



ビジネスパーソンのための データ分析

Hands-on入門

講師紹介 小池 歩

こいけ あゆむ

経歴

元俳優、大学卒業後、AI × 教育のベンチャー企業に入社。

AIやDLの講師を務め 10→50 人のフェーズを経験。

クラウドチームリーダーを務める。

製造業の物体検知モデルの作成や医療の子宮がんの画像分類（論文）

など様々なプロジェクトに参画。

前職よりトレノケートにて業務委託での講師をしており、2024年1月よりJOIN

主な経歴

- ・ マイクロソフト認定トレーナー
- ・ AWS 認定トレーナー
- ・ Ideactive Japan Project Professional Menta
- ・ AI-900 資格対策講座開発
- ・ 日本初データエンジニアリング講座開講
- ・ その他：AWS, MS の 8 資格



第0章 はじめに

コースの目標

- Excel でのデータ分析の手順を理解し、自ら実装ができる
- データから課題を発見し、適切なデータ分析、仮説の検証ができる
- 適切な情報取得ができ、データから次の施策を立案できる



本コースで学べること、学べないこと

データ分析の流れ	○	● データ分析の全体像と各ステップを学びます。
Excel での実装方法	○	● Excel を用いて、データ準備、前処理 分析の実装方法を学ぶことができます。
統計学の数学・理論	△	● 必要最小限にとどめます
データ分析の活用方法	△	● 活用例や開発の流れを扱いますが、代表的な ものに絞っています。

コース目次

第0章

はじめに

第1章

データ分析の概要

第2章

データの収集・探索的データ分析

第3章

データに基づく分析、処理

第4章

分析結果の考察、結論

第5章

演習

第1章 データ分析の概要

章の概要

この章では、データ分析とは、分析の流れについて紹介します。

実際のユースケースを通してデータ分析の流れを理解しましょう。

💡この章の目標

- データサイエンティストとはなにかがわかる
- なぜデータ分析を学ぶ必要があるのかわかる
- データ分析の流れを理解し説明できる



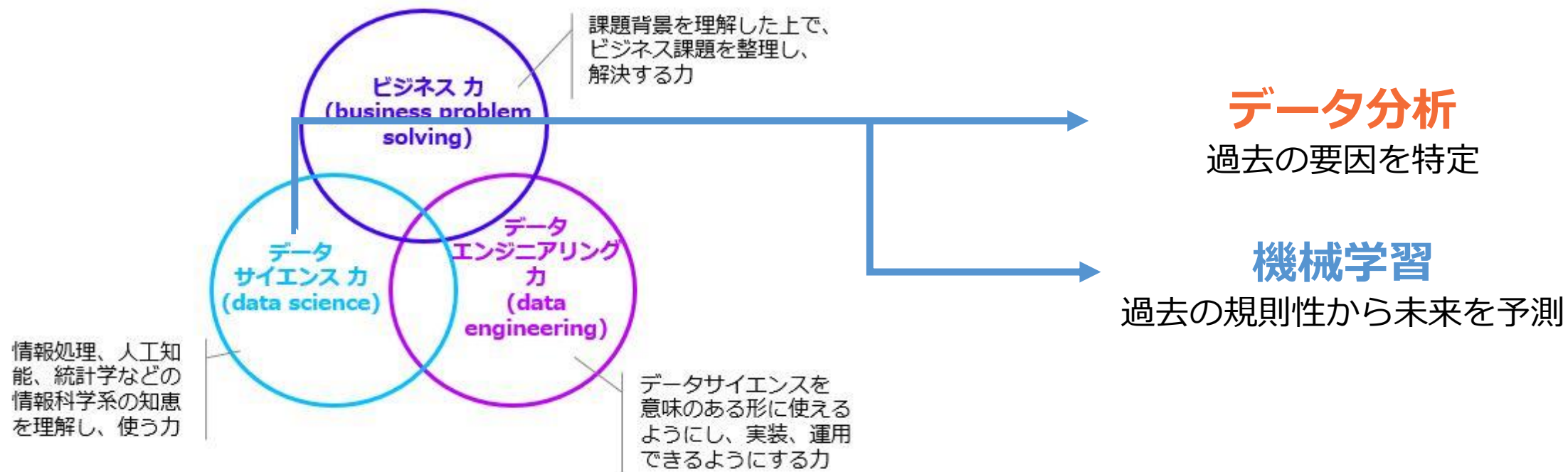
第1章 データ分析の概要

章の目次

- ✓ データ分析とは
- ✓ データ分析のフレームワーク（PPDAC）

データサイエンスとは

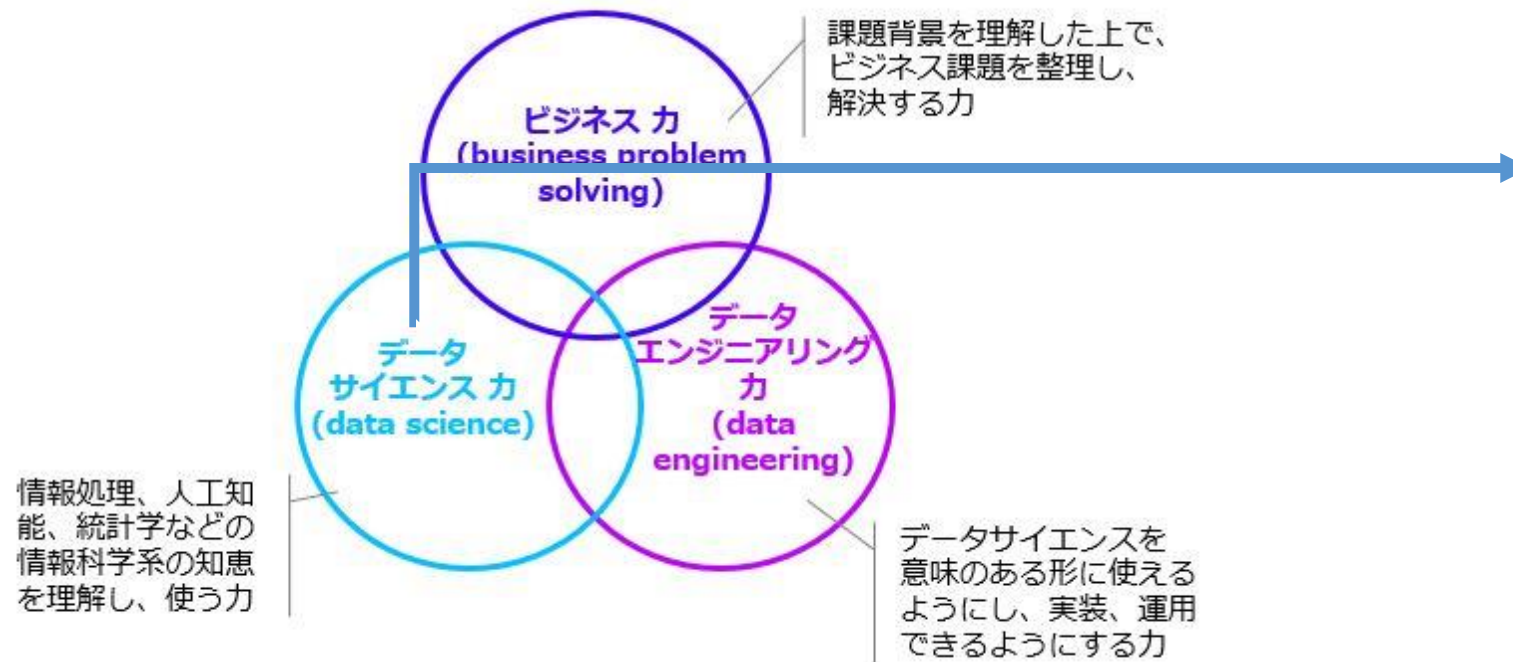
データサイエンティストに求められるスキルセット



出典：データサイエンティスト協会

データ分析とは

データサイエンティストに求められるスキルセット



出典：データサイエンティスト協会

データ分析
過去の要因を特定

問題の発見・課題の設定



問題の解決・課題の達成

データ分析を学ぶ背景

IT人材の不足規模に関する予測



- ✓ 少子高齢化社会による人口減少
- ✓ テクノロジーの発展によるニーズの拡大

IT やデータ活用が当たり前な時代

エンジニアでない人
にもデータ活用力が求められる

技術スキル

- ✓ データ収集 (SQL)
- ✓ Excel 操作
- ✓ 統計..など



データ分析を行う理由

成し遂げられる事

- ✓ 問題の原因を究明し、課題を達成する
- ✓ 自分の思考を裏付ける
- ✓ 統計的に、定量的に相手を説得する

メリット

- ✓ 仕事への**理解度が深まる**
- ✓ **視点が多角化**し 1 データから複数の情報が引き出せる
- ✓ **経験と勘**に磨きがかかる

第1章 データ分析の概要

章の目次

- ✓ データ分析とは
- ✓ データ分析のフレームワーク（PPDAC）

本研修のシナリオ

あなたの銀行では過去に取引のあった顧客に対して積立株式口座の電話による**販促キャンペーン**をおこなっています。

しかし営業マンは5名に限られており、残りのキャンペーン期間中にすべての顧客に架電することができません。

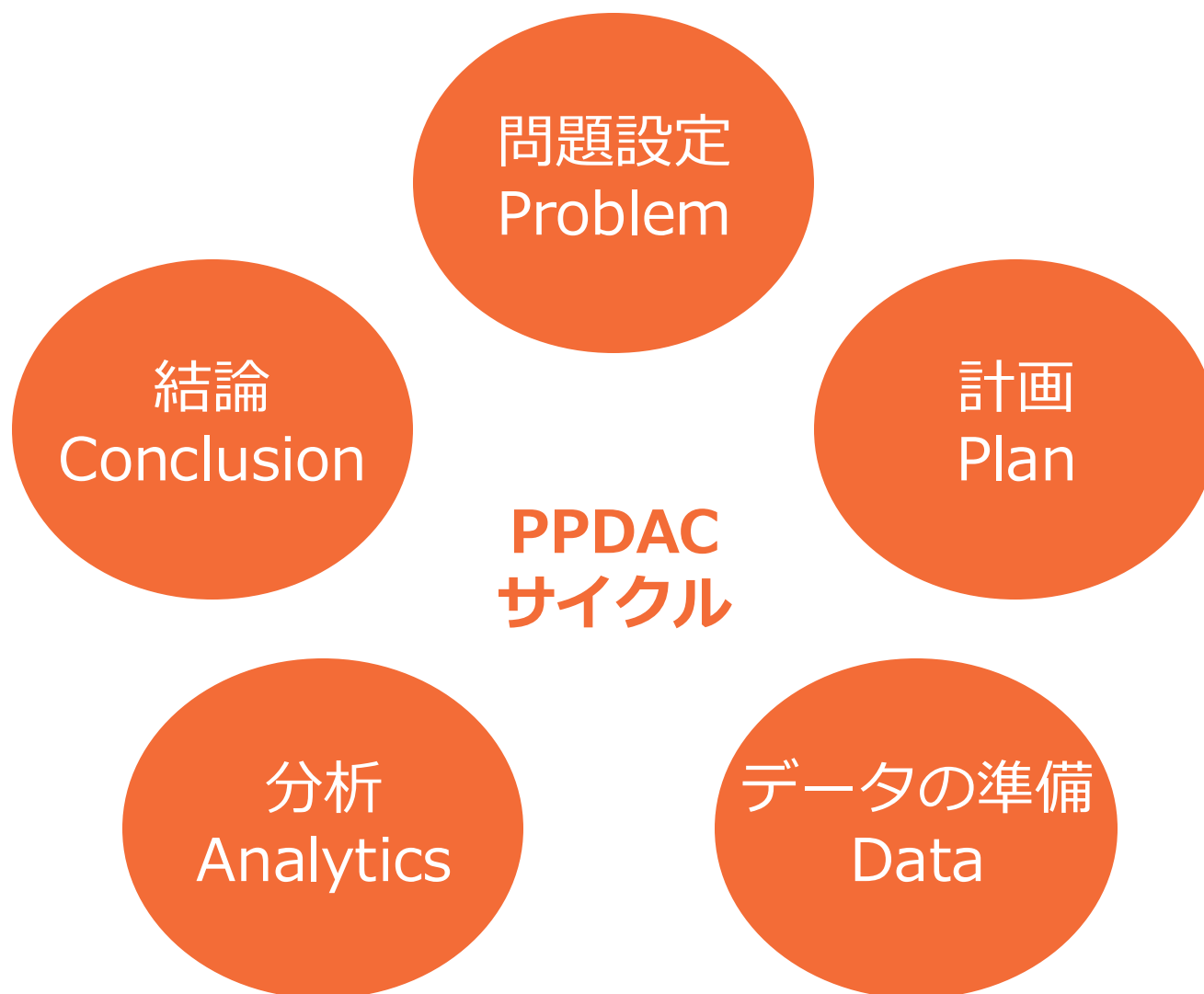
そこで、あなたはより効率よく営業ができるよう営業施策をしようと考えました。

手元には**直近 1 ヶ月分の販促キャンペーンのデータ**の一部があります。

これらのデータを活用し営業支援策を立案しましょう。



データ分析のフレームワーク（PPDAC）



問題設定

- ビジネスの理解
- 分析目標の設定

計画

- 二次データの収集
- データの理解
- 分析の指針の計画

データの準備

- 一次・二次データの収集
- データの加工

分析

- 統計仮説検定
- 統計分析
- 探索的データ分析

結論

- とりあえずの結論付け
- 課題達成

繰り返す行う

データ分析の簡略化フレームワーク

問題設定

- 分析目標の設定
- 仮説設定

データの準備

- 一次・二次データの収集
- データの前処理

分析

- 可視化分析
- 統計仮説検定
- 統計分析

結論

- とりあえずの結論付け
- 課題達成

課題設定→分析→結論の繰り返し

問題設定

- 成約率向上のための施策を立てる

データの準備

- 過去の売上データを集める

分析

- 可視化分析
- 集計分析

結論

- 〇〇な人が成約が高そう

問題設定

- どの列データと成約は関係がある？

データの準備

- 分析のためのデータ加工
- 不足データの収集

分析

- 統計仮説分析
- 統計分析

結論

- △△列と成約有無には関連がある

問題設定

- △△列と〇〇な人が成約し易いのではないかな？

データの準備

- 分析のためのデータ加工
- 不足データの収集

分析

- 統計仮説分析
- 可視化・集計

結論

- 〇〇な人が成約率が高い
- X〇な新たな仮説

データ分析の簡略化フレームワーク

例) あるホテルの満足度を向上させたい

性別	年齢	個人/ グループ	旅行の 目的	Wi-Fi設備	送迎の 便利さ	オンライン 予約の しやすさ	ホテルの 立地	食事の 満足度	滞在時の 快適さ	チェックイ ン・アウト の接客	その他の サービス	清潔さ	満足度
女性	50	グループ	学生	4	4	4	4	3	4	4	4	4	4
女性	35	個人	ビジネス	4	5	4	4	1	1	5	4	1	2
女性	22	グループ	その他	1	1	1	4	1	5	1	3	1	3

問題設定

- 満足度向上のための施策を立てる

データの準備

- 過去のアンケートデータを集める

分析

- 可視化分析
- 集計分析

結論

- Wifi 設備の満足度が高いと満足度も上がりそう

問題設定

- Wifi設備列データと満足度は関係がある？

データの準備

- 分析のためのデータ加工
- 不足データの収集

分析

- 統計仮説分析
- 統計分析

結論

- Wifi設備列と満足度には関連がある

問題設定

- Wifi設備を重要視している人はどのような人

データの準備

- 分析のためのデータ加工
- 不足データの収集

分析

- 統計仮説分析
- 可視化・集計

結論

- 年齢層の若い、グループ旅行の人がWi-Fi設備の評価が低い

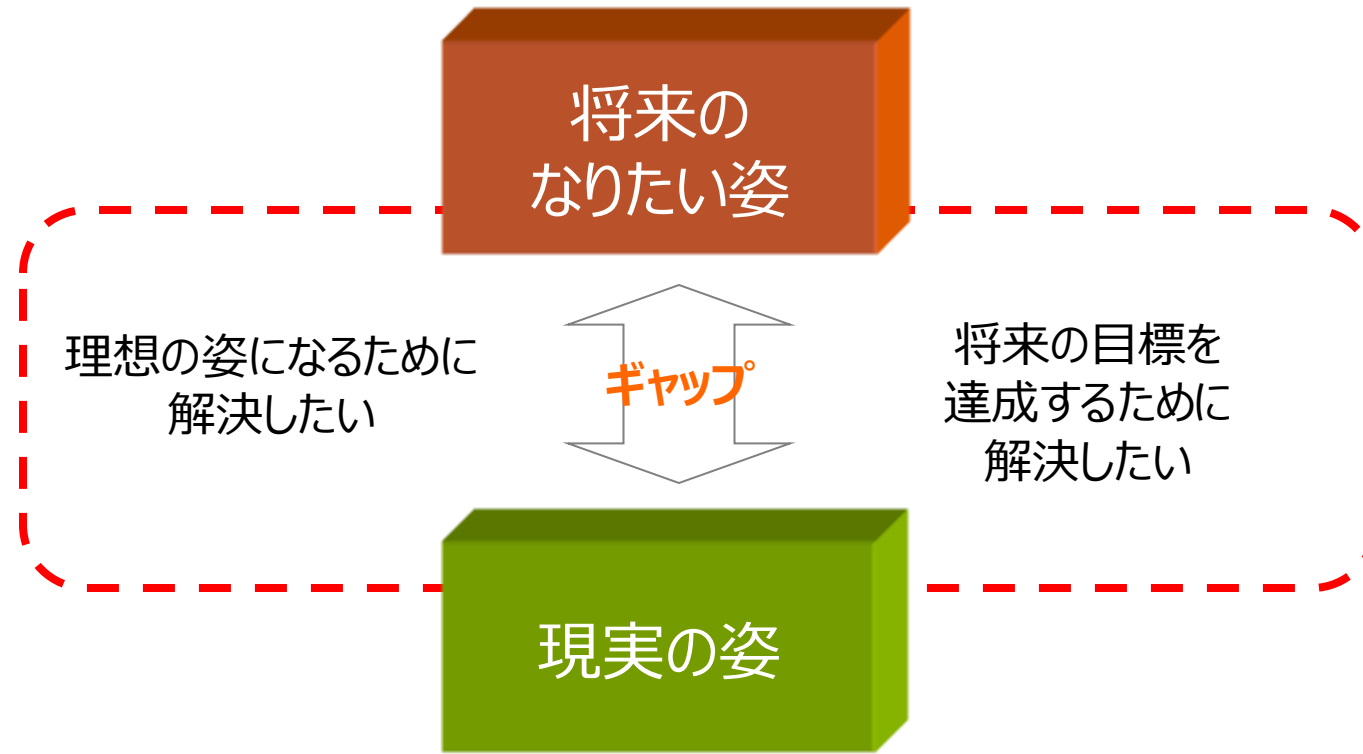
第1章 データ分析の概要

章の目次

- ✓ データ分析とは
- ✓ データ分析のフレームワーク（PPDAC）

問題とは

- 問題とは、「あるべき（Shoud）姿」と、「実際（Actual）の姿」とのギャップである



問題の設定



データを活用し
効率的な営業をしたい

キャンペーンによる売上最大化

イシューツリー

1 成約の売上を上げる

成約率を向上

↑
定期預金金額×10%

↑
成約数÷テレアポの回数

本課題での目的

成約しやすい顧客の特徴をみつけだし、
マネージャーに支援案を提案する

✓ どのような顧客にアプローチすると成約数を高められるか

第2章 データの収集・前処理

章の概要

この章では、データの収集と理解について紹介します。

データの集め方や特徴を理解し適切な情報の引き出し方を学びましょう。

この章の目標

- データ収集の流れがわかる
- データの違いがわかる
- データから情報を抽出できる



第 2 章 データの収集・前処理

章の目次

- ✓ データの収集と種類
- ✓ 探索的データ分析

データの収集



年齢	仕事	結婚状況	最終学歴	債務不履行の有無	銀行残高	持ち家の有無	ローンの有無	成約確度	担当営業マン
31	失業者	1	1	0	302	0	0	10	E
28	工場作業員	1	3	0	54	1	1	3	B
46	マネージャー	3	2	1	7331	0	0	7	A

* 数値の詳細はExcelの方に記載してます

データの種類

データは集めやすいデータ（すでに活用可能なデータ）から集める

一次データ

目的に沿って集められるデータ



- ✓ 観察
- ✓ 実験
- ✓ アンケート
- ✓ インタビュー

二次データ

他の目的によってすでに取得済みのデータ

外部データ

- ✓ 統計情報
- ✓ 新聞・雑誌・天気
- ✓ 民間調査会社

内部データ

- ✓ 顧客の購買履歴
- ✓ 製品情報
- ✓ POS データなど

データの種類

二次データ

他の目的によってすでに取得済みのデータ

外部データ

- ✓ 統計情報
- ✓ 新聞・雑誌・天気
- ✓ 民間調査会社

内部データ

- ✓ 顧客の購買履歴
- ✓ 製品情報
- ✓ POS データなど



- ✓ 分析の切り口にもなるため、属性データも含めた収集を行う。
- ✓ 不要な外れ値のデータは削除する。散布図や3σ法を活用する。

第 2 章 データの収集・前処理

章の目次

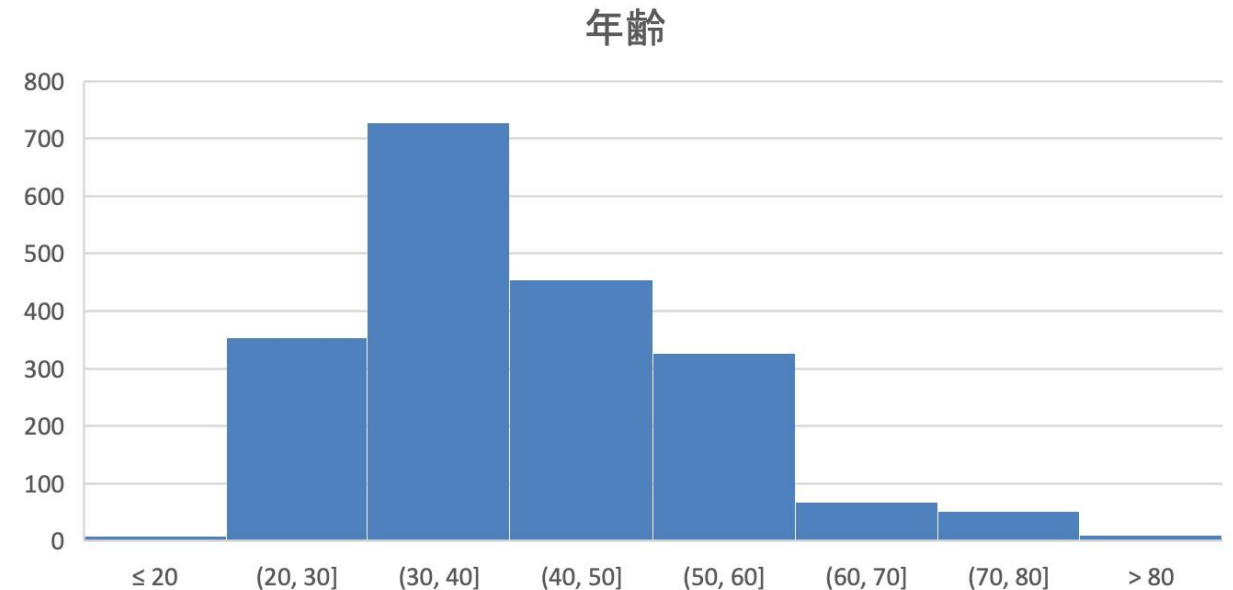
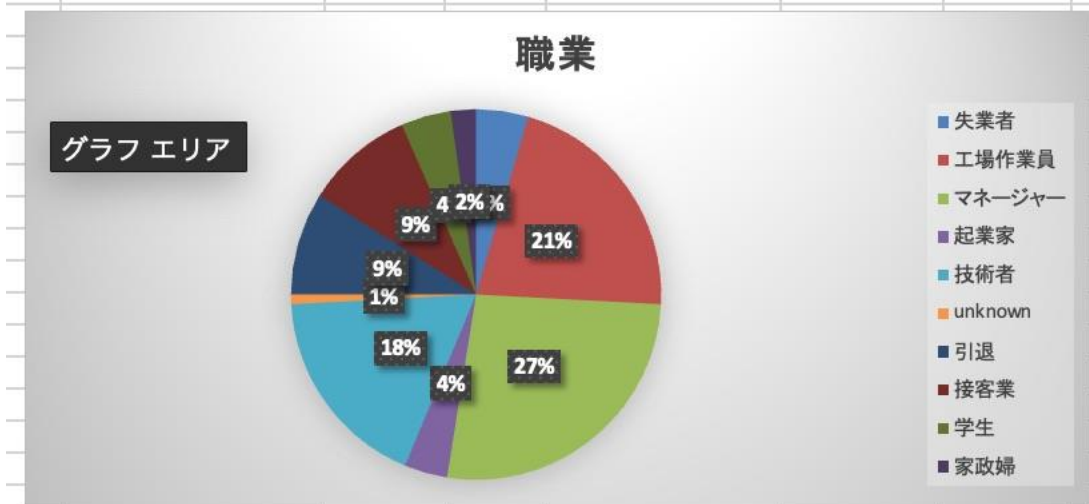
- ✓ データの収集と種類
- ✓ 探索的データ分析

データを理解する

- ✓ 可視化を行い、データの特徴を理解する
 - ✓ **散布図**を用いて外れ値を確認する
 - ✓ **円グラフ**で全体の割合を確認する
 - ✓ **ヒストグラム**でばらつきを確認する
 - ✓ **代表値**を確認する

代表値の確認

年齢の平均値	41.471
年齢の最頻値	31
年齢の中央値	39



実装：データの準備を行う（可視化）

実装資料：[3. データの収集・整理 — Sphinx ドキュメント \(d27upjc0zlhs3v.cloudfront.net\)](https://d27upjc0zlhs3v.cloudfront.net)

ワークシート：ワークシート_1_グラフ化

- ExcelのバージョンやOS の差異があるため資料と実装画面が異なる場合があります。
適宜お調べorお聞きください
- 学歴列の0は「未回答」「不明」のものが0とふられています

例)

- 成約と結婚、既婚者（3）がアポイントメント（7－1 0）が多い
- ローンや債務不履行があると成約にいたらない
- 持ち家の有無があると成約する数がおおい、0が少ない
- マネージャーとしても既婚者の成約が多い

■ などなど

- ・ 成約とローン以外の関係はないの？他の列はどのくらい成約に関連しているの？
- ・ この結果は全体に当てはまるの？

第3章 データに基づく分析

章の概要

この章では、データ分析の手法について代表的なものをご紹介します。

使い所と実装方法を抑えて実務でも活用できる用に引き出しを増やしましょう。

＊すべてのデータ対しこれから学ぶ手法が当てはまるわけではありません。特徴を抑えて適宜活用可能な状態にしましょう。



この章の目標

- 重回帰分析・相関分析の違いがわかる
- 統計仮説検定とはなにか、なぜ行うのか理解できる



第3章 データに基づく分析

章の目次

- ✓ 相関分析
- ✓ 重回帰分析
 - ✓ データの前処理
- ✓ 統計仮説検定
 - ✓ T 検定
 - ✓ カイ二乗検定

* 一部の手法では別データを使用し紹介します。

要因分析

ある事象が起こる要因を計算式で分析し、その事象に与える影響度を数値で表す
例) レストランの満足度は何が影響している？
食事の満足度、サービスの満足度、レストランの清潔度、スタッフの清潔度、、、

代表的な手法

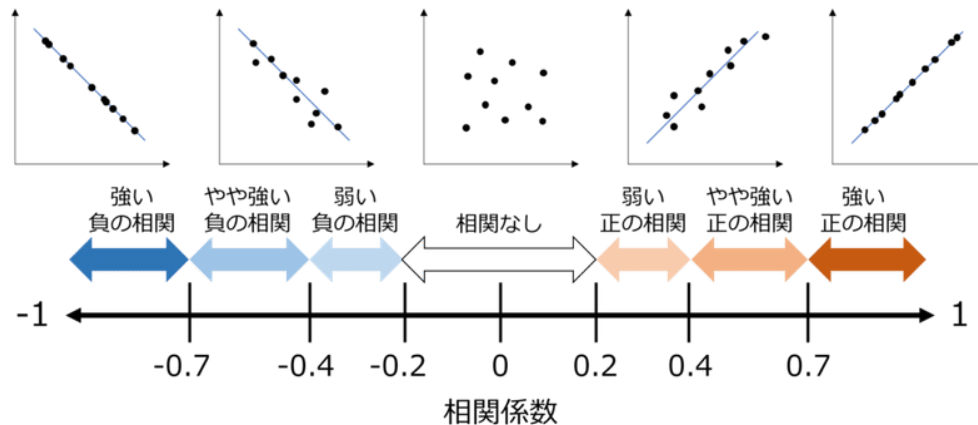
手法	説明
相関分析	ある変数が上がるともう一方もあがるような相関関係を見つけだすときに使用
重回帰分析	売上や家賃など連続値の要因分析に使用。
ロジスティック回帰	成約の有無や5択の満足度アンケートなどカテゴリの要因分析に使用

数値データの分析

相関分析

1つのデータに対し、もう一方の変数の動きを分析すること。（1対1の関係）
変数同士に相関関係があるかどうかを確認できる。

例）喫煙率と肺がん疾患率・営業の訪問数と受注件数



✓ **Positive Correlation : 正の相関**
一方の変数が増えともう一方の変数が増える関係

✓ **Negative Correlation : 負の相関**
一方の変数が増えともう一方の変数が減る関係

出典：[相関関係の見方 | tom \(note.com\)](https://tom(note.com))

相関分析の注意

因果関係

一つの事象（原因）が他の事象（結果）を直接的または間接的に引き起こす関係。この関係により、原因がなければ結果が発生しないという依存性が生まれます。

例）ウォーキングの時間と体重の増加には相関関係があることが知られている。因果関係は？



第三の要因を媒介した相関が出てしまう例

ウォーキングする

お腹がすいて
間食をする

体重が増える

相関関係の結果

相関係数：相関関係の強さを数値化したもの。1に近づくほど強い相関がある

	年齢	結婚状況	最終学歴	務不履行の有	銀行残高	持ち家の有無	ローンの有無	成約確度
年齢	1							
結婚状況	0.20803784	1						
最終学歴	-0.2350093	-0.0958138	1					
債務不履行の有	0.00901999	-0.0603545	-0.0258823	1				
銀行残高	0.09380156	0.01164697	0.03183934	-0.041339	1			
持ち家の有無	-0.2238899	-0.0485949	-0.0845817	-0.0221365	-0.035422	1		
ローンの有無	-0.0470965	0.00449499	-0.0171739	0.004585	-0.0692417	0.08358944	1	
成約確度	0.34930629	0.60162954	0.12433084	-0.1579281	0.24628502	0.33897082	-0.2878897	1

結婚状況と成約確度には数値上、やや相関があることがわかった
→結婚状況が既婚のほうが成約確度が高い

データの種類

データとデータから新しい情報を得る：データの種類

統計データの種類		
データの種類	例	説明
質データ	性別・電話の種類	ほかと区別するものの 順序に意味がない
	ランキング	ほかと区別するものの 順序・大小に意味がある
量データ	年齢・売上	順序・大小・値の差に意味 がある。 0を起点とする

探索的データ分析

統計や分析手法を使用せずにデータごとの特徴を明らかにする

データの比較方法		
	質的データ	量的データ
質的データ	クロス集計 代表値の比較	平均値の比較
量的データ	平均値の比較	相関係数

第3章 データに基づく分析

章の目次

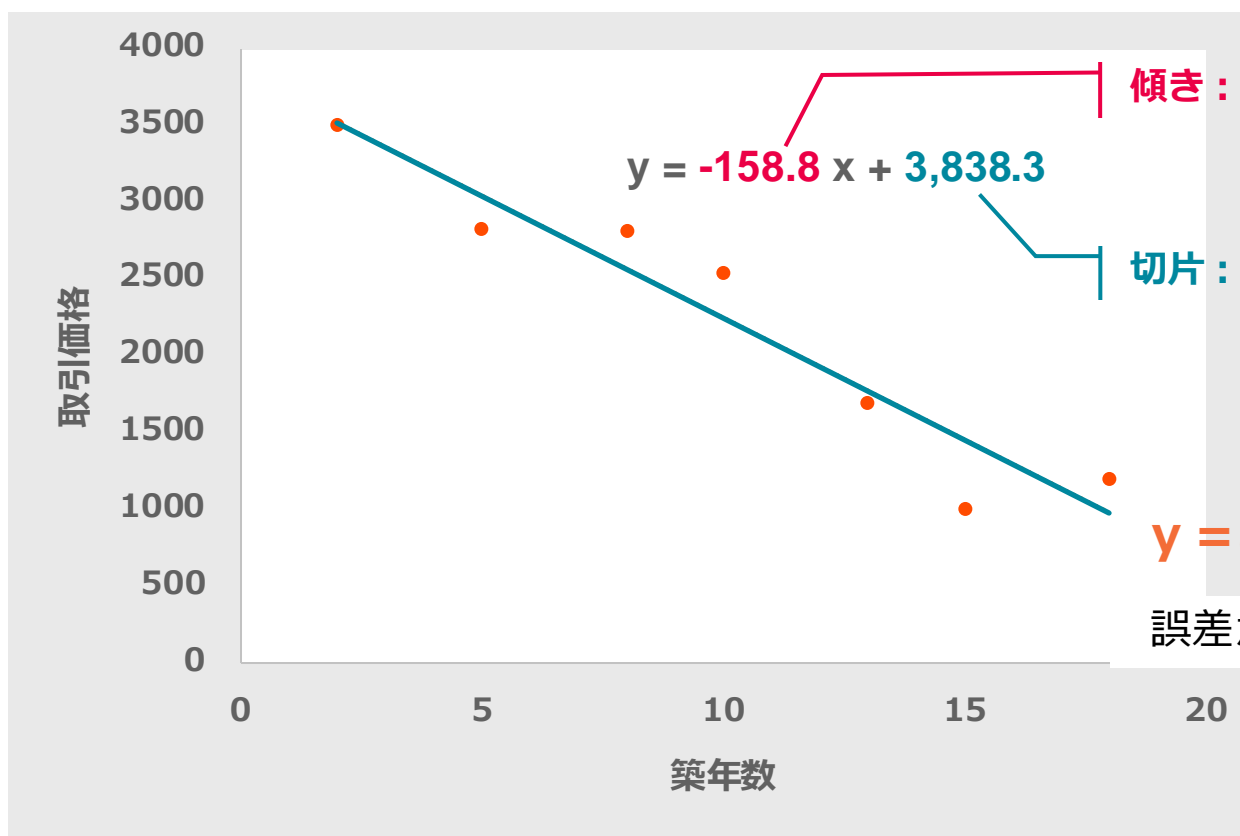
- ✓ 相関分析
- ✓ 重回帰分析
 - ✓ データの前処理
- ✓ 統計仮説検定
 - ✓ T 検定
 - ✓ カイ二乗検定

* 一部の手法では別データを使用し紹介します。

単回帰分析

2つの変数を目的変数（結果）と説明変数（原因）で考え、回帰直線を求める分析

例）築年数を説明変数にとって線形回帰を行うと下のグラフのように、線形的に特徴をとらえた方程式を作成できます。



傾き：築年数が1年増えるごとにどう価格が変化するか

切片：調整項

$$y = ax + b$$

誤差が最小化するように a の値を調節しながら直線を求める

重回帰分析

複数の変数を目的変数（結果）と複数の説明変数（原因）で考え、回帰直線を求める分析
 例）ある家の家賃（結果）を家の広さ、駅からの距離、築年数・・・で分析する
 ＊Excelの場合、入力変数は最大 16 列まで含めることが可能。

売上の分析

年齢	仕事	結婚状況	最終学歴	債務不履行の有無	銀行残高	持ち家の有無	ローンの有無	担当営業マン	成約確度
31	失業者	1	1	0	302	0	0	E	10
28	工場作業員	1	3	0	54	1	1	B	3
46	マネージャー	3	2	1	7331	0	0	A	7

説明変数

目的変数

$$y = a1 \times \text{年齢} + a2 \times \text{仕事} + a3 \times \text{結婚状況} \cdots + b$$

Y（成約確度）に影響度の高い a（要因）を調べることが出来る

＊文字列→数値化が必要、入力変数も連続値のほうが○

代表的なデータ前処理の手法

欠損値(欠落している情報)

- ・ 平均値や中央値などに置き換え
- ・ エントリ自体の削除

次元削減

相関係数の低い変数などノイズになりやすい情報の削除または変数の集約

来訪者	照会物件 (万円)	年収 (万円)	自己資金 (万円)	電話番号	成約
ID1	3,260	800		090-1111-2222	Yes
ID2	2,850	650	500	03-3333-4444	Yes
ID3	192,500	6,000	2,000	06-4444-5555	No
...

特徴量エンジニアリング

機械学習モデルに入力される特徴量（データの属性や情報）を適切に設計、変換、選択するプロセス。

例)

りんご、ばなな、みかん→一般果物
マンゴー、グアバ→南国果物

外れ値(規格外の値)

外れ値の範囲を定義したうえで、データの削除や値を変換

ダミー変数

文字列カテゴリカルデータを数字へ変換
例: Yes=0, No=1

第3章 データに基づく分析

章の目次

- ✓ 相関分析
- ✓ 重回帰分析
 - ✓ データの前処理
- ✓ 統計仮説検定
 - ✓ T 検定
 - ✓ カイ二乗検定

* 一部の手法では別データを使用し紹介します。

実装：データの分析を行う

実装資料：[4. データに基づく分析・処理（分析） — Sphinx ドキュメント \(d27upjc0zlhs3v.cloudfront.net\)](https://d27upjc0zlhs3v.cloudfront.net)
ワークシート：ワークシート_2_相関分析
 ワークシート_3_重回帰分析

第3章 データに基づく分析

章の目次

- ✓ 相関分析
- ✓ 重回帰分析
 - ✓ データの前処理
- ✓ 統計仮説検定
 - ✓ T 検定
 - ✓ カイ二乗検定

* 一部の手法では別データを使用し紹介します。

統計仮説検定

有意差があるかどうか判断をすること

事例

薬Aを開発しました。この新薬が病気Bに対して有効であるかを確認するために、実験を行います。実験の対象は100人で、50人には新薬Aを投与し、残りの50人には従来の治療法を行います。

- ✓ 従来の治療法グループ：50人中10人が病気Bから回復
- ✓ 新薬Aグループ：50人中20人が病気Bから回復



この差は意味のある差なのか、たまたま出ている差なのか

仮説検定を用いて意味のある差かどうか判断を行う

検定の流れ

1. 例) じゃんけん

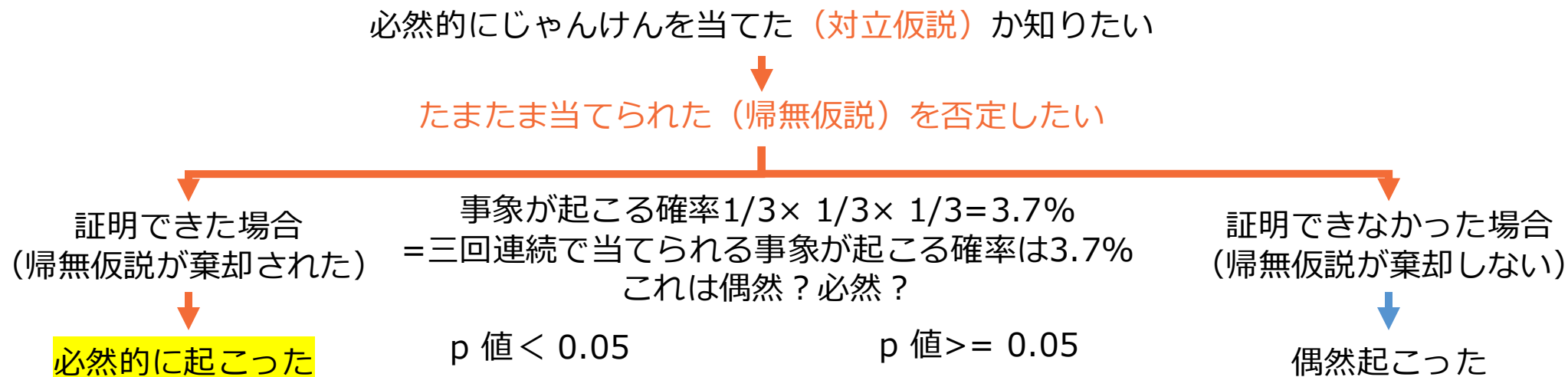
- ・じゃんけんで相手が出手を三回連続で当てた！！

2. 仮説を考える

- ✓ 何らかの要因（未来が見える・相手のクセがわかる）によってあてられた！！→対立仮説
- ✓ たまたま当てられただけ→帰無仮説

3. 検証する

じゃんけんで相手の出手を当てられる確率はどのくらい



→3.7 %しか起こらない事象がたまたま起こるわけない！

検定の流れ

1. データを確認した気づき

- ✓ 既婚の顧客は確度が高い傾向にある
- ✓ 最終学歴が中等教育だと成約確度が低い。学歴が上がるほど成約率が上がっている。
- ✓ 例) キャンペーン前と後で売上に差がある

2. 仮説を考える

- ✓ 結婚状況と成約の有無には差がある？→優位な差があるか調べる
- ✓ 例) キャンペーン前と後で差がある？→優位な差があるか調べる

3. 検証する

キャンペーン前とキャンペーン後の売上の平均に差がある（対立仮説）か知りたい

↓
キャンペーン前とキャンペーン後の売上の平均が同じ（帰無仮説）であることを否定したい

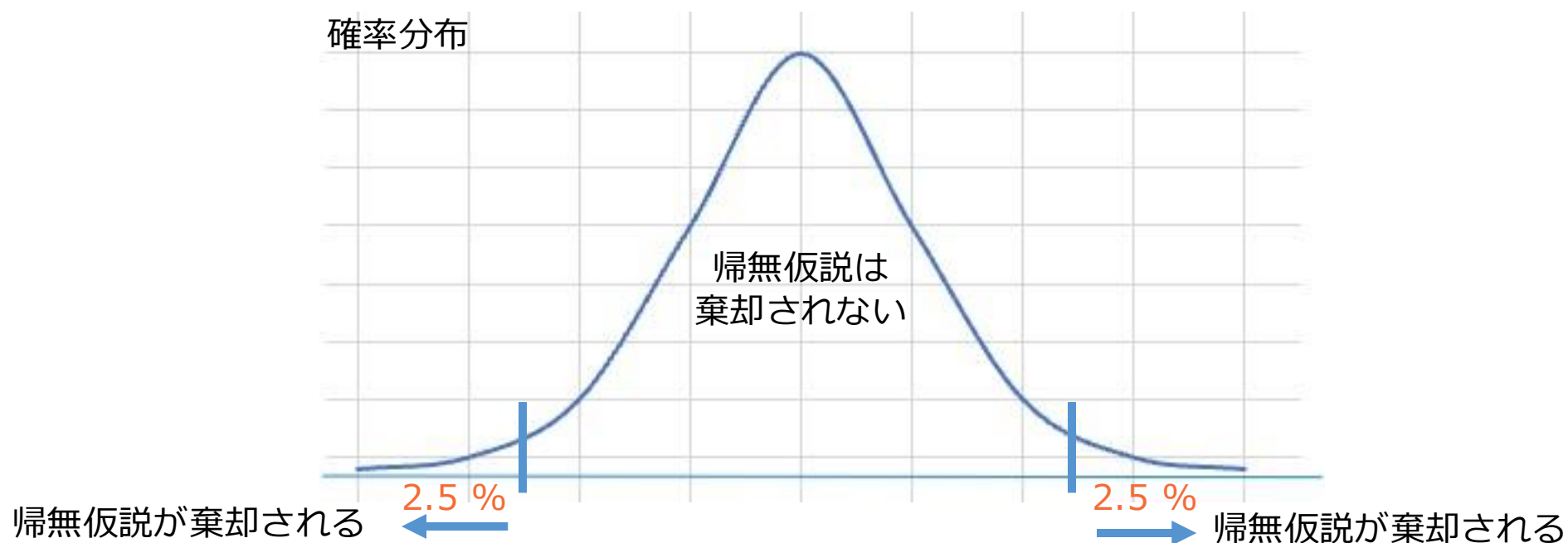


結果の解釈

結果の解釈 : p 値を確認

p 値: p 値は、帰無仮説が正しいという条件の下、今回の統計量の値、またはそれ以上に極端な値が観測される確率のことです。

p 値が 5% の場合、100 回に 5 回しか、平均が同じにならない事 (t 検定の場合) となり、平均に差があることは偶然ではなさそうということが出来ます。(有意な差)
この偶然を判断する閾値 (有意水準) は 5% を用いることが一般的



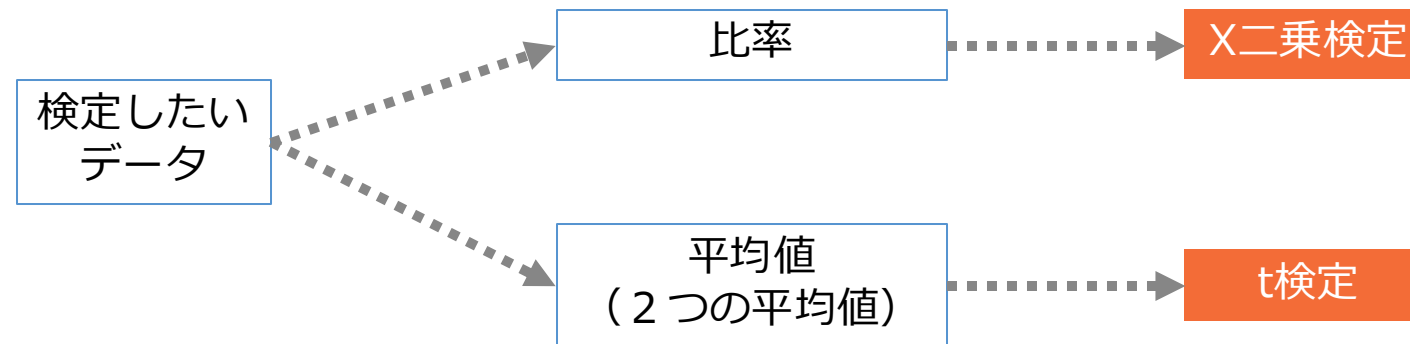
代表的な検定の手法

✓ **t 検定**：平均値を対象とした検定手法

【事例】 クラスAの平均スコアとクラスBの平均スコアの比較

✓ **X二乗検定**：割合や分布を対象とした検定手法

【事例】 キャンペーンXの購入者数と非購入者数、キャンペーンYの購入者数と非購入者数の比較



t 検定

t検定 (t-test) は、2つのグループの平均値が統計的に有意に異なるかを評価する方法です。
例えば、ある小売店のキャンペーン効果の比較をする場合に使います。
平均値の差を考慮し、差が偶然ではなく有意なものかどうかを判断します。

キャンペーン前	売上 (万円)
A 店	19
B 店	15
C 店	17
D 店	19
平均	17.5

キャンペーン後	売上 (万円)
A 店	18
B 店	20
C 店	22
D 店	21
平均	20.25



ランダムに 4 サンプル比較したが、
本当にキャンペーンには効果があったの？

t 検定

t 検定の使い分け

個々の変数に対応がない
(別々の標本)



分散 (ばらつき) に違いがあるか
(F検定)

等分散



不等分散



t 検定: **等分散**を仮定した
2 標本による検定

t 検定: **分散が等しくない**と
仮定した 2 標本による検定

個々の変数に対応がある
(1 つの標本の反復測定)



t 検定: **一対の標本**による平均の検定

対応がある: キャンペーン前後の効果を同じ店舗で検証

対応がない: 分散が等しくない: キャンペーン実施店舗と未実施店舗の売上の検証

等分散: 同じ規模の 2 店舗の売上データの検証

第4章 データに基づく分析

章の目次

- ✓ 相関分析
- ✓ 重回帰分析
 - ✓ データの前処理
- ✓ 統計仮説検定
 - ✓ T 検定
 - ✓ カイ二乗検定

* 一部の手法では別データを使用し紹介します。

χ 二乗検定

χ 二乗検定は独立性の検定とも呼ばれ、A と B （2 つの変数）
には関連があるかないかを調べるために使用します

	感染あり	感染なし	合計
男性	180	70	250
女性	120	30	150
合計	300	100	400



ワクチン効果には性別による違いがあるでしょうか？

医療用のワクチンを開発し、その後の結果から
男性と女性で効果が異なるような考察をえました。

χ^2 乗検定の実装

ピボットテーブル

	感染あり	感染なし	合計
男性	180	70	250
女性	120	30	150
合計	300	100	400

期待値

例) $300 \times 250 \div 400 = 187.5$ (期待値)

	感染あり	感染なし	合計
男性	187.5	62.5	250
女性	112.5	37.5	150
合計	300	100	400

実際の値と期待値を比較し検定を行う

- ✓ P値が 5 % 未満 : 男女に差があるといえる
- ✓ P値が 5 % 以上 : 男女に差があるかどうかわからない

(補足) カイ二乗値の求め方

感染あり	感染なし	合計	
男性	180	70	250
女性	120	30	150
合計	300	100	400

感染あり	感染なし	合計	
男性	187.5	62.5	250
女性	112.5	37.5	150
合計	300	100	400

$$\text{カイ二乗値} = \frac{(\text{元データ} - \text{期待度数})^2}{\text{期待度数}} = \text{CHISQ.TEST}(\text{実測度数}, \text{期待度数})$$

実装：統計仮説検定を行う

実装資料：4-2データに基づく分析・処理（統計仮説検定）

ワークシート：ワークシート_4_t検定

ワークシート_5_x二乗検定

15：15まで実装

分析・検定によって得られた結果のまとめ

分析・検定の手法を用いることによりデータから新たな気づきを得ることができました。

【注意点】

- ✓ Excel 上で使用できる分析手法には限りがある
- ✓ データはすべて数値化する必要がある
- ✓ 分析における入力変数の上限は 16

相関分析

- ✓ 銀行残高と年齢に相関はない、結婚状況と成約確度には相関がありそうだが数値変換が必要

重回帰分析

- ✓ 持ち家の有無・結婚状況のT値が高く成約確度に大きな関連がある、他の変数もP値から成約確度に関連している

検定

- ✓ 成約の確度と結婚状況には統計的に関連がある

第5章 分析結果の考察

章の概要

この章では、分析結果の可視化・考察ツールとして Power BI について紹介します。

一般的にはデータの可視化のツールとして利用されますが、データ分析においても広く使用することができます。



この章の目標

- Power BI と Excel の違いがわかる
- PowerBI を持ちてグラフの作成ができる
- Power BI を用いて分析結果の考察ができる

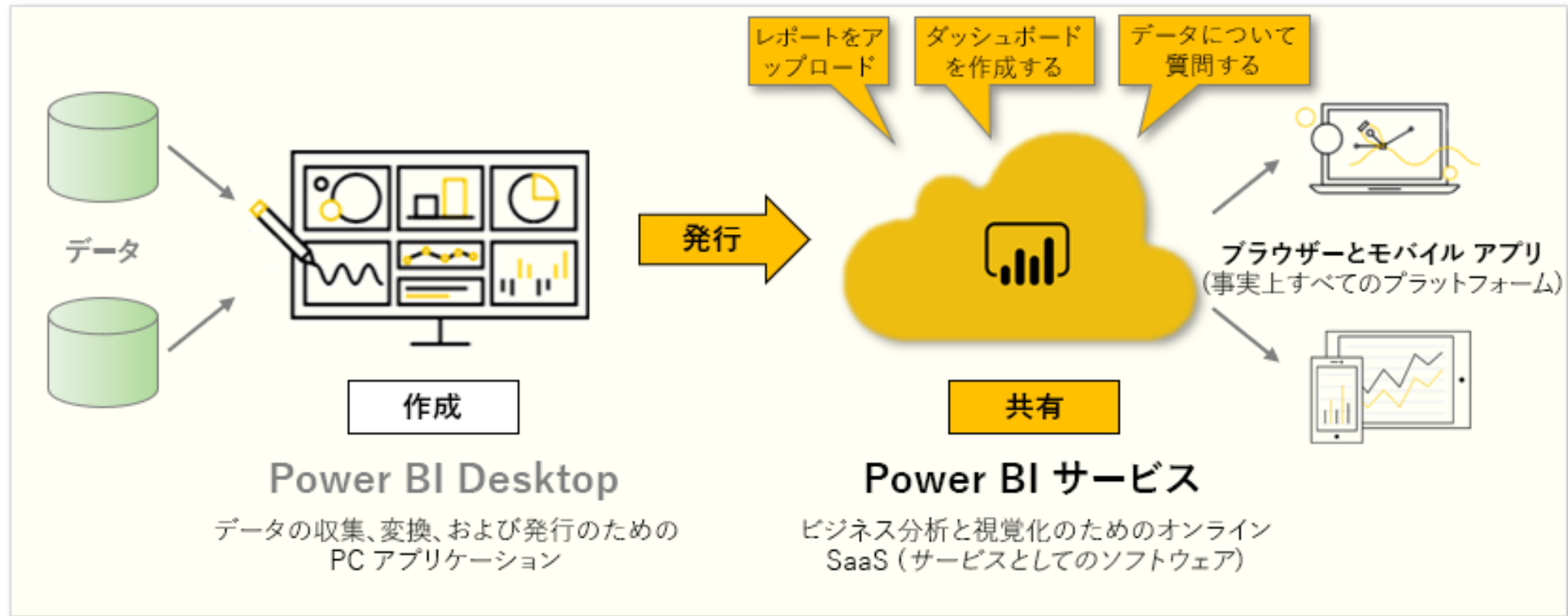


第4章 分析結果の考察【Power BI】

章の目次

- ✓ 結果の共有・可視化（Power BI）
- ✓ Power BI の実装

PowerBI



出典 : [Power BI とは? - Power BI | Microsoft Learn](#)

データの可視化・分析・共有等を大規模に実施する SaaS サービス

Excel と PowerBI

特徴	Excel	PowerBI
主な用途	小規模なデータ処理や個人の分析	大規模なデータやビジネス分析
データの規模	小規模なデータ処理に適している	大規模なデータ処理に適している
柔軟性	柔軟性が高く、さまざまなデータ処理や分析が可能	データの視覚化や分析に特化し、リアルタイムでの分析が可能
自動化の容易性	関数やマクロを使って自動化することができる	リアルタイムデータ分析や自動化が可能
データソースの統合	一つのファイルにデータが含まれるため、統合が容易	複数のデータソースからデータを統合し、関係性を可視化
リアルタイム分析	不可能	リアルタイムでのデータ分析が可能

PowerBI の実装

あるホテルのレビューデータを可視化と分析を行う

性別	年齢	個人/ グループ	旅行の 目的	Wi-Fi設備	送迎の 便利さ	オンライン 予約の しやすさ	ホテルの 立地	食事の 満足度	滞在時の 快適さ	チェックイ ン・アウト の接客	その他の サービス	清潔さ	満足度
女性	50	グループ	学生	4	4	4	4	3	4	4	4	4	4
女性	35	個人	ビジネス	4	5	4	4	1	1	5	4	1	2
女性	22	グループ	その他	1	1	1	4	1	5	1	3	1	3

目標

- PowerBI で何ができるか説明出来るようになる
- PowerBI で基本的な可視化が出来るようになる
- PowerBI とExcelの違いが説明できるようになる

【補足】 C(conclusion、結論) : 分析結果の考察、結論 — Sphinx ドキュメント (d27upjc0zlhs3v.cloudfront.net)

第 6 章 演習

第6章 実習

章の目次

- ✓ 結果による気づき
- ✓ 実習案内
- ✓ 結果の共有
- ✓ FB

データの可視化による気付き

単一のデータの可視化による気づき

- ✓ リストの平均年齢41歳
- ✓ リストの平均残高1790ドル
- ✓ 顧客のボリュームゾーンは 40 代 など

複数データの集計・可視化による気づき

- ✓ 「工場作業員」や「接客」の仕事に就いている顧客は、断られることが多い
- ✓ 既婚の顧客は成約率が低い傾向にあるがデータ数は一番多い。
- ✓ 最も成約が多いのはマネージャー など

分析・検定

- ✓ 結婚状況と成約確度には関連性がある
- ✓ 持ち家の有無と成約確度が関連している

演習

ここまでデータ全体の分析を行いさまざまなことがわかりました。
具体的な施策を立てるうえでさらなるデータ分析をおこない施策の立案を行いましょう。
また、時間の余った人は検定を用いて今回のキャンペーンに意味があったのかどうかを説明しましょう。

イメージ)

- 1.結婚状況と成約確度が関連していることがわかった。→実装完了
2. 結婚状況があがると成約確度はどう変化するの？
3. 仮) 結婚状況が3のどのような人が成約しやすいだろうか？
結婚状況が3になるには何の関係している？
・・・→ 施策の立案

施策のイメージ)

結論：既婚者で〇〇職の〇〇の特徴の成約率が高い（統計的に証明できると〇）

施策1・上記に当てはまる人に優先的に営業をする

施策2・〇〇向けの新たなキャンペーンを作成・案内する

タイムスケジュール（講師より案内があります。）

	分析の道筋だて
17:20 まで 成果物をZoomで提出	データ分析

結論

最終的な施策案となぜその施策が有効であるかの証明を下記におこなってください。

持ち家と成約確度にも統計的に優位性がある。（分析→結論）
結婚状況 3 と持ち家有りのどんな人が成約しやすいんだろう（仮説）
年代列の作成（データの収集）
可視化（分析）

結論
40代-60代の既婚者で持ち家のある人（データの半数）は職種に問わず成約確度が高いことがわかった
→

推察
・子供・孫に対して資金を残したい？
・老後の資金をためている？
→

施策
上記の条件の人に対して優先的に営業を行う
「子供・孫に対して資金」「老後の資金」をキーワードとした営業スクリプトを使用し、成約数の変化を図る
より最適な分析をするため、成約の目的のヒアリングを行いさらなるデータを収集する

回答例

- ✓ 既婚者の仕事を引退した人に対して預金の案内をすることで成約確度が上がる傾向がある
 - ✓ 結婚状況と成約確度の関連性は統計的に証明済み



施策

- ✓ Web マーケティングではなくチラシなど年配の方が把握し易い販促活動を検討
- ✓ 年配の既婚者を優先的にテレアポする。
- ✓ 相続関連のマニュアルを作成し案内する



- ✓ 実現場で施策を適用し効果検証を行いさらなる利益追求を行う

まとめ

問題設定
Problem

計画
Plan

結論
Conclusion

分析
Analytics

データの準備
Data

講座の目標

- Excel でのデータ分析の手順を理解し、自ら実装ができる
- データから課題を発見し、適切なデータ分析、仮説の検証ができる
- 適切な可視化を行うことができ、結果から次の施策を立てることができる