

# データ分析の活用

～情報の徹底活用～

渡邊研究室4年

110425300 三輪卓也

# 書籍紹介

- 題目：これからデータ分析を始めたい人のための本
- 著者：工藤 卓哉
- 発行：2013年12月6日
- 発行所：PHPエディターズ・グループ

# はじめに

## ■ データ分析とは

- 様々な膨大なデータから相関を見つけ、仮説を立てることである

## ■ データ分析による恩恵

- 意思決定プロセスの最適化

## ■ データ分析はもはや常識であること

- 2000年代中頃からアメリカの企業を中心に、ビッグデータをビジネスに活かした成功事例が注目されるようになった

# データ分析による成功事例

## ■ Amazon

- 購入確率が高そうな商品を勧めるシステムをデータ分析の結果から構築し、成功を収めた

## ■ ディズニー

- 混雑解消やアトラクションの待ち時間による心理的ストレスの緩和をする目的でデータ分析を活用

## ■ Google

- ユーザの入力する数十億にもものぼる検索語と季節性インフルエンザの統計情報から、流行地域をリアルタイムで特定できることを証明

# Amazonの成功(1)

よく一緒に購入されている商品



合計価格: ¥ 5,486

3つともカートに入れる

在庫状況の表示

- ✓ 対象商品:これからデータ分析を始めたい人のための本 工藤卓哉 単行本(ソフトカバー) ¥ 1,620
- ✓ データサイエンス超入門ビジネスで役立つ「統計学」の本当の活かし方 工藤卓哉 単行本 ¥ 1,728
- ✓ データサイエンティスト養成読本 [ビッグデータ時代のビジネスを支えるデータ分析力が身につく!](Software Design plus) 佐藤 洋行 大型本 ¥ 2,138

この商品を買った人はこんな商品も買っています



データサイエンス超入門ビジネスで役立つ「統計学」の本当の活かし方

工藤卓哉

★★★★★ (6)

単行本  
¥ 1,728



問題解決のためのデータ分析

齋藤 健太

★★★★★ (5)

単行本(ソフトカバー)

¥ 1,706



会社を強くするビッグデータ活用入門 基本から実践まで

網野 知博

★★★★★ (3)

単行本  
¥ 1,944



とある弁当屋の統計技師(データサイエンティスト)の統計学入門

石田 基広

★★★★★ (43)

単行本  
¥ 1,404

# Amazonの成功(2)

購買・サイト  
閲覧

- コンテンツ情報
- 顧客属性情報
- 時間情報

レコメン  
デーション  
システム

- 様々な分析モデルによる  
データ分析

商品のレコ  
メンド

- おすすめ  
商品
- 関連商品

# Amazonの成功(3)

## ■ 商品を勧めるシステム

ユーザの属性、購買行動、検索、ページビュー、商品評価等の情報を分析

購入確率が高そうな商品をユーザーに勧める

## ■ 在庫管理

過去履歴データから売れ行きを予測

売り逃しや過剰な在庫確保を極小化し、無駄を省く

# ディズニーの成功(1)

混雑解消のための  
待ち行列理論

- 待たずにサービスを受けられる確率
- 行列の平均の長さ
- サービスを受けられるまでの待ち時間の平均
- サービスの提供時間

この様々な要素から待ち時間を算出



# ディズニーの成功(2)

- 行列によるストレスの緩和

データ分析により顧客の導線を先読み



混雑する気配の出始めたアトラクションの導線経路に対してブランドキャラクターを効果的に配置



過剰な混雑の回避と顧客の満足度向上

# データ分析に必要な基礎知識

## ■ 平均

## ■ 分散・標準偏差

- データのばらつき度合いを示す
- 応用例（営業成績の評価）

単位：百万

	1月	2月	3月	合計	月平均
Aさん	100	110	120	330	110
Bさん	50	300	40	390	130

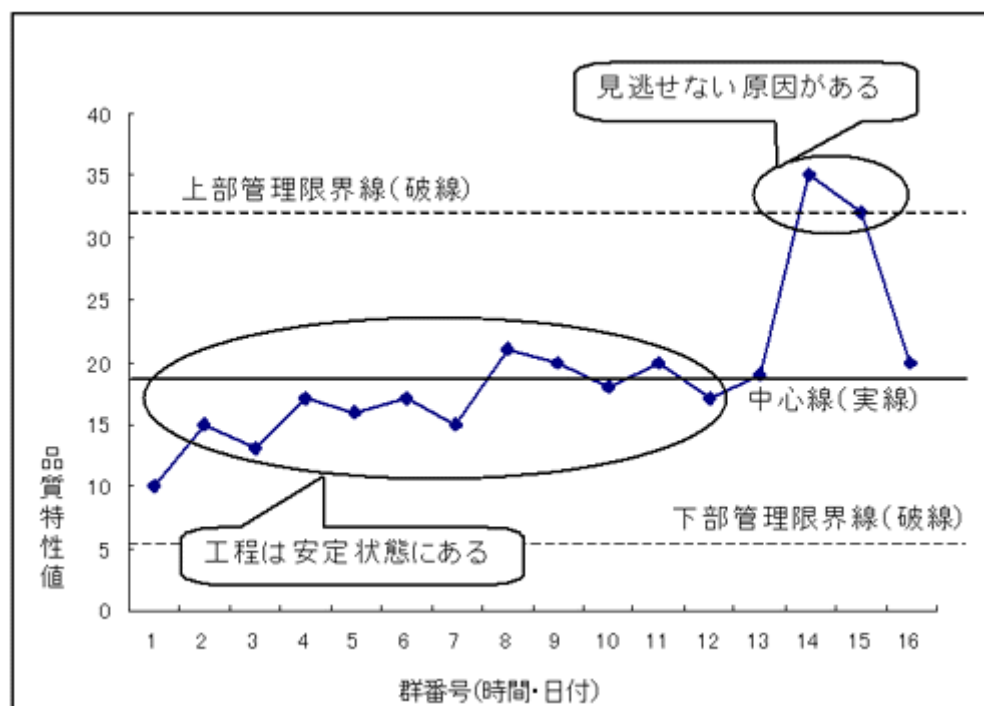
Aさんの売上の標準偏差：8.16

Bさんの売上の標準偏差：120.28

# 標準偏差を用いた分析手法

## ■ シューハート管理図

### □ 製品の品質管理や売上の状況把握



# 分析手法（ベイズ確率）

- 与えられた情報を基に、その事象が起こりうる確率の分布を更新
- 迷惑メールのフィルタリング等に利用
- 例：知人のAさんが関西出身である確率が40%であるという情報に、新たにAさんがプロ野球の巨人ファンであるという情報を得た場合
  - 関西出身者が巨人ファンである確率：20%
  - 他の出身者が巨人ファンである確率：70%

$$\frac{40[\%] \times 20[\%]}{40[\%] \times 20[\%] + (100 - 40)[\%] \times 70[\%]} = 16[\%]$$

# 分析手法（協調フィルタリング）

- 多数のユーザーの行動履歴を基に、他ユーザーの嗜好を予測する分析手法
- 商品のレコメンドシステムに利用
- 例：商品の満足度結果を基にした手法

①似た嗜好の顧客を探す	デジカメ	一眼レフ	ノートPC	タブレット端末
Aさん	1	4	?	?
Bさん	5	3	2	4
Cさん	1	5	5	1
Dさん	3	2	2	5

②商品の満足度を予想

# データ分析を現場で活かす

## ■ 知識の他に重要な要素

- コミュニケーション能力によって影響度合いは大きく変わる
- 最終目標に見合った的確な状況判断を下し、それに向けた行動力

## ■ 目的を見定め段階を踏まえた分析

- データを「とりあえず探索する」では失敗する
- たった一度の分析で作業完了ではない

# まとめ

- データ分析はビジネスの成功には欠かせない
  - Amazonとディズニーの例
- データ分析に必要な基本概念
  - 平均と分散・標準偏差
- データ分析を現場で活かすには
  - 知識のみならず、コミュニケーション能力や的確な状況判断も重要
  - データ分析自体を目標としないこと

# 参考資料

## ■ 初級シスアド取扱説明書

(<http://www.geocities.co.jp/SiliconValley-Cupertino/2190/index.html>)