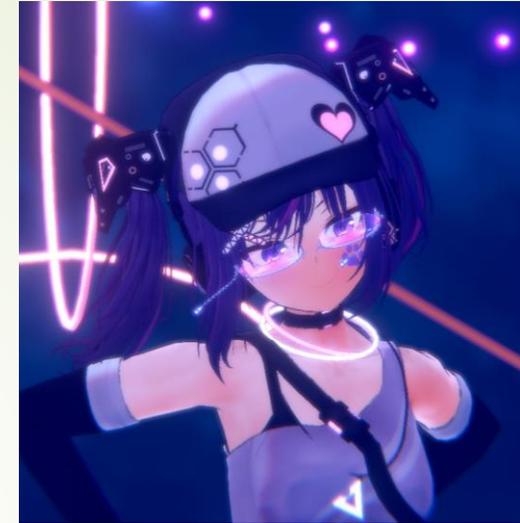
DS初心者向け
知名度は低いけどオススメの良書紹介 2
～データ分析業務 技術の外側入門～

1

ぶんちん

自己紹介 ぶんちゃん

- ▶ 複合経営が特徴の企業（製造業）に所属
- ▶ データ分析担当者だったが。。。



良い書籍はたくさんあるけど。。。。

▶ 良書は人気になりやすい



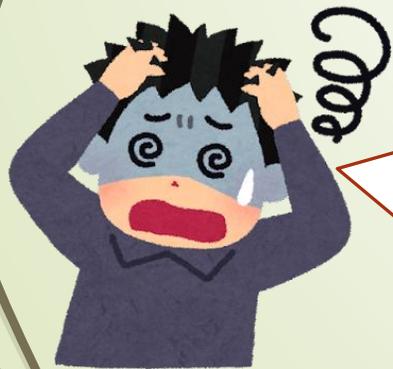
知られていない本でも良書がある
今回はそれを紹介

初心者が直面する壁

将来のためにデータサイエンスの勉強するぞ！
数学！機械学習！プログラミング！etc



勉強したことが仕事で使えない。
使ってみたけど成果につながらない。
実力不足。もっと勉強しないと！



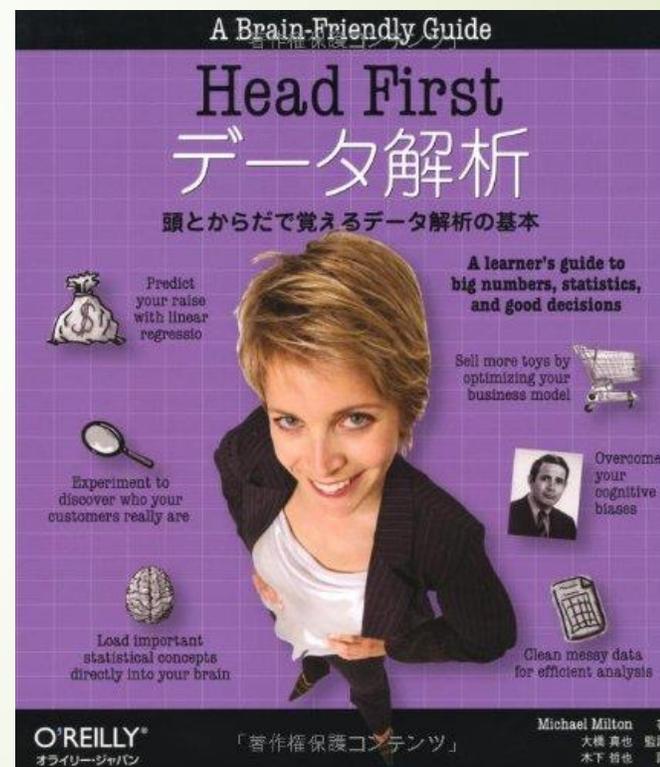
Q.いつになったら
仕事に生かせるの??
成果を出せるの??

A.不足しているのは技術ではなく、
仕事の進め方（技能）かも？

Head First データ解析

- ▶ オライリー本の超初心者向けシリーズ
- ▶ データ分析技術がメインの本ではない
- ▶ ゆるーい感じの本
- ▶ イラストまみれ

WOW!



技術面だけ見ると高度な話はない

目次

- ▶ データ解析入門—分析する
- ▶ 実験—持論を検証する
- ▶ 最適化—最大にする
- ▶ データの可視化—図を使うと賢くなる
- ▶ 仮説検定—否定する
- ▶ ベイズ統計—基準を活用する
- ▶ 主観確率—数値で表した信念
- ▶ 経験則—人間らしい解析
- ▶ ヒストグラム—数値の形状
- ▶ 回帰：予測
- ▶ 誤差：誤差を適切に示す
- ▶ リレーショナルデータベース：関連付けられませんか？
- ▶ データクリーニング：秩序を与える

基本的にExcel（一部R）で
できる程度の技術レベル

本書の本編は技術以外の部分

事例を用いて、仕事の進め方を説明

例えばヒアリングの観点や読み解き方を事例を用いて解説

課題は「売上を増やす方法を考え出す」ことであると定義しました。しかし、この課題ではあなたがどのように対処することを期待されているのかがほとんどわからないので、CEOから多くの有益なコメントを引き出しました。

このコメントから、化粧品事業の状況に関する重要な**基準となる仮定**がわかります。これが解析の**根幹**となるため、CEOが正しい仮定を行っていることが望まれます。CEOが述べた最も重要な点は何でしょうか？

このコメント自体が一種のデータです。どの部分が最も重要でしょうか？

これは「数量や状況」を聞く質問です。

最も役立つ情報は何か？

From: CEO, アクメコスメティックス
TO: Head First
Subject: Re: 課題の定義

売上をいくら増やしたいのですか？
この表の目標売上にまで戻す必要があります。当社のすべての予算はこの目標を基に作られているので、到達できないと困ったこととなります。

どのように対処すればよいと思いますか？
それを考えるのはあなたの仕事です。ただし、戦略としては人々にもっと購入してもらう必要があります。ここでの「人々」とはローティーン(11才から15才まで)の女子です。何らかのマーケティングを行って売上を上げることになるでしょう。あなたはデータ担当者です。考え出してください！

いくらの上増が実現できると思いますか？目標売上は適切ですか？
ローティーンの女子はお金を持っています。ベビーシッターでの収入やお小遣いなどです。彼女たちへのモイスチャブラスの売上には上限はないと思います。

競合他社の売上はどうなっていますか？
具体的な数字は知りませんが、わが社が圧倒されているという印象があります。他社は保湿剤の総収入で50%から100%上回っているでしょう。

広告費やソーシャルネットワークマーケティングの予算はどうなっていますか？
新しい方法を試しています。総予算は最初の月の収入の20%です。以前はこのすべてを広告に使っていましたが、ソーシャルネットワークに振り替えているところです。広告費を同じ水準にし続けていたらどうなっていたかを考えるとぞっとします。

コラムに重要情報もりだくさん

素朴な疑問に答えます

Q: 質問に答えるためにもっと詳細なデータが必要な場合、どこまで調べればいいのでしょうか？自分で顧客に聞き取り調査までしなければいけませんか？

A: 新たな詳しいデータをどこまで追い求めるかというのは最善の判断を下すための究極の課題です。ここでは新しい市場区分を見つけるまで調べましたが、説得力のある新たな販売戦略を打ち立てるにはそれで十分でした。データ収集をどこまで行うべきかについては、後の章でもさらに取り上げていきます。

Q: 最初にメンタルモデルが間違っていたことが最初に行った解析にダメージを与えたように思えるのですが。

A: そうです。最初に仮定が間違っていたことが解析を間違った結果へ導きました。そのため、モデルが最初から正しい仮定に基づいているかを確認し、仮定を覆すデータがあればすぐに戻って見直すことがとても重要なのです。

Q: これで解析は終わりですか？何らかの最終状態が欲しいのですが。

A: データ解析で重要な課題に答えを出すことは確かにできますが、すべてを知ることには決してできません。また、たとえ今日すべてのことがわかって、明日は異なっているでしょう。年配の男性に販売するというあなたの提言は現在は有効かもしれませんが、アクメ社は売上を追跡する解析が常に必要となるでしょう。

Q: 気が重くなりそうですね。

A: むしろ逆です！アナリストは探偵のようなものであり、解決すべき謎は常に存在します。だからデータ解析はとても楽しいのです！データアナリストとしての仕事の基本的な作業としては、規則の例外とみなすのではなく、立ち返り、モデルを改良し、新しいモデルを通して世界を見ることだけを考えてください。

どんな人にオススメなのか

周りに頼れる人がいない初心者・新入社員

- ▶ 技術以外に求められる業務能力があります。
- ▶ その一端を本書で学んでみましょう。

部下や後輩の指導をする人

- ▶ これまで経験して覚えてもらっていたことを言語化されている書籍です。
- ▶ 本書を効率よく指導する際の参考にしてみてください。