

データマーケティングの基礎①



マーケティングの重要性

学習目標

1. マーケティングの重要性を理解する
2. マーケティングにおける「ニーズ」、「ウォンツ」を理解する
3. マーケティングのプロセス、全体像を理解する

マーケティングの重要性

■ 観光客ニーズを例として、マーケティングの重要性を確認する

観光客ニーズとは、

観光・旅行を実施しようと思った「きっかけ、動機」

日常生活では何か不足していて、それを埋めたいと思う欲求のことです。

！ 観光客のニーズを探ることはマーケティングの第一歩です。
例を考えて、記入しましょう。

マーケティングの重要性

■ 「観光」分野でマーケティングする場合の定義

観光マーケティングとは、

自地域に呼びたい**観光客のニーズ・ウォンツを探り**、
他地域ではなく、自地域を訪問し、自地域のファンになって頂くための
活動全般を指します。

【どのような活動を想定するか？】

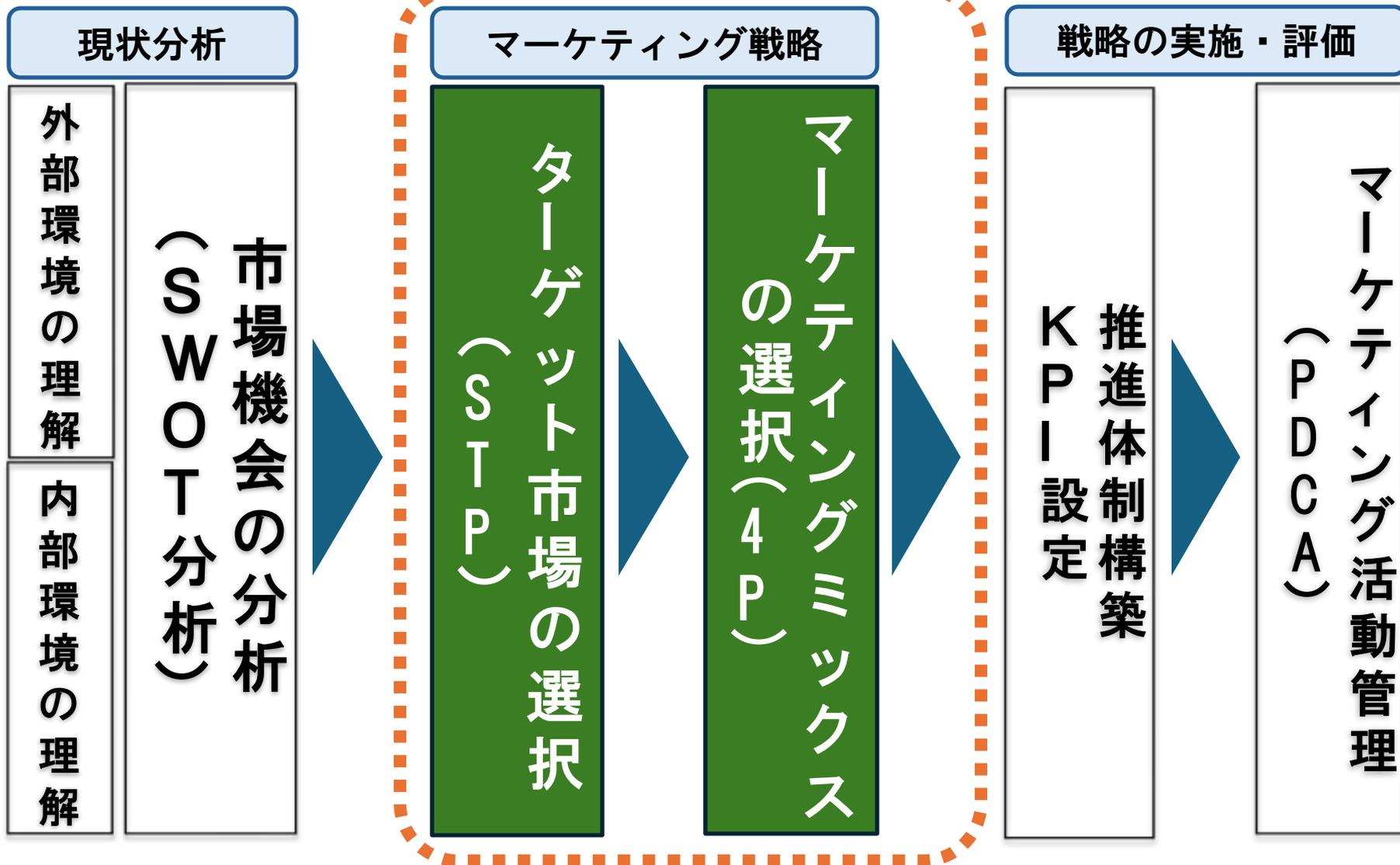
- ・ 顧客嗜好の調査・研究
- ・ ライバルとなる観光地域の特定・調査
- ・ 顧客ニーズを満たす観光資源の開発
- ・ 観光商品・サービス作り
- ・ 観光商品の販売方法の検討
- ・ 観光商品のPR、プロモーション など

マーケティングとは、

“ **誰に** × **何を** × **どうやって** ” 提供するか

マーケティングの重要性

プロセス



ニーズとウォンツ

ニーズ

人々が生活する上で感じる
現状に対する不満や欠乏
(具体的に何が欲しいかは
わからない状況)



仕事に追われて
ストレスが溜まっている。
リフレッシュしたいな
あ・・・

ウォンツ

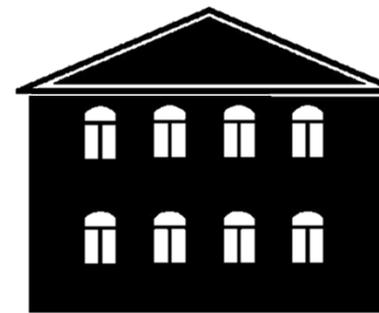
本質的ニーズを満たすための
具体的な手段への欲求
(具体的な商品・サービスが
欲しい状況)



休暇を取って
温泉に旅行に行きたい
美味しい食事も食べたい

購買行動

ウォンツに対して、実際に商品・サービスを選択し、購入する活動



〇〇温泉の〇〇旅館を
申し込もう

ニーズとウォンツ

顧客の購買行動には、顧客のニーズやウォンツが隠されています。
顧客の購買行動をよく観察することで、ニーズやウォンツを辿ることが可能となります。



ニーズ



購買動機

ウォンツ



購買目的

ニーズとウォンツ

顧客のニーズ、ウォンツは「心理的なもの」で目に見えません。
それでは、どのように確認すべきでしょうか？

1. 既にあるデータを確認する

【顧客ニーズ】

旅行動機に関する各種オープンデータ、調査結果を確認

【顧客ウォンツ】

旅行目的に関する各種オープンデータ、調査結果を確認

2. 直接お客様に聞いてみる

【顧客ニーズ】

今回旅行に行こうと思ったきっかけは何かをインタビューやアンケート調査で確認

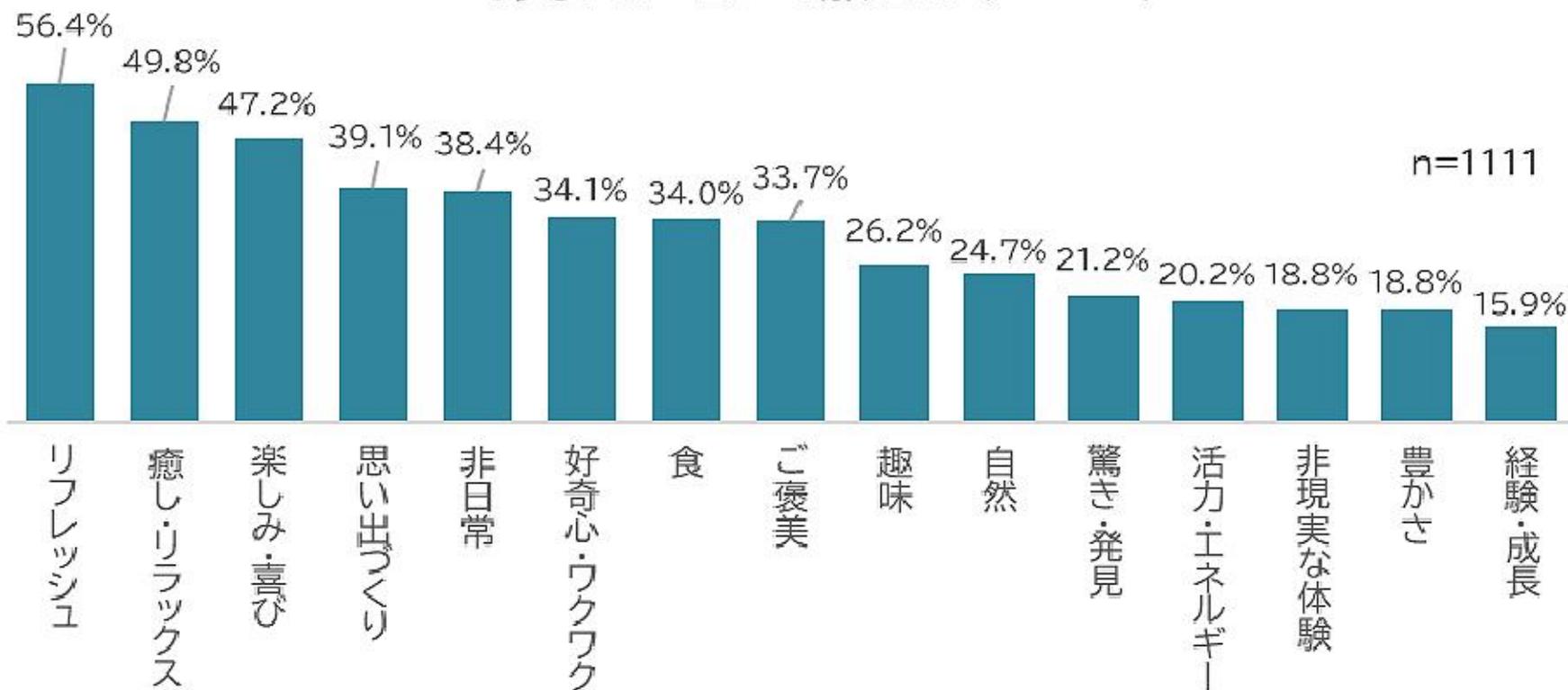
【顧客ウォンツ】

今回の旅の目的、やりたいことをインタビューやアンケート調査で確認

ニーズの確認

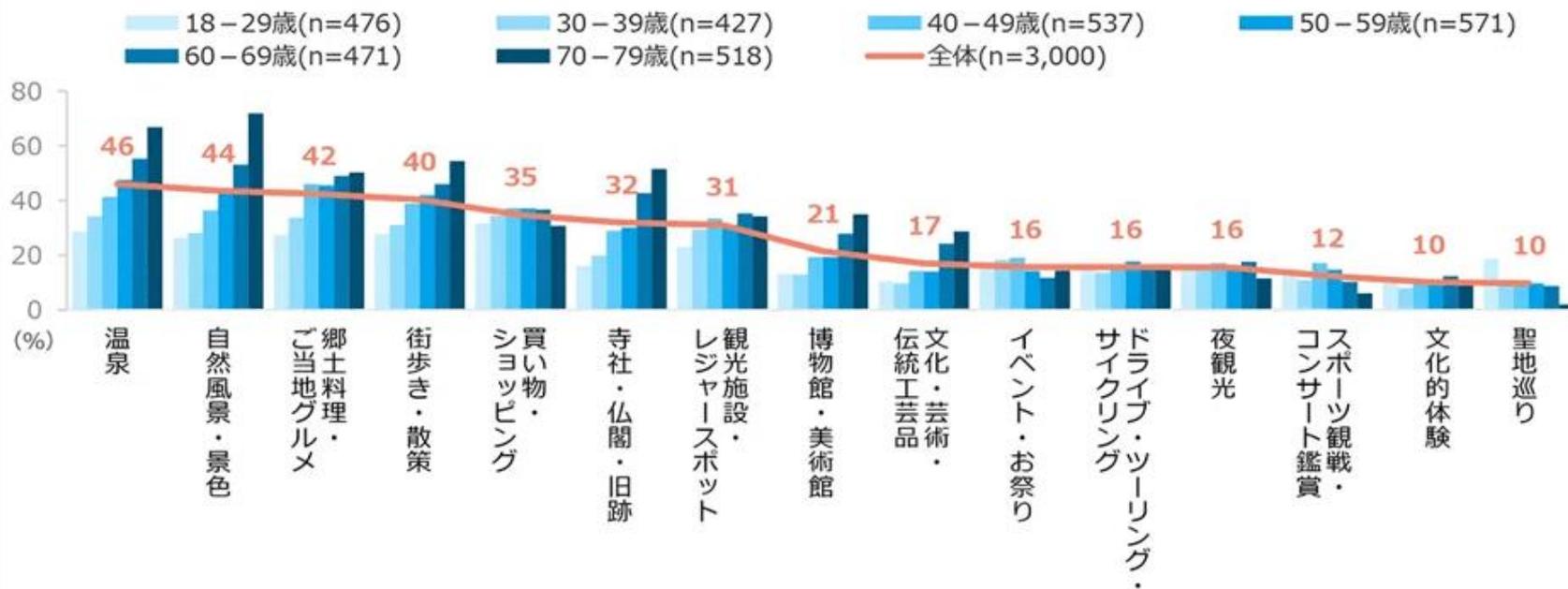
JTB総合研究所

あなたにとって旅とは(TOP15)



ウォンツの確認

＜図4＞ 旅行先で興味・関心があること（複数回答）※上位15項目を抜粋



＜図5＞ 今後訪れたい場所と理由（各3つまでの自由回答抜粋）

日帰り旅行・レジャー

- 温泉でのんびり体をリフレッシュしたい
箱根、熱海、草津、別府、有馬
- テーマパークで遊びたい
ディズニー、USJ、ハウステンボス、ジブリ
- 観光地で散策や食べ歩きをしたい
横浜中華街、鎌倉、軽井沢、日光

宿泊をともなう国内旅行

- 温泉地で普段の疲れをいやしたい
- 北海道で海鮮を堪能したい
- 沖縄で綺麗な星や海を見たい
- 京都で寺社・仏閣巡りをしたい
- 東京で美術館・博物館を見て回りたい
- ディズニーオフィシャルホテルに泊まりたい

海外旅行

- ハワイで泳いだり買い物をしたい
- 台湾で夜の屋台の食べ歩き
- イタリアで世界遺産に触れたい
- 韓国で食べ歩きやショッピングをしたい
- アメリカで野球観戦をしたい
- フランスで美術館巡りをしたい

マーケティングの力

マーケティングを理解すると商品を売ろうとしなくても、勝手に売れる仕組みを作ることが可能となります。

購買者心理と購買者行動をマーケティングの観点で分析することで、扱う商品・サービスの販売力が大きく向上します。

個人演習

個人演習①

皆さんの施設・お店によく来店するお客様を詳しく調べてみます。

よく来店するお客様をイメージして、

①年代

②性別

③居住地（県内在住または県外在住）

④同行者（誰と訪問するか）

といった基本属性を考えてみましょう。

良く来るお客様層ベスト5を次ページのシートに書いてみます。

個人演習②

年代	性別	居住地	同行者
(例) 30代	(例) 女性	(例) 県外 (東京都)	(例) 家族 (夫、子ども2人)

個人演習③

<顧客ニーズ／ウォンツの把握>

個人演習②で分けた顧客属性ごとの
「ニーズ」、「ウォンツ」を考えてみましょう。

個人演習④

顧客基本属性	ニーズ	ウォンツ
<p>(例) 県外在住30代女性子供有 家族旅行</p>	<p>(例) 家族みんなでリゾート気 分を味わいたい！</p>	<p>(例) 沖縄のビーチリゾートホテルに泊まりたい。</p>

観光データの活用方法

観光データ活用の意義

＜考えてみよう＞

山口県を観光で盛り上げたい！
もっと観光客に来てもらいたい！
国内No. 1の観光県にするぞ！

ところで山口県にそもそも年間何人の観光客が来ている
のだろう・・・

観光データ活用の意義

＜考えてみよう＞

山口県の経済を活性化したい！
もっと観光客に満足してもらい、
たくさんお金を使ってもらいたい！

でも、いったいいくら使ってもらえばいいのだろうか？

データ活用の意義

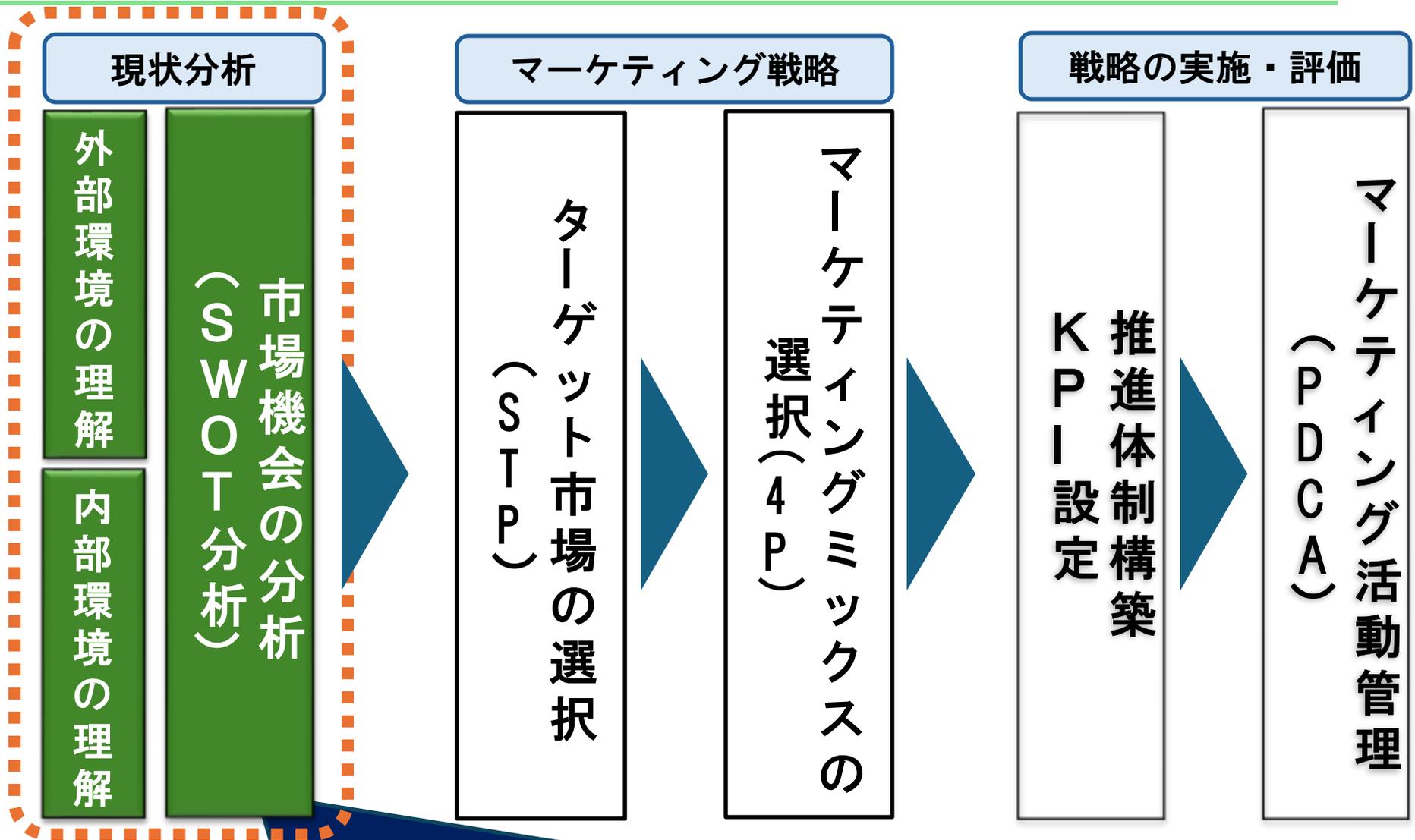
マーケティングスキルを活用してビジネスを推進するには、適切なマーケティングデータを活用することが大切です。



- ① なりたい姿、ビジョン、目標を明確にする
- ② その目標を実現するために、現状を正しく理解する
- ③ 未来、現状がわかったら、その間にある「やるべきこと = アクション項目 (課題) を明確にする

現状を正しく理解し、アクション項目、課題を明確にするには
「調査・データ」が必要

マーケティングデータの活用



マーケティング戦略を立案・実行するために“現状を分析”する必要がある
→ マーケティングデータを収集・分析し、活用する！

＜具体的な目標例＞

- 国内観光客を呼びたい
- ツアー以外の外国人観光客を呼びたい
- リピーターを現状の3倍にしたい
- 地域内を回遊してもらいたい
- 宿泊してもらいたい
- 観光体験で消費をしてもらいたい 等

現状は？課題は？ → 調査で確認

マーケティング調査とは

マーケティング調査とは、

「マーケティング活動を推進する上での現状や課題に対して、有効な意思決定をするための科学的かつ合理的な調査・分析」のことです。

マーケティング上の「現状」の例

- ・ どの国・どの地域から観光客が来ているのか？
- ・ どんな観光地を回っているのか？
- ・ 観光客は満足しているのか？

マーケティング上の「課題」の例

- ・ 自地域はどんな人をターゲットとすべきなのか？
- ・ 自地域と競合地域ではブランドイメージはどう違うのか？
- ・ 実施したプロモーションの効果はどの程度あったのか？

このようにマーケティング上の現状把握、課題の抽出は、事業活動全体に大きな影響を与えられるので、それらを解決する手助けとなるマーケティング調査は非常に重要なものです。

データの種類

調査で収集・活用するデータには2つの種類があります。

1. 1次データ(プライマリーデータ)

1次データとは、**調査実施者が自らその調査目的のために独自の方法で収集したデータ**のことをいいます。例えば、観光客アンケート調査や観光客へのインタビューなどがこれに該当します。また、現場へ出かけて、市場、対象者を観察し得られるデータも1次データに分類されます。

2. 2次データ(セカンダリーデータ)

2次データとは、その調査目的のためにゼロから独自の方法で収集したものではない、**既に収集されて存在するデータ**のことです。例えば、官公庁による統計データ、研究機関の調査レポートなどがあります。

【参考】ビッグデータ

デジタル化の更なる進展やネットワークの高度化、またスマートフォンやセンサー等IoT関連機器の小型化・低コスト化によるIoTの進展により、スマートフォン等を通じた位置情報や行動履歴、インターネットやテレビでの視聴・消費行動等に関する情報、また小型化したセンサー等から得られる膨大なデータをビッグデータといいます。

1. 1次データ

1次データのメリット、デメリットを確認します。

メリット

1. 調査実施者が欲しいデータを欲しい状態で入手できる

2次データでは、データの対象や取得方法、分類方法などが必ずしも調査実施者が求めるものと合わない場合があります。一方、1次データは調査実施者が求めているデータを収集することが出来ます。

2. 様々な調査手法を組み合わせて取得できる

必要なデータを独自の方法で収集するので様々な調査手法を組みまわせることができます。このことによりよい精度の高い、深みのあるデータを収集することが可能となります。

デメリット

1. データ収集にお金がかかる

2. データ収集に時間がかかる

3. 結論、仮説ありきの分析が行われる可能性がある

調査実施者が設計、実施、収集したデータなので、活用・分析には、調査実施者の考え、仮説などが影響しやすくなります。

アンケート調査票

大子町 アンケート調査

大子町にお越しいただきましてありがとうございます。以下のアンケートにご協力をお願いいたします。
チェック及び回答のご記入をよろしくお願いいたします。

Q1 居住地 _____ 都道府県 _____ 市町村 _____

Q2 性別 男性 女性

Q3 年齢 19歳以下 20歳～29歳 30歳～39歳 40歳～49歳 50歳～59歳
60歳～69歳 70歳以上

Q4 今回の旅行の目的は何ですか

観光 ビジネス 帰省 知人宅等訪問 その他 (_____)

Q5 大子町へは今回を含め、何回来たことがありますか

初めて 1回 2回 3回 4回 5回以上

Q6 大子町にはどのくらいの期間滞在する予定ですか

1日 2日 3日 3日以上 1週間 1ヶ月 その他 (_____)

Q7 宿泊施設について教えてください。(複数回答可)

ホテル 旅館 民宿 公共の宿 キャンプ 知人友人宅等 その他 (_____)

Q8 大子町まで何で来ましたか。

車 バス 電車 レンタカー その他 (_____)

Q9 今回の旅行の同伴者について教えてください

1人 夫婦 家族 同僚 友人 団体旅行 その他 (_____)

Q10 今回の旅行はどのような方法で手配しましたか

ツアー利用 個別手配

Q11 大子町でどのような観光地を訪問しましたか(複数回答可)

袋田の滝 八溝山 男体山 湯沢峡 奥久慈茶の里公園 大子ふれあい牧場 大子広域公園

奥久慈憩いの森 道の駅奥久慈だいが 大子おやき学校 オートキャンプ場グリーンヴィア

永源寺 クライミングウォールやみぞ 月待の滝 旧上岡小学校

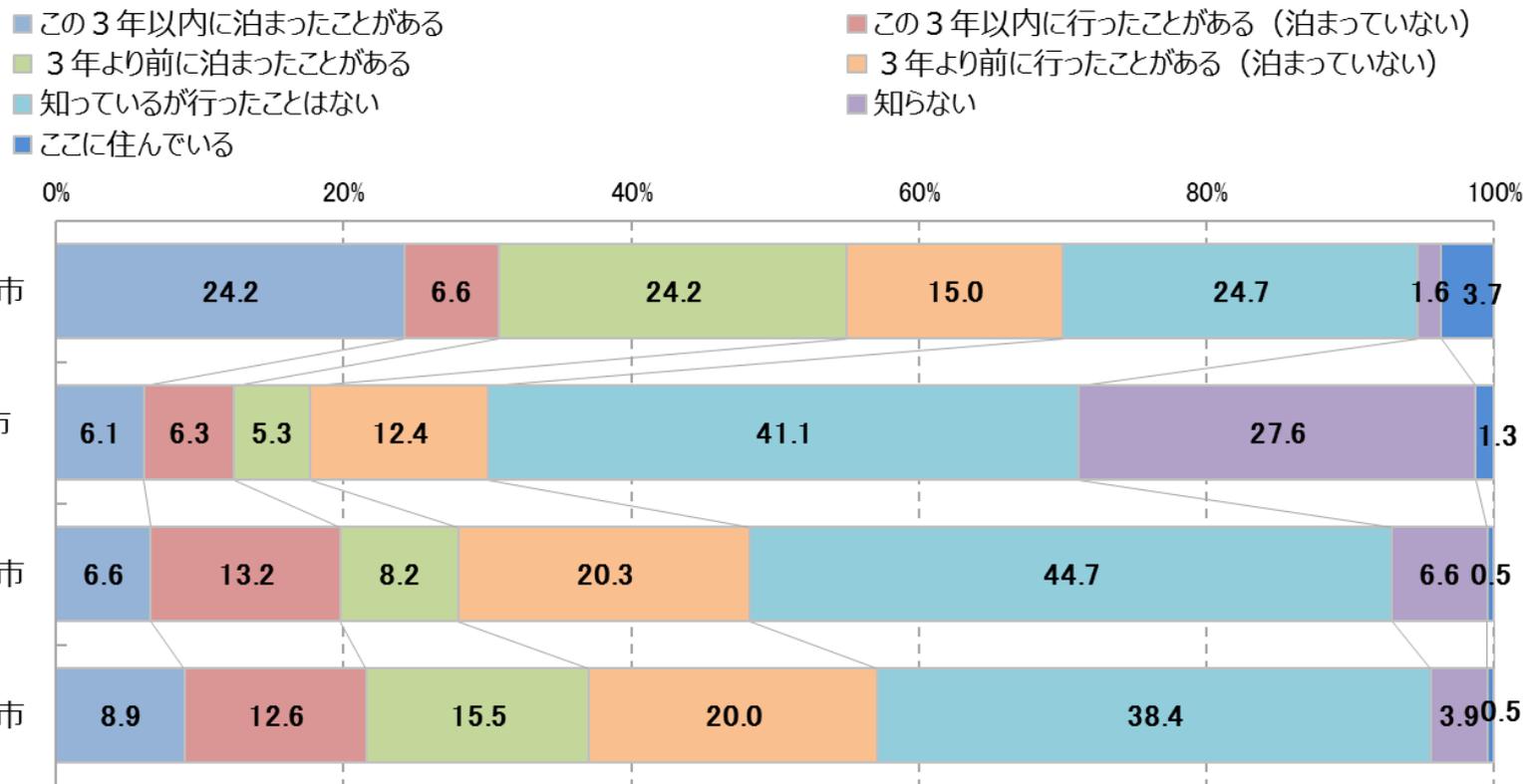
その他 (_____)

アンケート調査票 (参考例)

⇒調査者の聞きたいことが
聞ける

アンケート調査票の集計例

認知度・訪問経験



定量分析/クロス分析の例

2. 2次データ

2次データのメリット、デメリットを確認します。

メリット

1. アクセスが容易である

インターネット環境であれば誰でも容易に沢山の2次データにアクセスすることが可能となります。また、収集のためのコストや時間の節約ができます。

2. データの信頼度が高い

政府系機関等が収集しているデータであれば、そのデータの信頼度が高いといえます。

3. 分析がしっかりとされている

データ収集のみならず、そのデータに対する分析も適切にされていることが多く、調査結果の信憑性が増します。

デメリット

1. 誰もがアクセス・活用している

多くの方がアクセス可能なので、データ自体の独自性はないです。

2. 調査実施者が本当に知りたいデータが完全には揃っていない

調査実施者の知りたい形でデータが揃っていないことが多いです。それゆえ、深い分析ができないケースもあります。

【参考】 観光ビッグデータ

ビッグデータの中で、観光客の動向に特化したビッグデータを「観光ビッグデータ」と呼びます。

観光ビッグデータの例

- ・ 位置情報のデータ
- ・ SNSでの投稿データ
- ・ 交通系ICカードによる移動、消費のデータ 等

例えば、外国人観光客がどこから移動してきたか、どのお店で何を買ったのか、観光地に満足しているのか、不満を感じているのかなどをデータとして収集することで判断することが可能となります。

つまり、このようなデータを分析することで、

- ・ インバウンド向け観光周遊コースの設定
- ・ ターゲットとなる国・地域の設定
- ・ 価格の設定
- ・ 効果的なプロモーションの方法

など、マーケティング戦略を実施する上で重要な要素を決めるヒントにすることが可能です。

観光ビッグデータの種類

ビッグデータの種類と主な特徴を整理します。

1. スマホアプリによる位置情報データ

Google検索やFacebook、Twitter、電車の乗り換えアプリ、マップアプリ、ゲームなど位置情報を利用したアプリが多数あります。これらのアプリから取得するデータにより、

- ・ 観光客がどのように移動したのか
- ・ どこでどのようなアプリを起動し、何をしたのか
- ・ 移動してアプリを使った観光客の属性は何なのか

などのデータを取得することが可能です。

観光ビッグデータの種類

2. SNSでの投稿データ

SNSでの投稿データも観光ビッグデータの1つです。SNSで投稿される内容は主に以下のようなものです。

- 観光地への興味・関心
- 旅館ホテルの感想
- お土産などに対する評価
- 観光地やサービス、商品に対する不満

などの情報がダイレクトに取得できます。よりリアルな観光客の生のデータを取得できます。また、SNSによっては、投稿者の属性や位置情報なども同時に公開されるので、クロス分析等が容易になり分析の幅が広がります。その他、特定の観光地や商品、サービスに対する投稿数が増えることで、認知度が上がったなどの情報も入手することが可能となります。

観光ビッグデータの種類

3. 交通系ICカードによる移動、消費のデータ

交通系ICカードから様々なデータを取得することも可能です。例えば、交通系ICカードでは買い物もできるため、移動だけでなく消費行動のデータもあわせて取ることも可能となります。

- ・ 観光客の移動に関するデータ
- ・ 観光客の消費行動に関するデータ

しかし、交通系ICカードが買い物で利用できるのはまだ一部の店舗に限られているので、全ての消費行動を入手することはできません。

マーケティング調査の手法①

主なマーケティング調査の手法は以下の通りです。

① 定性調査

調査対象者の考えや行動、状態や、観察者の印象などを収集する調査手法を定性調査といいます。対象者の深層心理を探る調査ともいえ、定量調査の答えの奥にある真意（＝なぜ）を深く掘っていく調査です。

・ 代表的な手法

① フォーカスグループインタビュー

座談会形式でファシリテーターが対象者に商品選択の理由や感想、改善希望などを聞いていく方法

② デプスインタビュー

面談形式で調査者が対象者に個別に聞いていく方法

③ オブザベーション調査

対象者がどのような行動をしているかどのような購買をしているかなどを、調査員が実際の現場に出向いて観察する方法

マーケティング調査の手法②

② 定量調査

数や割合、変化などを「数字＝量」で結果を表すことが可能な調査を定量調査といいます。

表やグラフにできるのが利点です。

・ 代表的な手法

① インターネット調査

インターネット上で行うアンケート調査。容易に沢山のサンプル数入手することができる方法

② 聞き取り調査

調査員が調査票を持参し、観光地等で対象者（＝観光客）にアンケート記入してもらう方法

③ 電話調査

対象者に電話をかけ、口頭でアンケートに答えてもらう調査方法

公開されているビッグデータ

公開されている主なビッグデータを確認しましょう。

① RESAS（リーサス）

経済産業省・内閣府が提供している地域経済に関する様々なビッグデータ（産業の強み、人の流れ、人口動態など）のこと。地図やグラフで分かりやすく「見える化（可視化）」したシステムで地域経済分析システムとされています。

② 観光予報プラットフォーム

宿泊を基軸にした「観光」に関連するデータを提供するデータプラットフォームが「観光予想プラットフォーム」です。観光に関する各種ビッグデータをエリア単位でまとめて提供しています。

③ Googleトレンド

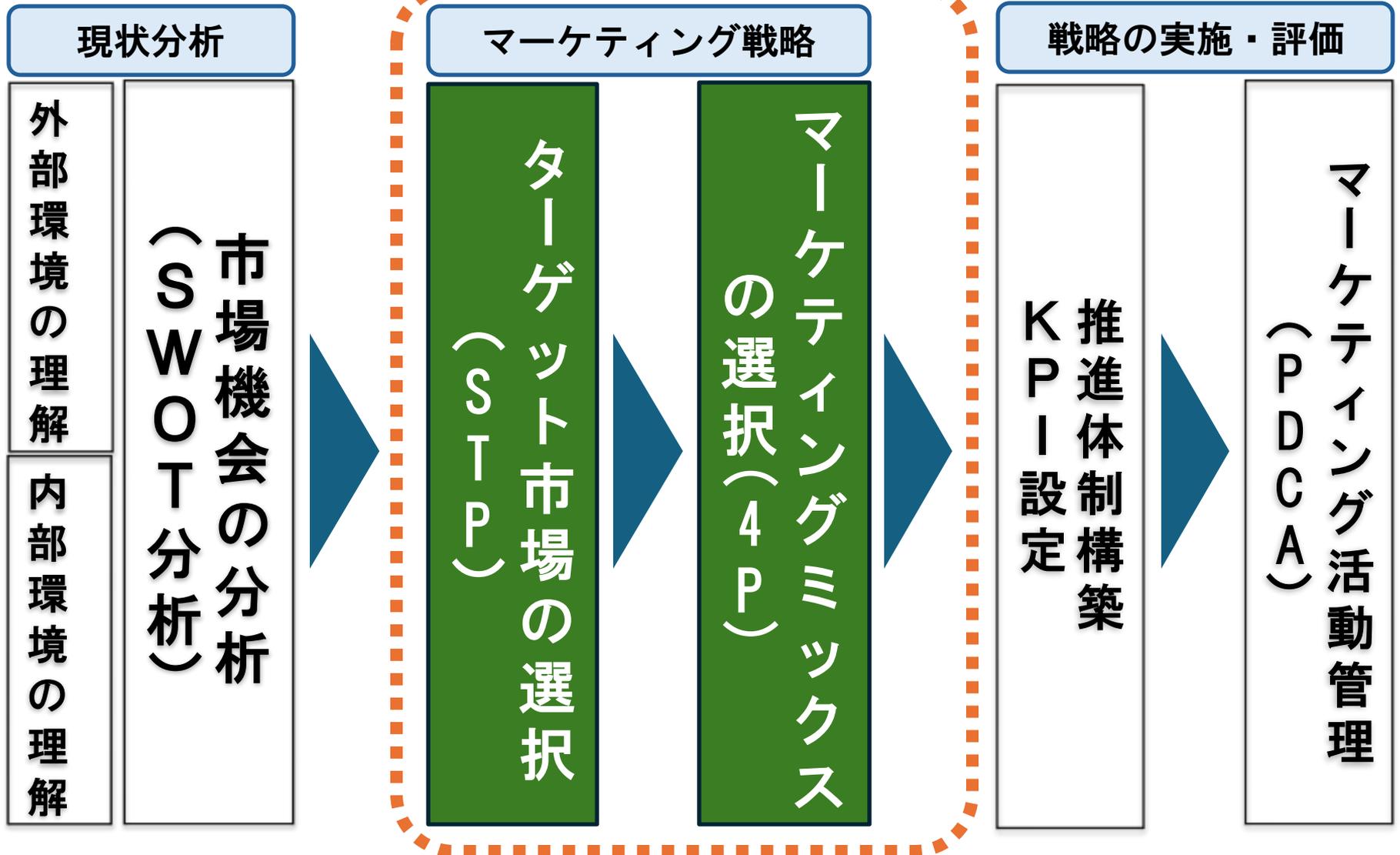
Googleが提供している“キーワードやトピックの検索回数トレンドを確認できるWebベースのツール”のこと。どのようなキーワードがマーケットで注目されているか確認するの役立ちます。

データマーケティングの基礎②



観光マーケティングのプロセス

マーケティングのプロセス



市場機会の分析

自社、自地域を取り囲む外部環境と内部環境を整理してみましょう。
その際に便利な方法が**SWOT分析**です。

■ SWOT分析

自社、自地域の強み・弱み(**内部要因**)と自地域を取巻く機会・脅威(**外部要因**)の4つを整理し、それらを組み合わせて考えることで現状を把握する手法です。

自地域の強み(Strengths)	自地域の弱み(Weaknesses)
自地域を取り囲む機会(Opportunities)	自地域を取り囲む脅威(Threats)

ニーズ理解の重要性

■ 観光客ニーズの場合

観光客ニーズとは、**観光・旅行を実施しようと思ったきっかけ、動機**。
日常生活では何か不足していて、それを埋めたいと思う欲求のこと。

→ 顧客のニーズを探ることがマーケティングの第一歩となります！

■ 例を考えてみましょう！

温泉旅館にやって来ました・・・

Aさん：一人で非日常の空間をゆっくりと過ごしたい

Bさん：両親の古希を3世代でお祝いしたい

Cさん：体調が悪いので身体に効く温泉に何度も入りたい

Dさん：美味しい会席料理を堪能したい

Eさん：会社の仲間と来ました



顧客のニーズを掴まないと、どのようなおもてなしが的k説かわからない・・・

観光マーケティングの定義

観光マーケティングとは、自地域に呼びたい
観光客のニーズ・ウォンツを探り、
他地域ではなく、自地域を訪問していただき、
自地域のファンになっていただくための活動全般を指す。

■ どのような活動を想定するか？

- ・ 顧客嗜好の調査・研究
- ・ ライバルとなる観光地域の特定・調査
- ・ 顧客ニーズを満たす観光資源の開発
- ・ 観光商品・サービス作り
- ・ 観光商品の販売方法の検討
- ・ 観光商品のPR、プロモーション など

観光マーケティングとは、

どのような観光客に、自地域のどのような観光資源を、どうやってPR・提供するかというプロセス

“誰に × 何を × どうやって” 提供するか

STP

Three horizontal lines are positioned below the text 'STP'. The top line is blue, the middle line is red, and the bottom line is green. They are all of equal thickness and extend across most of the width of the page.

ターゲット市場の選択＝STP

「誰」をターゲットとするか決める → 「STP」という考え方を活用しましょう！

・S = セグメンテーション

機会があると考えられるマーケットを共通の切り口で細分化していきます

・T = ターゲティング

細分化した細かい市場の中から最も魅力のあるセグメントにターゲットを絞り込みます

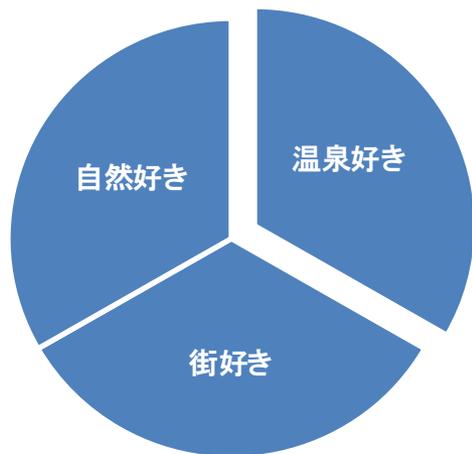
・P = ポジショニング

そのターゲットに対して、競合先の観光地に勝てるポジション（観光客からの見え方）を探します

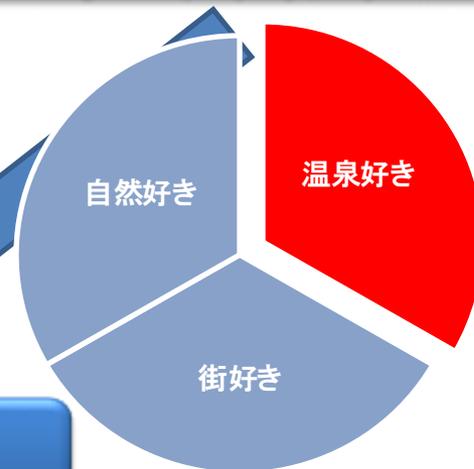
The logo consists of the letters 'STP' in a bold, blue, sans-serif font, centered within a solid yellow square background.

ターゲット市場の選択=STP

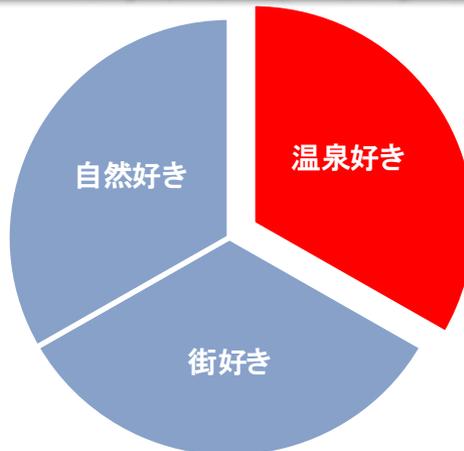
STEP 1
セグメンテーション



STEP 2
ターゲティング



STEP 3
ポジショニング



ライバルに勝てる
立ち位置を決める

The STP Process



セグメンテーションの実施方法

4つの基準を活用

(1) 人口属性基準

性別、年齢、収入、家族構成、職業など、基本的な基準

(2) 物理的基準

どこに住んでいるかという地理的な尺度。国、都道府県、市町村、地理条件などの基準

(3) 心理的基準

価値観、考え方、性格、ライフスタイルなどの基準
(家族との時間を大切にする、エコ志向など個人的な志向)

(4) 行動基準

商品の購買履歴、特定ブランドへのこだわりなど実際の消費行動に着目した基準



以上のような軸を活用してターゲットを細分化することで、
自地域の地域資源にマッチするターゲット顧客を設定します

ターゲット選定について

設定したターゲットについては、以下の3つの観点で実現可能性があるか確認をします。

[ターゲット選定を実施する際のポイント]

① 測定が可能であること

ターゲットとなるセグメントを「定量的に把握すること（＝測定すること）」が重要となります

※実際には数値で計るのが難しい場合もあります

② ターゲットセグメントへリーチができること

魅力的なセグメントが見つかってもしリーチ（＝コンタクト）が出来ないと現実的なビジネスへ発展しません

③ 需要量が相当数量あること

市場の細分化を図れば図るほど需要量が減少します。よってビジネスが成立する規模を見極めることが重要です

個人演習

個人演習①

SWOT分析の実践

自社または自地域を中心にSWOT分析を実施しましょう。

自社／自地域の強み(Strengths)	自社／自地域の弱み(Weaknesses)
自社／自地域を取り囲む機会 (Opportunities)	自社／自地域を取り囲む脅威 (Threats)

個人演習②

STPの実践

自地域に呼び込みたい「観光客」をセグメントしてみましょう。

セグメントする方法は、前頁の4つの基準を中心に、「居住地」、「旅行形態」、「年齢層」、「ニーズ」、「目的」など自ら考えられる指標で観光客を細分化しましょう。また、各セグメントについて詳細も記載しましょう。

セグメント例：

「職場旅行」、「修学旅行生」、「個人旅行」、「MICE」、「視察旅行」、
「小中学生連れ家族旅行」、「新婚旅行」、「シルバーご夫婦の旅」、
「3世代旅行」、「乳幼児連れ家族旅行」、「一人旅」、「若者グループ」、
「関東圏居住者」、「中京圏居住者」、「関西圏居住者」、「外国人」、
「台湾のリピーターFIT客」、「中国のクルーズ客」、「タイのグループ客」、
「欧米のスキーヤー」、「バス旅行者」、「温泉めぐり」、
「自然景観」、「歴史的建造物」、「その土地ならではの体験をしたい」、
「恐竜博物館に行きたい」、「蟹を食べたい」、「海の幸を満喫したい」、
「地元民と交流したい」、「パワースポットめぐり」、「女子旅」、「歴女」 など・・・

個人演習③

STPの実践

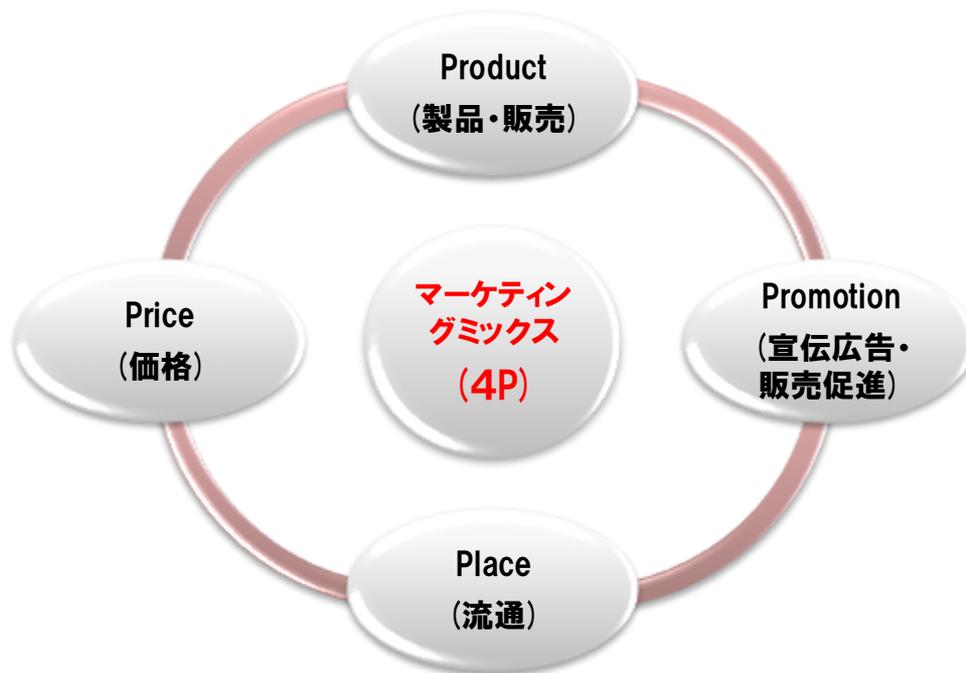
セグメント名	ニーズ／特徴など

マーケティングミックス

マーケティングミックスの理解

「何」を「どうやって」提供するか決めましょう。

ターゲットが決まったら、自地域の「どのような観光資源」を「どのような手段」で提供していくかを決めます。これらを決定する事が、マーケティングの主活動になります。具体的には、**マーケティングの4つの要素に従って、戦略を立案**します。つまり、マーケティング活動とは、**4つの要素 (= 4P/マーケティングミックス) をどのように組み合わせ**て顧客へ訴求するかを考えるシナリオ作りのことです。



最初のP=PRODUCT

■ **Product**=商品・サービスのこと

- ・ ターゲットとなる観光客に提供する商品・サービスのこと。
- ・ 商品・サービスの概念に含まれるものは、「ブランド」、「機能・性能」、「品質」、「接客方法」など商品・サービスに関連するあらゆるもの。

■ 観光関連商品の特徴

1. 形がない

観光商品は体験して頂くもの。形がない商品となります。

2. 返品、やり直しができない

観光商品を体感して頂くのは1回のみ。基本的には返品が出来ない商品となります。

3. 季節変動がある

同じ商品・サービスでも季節により金額が変わるシーズンリティーが存在します。

最初のP=PRODUCT

観光商品のもととなるもの = **地域資源**

地域資源とは、**地域に眠る**様々な「リソース」を指します。

地域におけるリソース

自然資源 山／高原／原野／湿原／河川／滝／海岸／動物／自然現象 など

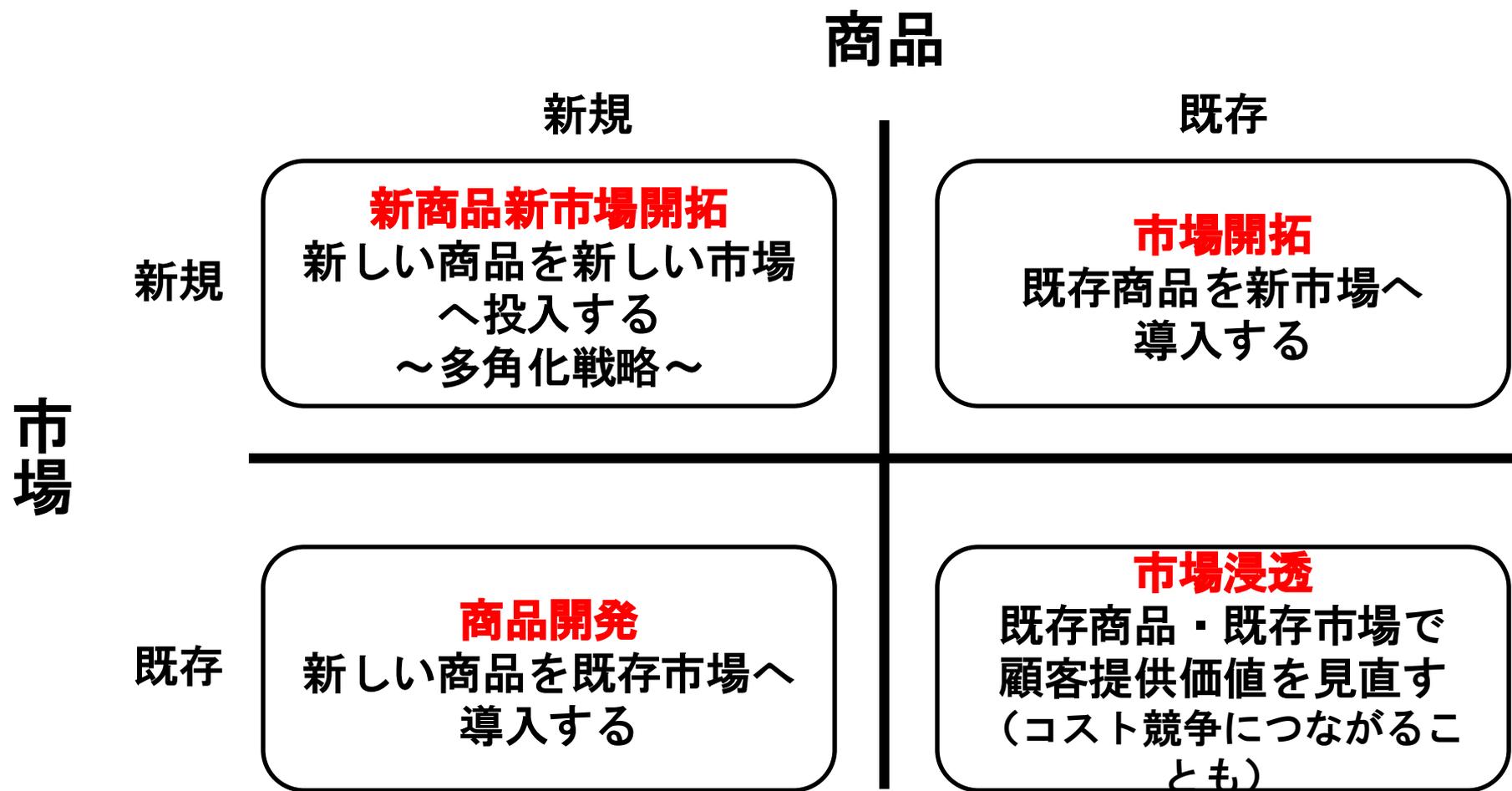
物的資源 名所／史跡／建築物／街並み／商店／地元の食 など

人的資源 祭・伝統／イベント／体験／自然現象／文化／人／人の技術／
産業資源／歴史 など

地域資源を発掘する際のポイント

他の地域では味わえない、「**その地域ならではの**」のもの

商品展開・市場開発＝アンゾフの成長マトリックス



どの軸で商品開発をしていきますか？

2つ目のP=PRICE

■ Price=価格

- ・ 観光客が商品・サービスを手に入れるために支払うべき対価
- ・ 旅行全体にかかる金額を掴むこと、お値打ち感、競合商品の金額の意識などが大切

ナチュラルパークツアー<<パート1>> 大好評運行!!

期間: 2014年4月1日～2014年11月30日
バスで行くゆったり快適な半日観光
奥日光の大自然をリーズナブルな料金で半日観光!!

1日2便運行(午前便/午後便)

料 金: 大人お一人様3,200円 子供1,600円(3歳以上12歳未満)

コース:

東武日光駅観光協会前発[09:20/13:45]～神橋板垣退助銅像前[09:25/13:50]～ホテル清晃苑前[09:30/13:55]～鶴亀大吉前[09:33/13:58]～日光千姫物語前[09:35/14:00]～懐かし家風和里前[09:40/14:05]～いろは坂～明智平(トイレ休憩※お申し出が無い場合は道路が渋滞の場合通過致します。)～中禅寺湖ボートハウス～竜頭の滝～立木観音～華厳の滝～いろは坂～方等の滝・般若の滝(車窓より)～懐かし家風和里前[12:30/16:55]～日光千姫物語前[12:35/17:00]～鶴亀大吉前[12:37/17:02]～ホテル清晃苑前[12:40/17:05]～神橋板垣退助銅像前[12:45/17:10]～東武日光駅[12:50頃/17:15頃]

感動の瞬間! 早朝にガイドと行く小田代ヶ原の貴婦人

ゆったりバスで行く早朝トレッキング(ガイド付)
小田代ヶ原に「貴婦人」と呼ばれる一本の白樺の木があります。

設定日: 2014年9月7・14・15・21・23・28日

料 金: お一人様2,000円(バス代・ガイド料含む)

●コース

東照宮昇陽苑(4:50発)～神橋板垣退助銅像前(5:00発)～ホテル清晃苑前(5:05発)～鶴亀大吉前(5:10発)～日光千姫物語前(5:12発)～懐かし家風和里前(5:20発)～いろは坂～赤沼…(徒歩約40分)…小田代ヶ原(ハイブリットバス7:15分発)～赤沼(7:27着/7:40発)～いろは坂～懐かし家風和里前(8:30頃着)～日光千姫物語前(8:35頃着)～鶴亀大吉前(8:38頃着)～ホテル清晃苑前(8:40頃着)～神橋板垣退助銅像前(8:45頃着)～東照宮昇陽苑(8:55頃着)

出典: 日光おでかけサロンホームページ

商品・サービスには必ず金額がある。ポイントは観光客がその金額をどのように捉えるかである

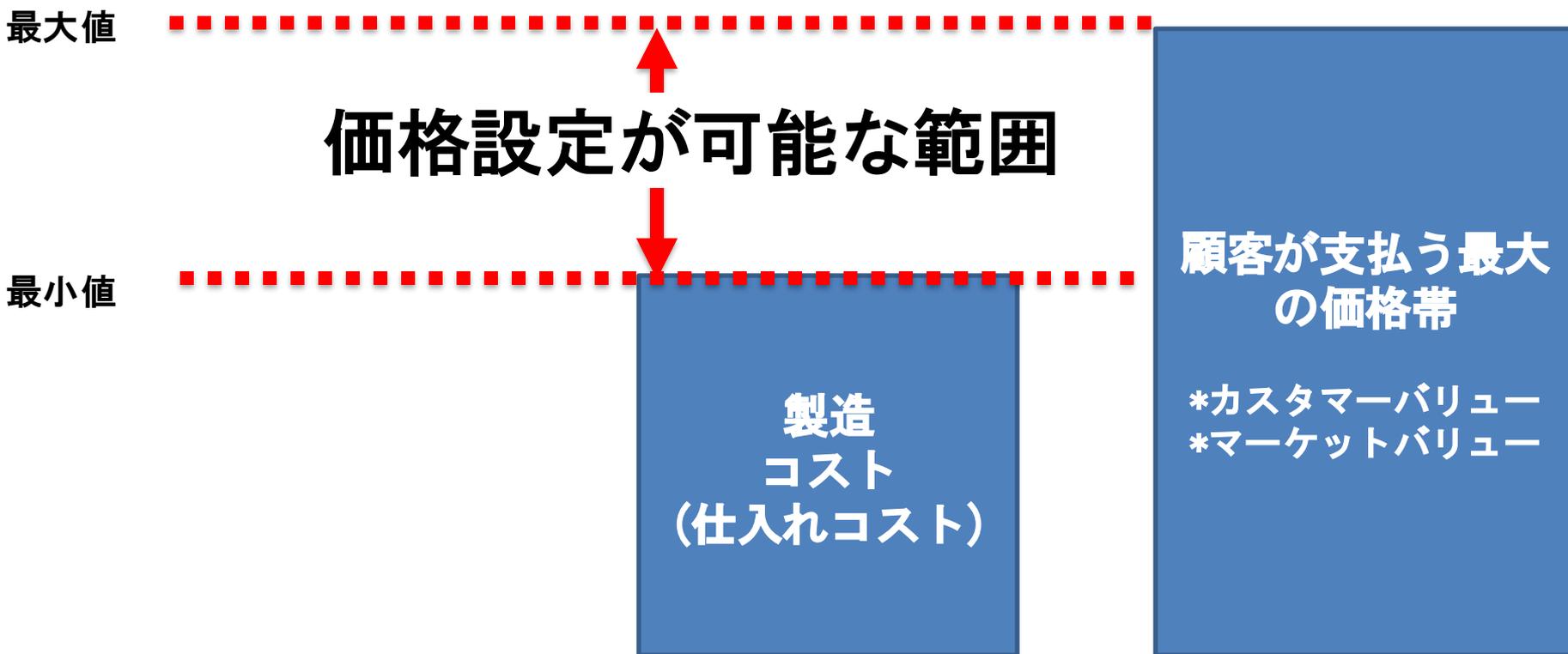
- ◆ 観光客の期待値 > 提供側のパフォーマンス → 高いと感じる。場合によってはクレームへ
- ◆ 観光客の期待値 < 提供側のパフォーマンス → お得と感じる。満足、感動の一要素へ

但し、旅行商品は気象条件、同行者、受益者の感情など多くの複合的要素で満足度を形成する為一概にいけない

2つ目のP=PRICE

■ 値決めの方法

商品・サービスの値段の決め方



観光客の満足度、ライバルの価格設定などを考慮し、**最小値から最大値の間**で「**値決め**」を実施する。原価が捉えづらい観光関連商品であるが、**一定の値決めのルールを決める**ことはマーケティング戦略上、大切になってくる。

3つ目のP=PLACE

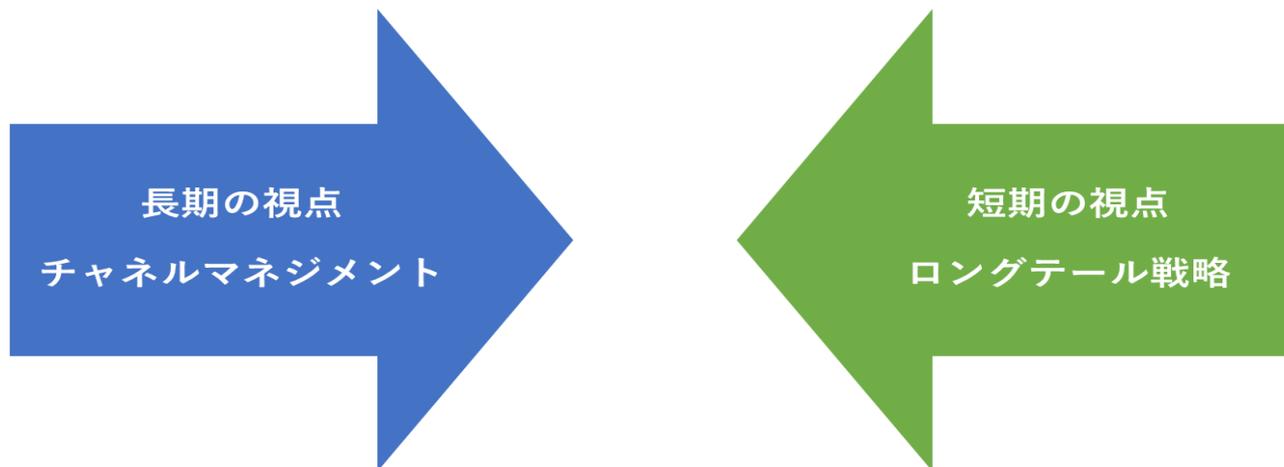
■ Place

Placeとは「流通」のこと。ターゲット観光客へ商品・サービスを提供するための手段・方法を言います。ターゲット顧客が購入できる的確な流通チャネルを探ることが大切です。

■ 流通チャネルの例

販売チャネル
ホテル・旅館のフロント
ホテル・旅館のホームページ
旅行会社窓口
オンライントラベルエージェント
旅行会社のホームページ
観光協会のホームページ
観光案内所
コンビニエンスストア
交通事業者のホームページ
旅行の雑誌・TVでの申し込み

チャンネル戦略における基本的な考え方



チャンネルマネジメント

販売チャンネルを主体的に管理、運営。自社の競合優位性を確立するために、構築したチャンネルを適切に管理、運用。将来的なロイヤルカスタマーの育成、自社ブランドの構築を図るべく、お客様とどのような関係を構築するかを考えてからチャンネルを管理する。

ロングテール

インターネット等を用いた販売の手法。販売機会の少ない商品・サービスでもアイテム数を幅広く取り揃えること、または対象となる販売機会・顧客総数を増やすことで、総体としての売上げを大きくすることが可能となる。

4つ目のP= PROMOTION

■ Promotion=宣伝広告・販売促進のこと。

ターゲットと設定した顧客に対して、商品・サービスの特徴・価値を伝えて、**購買行動を促す一連の活動**をプロモーションといいます。

宣伝広告、セールスプロモーション、新聞・雑誌・TV・WEB等での告知など、様々な販売促進活動があります。提供者はこれらの活動を効果的に計画します。

■ 顧客購買心理からプロモーションを考察する2つの手段方法

1) AIDMA



こんな面白そうな
観光地があるか・・・

子供も喜ぶかもなあ・・・

次の休みには
行ってみようかな・・・

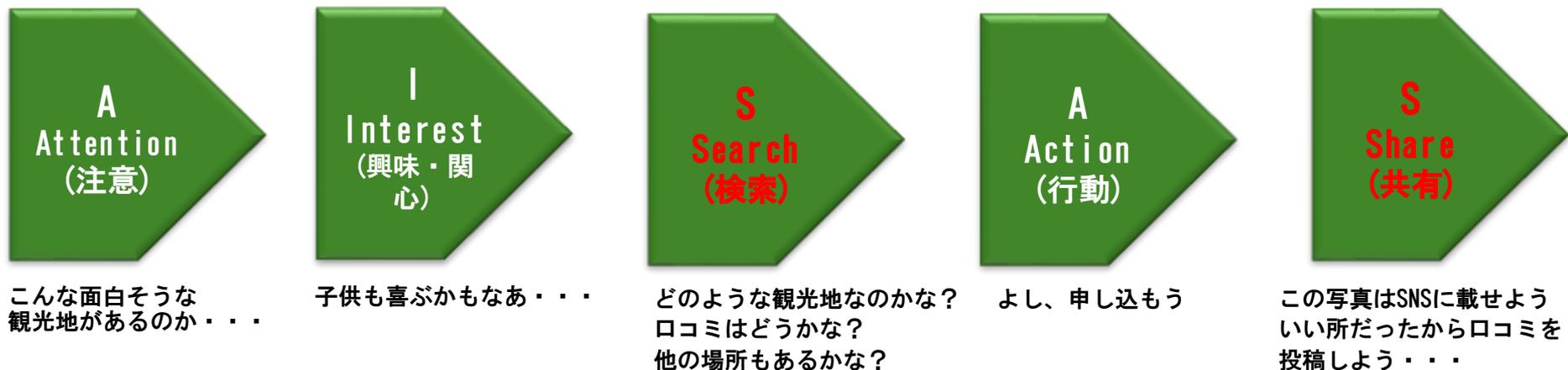
夏の家族旅行はどこにす
んだっけ・・・
あ、あそこだ！

よし、申し込もう

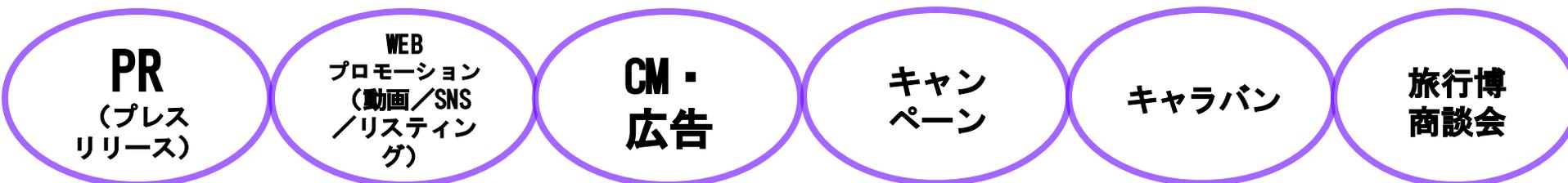
4つ目のP= PROMOTION

2) AISAS

インターネットの登場により顧客の購買心理と行動にも変化が生まれました



◆プロモーションミックス例



顧客の購買心理・行動の各フェーズによって取るべき
プロモーションの施策は異なってきます

4Pに追加する3つのP

5つ目のP=PERSONNEL

■ 人員 (=ヒューマンリソース)

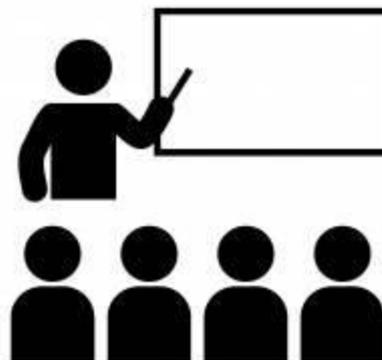
従業員（正社員、臨時雇用含む）、関係会社等までを含めた顧客にサービスを提供する**全ての人員 (=ヒューマンリソース) を整える**必要があります。

顧客にとっては、サービスの提供者が自社の社員なのか協力会社の社員かは関係ありません。どの人員においても、しっかりとしたマニュアルのもと、一定のサービスが提供できるようにし、提供価値を均一にする必要があります。

その上では人材トレーニングを定期的 to 実施する必要があります。



スキルアップ
モチベーションアップ



トレーニング
OJT & OFF JT

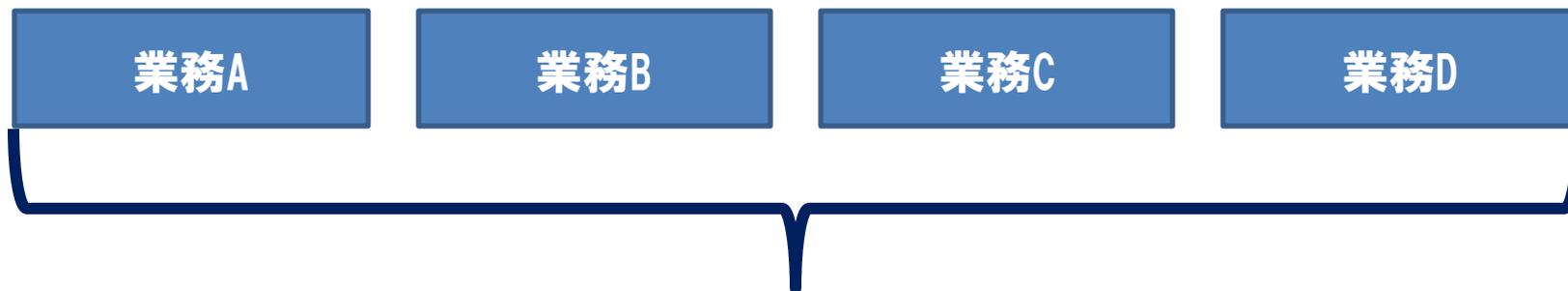
サービス産業においてはサービスを提供する人員=ヒューマンリソースの育成、管理は重要な要素を占めます

6つ目のP=PROCESS

■ 業務プロセス

顧客がサービスを受ける前後のプロセスを効果的に管理することも重要です。旅行客に対しては、旅マエでの顧客接点の作り方、旅ナカでの効果的な情報提供、旅アトでの満足度、推奨度の把握と共有といった業務をプロセスとして管理することが必要です。

昨今のデジタル化の加速により、これらプロセス管理が非常にやりやすい環境にあります。



顧客満足度を効率的に創出するためには全ての業務を管理することが必要です。まずは現状の業務を全て洗い出すことから始めます。

プロセス管理には、マニュアル作成、デジタル活用が有効です

7つ目のP=PHYSICAL ENVIRONMENT

■ 物理的環境整備

旅行やサービスは目には見えない商品です。それら目に見えない商品に対して、目に見えるもの（＝物的証拠）を追加することで、顧客の安心感・満足感を生み出すことが可能となります。

【例】

- ・ 旅行の申込書、旅のしおりなど、しっかりとした書類を充実させること
- ・ オリジナルグッズ、タグなどを作成し顧客に配布すること
- ・ リゾートホテルにおけるスタッフのユニフォーム統一
- ・ 旅館などで館内装飾品、アート品を展示すること



目に見えない「サービス」や「旅行」に物理的な環境を整備することで、商品としての付加価値が上がります

KPIとPDCAサイクル

KPIの設定

KGI (Key Goal Indicator)

重要目標達成指標 目標が達成したかどうか成果を測るものさし

KPI (Key Performance Indicator)

重要業績評価指標 プロセスが適切に実施されているかどうかを
中間的に計測するものさし

- 組織において、個人や部門の業績評価を定量的に評価するための指標。
- 達成すべき目標に対し、どれだけの進捗がみられたかを明確にできる指標が選択される。
- これをもとに、日々の進捗把握や業務改善などが行われる。

KPIの設定



KPIの例

- **観光地におけるKPI**
観光入込数／宿泊数／消費額／満足度／リピート率 など
- **WEBプロモーションにおけるKPI**
サイトへの訪問数／直帰率／平均ページビュー／サイト滞在時間
コンバージョン率／動画再生回数 など

PDCAサイクル

商品・サービスを造成し、市場に導入した後は、「PDCAサイクルをしっかりと回す」事でマーケティングの諸活動を管理します。最終的な目標は、リピーター（自地域のファン）を作ること、自地域のブランドを作ることです。



個人演習



個人演習①

観光産業の4P戦略

自社、自地域のターゲット顧客を想定してマーケティングミックスを作ってみましょう。

ターゲット顧客	
商品・サービス	価格
販売チャネル	プロモーション手法

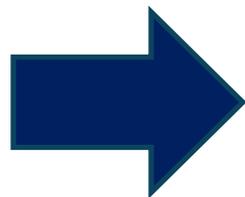
マーケティング理論の振り返り



デジタルマーケティングの基本1

学習目標

1. デジタルマーケティングの基本の理解
2. インターネット広告の理解



このコースを学習すると、観光マーケティングの基本とインターネット広告について理解することが出来ます

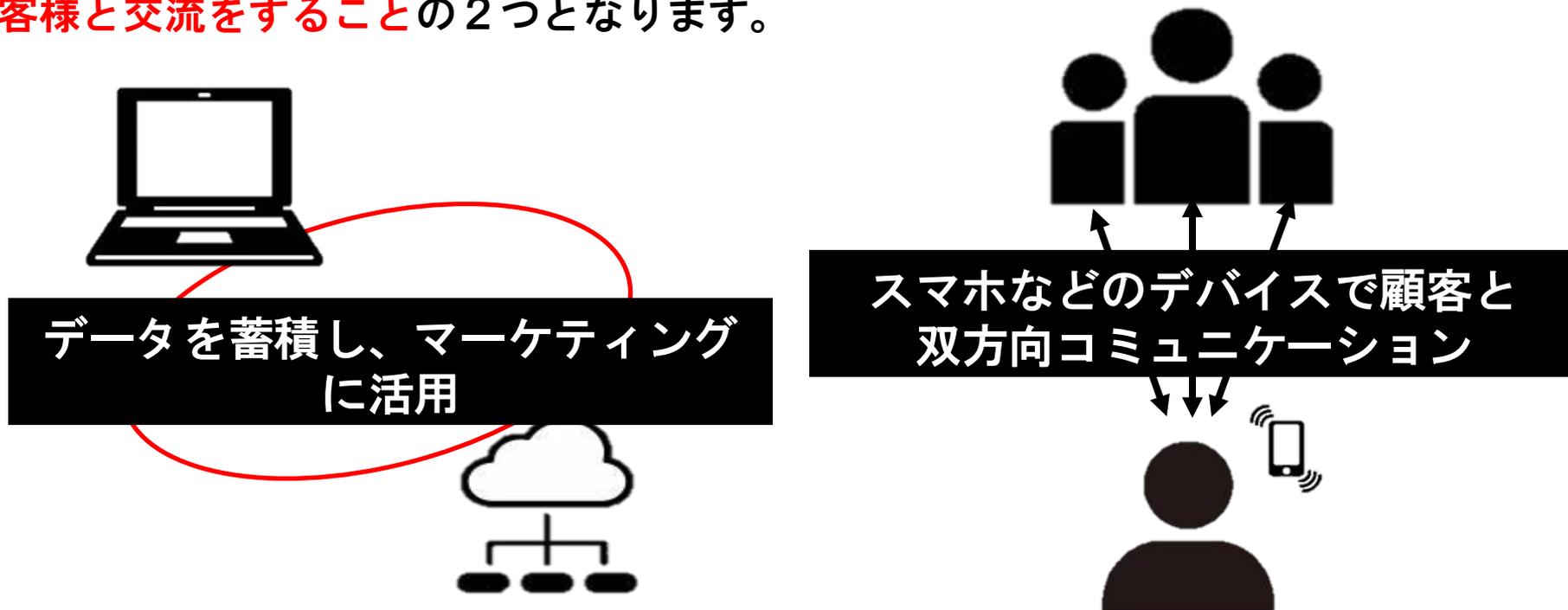
1 - 1 デジタルマーケティングの基本

学習目標

デジタルマーケティングの考え方

デジタルマーケティングとは、**デジタルの力を活用してマーケティングをしやすいようにという考え方**です。顧客ニーズ・ウォンツの把握やマーケティングミックスの考え方など、基本は全く一緒であり、単に手段としてデジタルを活用するということです。

デジタルマーケティングが特に力を発揮するのは、マーケティングミックスの4つ目のP=「**プロモーションのプロセス**」にデジタルを活用することです。また、デジタル活用で大切なことは、**データを蓄積し活用すること**と、**デジタルデバイスでお客様と交流をすること**の2つとなります。

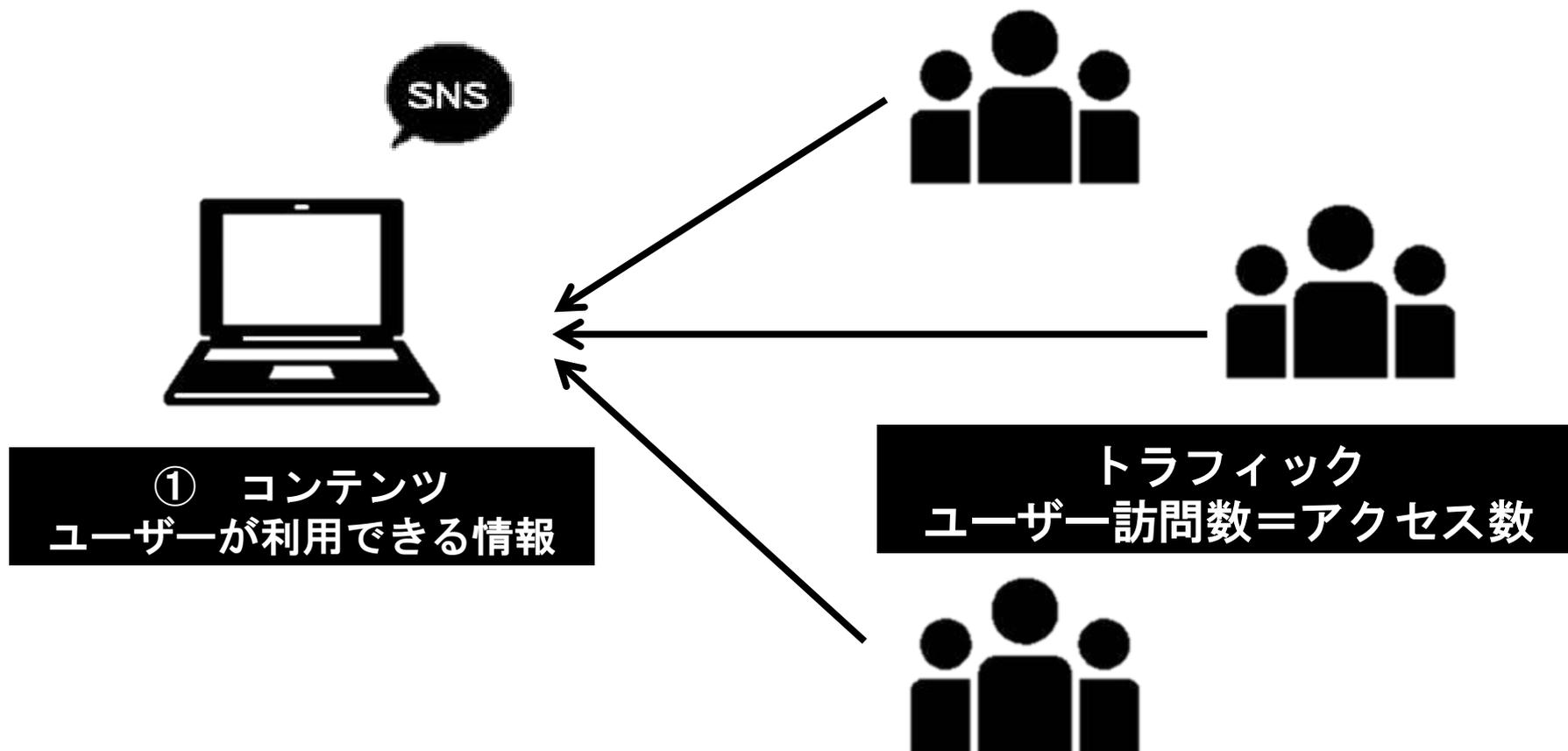


学習目標カスタマージャーニー

プロセス	知る/興味を持つ	調べる/比べる	予約/購入する	旅する	共有する		
顧客行動	<ul style="list-style-type: none"> ✓ 友達・家族からお勧めされる ✓ 広告を見る ✓ 動画を見る ✓ 記事を読む 	<ul style="list-style-type: none"> ✓ WEBでウェブサイト検索 ✓ 友達・家族に相談 ✓ 比較サイトで調査 ✓ 口コミサイトで比較 ✓ 旅行会社へ相談 	<ul style="list-style-type: none"> ✓ 事業者ウェブサイトで予約 ✓ OTAで予約 ✓ 旅行会社で予約 ✓ TELで予約 ✓ SNSで予約 	<ul style="list-style-type: none"> ✓ 乗り物に乗る ✓ 宿泊施設に泊まる ✓ 食事をする ✓ 体験をする ✓ 交流をする ✓ 観光地を訪問する 	<ul style="list-style-type: none"> ✓ 友達・家族に感想を話す ✓ SNSに掲載する ✓ お土産を食べる/配る ✓ 口コミを書く ✓ 次の旅行計画を作る 		
顧客心理	<ul style="list-style-type: none"> ✓ そんな観光地があるんだ ✓ 何だか楽しそうなところだなあ… ✓ 子供を連れて行ったら喜ぶかな 	<ul style="list-style-type: none"> ✓ 宿泊施設はどんな所があるのだろう ✓ 金額はいくらだろう ✓ 口コミはどうか ✓ 他地域よりもよさそうだな 	<ul style="list-style-type: none"> ✓ 安く購入したい ✓ 安心な買い物をしたい ✓ 失敗しない旅行にしたい 	<ul style="list-style-type: none"> ✓ この観光地は良かった ✓ 旅館もよかった ✓ 昼食は期待ははずれだった ✓ フォトジェニックだな 	<ul style="list-style-type: none"> ✓ 素敵な写真が撮れたのでSNSに乗せよう ✓ 良い施設だったので口コミを書こう ✓ 友達にも教えてあげよう ✓ また行こう 		
顧客との 接点例	SNS/ブログ	ウェブサイト	ウェブサイト	アプリ/デバイス	SNS/ブログ		
	動画	口コミサイト	Online Travel Agent	WEB検索	ウェブサイト/SNS		
	ウェブ広告 (ディスプレイ)	Online Travel Agent	SNS	SNS	口コミサイト		
	広告 (マスコミ4媒体)	航空会社/ホテル等サイト	クーポンサイト	ガイドブック	Online Travel Agent		
	ガイドブック	SNS/ブログ	航空会社/ホテル等窓口	パンフレット/クーポン	メールマガジン		
	交通・野外広告	ウェブ広告 (リスティング)	事業者 (TEL)	事業者スタッフ/観光協会	事業者アンケート		
	キャラバン/旅行博	航空会社/ホテル等 (パンフ)	観光協会/DMD (TEL)	ガイド/住民	DM		
役割	顧客	認	知	購	入	促	進

デジタルプロモーションの基本

デジタルプロモーションを実施する上でまず理解しなくてはならないものは、「コンテンツ」と「トラフィック」です。ユーザーが利用できる情報を「コンテンツ」と言い、訪問ユーザー数(=アクセス数)を「トラフィック」と言います。



デジタルプロモーションの基本

デジタルマーケティング成功のポイントは、**魅力的なコンテンツを作り、トラフィックを増やすこと**です。

① 魅力的な情報を揃えて
オンライン上で準備すること



魅力的なコンテンツ作り

② 沢山の人の閲覧してもらえるように
様々な工夫を凝らすこと



トラフィックを増やす

トラフィックの増やし方

トラフィックをる増やすためには、Googleやyahoo! JAPANといった検索サイトで 沢山検索されるように仕組みを作ることです。そのために理解する必要があるのが、「SEO」と「SEM」という考え方です。

SEO

(Search Engine
Optimization)

検索エンジン最適化

GoogleやYahoo! JAPANなどの検索エンジンで検索した際、検索結果ページの上位にWEBサイトやコンテンツを表示させること。いわゆる自然検索で上位表示させることです。

SEM

(Search Engine Marketing)
検索エンジンマーケティング

検索エンジンを効果的に活用してマーケティングを実施すること。主なSEMには有料広告である「リスティング広告」を駆使してトラフィックを増やすことが挙げられます。なお、SEOはSEMに含まれます。

トラフィックの増やし方

ユーザーに検索してもらうために、SEMを実施し、多くのトラフィックを獲得することを目標とします。リスティング広告、バナー広告の活用やSEO対策実施などで、検索画面の上位表示を目指します。

The image shows two side-by-side screenshots of Google search results. The left screenshot is for the search term '温泉旅館' (Hot Springs Ryokan) and shows approximately 219,000,000 results. The right screenshot is for '福井県観光' (Fukui Prefecture Tourism) and shows approximately 74,100,000 results. Both screenshots have red dashed boxes highlighting specific search results.

Left Screenshot (Search: 温泉旅館):

- 約 219,000,000 件 (0.68 秒)
- 広告・travel.rakuten.co.jp(楽天トラベル温泉宿検索) >
全国各地の人気温泉宿を検索 - 毎月5と0のつく日は5%オフ
15,000円以上のご予約がクーポン利用で5%オフ。温泉宿と高級宿が対象。人気温泉宿のプランが満載。温泉宿を予約するなら5と0のつく日がお得。30,000軒の登録宿泊施設・人気！楽天限定プラン・1,000万件のお客さまの声・秋&冬旅SALE開催中。
- 人気温泉ランキング
2020年上半期
最も多くの方が訪れた温泉地大発表
- 温泉宿検索
人気温泉宿簡単検索
オリジナルプラン満載
- 直前予約空室カレンダー
全国宿の空室簡単検索
今すぐ予約
- 週末空室検索
直前予約も可能
空室簡単検索
- 広告・www.jalan.net(じゃらん.net/宿・ホテル予約) >
温泉宿の宿泊予約なら じゃらん - GoToトラベル最大35%OFF
GoToトラベルキャンペーン/対象施設の宿泊料金が最大35%OFF jalan.net。今夜の宿をスグ探せる。憧れの高級宿をお得に予約・会員登録でポイント貯まる・エリア検索で簡単予約・日本最大級の宿予約サイト・当日予約OKの宿多数・お得なクーポン配布中。
★★★★★ Rating for jalan.net: 4.1 - 166 reviews
温泉・露天風呂の宿予約・土曜日に泊まれるホテル・憧れの贅沢ホテル・旅館・早割プラン

Right Screenshot (Search: 福井県観光):

- 約 74,100,000 件 (0.62 秒)
- www.fuku-e.com >
福井県観光情報ホームページ ふくいドットコム
福井県観光課が運営する福井県観光公式サイト「ふくいドットコム」。福井県内の各種イベント情報、恐竜、温泉、越前・若狭の美しい海岸線、グルメ、おもしろば、ソースカツ丼、ふくいゆかりの幕末備人、宿泊、アクセスランキングなど ...
- www.jalan.net | 観光 | 北陸 >
福井の観光スポットランキングTOP10 - じゃらん.net
おすすめの福井県立恐竜博物館や東尋坊などを口コミランキングでご紹介。福井周辺で観光スポットを探すならじゃらん.net。... 恐竜にどうしても興味があわなかったけど福井県イチオシの観光スポットだから行ってみました。恐竜の骨。
- travel.rakuten.co.jp | ... | 旅のランキング | 福井県 >
現地スタッフ厳選！福井県のおすすめ観光スポットBEST21 ...
2019/08/27 — 福井県現地スタッフおすすめの福井を訪れたら絶対に外せない観光スポットランキング！「永平寺」や「東尋坊」「福井県立恐竜博物館」などの定番以外にも、「熊川宿」や「きのこの森」など現地にいるからこそ知っている ...
福井県立恐竜博物館 周辺の宿・昔寺や深流の大自然など、全国 ... 東尋坊・三方五湖

画像出典：Google

リスティング広告

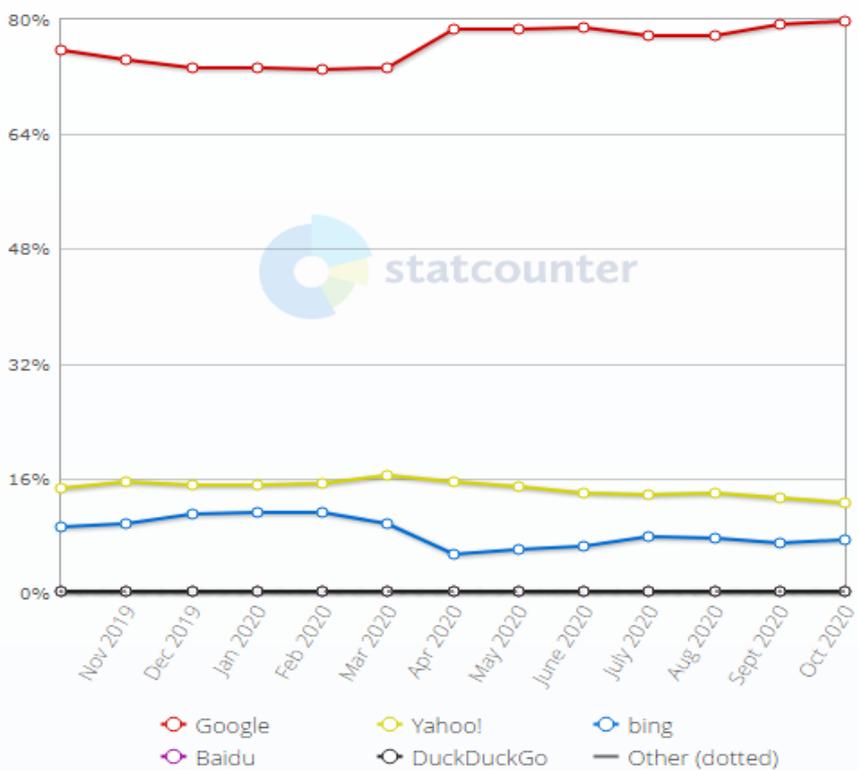
自然検索

検索エンジンの特徴

デバイスや国によっても、主に使われる検索エンジンが異なります。

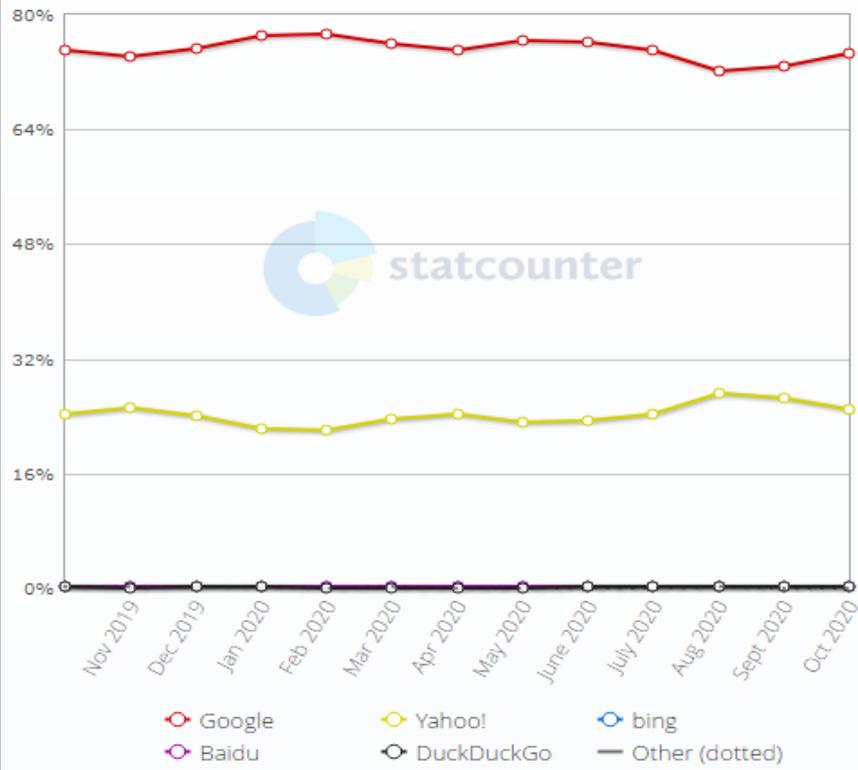
■ パソコンでの検索（日本）

Desktop Search Engine Market Share Japan
Oct 2019 - Oct 2020

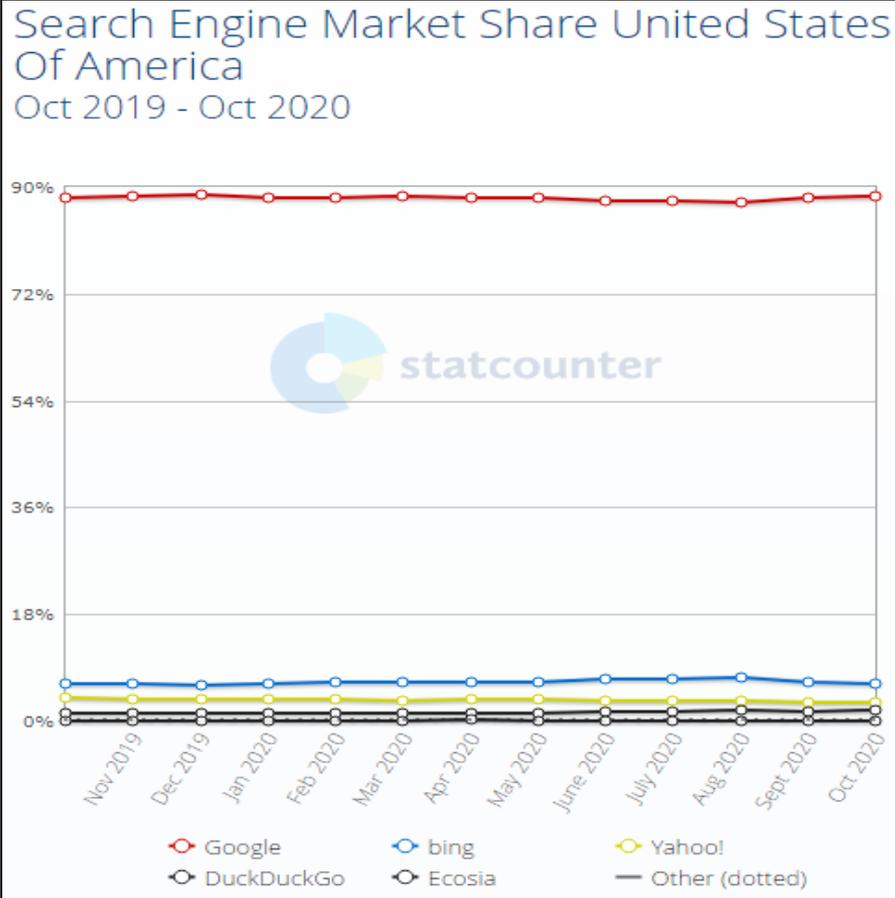


■ スマホでの検索（日本）

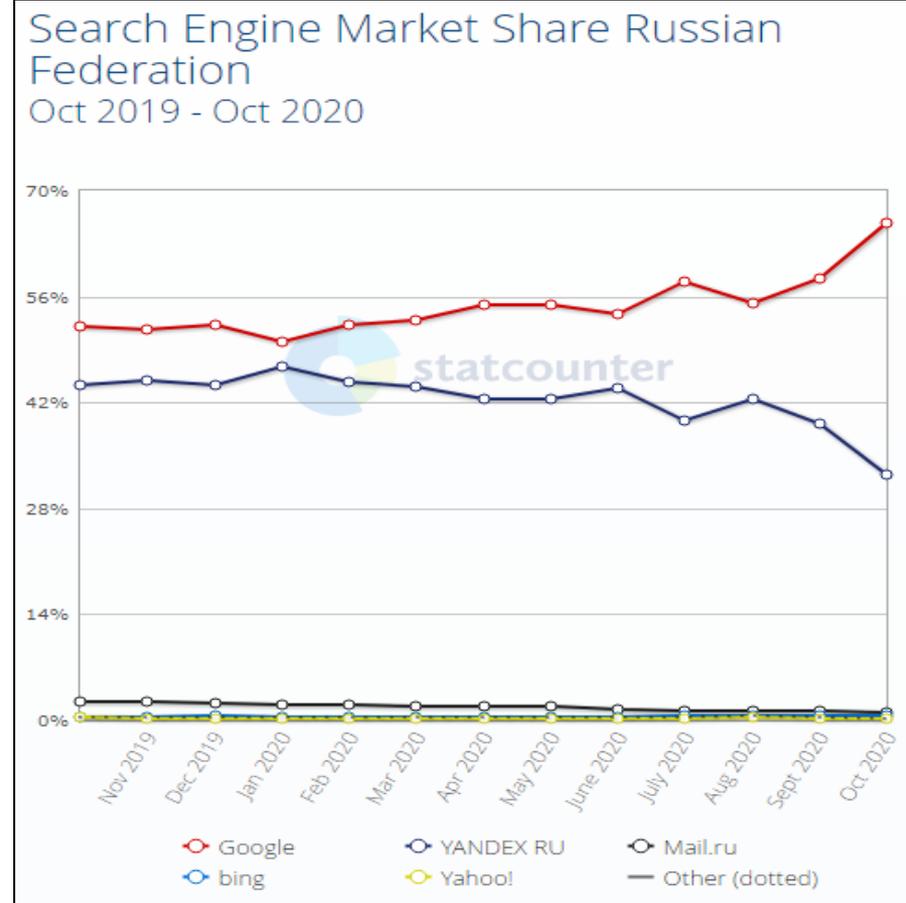
Mobile Search Engine Market Share Japan
Oct 2019 - Oct 2020



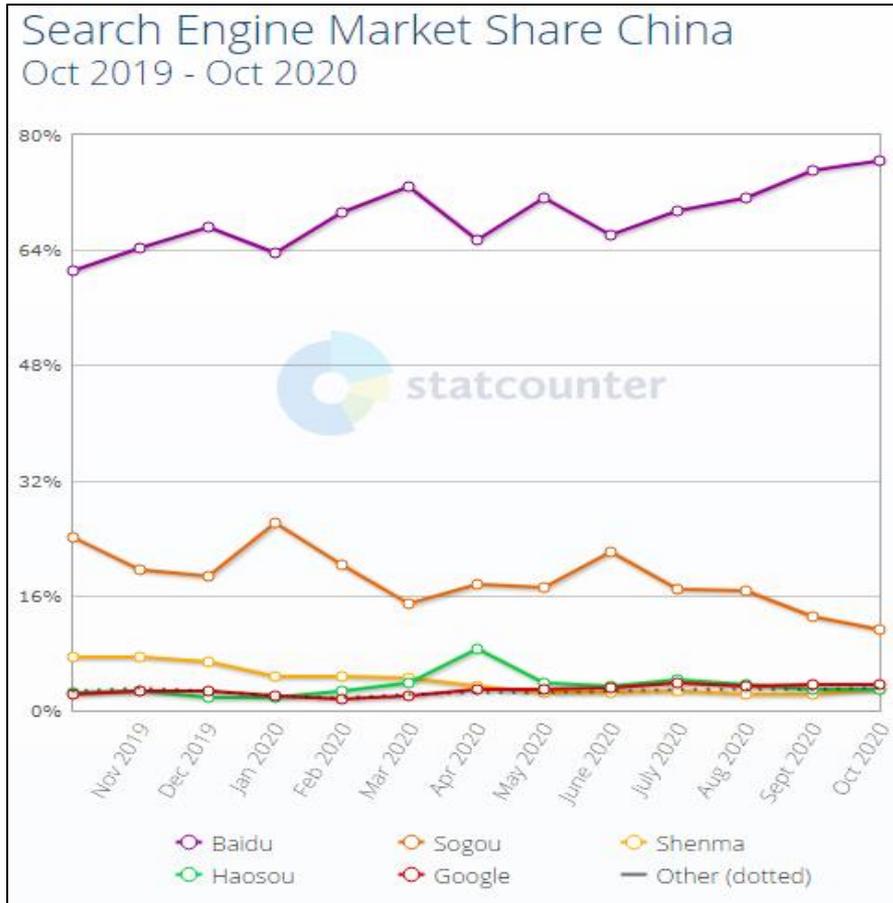
■ アメリカでの検索



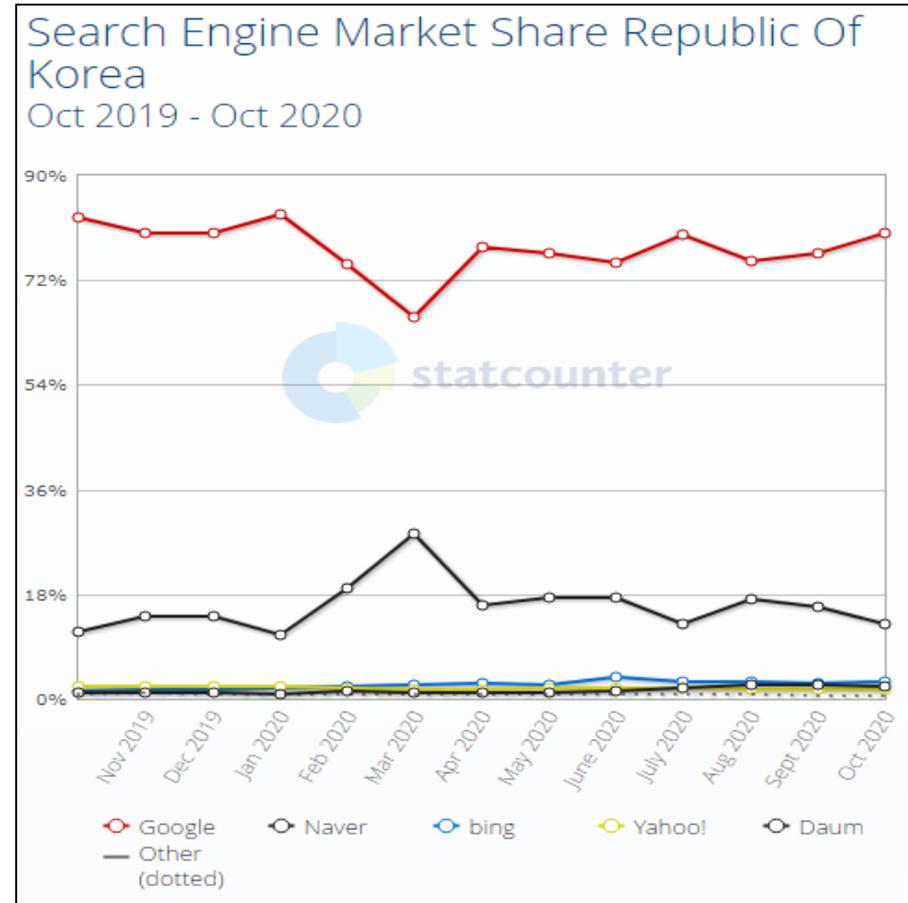
■ ロシアでの検索



■ 中国での検索



■ 韓国での検索

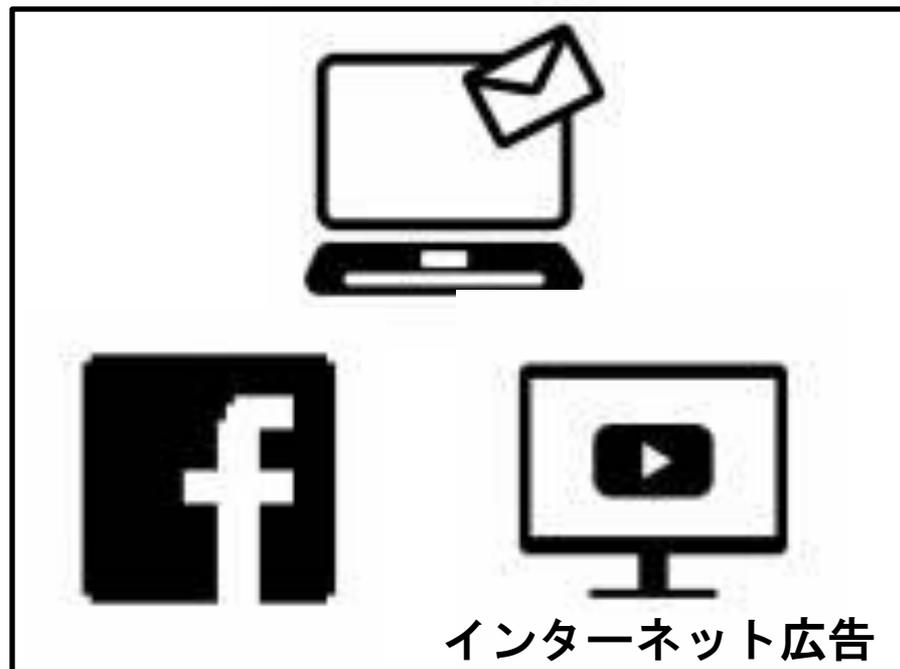


1-2 インターネット広告

広告の種類

人々に関心を持たせ購入を促進するために各種媒体を用いて商品・サービスの宣伝をすることを「広告」と言います。一般的に有料のもの言い、媒体により広告料が決まってきます。

主な広告の種類は、テレビ、新聞、ラジオ、雑誌といったマスコミ4媒体に加え、バナー広告、リスティング広告、アフィリエイト広告、SNS広告、動画広告といったインターネット広告があります。



広告の役割

沢山の種類が存在する広告ですが、その役割も異なります。
広告の役割は主に2つあります。

- ① 商品・サービスの存在を知らしめる役割 = 顧客認知向上
 - ② 商品・サービスを買ってもらう役割 = 購入促進
- この2つの役割を理解して、効果的に広告を活用します。

顧客認知向上

商品・サービスの存在を
知ってもらうこと

例：

こんな観光地があるのか。
知らなかったなあ・・・

購入促進

商品・サービスを購入する
アクションを後押しすること

例：

〇〇観光地で2日目に
そば打ち体験を予約しよう。

インターネット広告①

① バナー広告（＝ディスプレイ広告）

あらかじめ決められた枠内で、リンク付きの画像を表示し、その画像をクリックすると広告主のウェブサイトへ誘導される広告手法。ネットユーザの約9割にリーチできると言われています。潜在的顧客の訴求に効果的です。

The image shows a screenshot of the Yahoo! Japan homepage. At the top, there are navigation icons for Gole, Yafoku, Shopping, Yahoo! JAPAN, PayPay STEP, Cards, and Mail. Below these is a search bar with the text 'ウェブ 画像 動画 知恵袋 地図 リアルタイム 一覧▼' and a search button labeled '検索'. A navigation menu on the left lists various services like Shopping, PayPayモール, Yafoku, PayPayフリマ, ZOZOTOWN, LOHACO, Trabel, Ichi.com, Ichi.comレストラン, ニュース, 天気・災害, スポーツナビ, ファイナンス, and テレビ. The main content area features a 'ニュース' (News) section with a list of articles and a featured article titled '日本一イチョウ' (Japan's Best Chestnut) with a thumbnail image. Two large blue boxes with the white text '広告' (Advertisement) are overlaid on the page: one at the bottom center and one on the right side, partially covering the news content.

インターネット広告②

② リスティング広告

検索キーワードに応じて検索結果の上位に表示される広告をリスティング広告と言います。キーワードに対し入札し、オークション形式で掲載順位が決定します。検索連動型広告（PPC広告）ともいい、クリックされた分だけ課金されます。

Google

温泉旅館

約 219,000,000 件 (0.68 秒)

広告 · travel.rakuten.co.jp/楽天トラベル/温泉宿検索

全国各地の人気温泉宿を検索 - 毎月5と0のつく日は5%オフ

15,000円以上のご予約がクーポン利用で5%オフ。温泉宿と高級宿が対象。人気温泉宿のプランが満載。温泉宿を予約するなら5と0のつく日がお得。30,000軒の登録宿泊施設・人気！楽天限定プラン・1,000万件のお客さまの声・秋&冬旅SALE開催中。

人気温泉ランキング 2020年上半年 最も多くの人々が訪れた温泉地大発表	温泉宿検索 人気温泉宿簡単検索 オリジナルプラン満載
直前予約空室カレンダー 全国宿の空室簡単検索 今すぐ予約	週末空室検索 直前予約も可能 空室簡単検索

広告 · www.jalan.net/じゃらん.net/宿・ホテル予約

温泉宿の宿泊予約ならじゃらん - GoToトラベル最大35%OFF

GoToトラベルキャンペーン/対象施設の宿泊料金が最大35%OFF jalan.net。今夜の宿をスグ探せる・憧れの高級宿をお得に予約・会員登録でポイント貯まる・エリア検索で簡単予約・日本最大級の予約サイト・当日予約OKの宿多数・お得なクーポン配布中。

★★★★★ Rating for jalan.net: 4.1 - 166 reviews

温泉・露天風呂の宿予約・土曜日に泊まれるホテル・憧れの贅沢ホテル・旅館・早得プラン

インターネット広告③

③ リマーケティング広告

一度サイトに訪問したことがあるユーザーに対して、再アプローチし購入を促進する広告をリマーケティング広告と言います。特定の観光地を検索したユーザーに対して、再度広告を掲示することで、購買促進につながる可能性があります。

後日別サイトで広告掲載

インターネット広告の特徴

理解しておきたいインターネット広告の特徴は、
潜在的顧客にはバナー広告（＝ディスプレイ広告）が有効という点と、
顕在的顧客にはリスティング広告やリマーケティング広告が有効という
ことです。

潜在顧客

バナー広告
（＝ディスプレイ広告）

顕在顧客

リスティング広告
リマーケティング広告

個人演習

【インターネット検索の実践】

自分の地域名やの観光施設、お店の名前をGoogleで検索してみましよう。

何件表示されますか？バナー広告はありますか？

上位10のサイトはどのようなサイトでしょうか。

気づいたことをノートにまとめてみましよう。



デジタルマーケティングの基本2

2-1 ニューメディア

メディアの種類

■ 三つのメディア



Owned Mediaや
Earned Mediaに
見込み客を誘導する



深いコミュニケーションにより見込み客を
顧客に転換したりロイヤルティを強化する

Paid Media や Owned Media が
有効に機能したときに世評が発生する

出典：横山隆治著「トリプルメディアマーケティング」

オウンドメディア

公式ホームページ



公式SNS



公式ブログ



発信内容・発信タイミングを自社でコントロールできる一方、
単体サイトでは集客力が弱いという弱点もある。
また、管理するにはある程度のITリテラシーがある人材が必要となる。

ペイドメディア



集客力が高く、一度にたくさんの人に対して情報発信が可能である。
一方、広告コストが高いことや掲載期間が限定されるといったデメリットも。
また、昨今、広告に対する嫌悪感が消費者間に存在し、
広告は敬遠される傾向にもある。

アーンドメディア

口コミサイト



第三者のSNS



メディアの記事



第三者による発信コンテンツのため信用度が高く
昨今特に注目されているメディア。話題性があり、拡散力も高い。
一方、発信される内容自体は基本的にはコントロールが不可能なため、
ネガティブな情報は一気に拡散し企業のブランド毀損につながることもある。

2 - 2 SNSの活用

注目すべきSNS

SNSとは、ソーシャルネットワーキングサービス (Social Networking Service) の略称です。人と人とのつながり・ネットワークを構築・促進するコミュニティ型オンラインサービスのことです。

		リリース	アクティブユーザー (国内月間)	ユーザー層	特徴
LINE		2011/6	7,800万	<ul style="list-style-type: none"> 幅広い世代に利用されている 40代の割合が一番多い 	<ul style="list-style-type: none"> モバイル、スマートフォンメイン 通話機能とメッセージ機能がある
Twitter		2006/7	4,500万	<ul style="list-style-type: none"> 20代の割合が一番多い 	<ul style="list-style-type: none"> テキスト(つぶやき)メイン 実名登録でなくてもいい 【リツイート】機能があり、【拡散】向き リアルタイム性 ハッシュタグ「#」を使用 複数アカウントを利用するユーザーが多いため、純粋なUU数ではない
Instagram		2010/10	2,900万	<ul style="list-style-type: none"> 10代、20代の若年層にも人気 40代の割合が一番多い 5大SNSの中で最もユーザー数の伸びが大きい 	<ul style="list-style-type: none"> 写真メイン 実名登録でなくてもいい 知り合いでなくても興味のあるユーザーと繋がることのできる ハッシュタグ「#」を使用
Facebook		2006/9	2,800万	<ul style="list-style-type: none"> 比較的年齢層が高い(40代~50代の割合が大きい) 10代、20代が少ない(特に10代) 	<ul style="list-style-type: none"> 実名登録 基本1人1アカウント リアルな繋がり重視 ビジネスでの利用も多い 投稿内容は量より質重視
TikTok		2016/9	950万	<ul style="list-style-type: none"> 10代、20代の割合が過半数を占める 	<ul style="list-style-type: none"> 動画メイン(15秒) ハッシュタグ「#」を使用

- ・LINE：「LINE アカウント 2018年12~2019年3月期 媒体資料」、「平成30年12月期 第3四半期決算説明会 プレゼンテーション資料」より
- ・Instagram：Instagram公式サイト(2018年11月時点統計)
- ・Twitter：(2017年10月時点統計)
- ・Facebook：Facebook公式サイト(2017年9月時点統計)
- ・TikTok：(2019年2月時点統計)

フェースブック

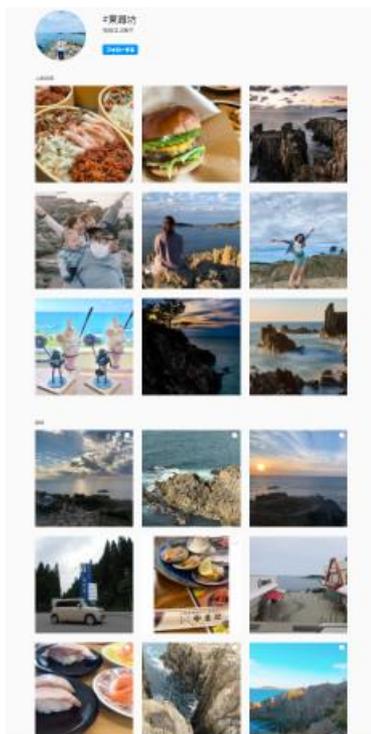
SNSはウェブサイトのような一方通行な情報発信とは異なり、ユーザー間の双方向なコミュニケーションや絞り込んだ相手とより深い関係が築きやすい媒体です。このことにより、観光マーケティング分野でも活用が可能であると注目されています。例えば、フェースブック広告では、日々公開される個人情報からターゲット属性を特定しやすく、また、低コストで顧客へ訴求することも可能です。



フェースブック広告で設定できるセグメント
地域（国、都道府県、市区町村）
性／年齢
学歴
家族構成
ライフイベント
勤務先
趣味・関心
行動（購入履歴等）
自社とつながりがある属人

インスタグラム

インスタグラム（Instagram）は写真や動画を投稿できるSNSサービスのことです。2015年から約3年で3倍のアクティブユーザーを獲得していて、国内のSNS市場でもっとも成長率の高いSNSといえます。「インスタ映え」という言葉が流行したように観光地をビジュアルで訴求するには優れたメディアと言えます。



東尋坊で検索すると、人々の東尋坊への興味関心をビジュアルで捉えることができます。

いくつかの検索ワードを#で検索すると、対象観光地の検索ボリュームがわかり、観光地の人気度を推測することも可能です。

#福井県
投稿63.1万件



YouTube

若い世代はテレビ視聴よりもYouTube視聴の時間の方が長い傾向にあります。今後は動画コンテンツを活用したプロモーションも有効な手段となり得るので検討が必要です。

1日あたりの平均視聴時間（単一回答）

TV視聴時間 (%)

		n=	3時間以上	30分~3時間未満	見ない/30分未満
全体		1,000	16.3	63.0	20.7
男性	13-15歳	100	11.0	72.0	17.0
	16-18歳	100	7.0	66.0	27.0
	19-22歳	100	8.0	55.0	37.0
	23-29歳	100	12.0	57.0	31.0
	30-34歳	100	18.0	68.0	14.0
女性	13-15歳	100	11.0	69.0	20.0
	16-18歳	100	14.0	68.0	18.0
	19-22歳	100	20.0	57.0	23.0
	23-29歳	100	28.0	61.0	11.0
	30-34歳	100	34.0	57.0	9.0

YouTube視聴時間 (%)

		3時間以上	30分~3時間未満	見ない/30分未満
全体		20.3	56.0	23.7
男性	13-15歳	24.0	64.0	12.0
	16-18歳	24.0	64.0	12.0
	19-22歳	33.0	54.0	13.0
	23-29歳	20.0	58.0	22.0
	30-34歳	11.0	58.0	31.0
女性	13-15歳	16.0	65.0	19.0
	16-18歳	25.0	60.0	15.0
	19-22歳	25.0	53.0	22.0
	23-29歳	15.0	45.0	40.0
	30-34歳	10.0	39.0	51.0

個人演習

【インターネット検索の実践】

自地域にある「観光地」や「観光施設」を一つ選び、インターネットで検索してみましょう。その際にオウンドメディア、ペイドメディア、アードメディアそれぞれどのようなサイトが存在し、どのように紹介されるいるかを確認しましょう。ワークには、次ページのシートを活用しましょう。



個人演習①

■ オウンドメディアの種類と紹介されている内容、特徴

■ ペイドメディアの有無とその内容、特徴

■ アーンドメディアの有無とその内容、特徴

【動画・画像検索の実践】

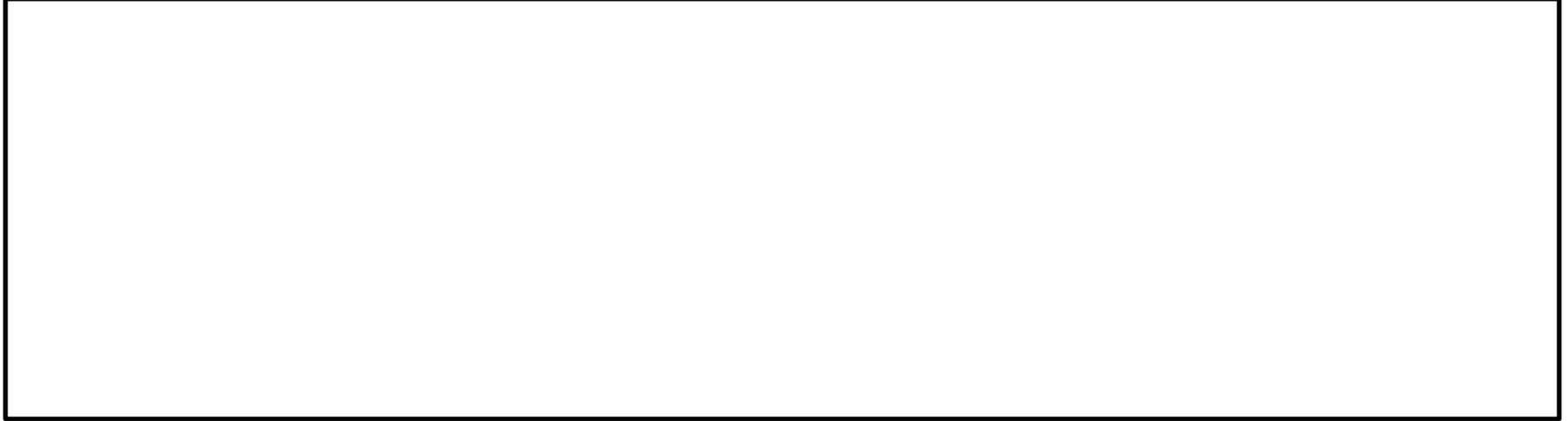
先ほど選んだ観光地、観光施設について、InstagramやYouTubeでも検索をしてみましょう。

どのような動画・画像が多く出てくるか、また出てきたビジュアルの特徴を次ページのシートにまとめましょう。



個人演習②

■ Instagramで検索される画像の数、内容、特徴



■ YouTubeで検索される動画の数、内容、特徴



オープンデータ・ビッグデータの活用方法



オープンデータの確認

公開されているオープンデータは多数存在しますが、官公庁や自治体が公開しているデータを紹介します。

- ① 山口県オープンデータカタログサイト
山口県及び県内市町が保有するオープンデータ
- ② 観光庁が集計するデータ
「観光白書」「宿泊旅行統計調査」「インバウンド消費動向調査」等
- ③ インバウンドの動向に関する調査報告書
JNTOデータ（例：訪日旅行DATAハンドブック）
- ④ 都道府県単位での観光に関する調査報告書
各都道府県発行「観光動態調査」 *都道府県により名称が異なります

山口県オープンデータカタログサイト

山口県及び県内市町が保有するオープンデータが登録されているサイトです。自治体単位、カテゴリ単位で選択ができるため、必要とする地域、分野に絞ってデータを確認することが可能です。

見たい分野、地域などを選択して、その数値を読み解く力を修得することが重要です。

<https://yamaguchi-opendata.jp/>

The screenshot shows the homepage of the Yamaguchi Open Data Catalog Site. The header features the site name '山口県オープンデータカタログサイト' in large white text on a dark red background. Below the header, there are three statistics: 759 データ一覧, 14 カテゴリ, and 289 タグ. A search bar with the placeholder 'キーワードで検索' and a magnifying glass icon is present. Below the search bar is a button labeled 'オープンデータマップ' with a location pin icon. At the bottom, there are four navigation buttons: 'データ一覧', 'ダッシュボード山口', '活用事例', and 'APIについて'.

The screenshot shows the search results page for 'データセット' on the Yamaguchi Open Data Catalog Site. The header includes the site name and navigation links like 'Top', 'AI検索', 'データセット', '団体', 'カテゴリ', 'タグ', 'About', and a search bar. The main content area shows a list of municipalities with their respective data set counts: 山口県 (99), 宇部市 (79), 下松市 (48), 周南市 (45), 山口市 (44), 下関市 (41), 岩国市 (39), 平生町 (36), 萩市 (35), and 上関町 (33). Below the list, there are sections for '検索結果を出力' and '759 件のデータセットが見つかりました'. There are also links to '【山口県】業務委託契約情報' and '【山口県】物品調達等契約情報', both with 'CSV' download options. At the bottom, there is a section for '【防府市】ボーリング柱状図' with a '防府市のボーリング柱状図' link.

出典：山口県オープンデータカタログサイト

山口県オープンデータカタログサイト

「山口オープンデータカタログサイト」<https://yamaguchi-opendata.jp/>で下記の項目について調べてみましょう。

・データ一覧から、「**山口市**」、カテゴリ「**くらし**」、「**【山口市】その他公共施設**」のデータを「**取得**」ボタンからプレビューをクリックし、どのようなデータが登録されているか確認しましょう。

・カテゴリから、自分が気になる分野（くらし、観光、福祉等）を選択し、気になる項目を選択して、データを「取得」ボタンからプレビューしましょう。

・その他に自分が気になる項目を調べてみましょう。

The screenshot shows the Yamaguchi Open Data Catalog Site interface. The header is red with the site name and navigation links. The main content area shows search results for '観光' (Tourism). A search bar contains '観光'. Below it, it says '4件のデータセットが見つかりました' (4 data sets found). The filters show '団体: 山口県 x' and 'カテゴリ: 観光 x'. There are two data set entries: '観光' with 0 followers and 4 data sets, and '観光' with 4 data sets. The first entry has a preview of '【山口県】都市計画基礎調査結果' (Yamaguchi Prefecture Urban Planning Basic Survey Results) and a download button for 'xlsx'. The second entry has a preview of '【山口県】観光施設一覧' (Yamaguchi Prefecture Tourism Facilities List) and download buttons for 'xlsx' and 'csv'.

観光庁が集計するデータ

日本の観光動向に関するオープンデータは、国土交通省 観光庁が公表している「観光白書」の調査結果が有効です。

観光に関するデータは、旅行業、宿泊業、飲食業、レジャー・娯楽産業、交通業、土産物産業など、多くの業種に関係する重要なものです。観光客の消費活動は、地域経済に大きく貢献するものですので、どのようなデータがあるかを把握することが重要です。



観光庁が毎年発行する「観光白書」は、観光立国推進基本法の規定に基づき、国土交通省が発行している資料です。

観光全般に関するデータや、過去に講じられた政策、今年度に講じられる政策がまとめられています。

近年の観光の動向や、観光がもたらす経済効果といったデータが多数のグラフと共に掲載されるとともに、幅広い観点から分析されています。

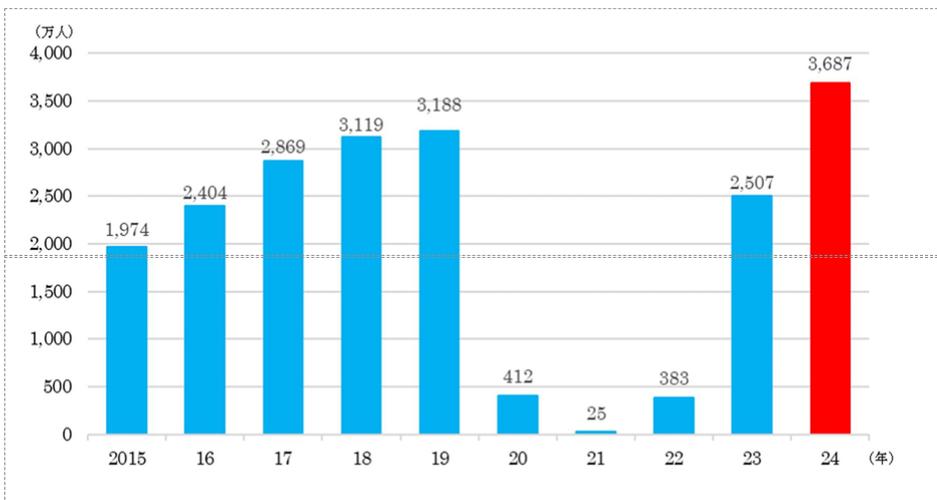
観光白書

「観光白書」の最新年度の「本文 第Ⅰ部 観光の動向」のページに掲載されているグラフから、読み取れることを確認しましょう。

まずは、訪日外国人旅行者のグラフを参照します。

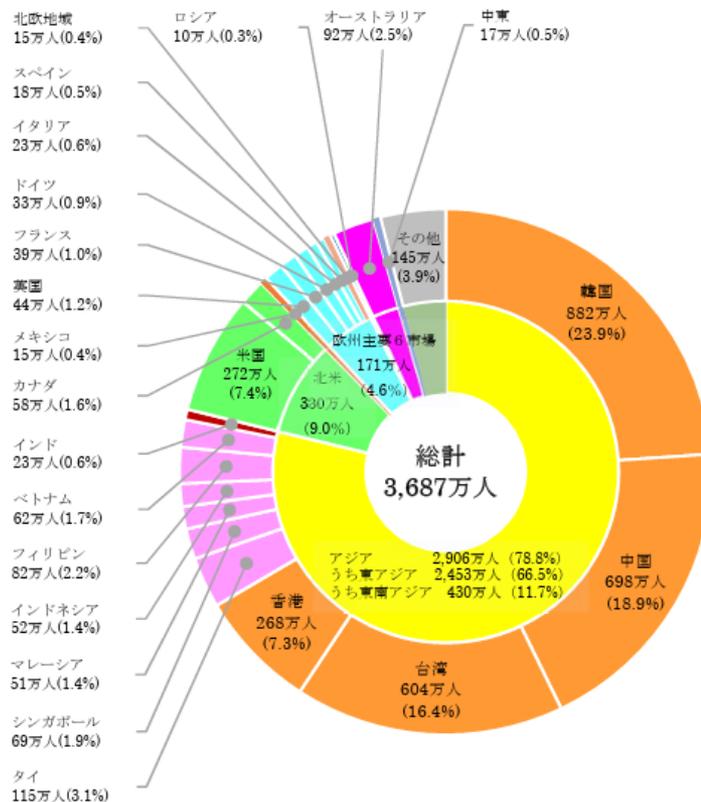
グラフには書かれていない、社会環境、社会動向についても確認することで、データから読み取れる内容が深くなります。

図表Ⅰ-7 訪日外国人旅行者数の推移



資料：日本政府観光局（JNTO）資料に基づき観光庁作成。

（図表Ⅰ-8）訪日外国人旅行者の内訳（2024年）



資料：日本政府観光局（JNTO）資料に基づき観光庁作成。

注1：（）内は、訪日外国人旅行者数全体に対するシェア。

注2：「その他」には、アジア、欧州等各地域の国であっても記載のない国・地域が含まれる。

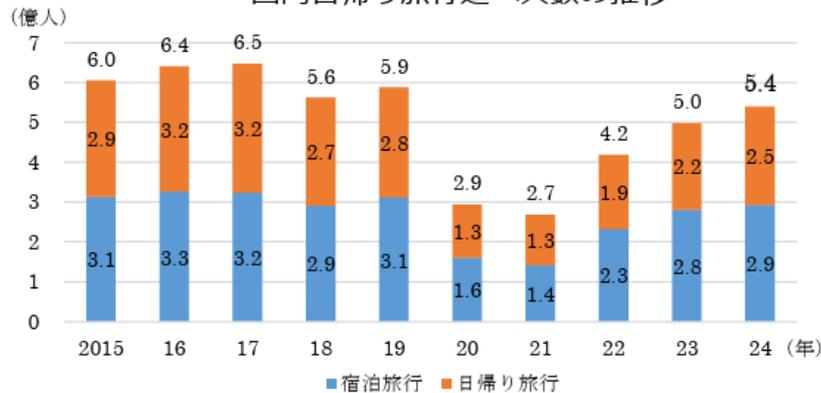
注3：四捨五入した値を表示しているため、個々の数値の合計は必ずしも総数と一致しない。

出典：観光庁「観光白書」

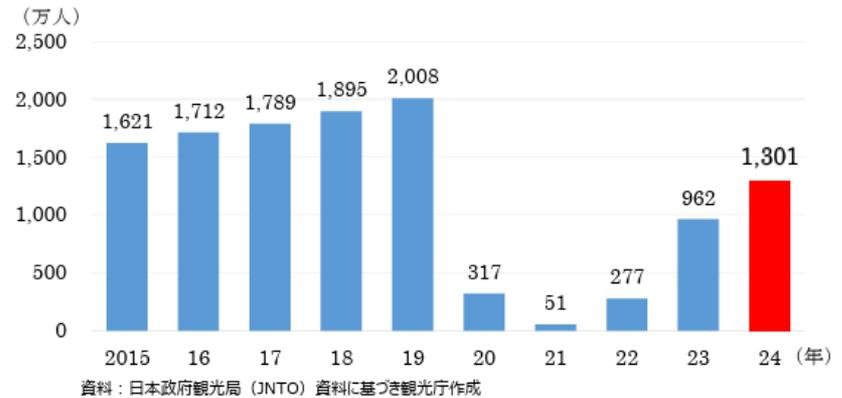
観光白書

次に、日本人の観光についてのグラフを確認し、そこから読み取れることを書き出しましょう。
以下に示すグラフの前後に掲載されている日本人の旅行に関するグラフも参照してみましょう。

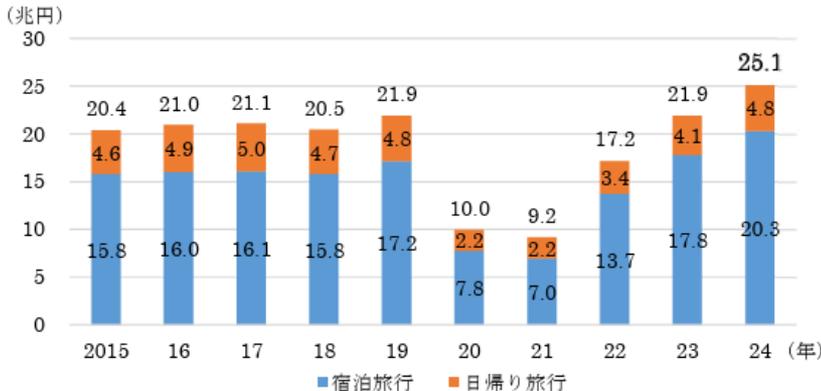
(図表 I -16) 日本人国内宿泊旅行延べ人数及び国内日帰り旅行延べ人数の推移



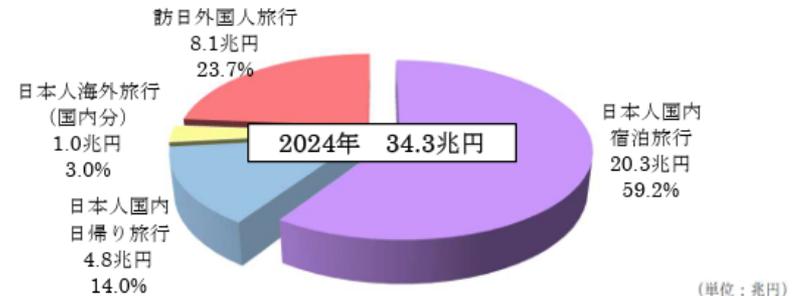
(図表 I -14) 出国日本人数の推移



(図表 I -17) 日本人国内旅行消費額の推移



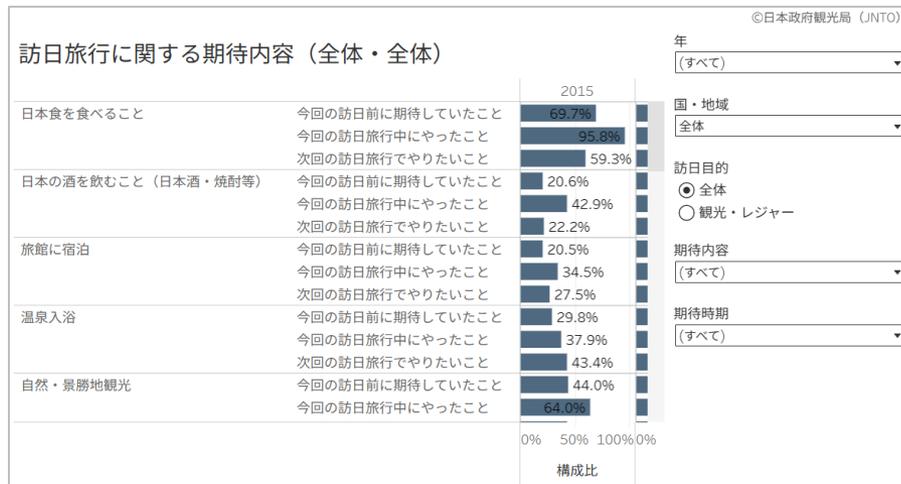
(図表 I -18) 日本国内における旅行消費額 (2024年)



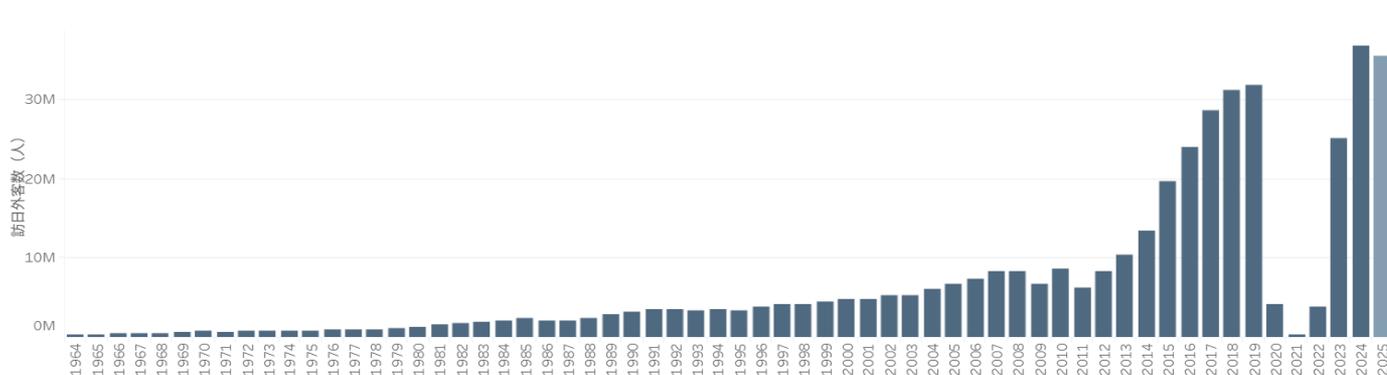
	2015年	16	17	18	19	20	21	22	23	24
日本人国内宿泊旅行	15.8	16.0	16.1	15.8	17.2	7.8	7.0	13.7	17.8	20.3
日本人国内日帰り旅行	4.6	4.9	5.0	4.7	4.8	2.2	2.2	3.4	4.1	4.8
日本人海外旅行 (国内分)	1.0	1.1	1.2	1.1	1.2	0.3	0.1	0.6	0.9	1.0
訪日外国人旅行	3.5	3.7	4.4	4.5	4.8	0.7	0.1	0.9	5.3	8.1
合計	24.8	25.8	26.7	26.1	27.9	11.0	9.4	18.7	28.1	34.3

日本の観光統計データ

日本政府観光局（JNTO）では、訪日外国人旅行者に特化した統計データを公表しています。2023年版までは、「訪日旅行データハンドブック」としてPDFで発行されていましたが、現在はWebサイト「日本の観光統計データ」（<https://statistics.jnto.go.jp/>）で公開しています。



年別 訪日外客数の推移



出典：「日本の観光統計データ」

RESAS（リーサス）

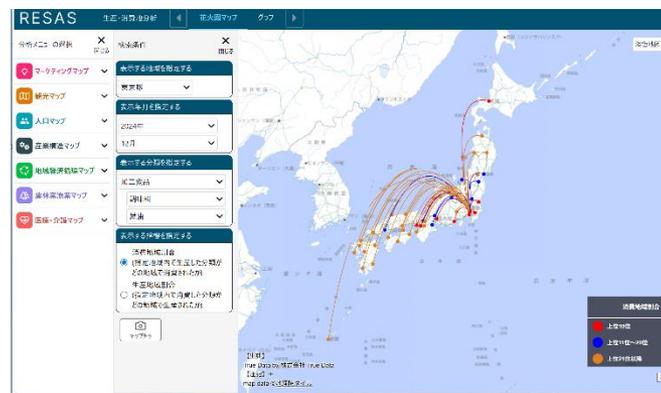
経済産業省・内閣府が提供している地域経済に関する様々なビッグデータ（産業の強み、人の流れ、人口動態など）のことをRESAS（リーサス）といいます。地図やグラフで分かりやすく「見える化（可視化）」したシステムで、誰でも視覚的に活用できる地域経済分析システムです。

RESAS 地域経済分析システム

仮説を根拠へ

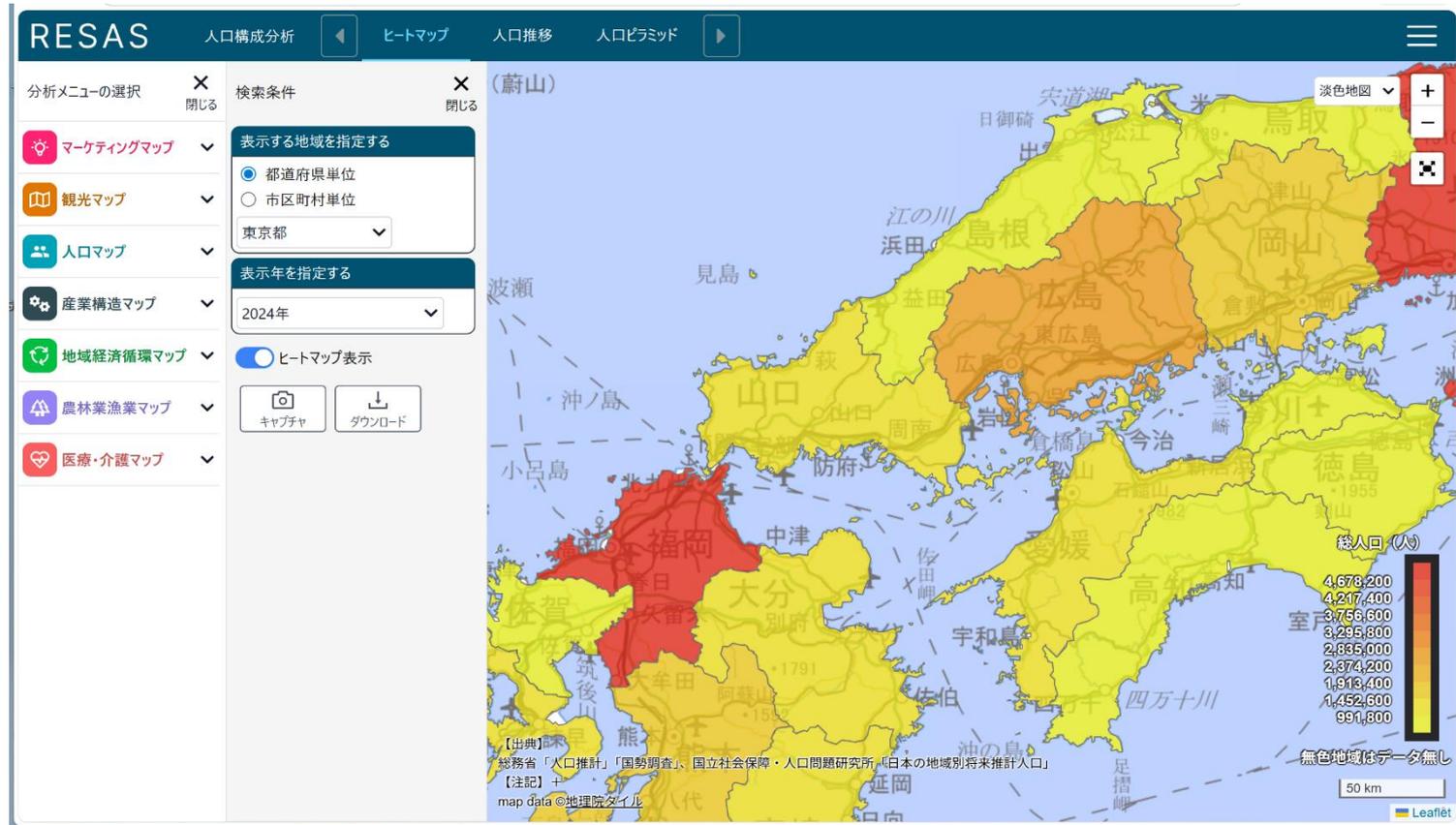
RESASは、地域経済に関するビッグデータを地図上やグラフで見える化できる政府のシステムです。

分析を開始する



RESASのデータ

人口マップ



出典：内閣府 地方創生推進室 ビッグデータチーム

指定した都道府県などの人口構成、増減、将来人口予測など、人口に関するデータを取得することが可能です。

観光予報プラットフォーム

観光事業者に対して、宿泊を基軸にした「観光」に関連する必要とされるデータを提供するデータプラットフォームが「観光予想プラットフォーム」です。観光に関する各種ビックデータをエリア単位でまとめて提供しています。会員登録をする必要があり、有料会員と無料会員の2タイプの会員種別が存在します。これらのデータを活用することで様々な観光マーケティングの戦略を作ることが可能となります。



資料ダウンロード

活用事例ダウンロード

観光予報
はじめました

◆新着情報

- 2020/10/2 観光予報10月更新情報のお知らせ
- 2019/12/19 STEAM学習サービス『観光予報DS (DataScience)』のお知らせ
- 2019/9/18 消費税法改正の対応に関するお知らせ
- 2018/12/19 観光需要予測モデル構築支援のお知らせ

The graphic features a central map of Japan with various icons representing different aspects of tourism: a sun, a heart, a person, a group of people, a person walking, a cloud, and a sun with a cloud. The text '観光予報はじめました' is prominently displayed in the center.

個人演習

個人演習①

パソコンでRESASを使ってみましょう。

RESASの「外国人訪問分析」を実施して、2023年に山口県を訪問した国・地域ベスト5を探してみましょう。

2023年山口県を訪問した国・地域ベスト5

1位	
2位	
3位	
4位	
5位	

個人演習②

RESASの「国内目的地分析」を実施します。

2022年3月（休日）に山口県内で自動車で訪問した目的地（観光箇所）ベスト5を探してみましょう。

2022年3月（休日）、山口県内を自動車で訪問した目的地ベスト5

1位	
2位	
3位	
4位	
5位	

オープンデータから分析する地域特性・顧客特性

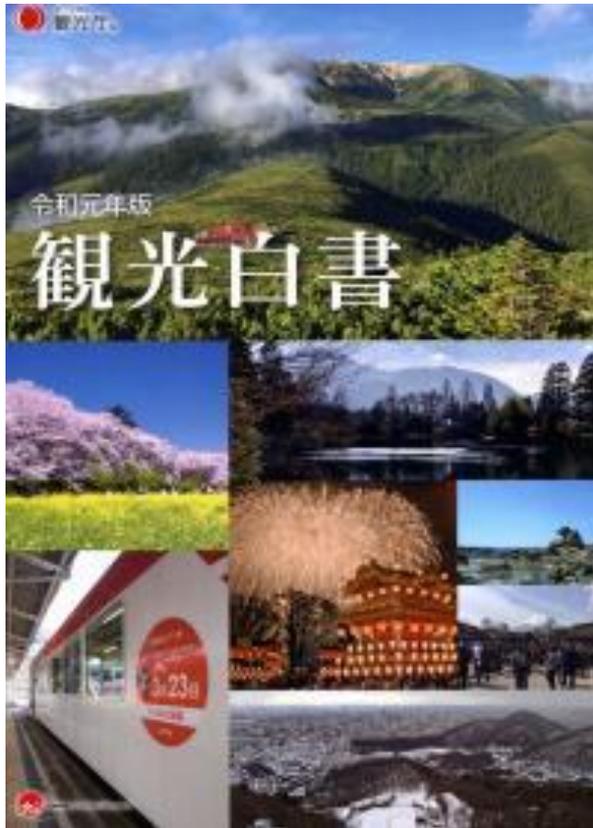


観光庁データから観光・地域に関する情報を分析する手法

観光庁データを活用した情報収集・分析

【使用するデータ＝観光白書】

観光庁が発行する観光白書を活用して、日本の観光や地域の観光に関する情報を分析するやり方を学びます。



出典：観光庁 「観光白書」

【観光白書とは】

観光庁が毎年発行する「観光白書」は、観光立国推進基本法の規定に基づき、国土交通省が発行している資料です。

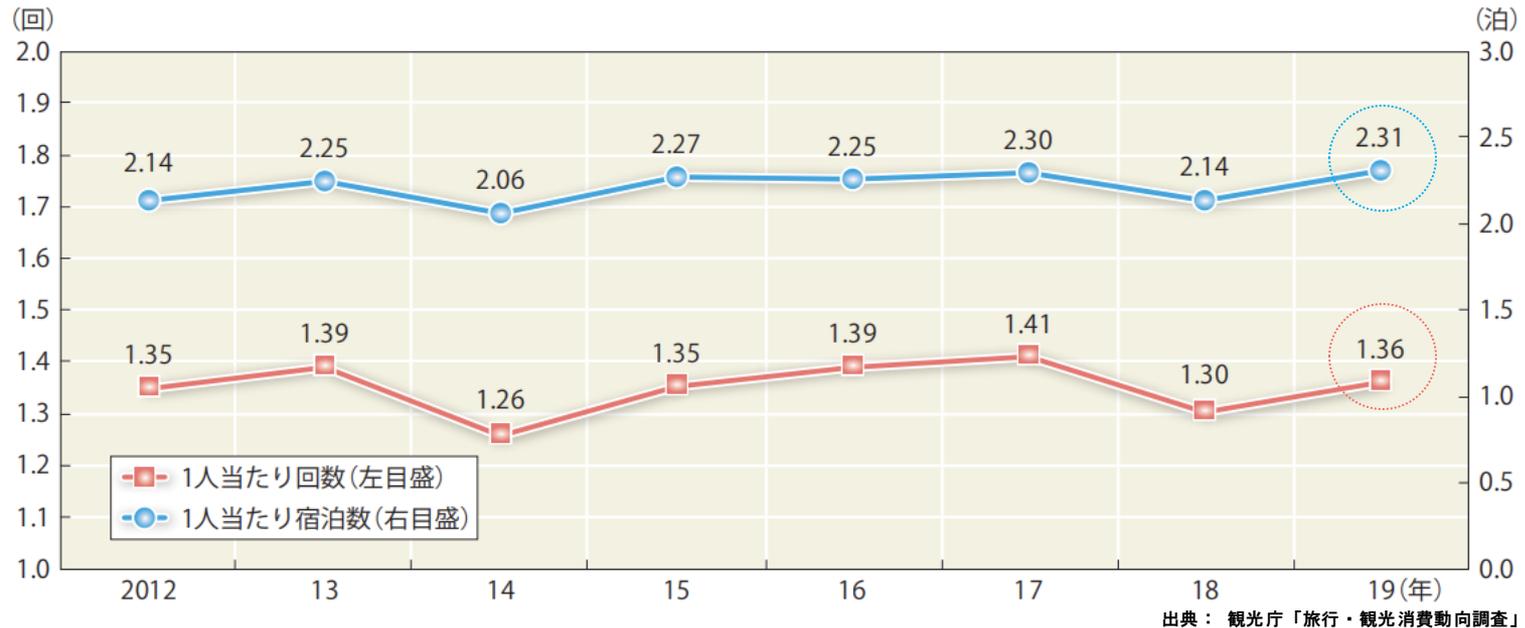
観光全般に関するデータや、過去に講じられた政策、今年度に講じられる政策がまとめられています。

近年の観光の動向や、観光がもたらす経済効果といったデータが多数のグラフと共に掲載されるとともに、幅広い観点から分析されています。

**①観光白書から日本の観光に関する
現状・課題を分析する方法
(P7～P11で実際に分析をします)**

①日本人国内宿泊観光旅行の回数及び宿泊数の推移

【日本人の1人当たりの旅行実態（宿泊数と旅行回数）を把握する方法】
観光白書から日本人の1人当たりの宿泊観光旅行の宿泊数と旅行回数を導き出します。



日本人の国民1人当たりの国内宿泊観光旅行の宿泊数を表しており、2019年（令和元年）では1回の旅行で平均2.31泊しました。



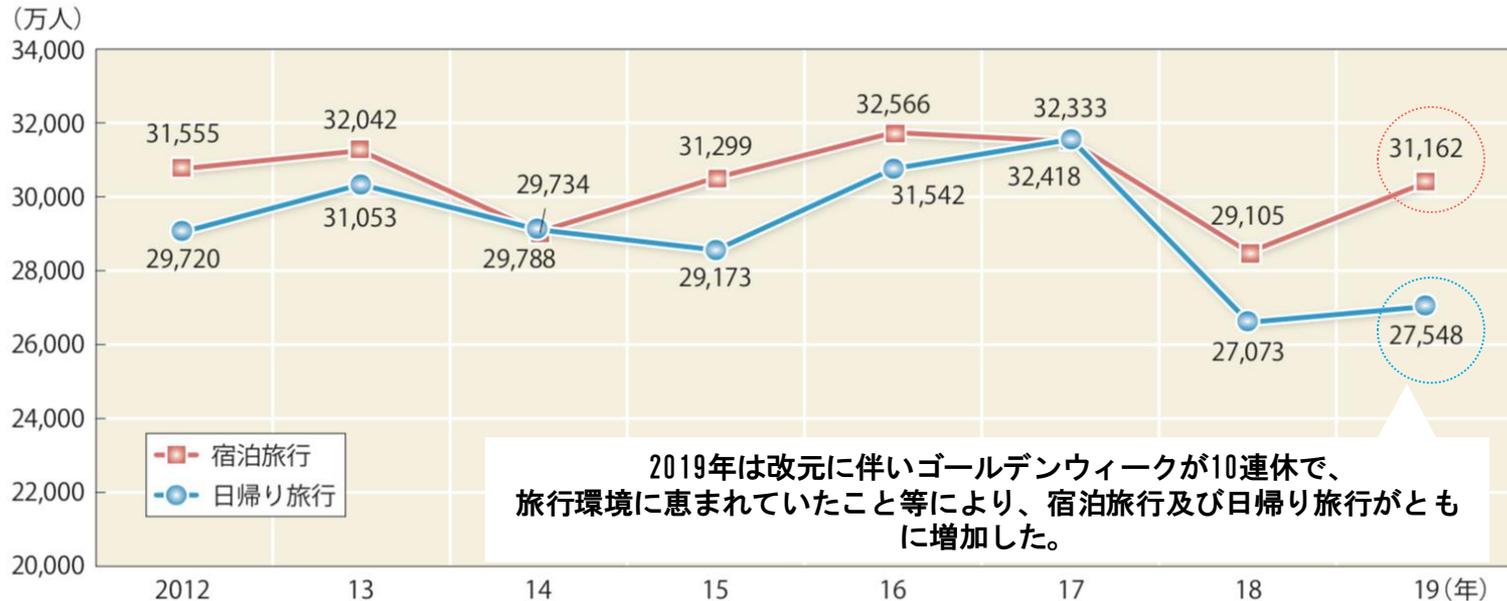
国民1人当たりの国内宿泊観光旅行の回数を表しており、2019年（令和元年）は平均で年に1.36回旅行しました。

日本人の宿泊数・宿泊回数は、過去5年以上、ほぼ横ばいで伸び悩んでいます。

② 日本人国内宿泊旅行延べ人数、国内日帰り旅行延べ人数の推移

【日本人の宿泊者数と日帰り旅行者数を把握する方法】

観光白書から日本人の国内旅行における宿泊旅行者数と日帰り旅行数の合計を導き出します。



出典：観光庁「旅行・観光消費動向調査」

● 日本人の国内宿泊旅行者数を示しており、
2019年は延べ3億1,162万人が宿泊しました。

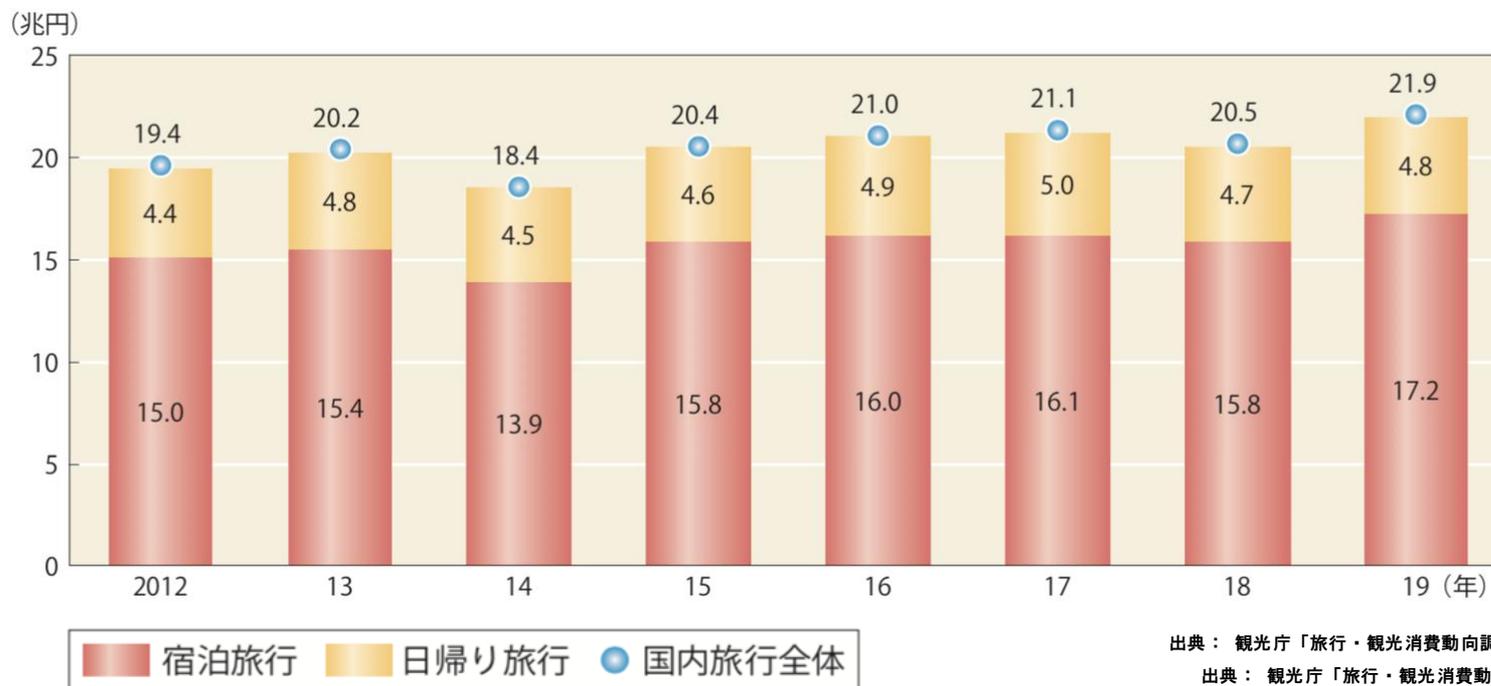
● 国内日帰り旅行者数を示しており、
2019年は延べ2億7,548万人が日帰りで観光しました。

宿泊・日帰り人数も多少の波はあるものの、全体的にほぼ横ばいで推移しています。

③ 日本人国内旅行消費額の推移

【日本人国内旅行消費額を把握する方法】

観光白書から日本人の国内旅行の各種別における消費額とその合計を導き出します。

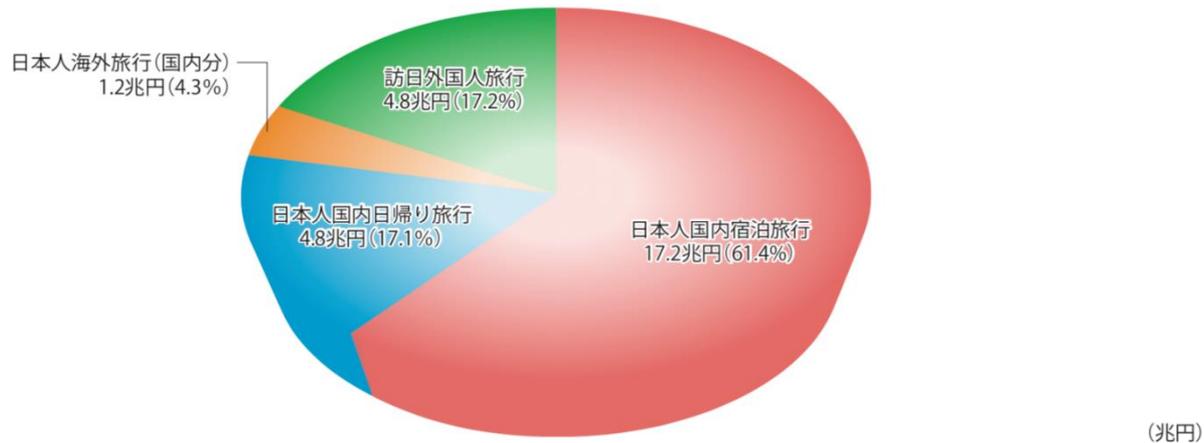


2019年の日本人国内旅行消費額は、宿泊旅行及び日帰り旅行ともに増加し、21.9兆円(前年比 7.1%増)でした。

国内旅行消費額はほぼ横ばいに推移していたが、2019年には増加傾向にあります。

④ 日本国内における旅行消費額

【日本人国内における旅行消費額を把握する方法】
観光白書から日本国内旅行における各種別の旅行消費額の合計を導き出します。



	2012年	13	14	15	16	17	18	19
日本人国内宿泊旅行	15.0	15.4	13.9	15.8	16.0	16.1	15.8	17.2
日本人国内日帰り旅行	4.4	4.8	4.5	4.6	4.9	5.0	4.7	4.8
日本人海外旅行(国内分)	1.3	1.2	1.1	1.0	1.1	1.2	1.1	1.2
訪日外国人旅行	1.1	1.4	2.0	3.5	3.7	4.4	4.5	4.8
合計	21.8	22.8	21.6	24.8	25.8	26.7	26.1	27.9

出典：観光庁「旅行・観光消費動向調査」

2019年の日本国内における旅行消費額は、27.9兆円となった。日本人による旅行消費額は23.1兆円、訪日外国人旅行者による旅行消費額は4.8兆円であり、訪日外国人による旅行消費額は3年連続で全体の15%を超えました。

日本人旅行者消費額よりも、訪日外国人旅行消費額の方が伸び率が高いです。

観光白書から日本の観光に関する情報を分析

日本人国内宿泊観光旅行の回数及び宿泊数の推移

横ばいで推移

日本人国内宿泊旅行延べ人数、国内日帰り旅行延べ人数の推移

横ばいで推移

日本人国内旅行消費額の推移

横ばいで推移

日本国内における旅行消費額

横ばいで推移

ポイント

少子高齢化による日本国内の旅行需要

減少傾向

ポイント

訪日外国人観光客の旅行消費による需要

増加傾向

日本国内における旅行消費額のうち訪日外国人旅行者の消費額のシェアは2012年では5.0%でしたが、2019年になると17.3%と増加しています。

分析結果

**訪日外国人観光客（インバウンド）は
増加傾向にあり、今後も期待が持てる**

**②観光白書から特定項目の現状や
特定の地域（＝沖縄）の観光に関する
現状を分析する方法
（P13～P22で実際に分析をします）**

⑤ 日本人・外国人の延べ宿泊者数の推移

【日本人と外国人の宿泊者数の推移を把握する方法】
観光白書から日本人と外国人の延べ宿泊者数の推移を導き出します。



日本国内のホテル・旅館等における延べ宿泊者数は、2019年は5億4,324万人泊でした。そのうち、日本人延べ宿泊者数は4億4,180万人泊、外国人延べ宿泊者数は1億143万人泊となり、外国人延べ宿泊者数は調査開始以降初めて1億人泊を突破しました。

訪日外国人宿泊者数よりも日本人宿泊者数の方が圧倒的に多いです。
一方、日本人宿泊者数よりも訪日外国人宿泊者数の方が増加率が高いです。

⑥ 三大都市圏及び地方部の外国人延べ宿泊者数の推移

【三大都市圏及び地方部の外国人延べ宿泊者数の推移を把握する方法】
観光白書から三大都市圏と地方部の外国人延べ宿泊者数の推移を導き出します。



出典：観光庁「宿泊旅行統計調査」

外国人延べ宿泊者数を三大都市圏と地方部で比較すると、
2019年は三大都市圏では 6,223万人泊、地方部では3,921万人泊となりました。

※三大都市圏とは、「東京、神奈川、千葉、埼玉、愛知、大阪、京都、兵庫」の8都府県を、地方部とは、三大都市圏以外の道県をいう。

外国人観光客は主に首都圏の宿泊が増加傾向にあり、地方部は横ばいとなっています。

⑦ 宿泊施設タイプ別の外国人延べ宿泊者数の割合の推移

【宿泊施設タイプ別の外国人延べ宿泊者数の割合を把握する方法】
観光白書から宿泊施設タイプ別の外国人延べ宿泊者数の割合推移を導き出します。



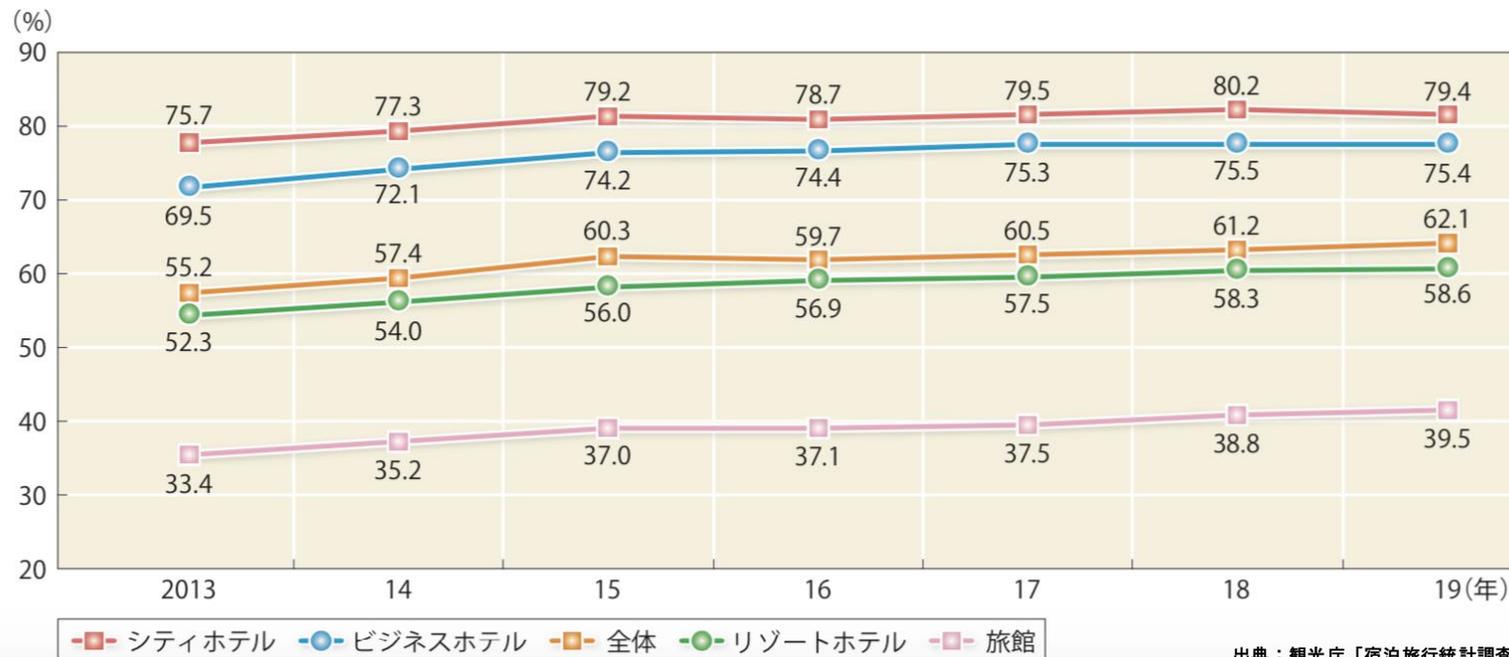
出典：観光庁「宿泊旅行統計調査」

延べ宿泊者数全体に占める外国人の割合をみると 2019年は18.7%でした。
宿泊施設タイプ別では、特にシティホテルにおける外国人の割合が高く、
2013年は20.4%であったが2019年には37.1%にまで上昇しました。

シティホテルの宿泊者の約4割が外国人観光客です。

⑧ 宿泊施設タイプ別の客室稼働率の推移

【宿泊施設タイプ別の客室稼働率の推移を把握する方法】
観光白書から宿泊施設タイプ別の客室稼働率の推移を導き出します。



宿泊施設タイプ別の客室稼働率は、特にシティホテル、
ビジネスホテルが高い水準にあり、2019はそれぞれ79.4%、75.4%となっています。

旅館の稼働率は他の宿泊タイプと比較して低い水準にあります。

観光白書から特定項目に関する情報を分析

訪日外国人宿泊者数よりも日本人宿泊者数の方が圧倒的に多い

ポイント

コロナ禍は国内市場注視

外国人観光客は主に首都圏の宿泊が増加傾向

ポイント

地方部として特色を出す必要がある

シティホテルの宿泊者の約4割が外国人観光客

ポイント

インバウンドに代わるターゲット選定が必要

旅館の稼働率は他の宿泊タイプと比較して低い水準にある

ポイント

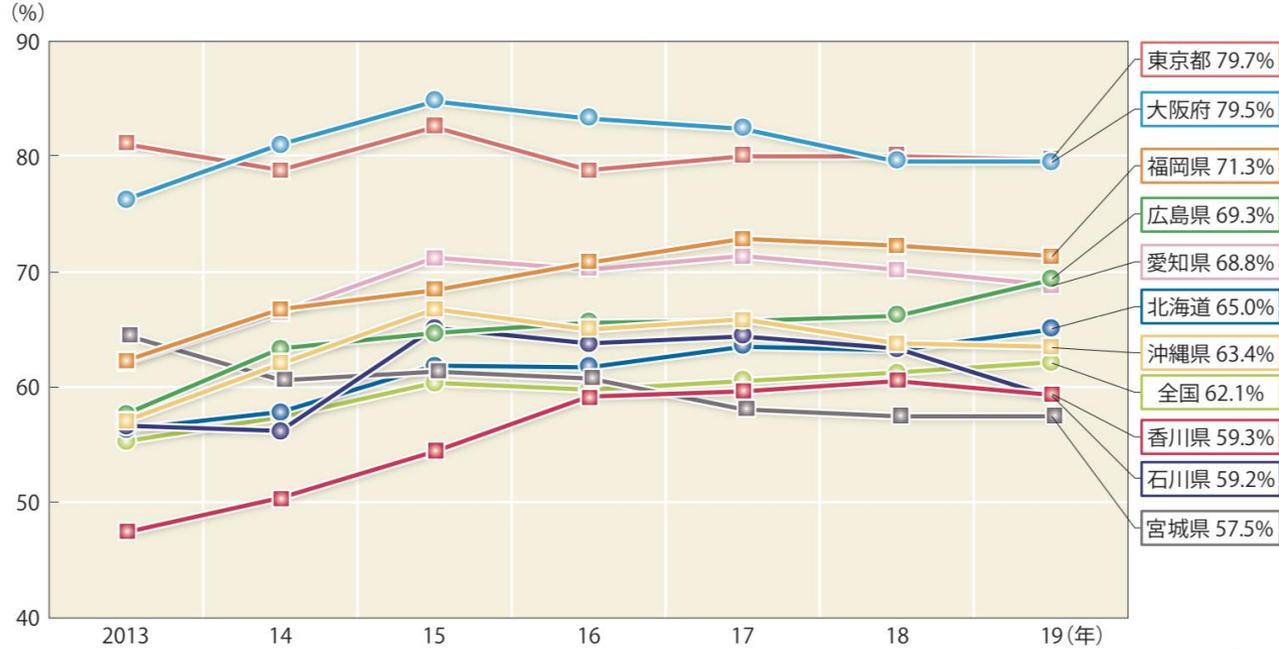
旅館のプロモーションが重要



自分が担当する観光ビジネスに関しても様々な観光データを分析することで傾向と対策が導き出せます。

⑨ 各県の客室稼働率の推移

【沖縄県の客室稼働を把握する方法】
観光白書から沖縄県の客室稼働率の推移を導き出します。



出典：観光庁「宿泊旅行統計調査」

客室稼働率は2015年の60.3%から2016年は59.7%に減少したが、2017年以降は増加傾向にあり、2019年は62.1%となりました。また、東京都と大阪府の客室稼働率は引き続き高い水準にあり、2019年はそれぞれ79.7%、79.5%となっています。

沖縄県の客室稼働率は東京都、大阪府などの都市圏と比べると低い状況にある。

⑩ 都道府県別・宿泊施設タイプ別客室稼働率

2019年

	宿泊施設タイプ					
	全体	旅館	リゾートホテル	ビジネスホテル	シティホテル	簡易宿所
全国	62.1	39.5	58.6	75.4	79.4	30.1
北海道	65.0	50.2	51.7	75.8	77.6	35.0
青森県	56.3	35.7	46.9	69.2	62.3	34.9
岩手県	55.4	39.3	52.8	72.9	73.1	23.7
宮城県	57.5	41.5	51.1	65.3	62.1	29.6
秋田県	52.1	36.8	43.3	67.8	71.8	18.9
山形県	51.7	39.2	31.4	69.3	68.9	11.6
福島県	51.0	42.6	52.1	66.5	68.7	15.4
茨城県	58.3	33.9	36.6	72.1	65.2	30.1
栃木県	50.8	42.7	51.9	68.2	66.0	14.4
群馬県	53.7	47.4	45.3	75.7	67.0	12.5
埼玉県	69.9	44.7	41.5	75.3	75.7	35.7
千葉県	71.2	40.5	85.4	73.9	82.0	23.1
東京都	79.7	44.9	65.9	84.2	84.2	50.8
神奈川県	70.7	51.4	67.1	83.4	84.6	52.7
新潟県	45.2	26.6	37.3	67.0	68.1	13.4
富山県	51.5	29.4	40.8	63.3	66.9	21.2
石川県	59.2	46.5	51.5	69.2	71.5	25.6
福井県	51.5	36.8	48.4	77.0	60.9	14.6
山梨県	46.7	42.5	59.3	73.0	65.6	21.4
長野県	38.5	28.1	41.6	71.0	75.0	13.1
岐阜県	55.5	43.3	49.0	77.3	75.2	28.6
静岡県	58.5	47.0	54.5	74.1	75.4	22.1
愛知県	68.8	30.3	49.6	75.6	76.9	37.9
三重県	54.2	39.5	49.8	72.1	72.0	10.6
滋賀県	53.2	38.7	57.4	66.0	75.0	16.7
京都府	67.2	43.9	56.2	80.5	82.2	41.7
大阪府	79.5	36.2	91.0	79.9	85.2	66.1
兵庫県	56.5	37.2	57.6	75.5	75.8	13.2
奈良県	42.4	28.3	64.4	57.8	70.5	24.7
和歌山県	48.7	36.8	56.5	72.4	73.2	21.9
鳥取県	50.5	45.4	45.1	69.0	76.6	13.9
島根県	56.3	46.5	41.0	73.5	72.2	12.3
岡山県	59.9	30.4	36.1	83.0	65.7	18.2
広島県	69.3	45.0	61.7	79.3	85.9	34.0
山口県	52.7	33.4	49.0	66.9	67.4	13.2
徳島県	50.4	25.2	47.3	69.4	63.7	21.6
香川県	59.3	47.2	68.0	68.4	65.7	30.2
愛媛県	55.9	49.4	61.3	65.8	72.8	14.9
高知県	50.9	37.9	48.5	68.4	71.0	12.9
福岡県	71.3	29.5	41.8	79.0	80.5	41.5
佐賀県	59.3	45.3	68.0	78.1	66.7	24.6
長崎県	55.3	39.9	57.7	74.4	72.3	14.6
熊本県	60.9	44.2	48.7	79.3	78.2	28.5
大分県	54.9	46.5	62.6	67.6	68.6	21.2
宮崎県	54.8	35.8	45.2	66.7	70.2	14.2
鹿児島県	56.1	38.1	48.0	74.3	71.5	23.4
沖縄県	63.4	22.3	69.6	74.4	80.0	30.0

【沖縄県の宿泊施設タイプ別客室稼働率を把握する方法】
観光白書から沖縄県の宿泊施設タイプ別客室稼働率を導き出します。

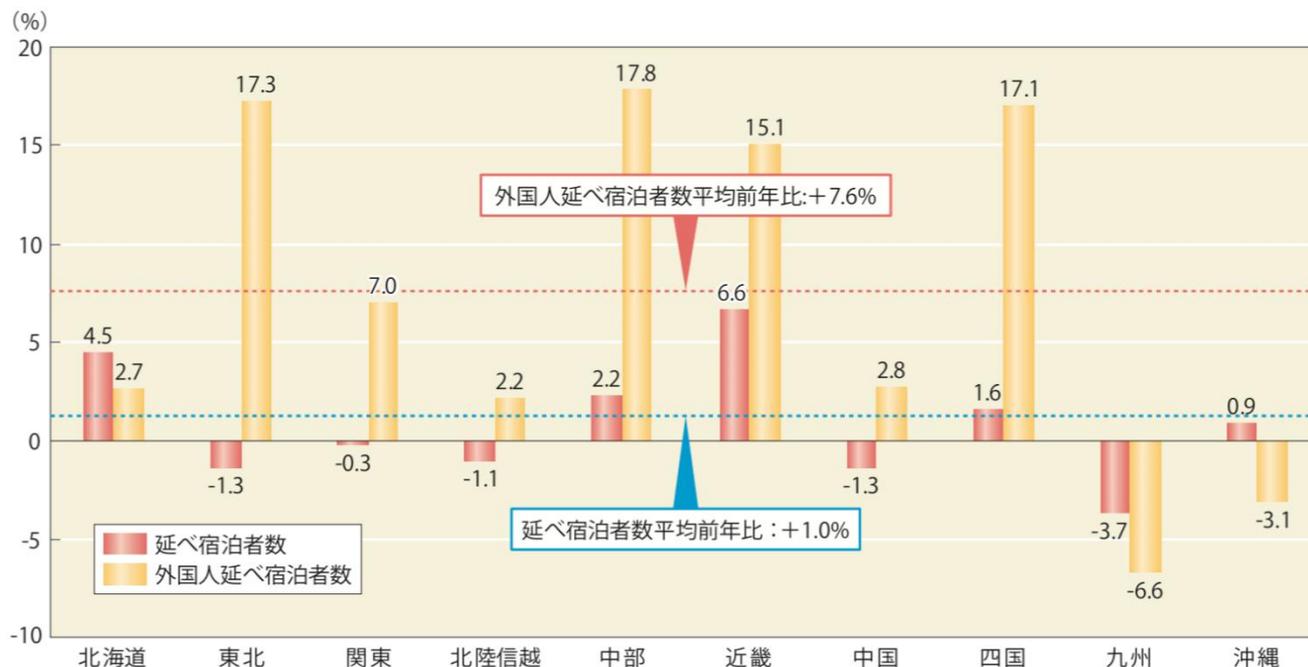
客室稼働率が80%を超えた都道府県は、
シティホテルは8箇所(2018年:10箇所)でした。
ビジネスホテルは4箇所(同:6箇所)、
リゾートホテルは2箇所(2018年:2箇所)でした。

沖縄の客室稼働率はシティホテルが
80%を超えています。
また、ビジネスホテルは74.4%、
リゾートホテルは69.6%でした
全体でも63.4%と、他の都道府県と比較して
高い客室稼働率になっています。

⑪ 延べ宿泊者数及び外国人延べ宿泊者数の地方ブロック別前年比

【延べ宿泊者数及び外国人延べ宿泊者数の地方ブロック別前年比を把握する方法】
観光白書から沖縄県の延べ宿泊者数及び外国人延べ宿泊者数の地方ブロック別前年比を導き出します。

(2019年)



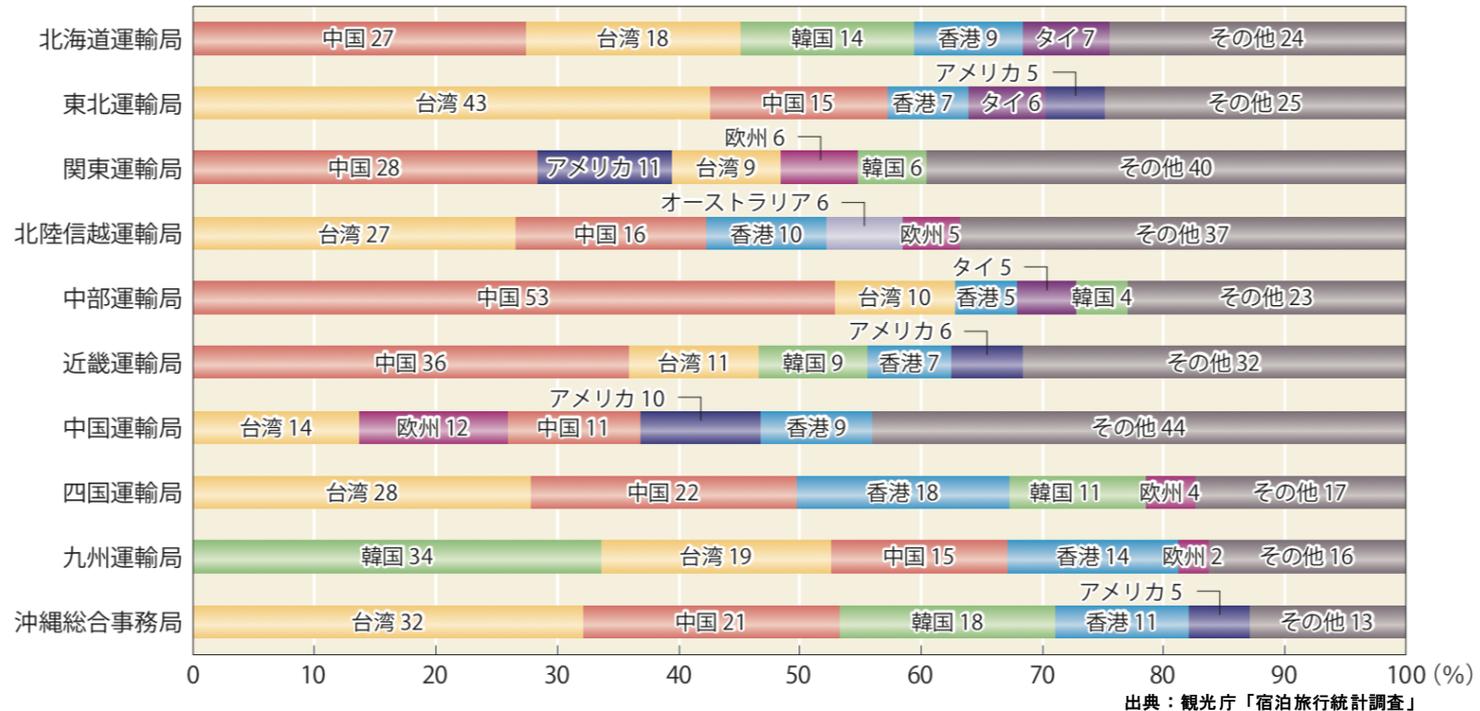
出典：観光庁「宿泊旅行統計調査」

2019年の延べ宿泊者数について、前年比で見ると北海道、中部、近畿、四国及び沖縄地方では増加となりました。外国人延べ宿泊者数は、10地方のうち8地方で伸びており、特に東北、中部、近畿及び四国地方の伸び率が高く推移しました。

九州・沖縄は、韓国からの訪日旅行控え等の影響などで外国人延べ宿泊者数は他地域と異なり、マイナスとなっています。

⑫ 地方ブロック別都道府県別外国人延べ宿泊者の国籍・地域別構成比

【地方ブロック別都道府県別外国人延べ宿泊者の国籍・地域別構成比を把握する方法】
観光白書から沖縄県の外国人延べ宿泊者の国籍・地域別構成を導き出します。
(2019年)



2019年における地方ブロック別外国人延べ宿泊者について、国籍・地域別にみると、中国からの宿泊者が北海道・関東・中部・近畿の4地方で、韓国からの宿泊者が九州で、台湾からの宿泊者が東北・北陸信越・中国・四国・沖縄の5地方で最も高い比率を占めました。

沖縄は、台湾・中国・韓国・香港と主に東アジアからの観光客が多い傾向。

観光白書から沖縄の観光に関する情報を分析

東京都、大阪府に比べて沖縄の客室稼働率は低い

ポイント

季節変動に対応して
平準化を目指す

沖縄のリゾートホテルは全国のリゾートホテルと比較すると高い客室稼働率を維持しているが、ビジネスホテルと比較すると低い

ポイント

ワーケーションなどで
ビジネスユースを狙う

九州・沖縄は、韓国からの訪日旅行控え等の影響でマイナスである

ポイント

ターゲット国の分散化

沖縄は台湾・中国・韓国・香港と東アジアからの観光客が多い

ポイント

アフターコロナの
東アジア戦略



沖縄県の観光情報を分析することで、
県としてのマーケティング戦略のベースを描くこともできます。

個人演習



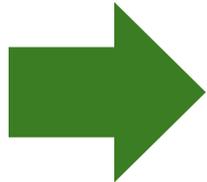
観光白書を見て、今後の沖縄県の観光で狙うべきターゲットは以下の中からどれになるでしょうか。理由と共に考えましょう。

1. **日本人観光客（国内観光）**
2. **訪日中国人観光客**
3. **訪日台湾人観光客**
4. **訪日韓国人観光客**
5. **訪日香港人観光客**
6. **訪日アメリカ人観光客**

JNTOデータからインバウンド顧客特性を分析

学習目標

1. JNTOのデータからインバウンド顧客の動向に関する情報を分析する手法を理解する。



このコースを学習すると、JNTOが提供するオープンデータからの顧客特性を理解し、分析することが出来ます

講義



**JNTO「訪日旅行データハンドブック」から
各国・地域の顧客特性を分析する方法
(P30～P38で実際に分析をします)**

JNTOのデータから訪日外国人旅行者の特性を分析

【情報収集・分析】

JNTOが公開する訪日外国人旅行者に特化した「日本の観光統計データ」のWebサイトで日本に旅行に来る国・地域の情報を収集・分析し、マーケットの特性をあぶり出してみましよう。



JNTOが出している調査・統計の中でも、「日本の観光統計データ」は集客を行う上で重要なデータです。

韓国、中国、台湾、香港、フィリピン、ベトナム、タイ、マレーシア、シンガポール、インドネシア、インド、豪州、カナダ、米国、ロシア、ドイツ、フランス、英国、イタリア、スペインといった世界20市場のデータが収集されています。

- ・ 訪日客の推移
- ・ 訪日客が訪日旅行中に役立った旅行情報源
- ・ 月別訪日客の推移
- ・ 訪日客の日本滞在中の人当たりの支出額

など、国や地域ごとに比較が可能となっています。

JNTOのデータから分析

- 各国・地域の基本属性データを比較しよう
- 1人当たり名目GDPの伸び率を比較しよう
- 海外旅行者数と訪日旅行者数の推移を比較しよう
- 訪日旅行者数の推移と旅行タイプの推移を確認しよう
- 訪日旅行に役立った旅行情報源を比較しよう

JNTOデータからわかるターゲット国・地域の顧客特性分析

ポイント

1. ターゲット国・地域の現状把握

JNTO「日本の観光統計データ」のデータを把握することで、ターゲット国・地域の観光に関する情報を入手することが可能です。現状よく来日するターゲット国・地域の特性はもちろん、今後狙いたいターゲット国・地域についてのデータも入手することが可能です。

ポイント

2. 観光マーケティング戦略作りに活用

今後狙うべき国・地域の訪日旅行者が何を求めて日本に来ているのか、どのような情報源から日本の観光・旅行の情報を入手しているかを把握するだけで、ターゲット国・地域への訴求方法のヒントが得られます。特にマーケティング戦略におけるプロモーション戦略を作る際に極めて重要な情報となります。



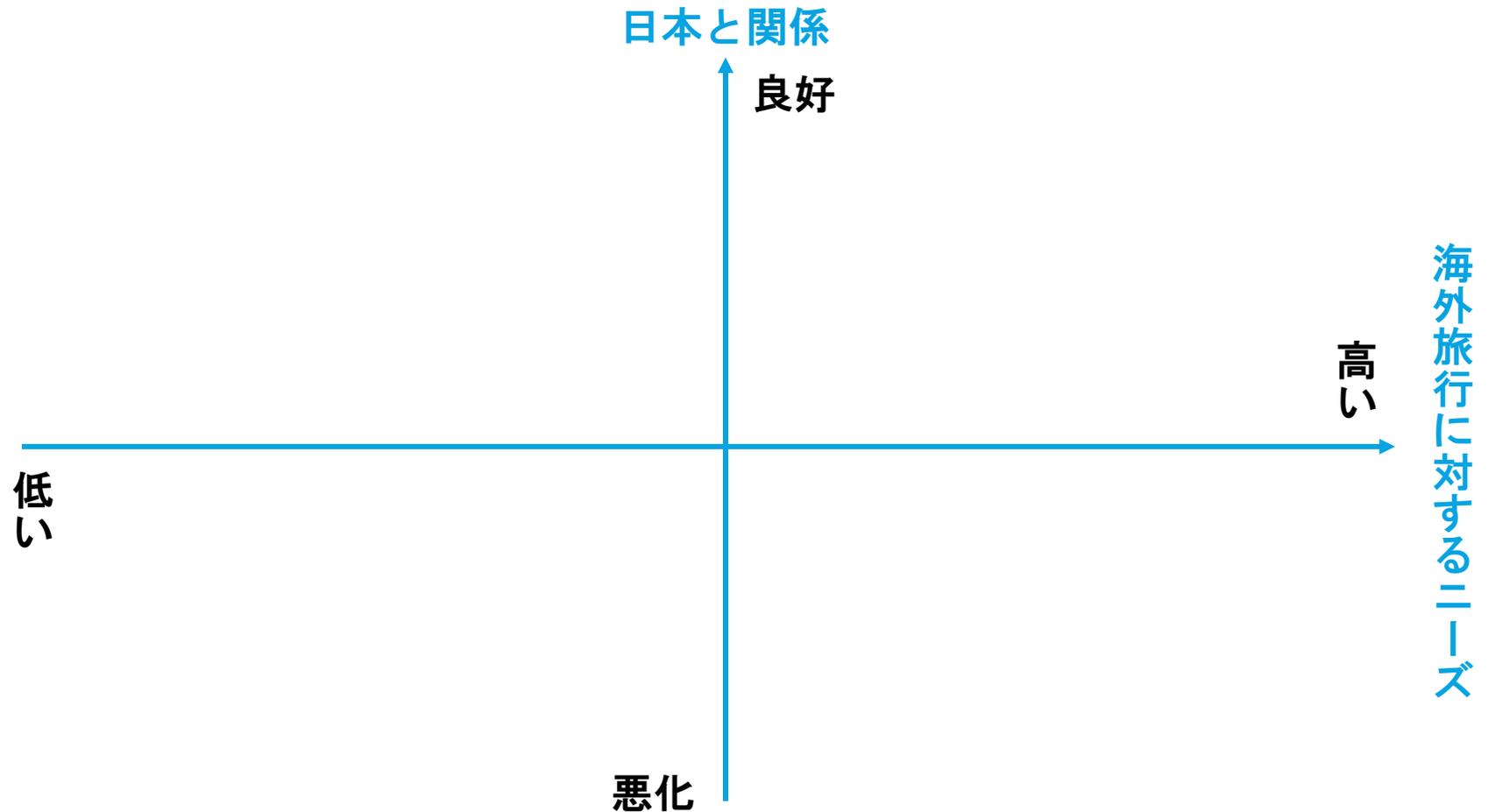
JNTO「日本の観光統計データ」を読み込み、分析するとインバウンド戦略作りに役立つことがわかります。

グループ演習

個人演習①

ポジショニングマップの作成

中国、韓国、台湾、香港4つの国・地域の特性を踏まえ、各国、地域を以下の2つの軸でプロットしてみましょう。



ビッグデータの活用①



RESASを使った分析演習のやり方

【復習】RESAS

RESASとは？

地域経済に関する様々なビッグデータ（産業の強み、人の流れ、人口 動態など）。インターネット上に公開されており、誰でも簡単に操作することが出来るデータ分析システムです。

RESAS 地域経済分析システム

仮説を根拠へ

RESASは、地域経済に関するビッグデータを地図上やグラフで見える化できる政府のシステムです。

分析を開始する

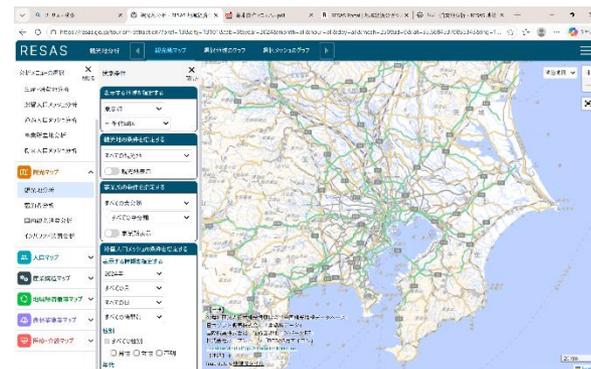
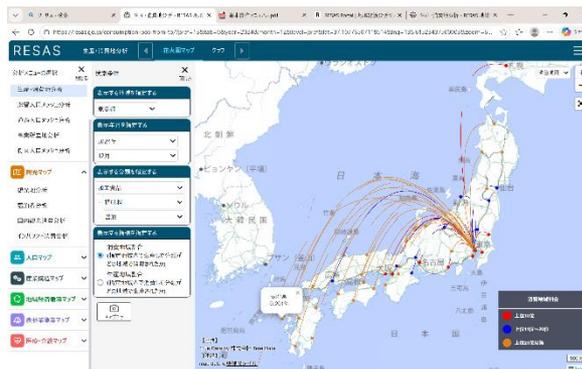
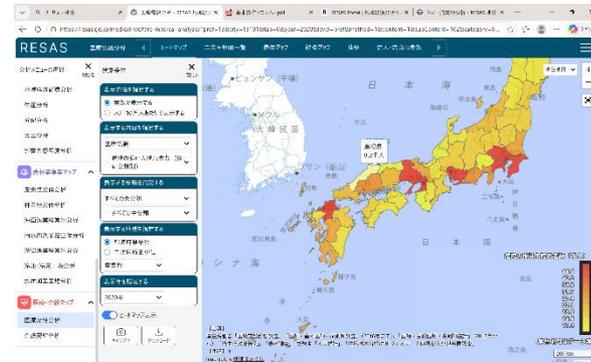
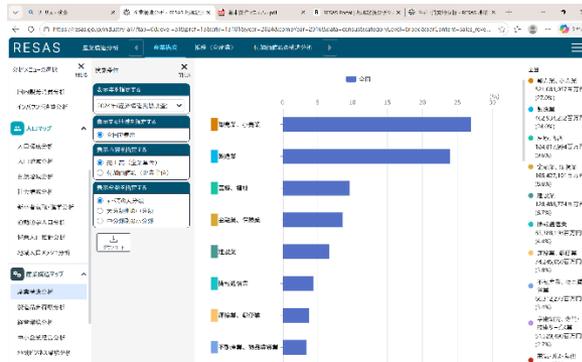
- マーケティング
- 観光
- 人口
- 産業構造
- 地域経済循環
- 農林業漁業
- 医療・介護

出典：内閣府 地方創生推進室 ビッグデータチーム

RESASは観光分野に限らずさまざまな領域のデータを搭載されています。地方創生戦略や地域活性化案など様々な戦略作成に活用できるシステムです。

“RESAS”（リーサス）の意味 頭文字（Regional Economy (and) Society Analyzing System）の略

マップ例



出典：内閣府
地方創生推進室
ビッグデータチーム

RESASの7つのマップ
「マーケティング」「観光」「人口」「産業構造」
「地域経済循環」「農林業漁業」「医療・介護」
で構成されています。

つまり…？

地域経済に関する様々なビッグデータ

（産業の強み、人の流れ、人口動態など）が

地図・グラフから分かりやすく「見える化（可視化）」したシステム

- インターネットから無料で誰でも利用可！
- 直感的に使える、見やすい画面操作！
- データから地域の実情を理解・分析できる！

RESASの使い方①

基本的な使い方

①RESASのページを閲覧する

<https://resas.go.jp/>

②必要な情報を探す

③情報を見る

どの情報を見るにも、基本的にはこの構造となります。

この①～③の流れを体感してみましょう。

RESASの使い方①

一緒にやってみましょう

① RESASのページを閲覧する

検索エンジンに「RESAS」

と検索して、

右にあるサイトを

クリックします。



The screenshot shows a search engine interface with the search term 'RESAS' entered in the search bar. Below the search bar, there are navigation tabs for 'すべて' (All), '画像' (Images), 'ニュース' (News), '動画' (Videos), 'ショッピング' (Shopping), '地図' (Maps), and 'ウェブ' (Web). The search results for 'RESAS 地域経済分析システム' are displayed, including the URL 'https://resas.go.jp' and a list of links to various analysis tools and reports.

RESAS

すべて 画像 ニュース 動画 ショッピング 地図 ウェブ : もっと見る

 RESAS 地域経済分析システム
https://resas.go.jp

RESAS 地域経済分析システム: トップページ

地域経済分析システム（RESAS：リーサス）は、地方自治体の様々な取り組みを情報面から支援するために、内閣官房新しい地方経済・生活環境創生本部事務局が提供する、...

目的地分析 >
どの地域から来る人が多く滞在しているかを平日・休日別、男女別、年代 ...

生産分析 / マップ >
地域のお金の流れを生産、分配、支出の三段階で表示するとともに、各 ...

データ分析支援機能とは >
Analysis Supportデータ分析支援機能：データ分析支援機能とは・分析 ...

地域経済循環図 / グラフ >
自治体の実際の人口に対し、一時間あたり月間平均で何倍の滞在人口 ...

CASE 自治体による利活用事例集 >
携帯電話のアプリ利用者の位置情報を用いて、月別・時間帯別の流動人 ...

[resas.go.jp のすべての検索結果を表示 »](#)

RESASの使い方①

一緒にやってみましょう

① 下記のサイトが表示されることを確認してください。

RESAS 地域経済分析システム

仮説を根拠へ

RESASは、地域経済に関するビッグデータを地図上やグラフで見える化できる政府のシステムです。

分析を開始する

- マーケティング
- 観光
- 人口
- 産業構造
- 地域経済循環
- 農林業漁業
- 医療・介護

RESAS 人口統計分析

人口推移グラフ

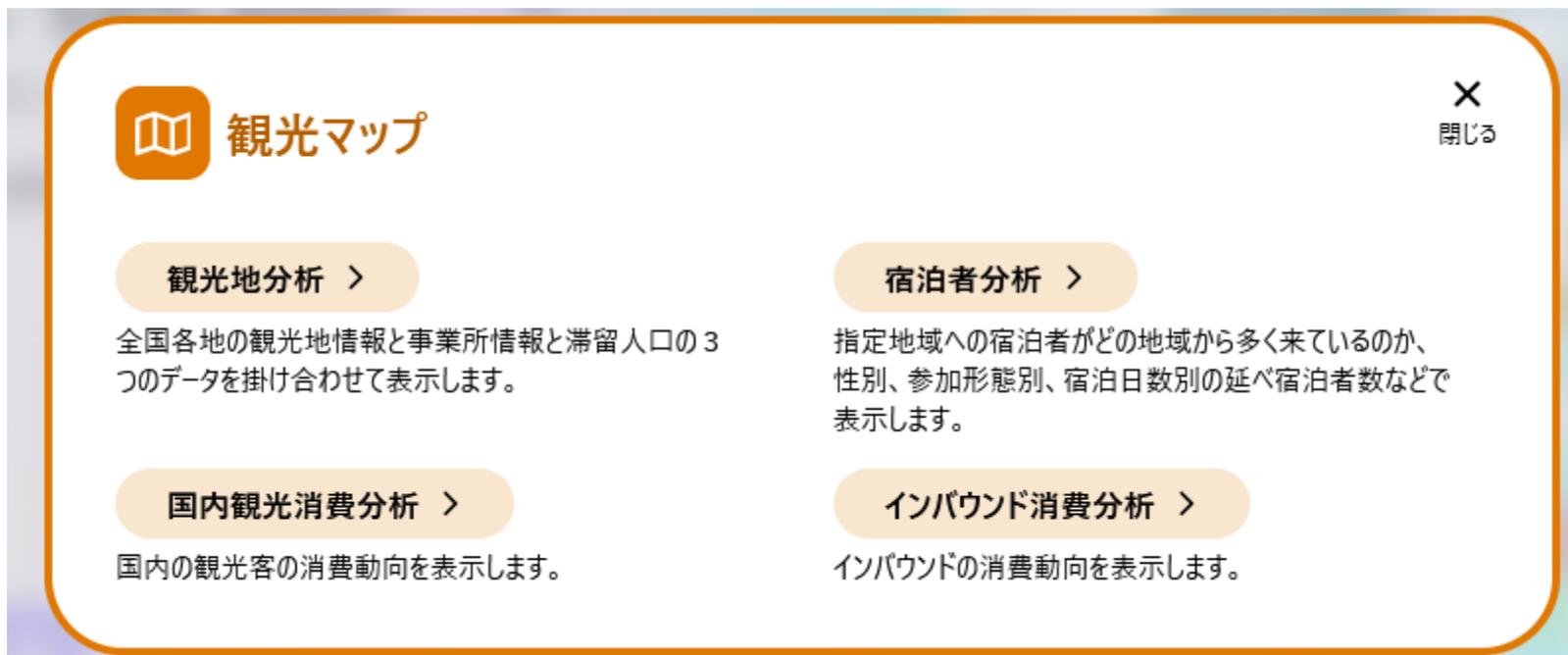
総人口 若年人口 生産年齢人口

RESASの使い方②

一緒にやってみましょう

②今回は観光マップを選択します。

すると、観光マップから選択できる分析一覧が表示されます。
ここから、宿泊者分析を選択してください。



観光マップ × 閉じる

観光地分析 >
全国各地の観光地情報と事業所情報と滞留人口の3つのデータを掛け合わせて表示します。

国内観光消費分析 >
国内の観光客の消費動向を表示します。

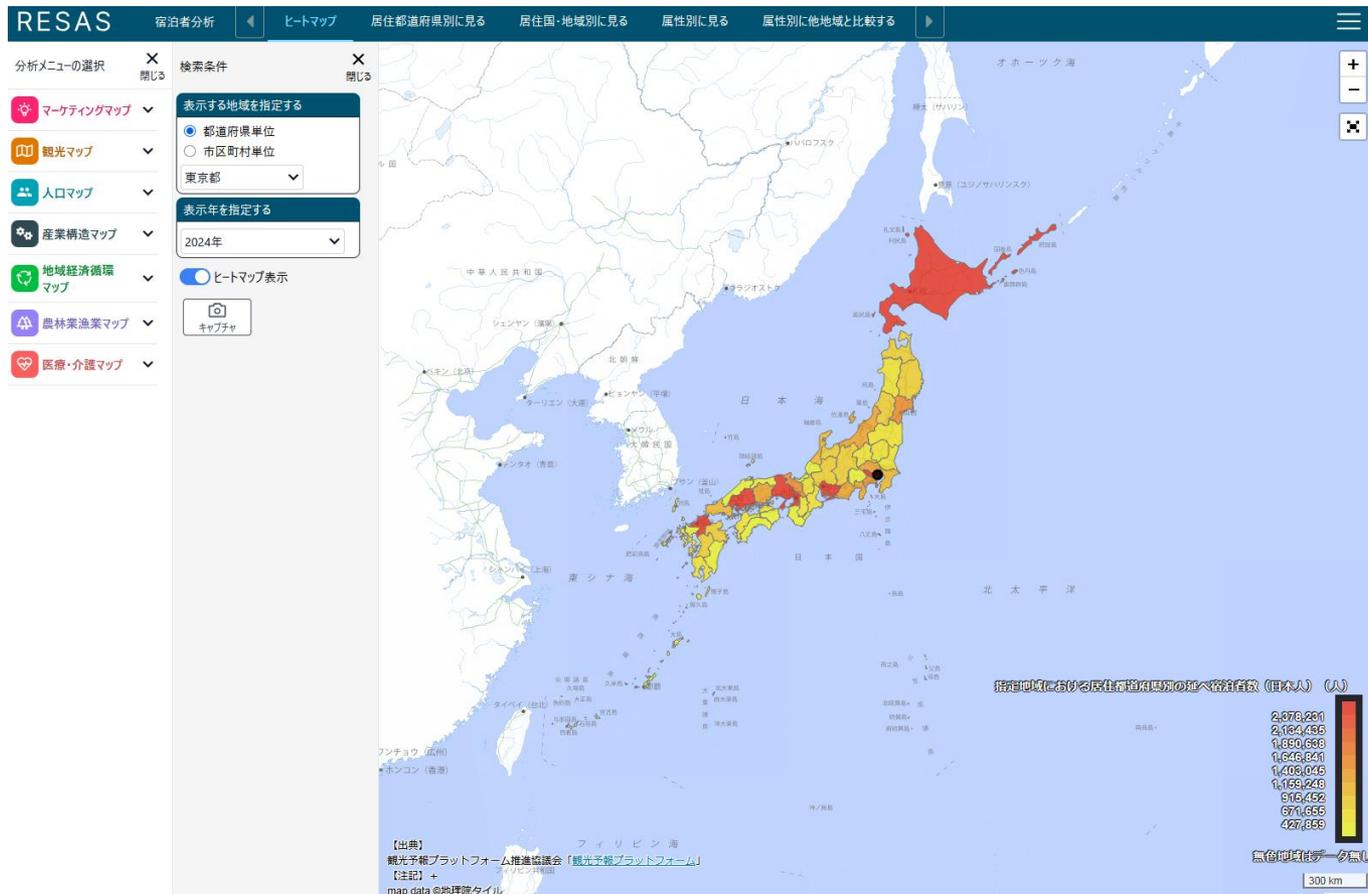
宿泊者分析 >
指定地域への宿泊者がどの地域から多く来ているのか、性別、参加形態別、宿泊日数別の延べ宿泊者数などで表示します。

インバウンド消費分析 >
インバウンドの消費動向を表示します。

RESASの使い方②

一緒にやってみましょう

②日本地図が表示されることを確認して下さい。



RESASの使い方③

一緒にやってみましょう

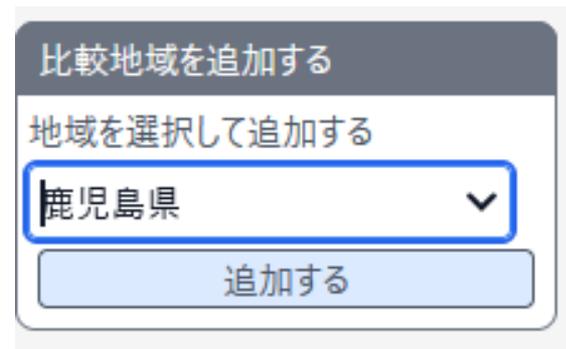
③山口県にどの県からの割合が多いかと、その推移を見ていきましょう。
地図が表示されたら、緑の箇所を自分の知りたい情報として
選択していきます。

今回は山口県の情報が見たいので、「県」を山口県に設定してください。
その後、「居住都道府県別に見る」をクリックします。

RESASの使い方③

一緒にやってみましょう

- ③これでは、上位しか表示されていないので特定の県の情報を見てみましょう。
「鹿児島県」を追加してみましょう。



比較地域を追加する

地域を選択して追加する

鹿児島県

追加する

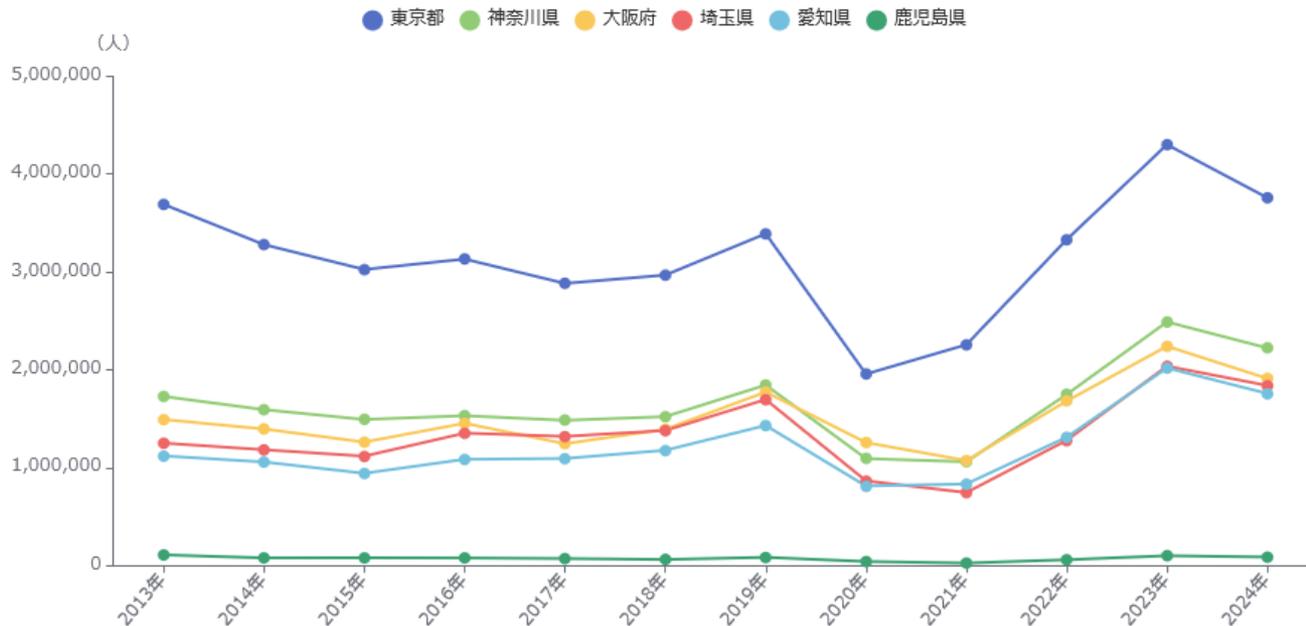
「比較地域を追加する」から、鹿児島県を選択し、下にある追加ボタンをクリックします。
表示地域を指定するに鹿児島県が入っていれば成功です。

RESASの使い方③

一緒にやってみましょう

③グラフにも鹿児島県が表示されました。

居住都道府県別の延べ宿泊者数（日本人）の推移
沖縄県
2024年



RESASの使い方③

一緒にやってみましょう

③この情報だと鹿児島県の情報が
見つらいので、グラフの加工を行います。

グラフにある上位5県をクリックすると、クリックした県のグラフを
削除できます。

居住都道府県別の延べ宿泊者数（日本人）の推移

沖縄県
2024年

● 東京都 ● 神奈川県 ● 大阪府 ● 埼玉県 ● 愛知県 ● 鹿児島県

RESASの使い方③

一緒にやってみましょう

③鹿児島県の情報の詳細に見えるようになりました。

居住都道府県別の延べ宿泊者数（日本人）の推移

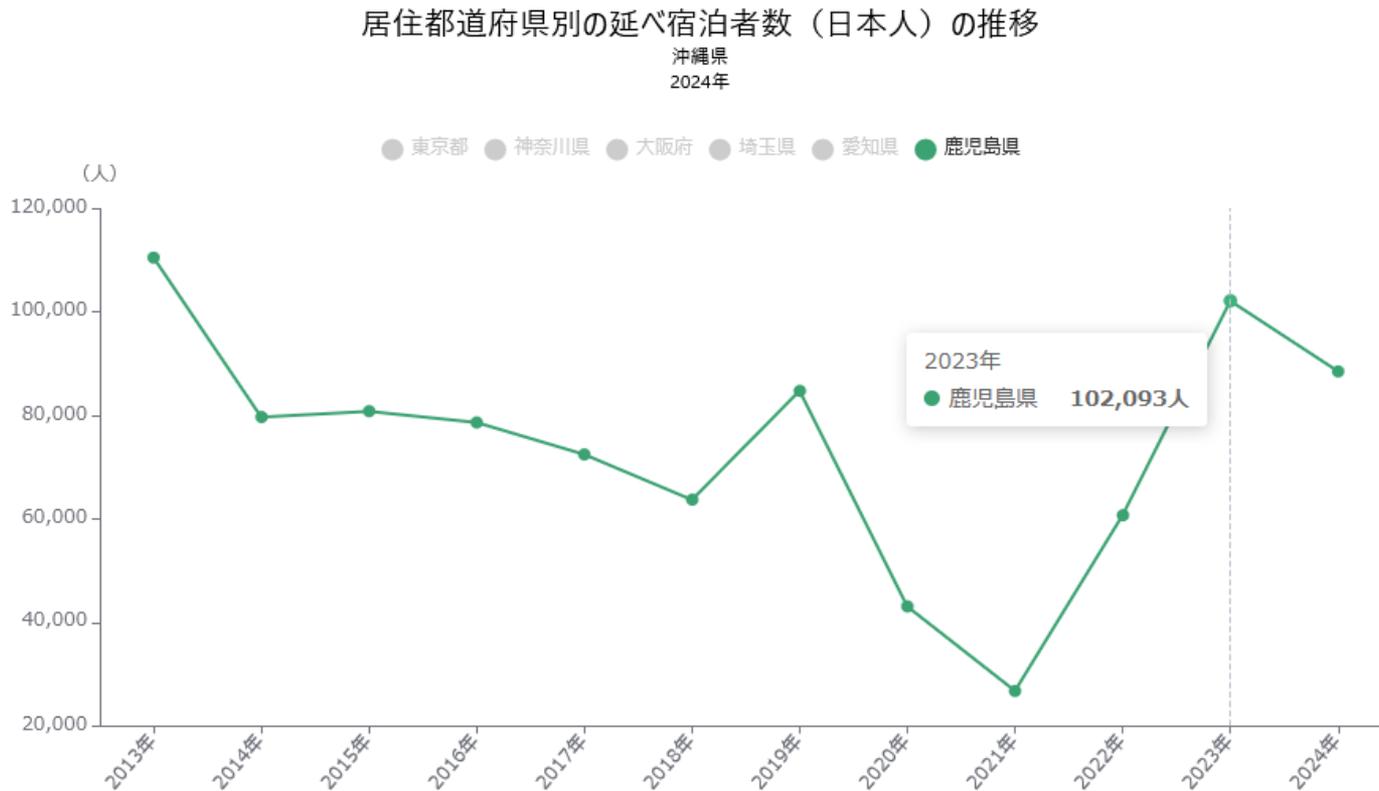
沖縄県
2024年



RESASの使い方③

一緒にやってみましょう

③特定の年にマウスカーソルを合わせると、詳細の人数が表示されます。



RESASの使い方③

一緒にやってみましょう

データを保存したい場合、それぞれのグラフの
右上にあるキャプチャボタンを押下します。

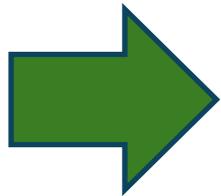
pngファイルでDLが出来ますので、後で使用が可能になります。



自ら選んだデータをもとに分析演習

学習目標

1. **実際に自らどのようなデータを取得したいか考えてRESASを使ってみる。**



**このコースを学習すると、
RESASを実際にも使えるようになります**

個人演習

個人演習①

山口県以外のエリア（都道府県または市町村）を選んでRESASの「観光マップ」から各種データを取得してみましよう。

どのようなデータを取得するかも自ら考えて選択します。
選択するデータは複数とします。その上でデータから見える選択したエリアの特徴をまとめましよう。

また、選んだエリアと山口県のデータを比較して、
両エリアの観光に関する現状を分析してみましよう。

個人演習①

- 選択したエリア（都道府県または市町村）

- 選択したデータ（複数のデータを選択）

- データの特徴（分析①）

- 山口県との比較（分析②）

ビッグデータの活用②



Google TrendなどGoogleのビッグデータの 入手方法・見方について

Googleとは

Googleとは主に検索エンジンなどのオンラインサービスを提供・運営している企業のことです。

ここでは、Google社が提供している検索サービスをGoogleと言います。

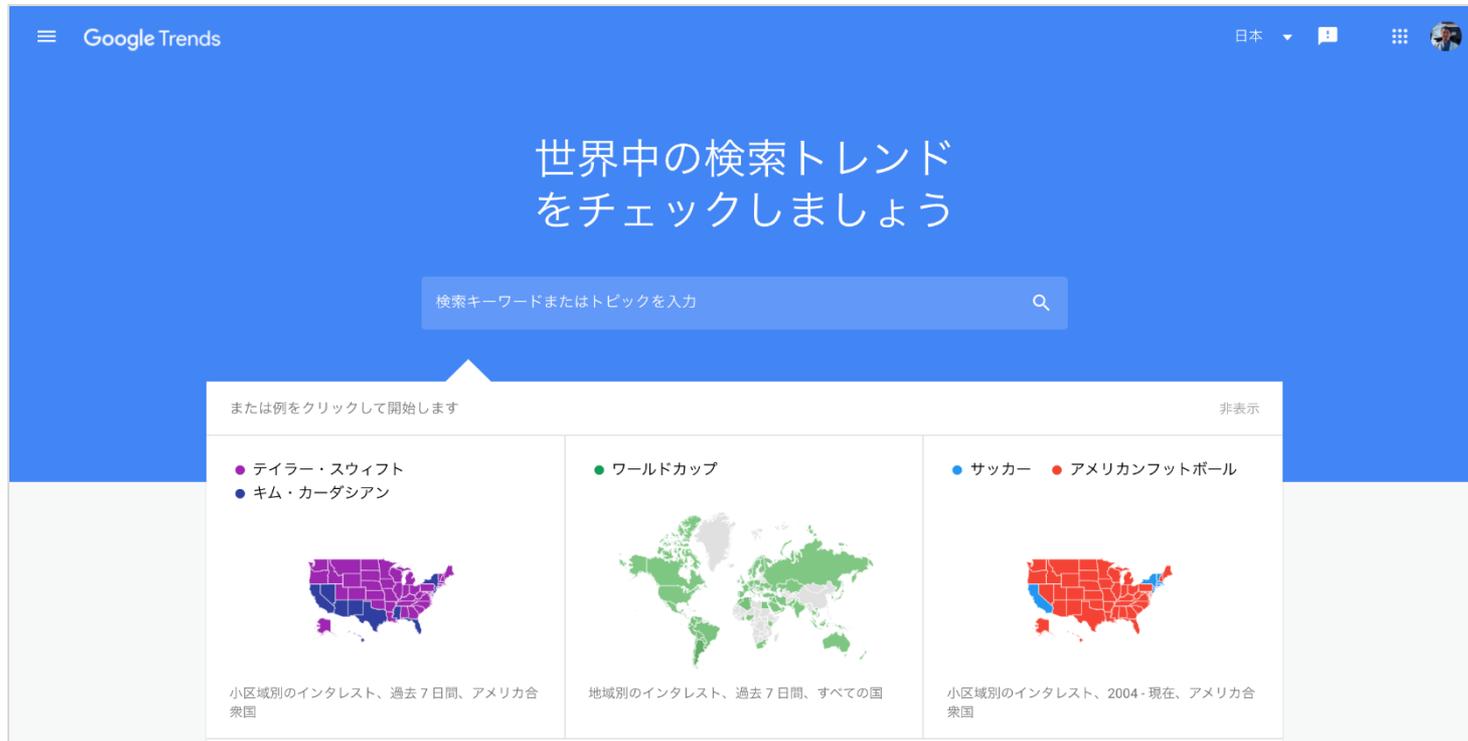
日本には、Google以外にも、Yahoo! Japanという検索エンジンが有名で、多くのインターネットユーザーはGoogleまたはYahoo! Japanでオンライン検索をすることになります。

Google社は検索エンジンサービスの他にも、インターネット広告である「Google AdSense」や「Google AdWords」、無料の地域検索サービス（オンライン地図）などのサービスを提供しています。

これらGoogle社が提供しているオンラインサービスは私たち観光産業のマーケティング活動にも非常に有用なので、使いこなせるようにしましょう。



Googleトレンドとは



画像出典：Google

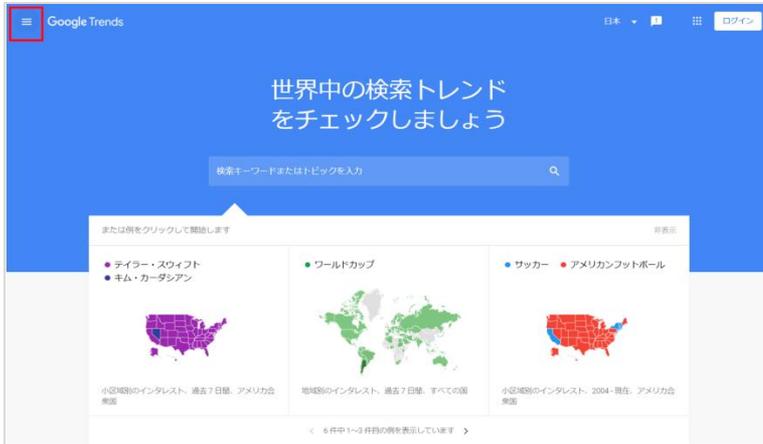
Googleトレンドとは、Google で検索されるキーワードの検索回数の推移がわかる、登録不要、無料で使えるツールです。

検索キーワードの検索回数が増加しているのか、減少しているのか、期間の指定をして調べることができるので、過去・現在・未来の市場調査をすることが可能です。

Googleトレンドの使い方

検索ボリュームの推移を知りたいキーワードについての調べ方と活用方法について

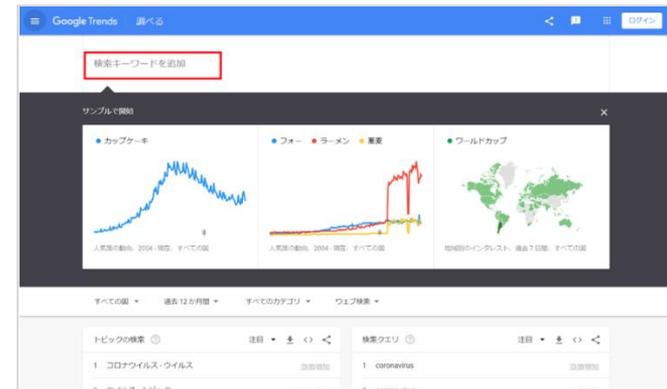
① 左上のメニューアイコンをクリックします。



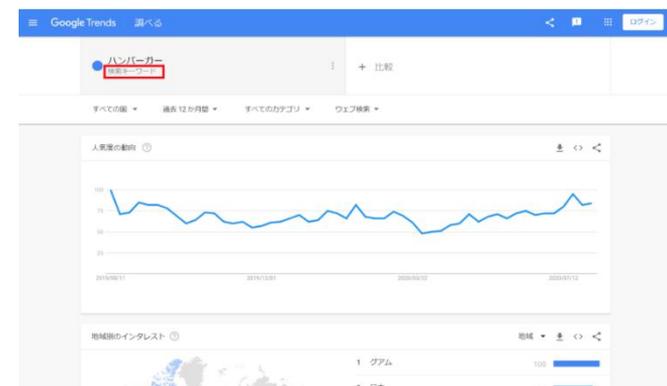
② [調べる]をクリックします。



③ [検索キーワードを追加]の欄に、検索ボリュームの推移を調べたいキーワードを入力します。



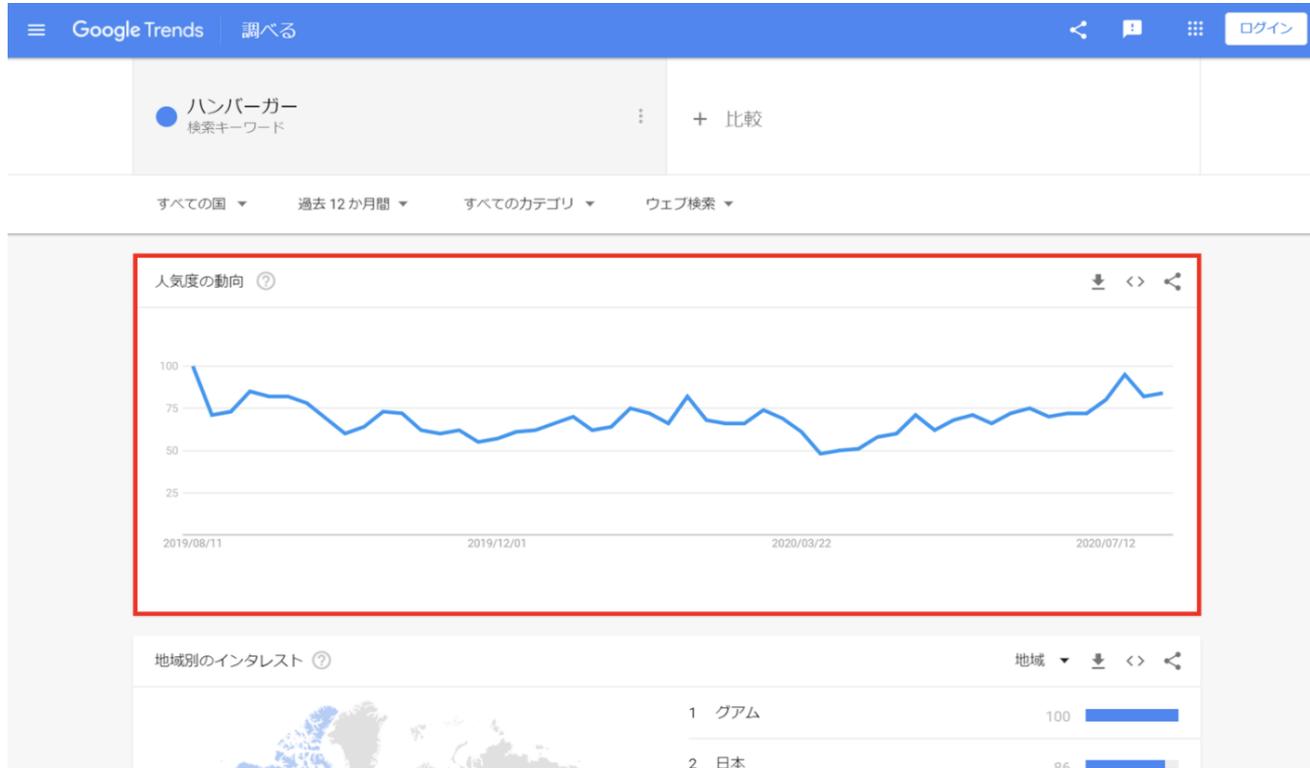
④ 例として「ハンバーガー」で検索していきます。



検索キーワードの場合、「ハンバーガーメニュー」や「ハンバーガーレシピ」などの「ハンバーガー」を含むキーワードも含まれます。

Googleトレンドの使い方

検索ボリュームの推移を知りたいキーワードについての調べ方と活用方法について



画像出典：Google

検索すると、検索推移のグラフが表示されます。Googleトレンドに表示される検索ボリュームは、相対数値となっていて、実際の検索数ではないので注意が必要です。チェックした期間の中で検索が最も多い時期が100となっていて、50だと約半分、0だと十分なデータがなかったことを意味します。

Googleトレンドの使い方

●関連トピックやキーワードを知る

キーワードを検索後、スクロールすると下部に関連トピックと関連キーワードが表示されます。検索した単語と同時に検索されることの多い項目が表示されます。



画像出典: Google

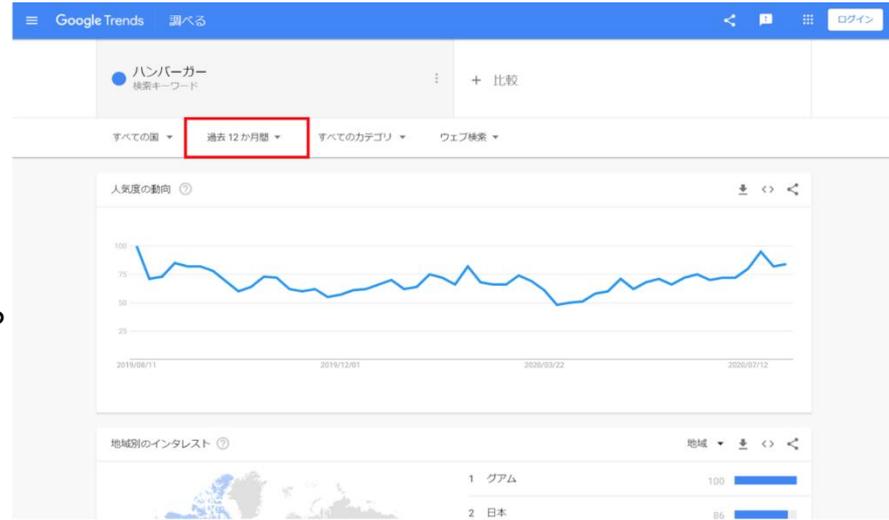
関連トピックの2番目に「NO. 18ハンバーガー」は何故、検索が急上昇しているのか？

「バーガーサミット」への出店を行い認知度をあげている、“ふきのとうバーガー”など限定のハンバーガーを販売するようになった。情勢に合わせてテイクアウトも始めるようになった。など、なぜ検索が上昇しているのか要因を考えながら活用します。

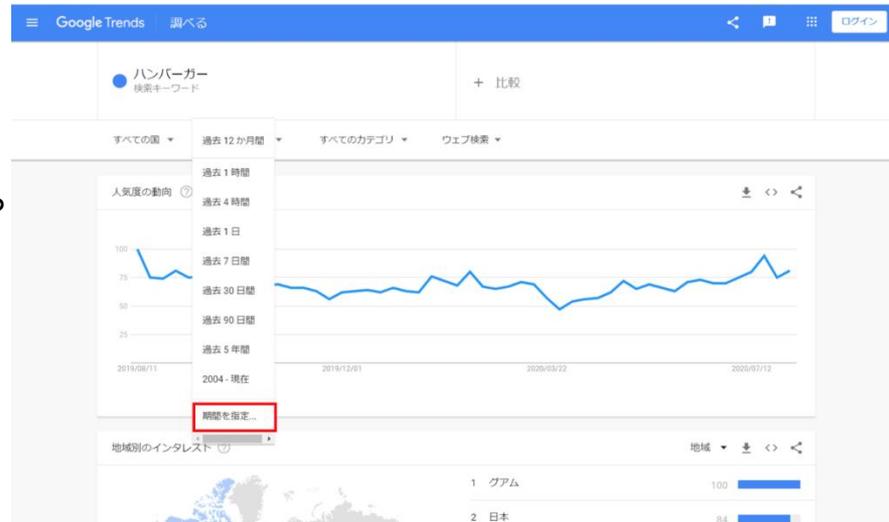
Googleトレンドの使い方

●期間を指定する

デフォルトでは、直近12ヶ月間の推移が表示されているため、[直近12ヶ月間▼]をクリックすると期間の変更が可能です。



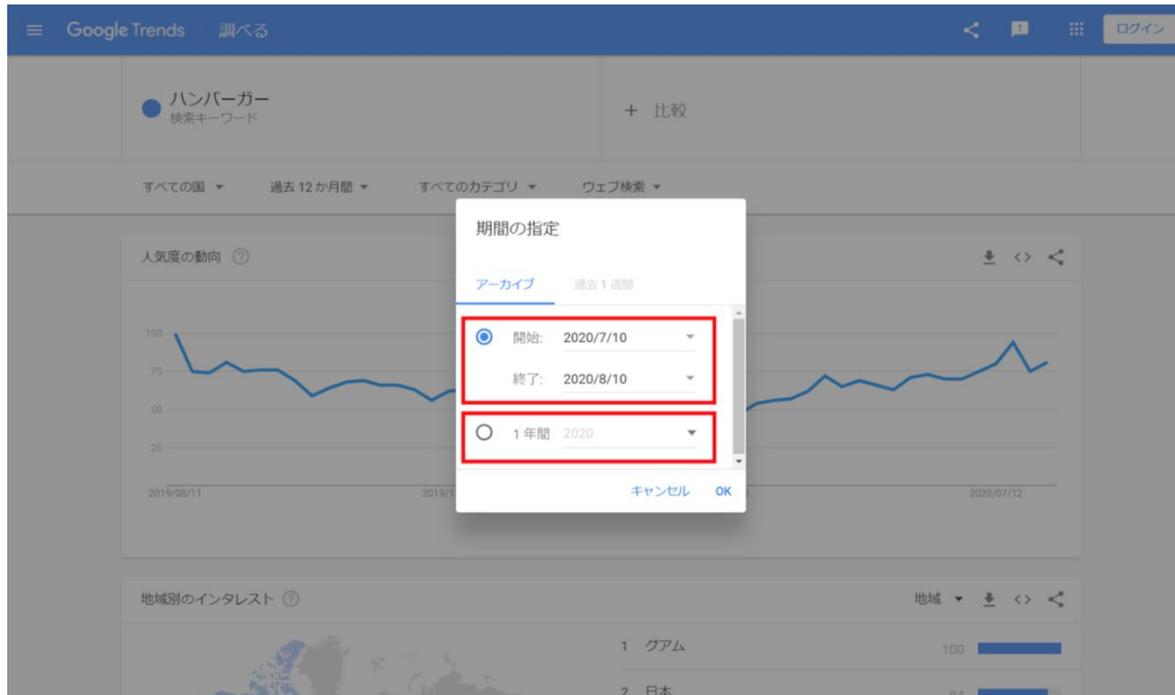
プルダウン上では、[過去1日]や[過去30日]などの候補がありますが、それ以外の期間も任意で選択することができます。



Googleトレンドの使い方

●期間を指定する

特定の期間(開始日～終了日)を指定する、もしくは西暦〇〇年の1年間を指定することができます。



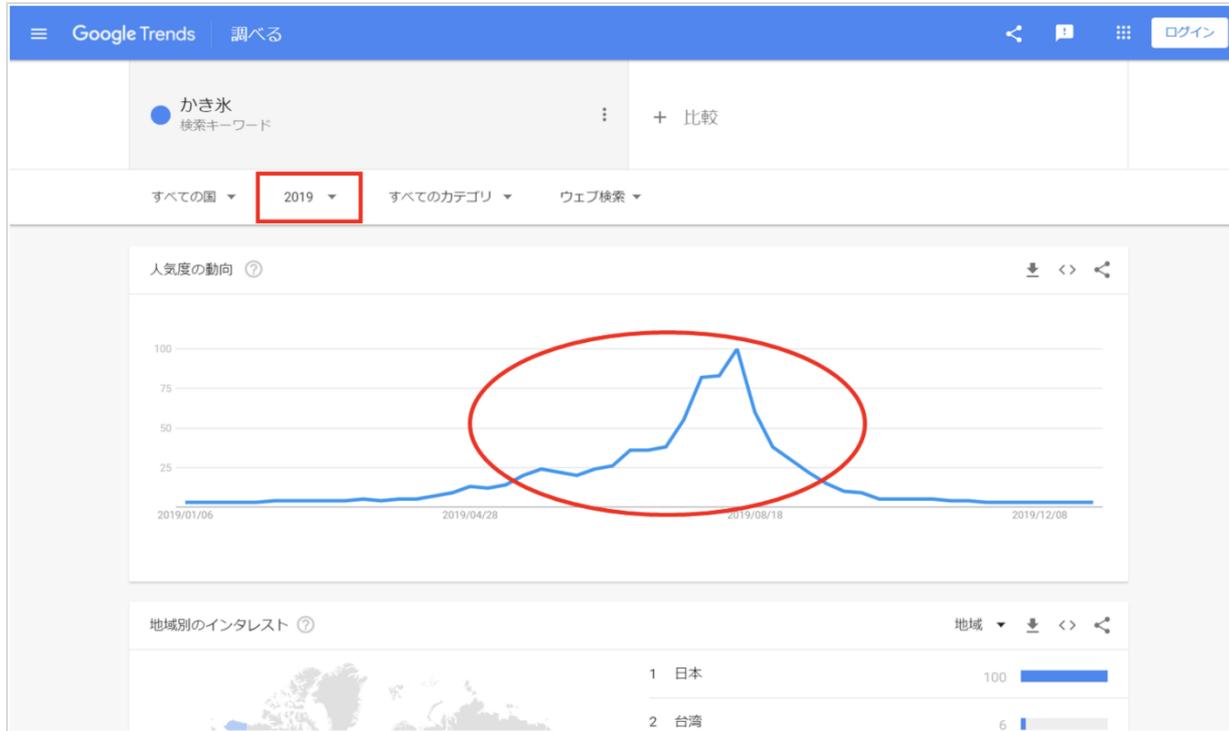
画像出典: Google

売上が拡大していた時期や、検索数が急上昇している期間を指定することで、その時期にどのような検索がされていたのかを確認し、その後マーケティング活動に活かすことが可能となります。

Googleトレンドの使い方

●期間を指定する

ある1年間の期間を指定することで、どの時期に人気上昇しやすいのかがわかります。



画像出典：Google

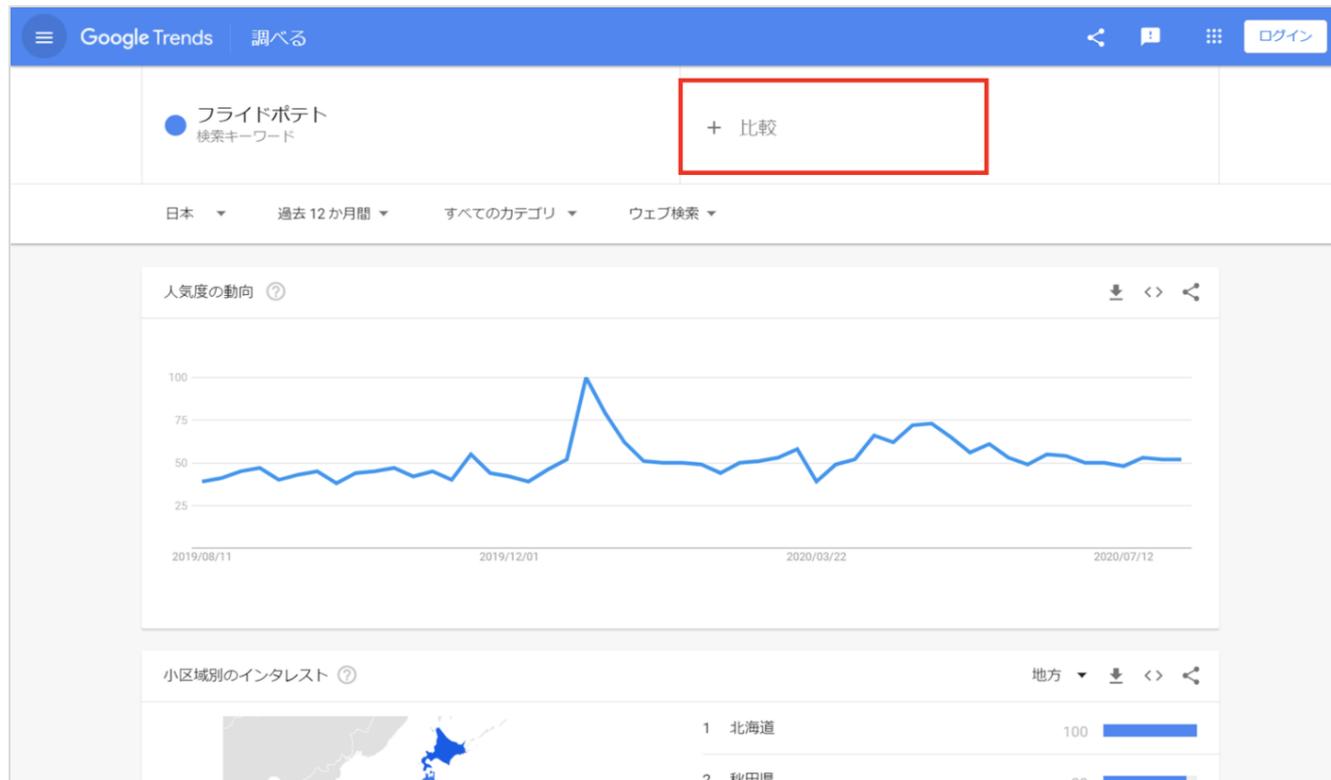
・季節性のキーワードを検索する際におすすめ

「かき氷」であれば1年間を通してみると、4月頃から需要が高まり8月にピークに達することが分かります。つまり、8月のピークに向けて広告戦略を企画するのが効率がよいことがわかります。さらに、5年分ぐらいの検索トレンドの推移と気温を比較すれば、気温が○度に達すると検索がピークに達するのか、などが見えてきます。

Googleトレンドの使い方

●キーワードを比較する

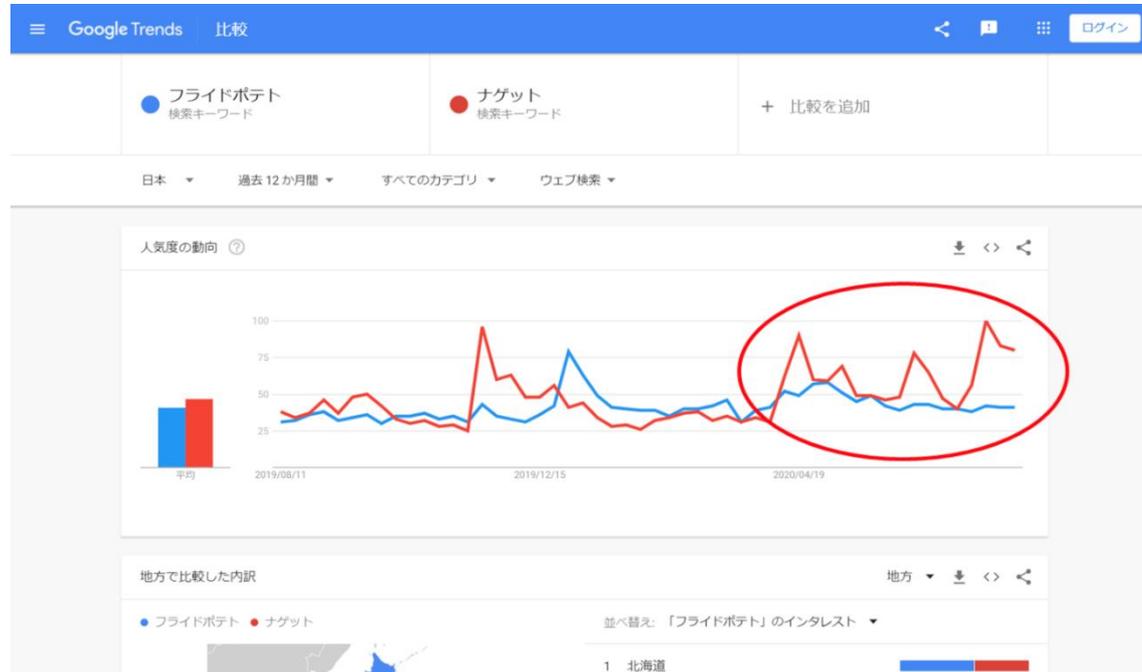
「+比較」をクリックして、検索キーワードを入力すると、同時に比較するキーワードを追加できます。



Googleトレンドの使い方

●キーワードを比較する

例えば、「フライドポテト」と「ナゲット」を比較した際に、1年程前はトレンドの差はでていなかったものの、直近ではナゲットの人気の向上していることが分かります。



画像出典：Google

人気の上昇している要因を関連キーワードを見て、確認しましょう。世間的にナゲットの需要が増加している、または競合会社がナゲットのキャンペーンを実施し、検索が増えているなど関連キーワードから想定できます。

前者であれば、ナゲットの訴求を実施すれば効果的ですが、後者であれば、競合にとられてしまわないよう対抗するか、異なる施策を検討する必要があります。

Googleトレンドの使い方

●急上昇キーワードを検索する

Google Trends

日本 ログイン

世界中の検索トレンド をチェックしましょう

検索キーワードまたはトピックを入力

または例をクリックして開始します 非表示

- テイラー・スウィフト
- キム・カードashian

小区域別のインタレスト、過去7日間、アメリカ合衆国

- ワールドカップ

地域別のインタレスト、過去7日間、すべての国

- サッカー ● アメリカンフットボール

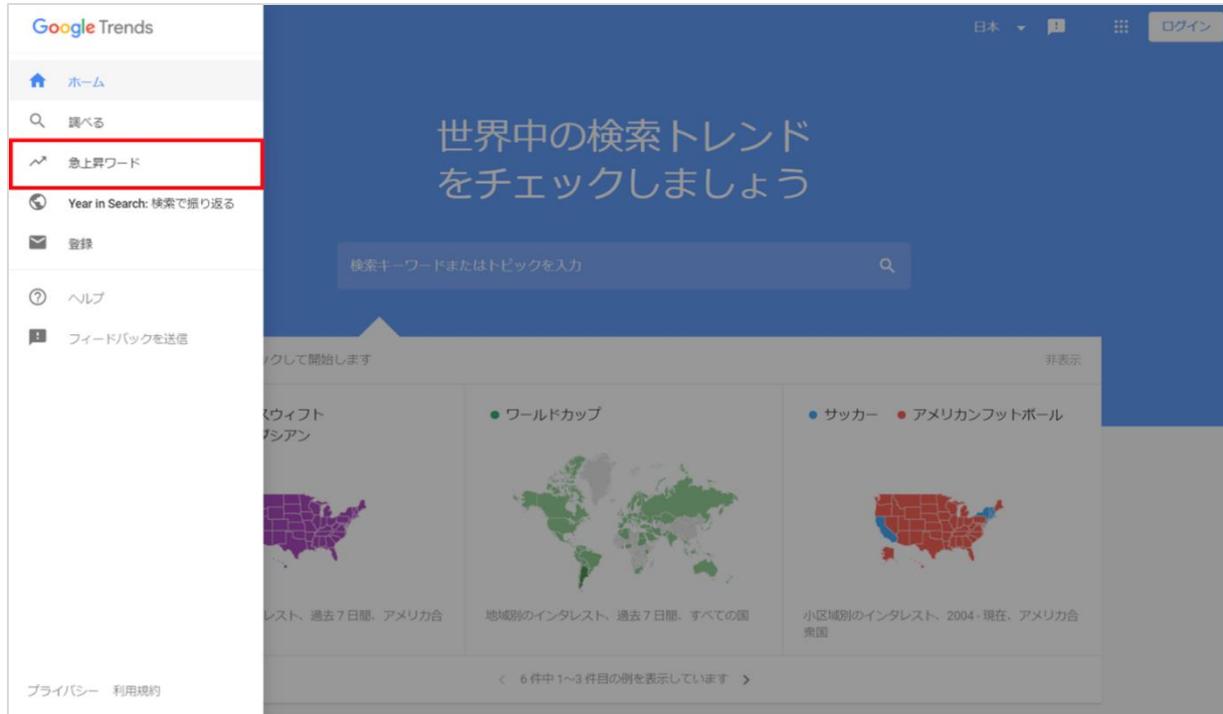
小区域別のインタレスト、2004 - 現在、アメリカ合衆国

< 6件中1~3件目の例を表示しています >

Googleトレンドの使い方

●急上昇キーワードを検索する

①左上のメニューアイコンをクリックします。



画像出典：Google

毎日の検索トレンド（日別の検索トレンド）が表示されます。さらにそれぞれのキーワードをクリックすれば、関連キーワードを確認したり、Web上の関連記事に飛ぶことができます。

Googleトレンドの使い方

●急上昇キーワードを検索する

②「急上昇ワード」をクリックします。



The screenshot shows the Google Trends interface for Japan. The 'Real-time search trends' tab is highlighted with a red box. The trending topics are as follows:

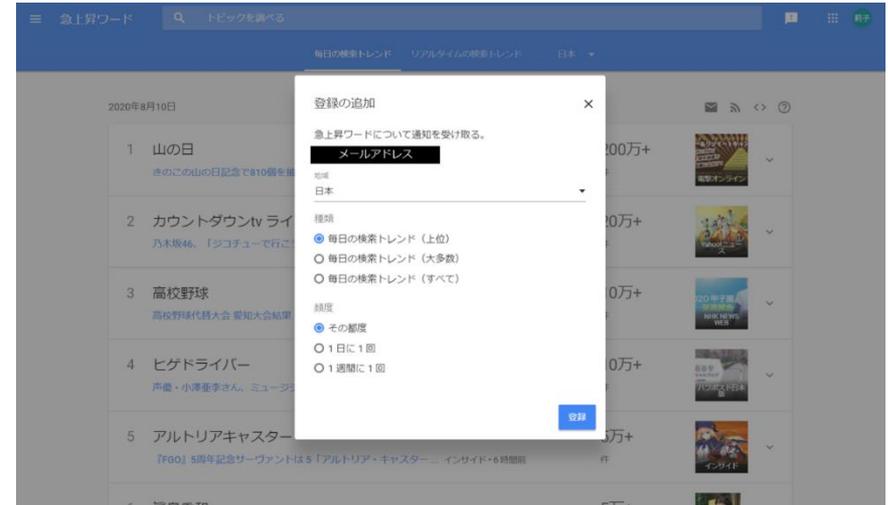
Rank	Keyword	Search Volume	Source
1	山の日	200万+ 件	電撃オンライン・17 時間前
2	カウントダウンtv ライブライブ	20万+ 件	リアルサウンド・11 時間前
3	高校野球	10万+ 件	Yahoo!ニュース・6 時間前
4	ヒゲドライバー	10万+ 件	Yahoo!ニュース・10 時間前
5	眞島秀和	5万+ 件	朝日新聞デジタル版・5 時間前
6	アルトリアキャスター	5万+	

画像出典：Google

横のリアルタイムの検索トレンドをクリックすると、現在の急上昇トレンドが表示されます。こちらはTwitterのトレンドに近いですが、Twitterはツイート内に記載のあるキーワードのトレンドですが、GoogleトレンドではGoogle上で検索されたキーワードです。そのため、母数もGoogleトレンドの方が多いため、よりデータの信頼性が高いと言えます。

Googleトレンドの使い方

●急上昇キーワードを検索する



画像出典: Google

- ・急上昇キーワードのお知らせ機能
メールアドレスを登録すれば、毎日、検索をかけずとも自分のメールを開けば、人気のトレンドを常にチェックすることが可能となります。

Googleトレンドの使い方

● Year in search:検索で振り返る

The screenshot shows the Google Trends homepage. At the top left, there is a menu icon (three horizontal lines) inside a red square box, followed by the text "Google Trends". To the right, there are options for "日本" (Japan), a notification icon, a grid icon, and a "ログイン" (Login) button. The main heading reads "世界中の検索トレンドをチェックしましょう" (Check search trends from around the world). Below this is a search bar with the placeholder text "検索キーワードまたはトピックを入力" (Enter search keywords or topics) and a magnifying glass icon. Underneath the search bar, there are three example cards. The first card is titled "テイラー・スウィフト" (Taylor Swift) and "キム・カーダシアン" (Kim Kardashian), with a purple map of the USA. The second card is titled "ワールドカップ" (World Cup), with a green world map. The third card is titled "サッカー" (Soccer) and "アメリカンフットボール" (American Football), with a red map of the USA. At the bottom of the grid, there is a navigation arrow and the text "6件中1~3件目の例を表示しています" (Showing 1-3 of 6 examples).

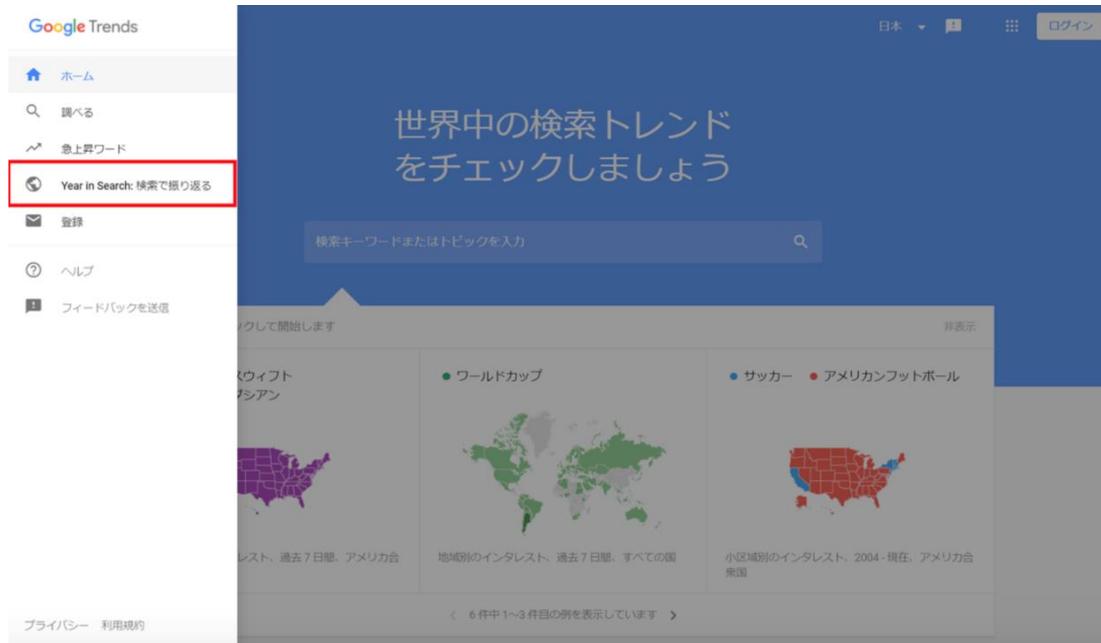
画像出典：Google

年単位の急上昇ワードをランキング形式で表示したものを「Year in search」といいます。

Googleトレンドの使い方

●急上昇キーワードを検索する

「Year in search:検索を振り返る」をクリック。



その年ごとに一番人気上昇したキーワード一覧を、「話題の人」、「ニュース」、「映画」、「曲」などのカテゴリごとのランキングをチェックすることができます。

その年に起きたことや関心の大きかった物事を把握することができるので、トレンドを振り返る際に便利です。

確認=Googleトレンドの使い方

Googleトレンドを活用することで、「検索キーワード」というユーザーのニーズが明確になります。

現在だけでなく、過去から長期間にわたって簡単に可視化できる素晴らしい検索ツールです。

検索キーワードの中には、「かき氷」などのある一定の時期になると関心が高まる季節性のものや、定期的に検索されるもの、ある出来事などが原因でトレンドが急上昇するものなど様々なものがあります。

トレンドの傾向をつかみ、トレンドを把握してマーケティング戦略の作成に活かすことが可能となります。

Googleデータを活用した分析演習

個人演習①

Googleトレンドの活用①

「山口」という単語をGoogleトレンドで調べてみましょう。その結果を見て、どのような傾向が見えるか書き出してみましょう。

【設定条件】

地域：日本

期間：5年間

カテゴリー：旅行

ウェブ検索

Googleトレンドの活用②

「yaaguchi」という単語をGoogleトレンドで調べてみましょう。その結果を見て、どのような傾向が見えるか書き出してみましょう。

【設定条件】

地域：全ての国

期間：5年間

カテゴリー：旅行

ウェブ検索

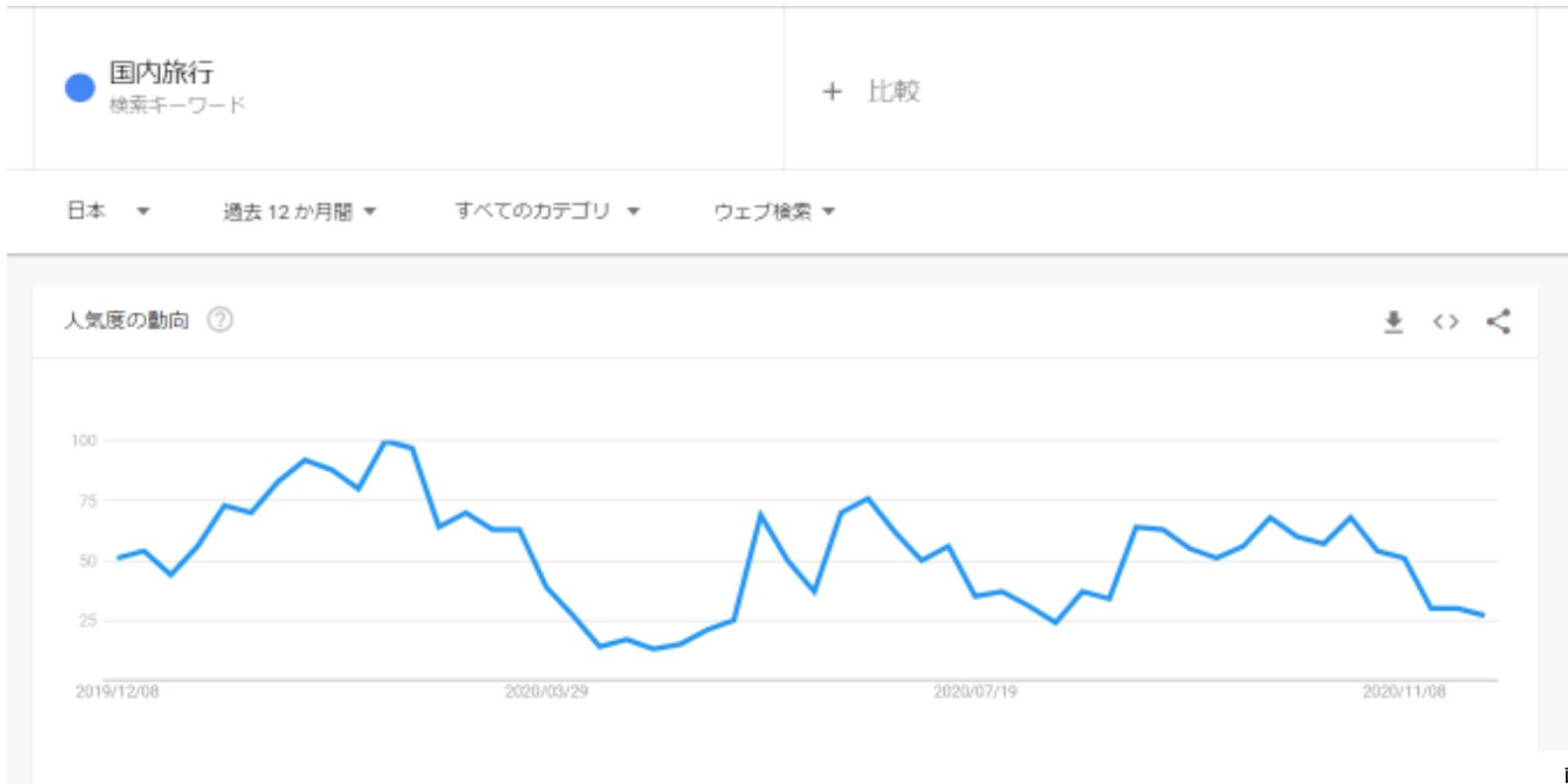
個人演習③

Googleトレンドの分析

Googleトレンドで「国内旅行」というキーワードを検索したら、昨年1年間のデータを出しましょう。

データから国内旅行について、どのようなトレンドがあると思うか自分の意見をまとめてみましょう。

※下記のグラフは、2019年12月ー2020年12月で調べた例です。



プライマリデータの活用方法①



データマーケティングの前提

■ データマーケティングのステップ

① データ収集

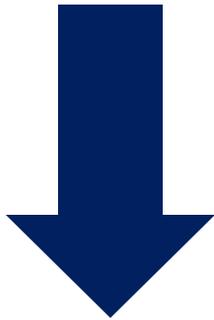


収集するデータについて

データには、「数値データ」「画像データ」「テキストデータ」「音データ」「動画データ」などがあります。

また、データの種類として「定量データ」と「定性データ」があります。

また、誰が収集したデータかによって、「プライマリーデータ」と「セカンダリーデータ」に分けられます。



② データ分析

「定量データ」・「定性データ」

定量データとは

商品の市場占有率や売上高など数値として把握できる情報を指します。数値データであるため、時系列や性別・年齢など顧客別に細かく加工し比較することが可能となり、あらゆるマーケティング戦略の基礎データとして活用されます。

定性データとは

「購入した商品のどこが気に入ったか」

「商品のパッケージデザインは使いやすいか」

「なぜこのお店でよく買うのか」

などのように、数値に表せない質的な情報を指します。

消費者の購買意識などを分析するデータとして活用されます。

消費者行動の広さを把握するのが定量データであり、消費者行動の深さを測定するのが定性データと位置づけることができます。

「プライマリーデータ」・「セカンダリーデータ」

プライマリーデータとは

自らリサーチを実施し収集したデータのことを指します。
アンケートやインタビューなどを実施して得たデータです。

セカンダリーデータとは

世の中に既に存在しているデータを指します。
過去の統計調査や市場調査データなどの
既存資料を調べて得たデータです。



プライマリーデータ

プライマリーデータの特徴

メリット

1. 調査実施者が欲しいデータを欲しい状態で入手できる

2次データでは、データの対象や取得方法、分類方法などが必ずしも調査実施者が求めるものと合わない場合があります。一方、1次データは調査実施者が求めているデータを収集することが出来ます。

2. 様々な調査手法を組み合わせて取得できる

必要なデータを独自の方法で収集するので様々な調査手法を組みまわせることができます。このことによりよい精度の高い、深みのあるデータを収集することが可能となります。

デメリット

1. データ収集にお金がかかる

2. データ収集に時間がかかる

3. 結論、仮説ありきの分析が行われる可能性がある

調査実施者が設計、実施、収集したデータなので、活用・分析には、調査実施者の考え、仮説などが影響しやすくなります。

セカンダリーデータ

セカンダリーデータの特徴

メリット

1. アクセスが容易である

インターネット環境であれば誰でも容易に沢山の2次データにアクセスすることが可能となります。また、収集のためのコストや時間の節約ができます。

2. データの信頼度が高い

政府系機関等が収集しているデータであれば、そのデータの信頼度が高いといえます。

3. 分析がしっかりとされている

データ収集のみならず、そのデータに対する分析も適切にされていることが多く、調査結果の信憑性が増します。

デメリット

1. 誰もがアクセス・活用している

多くの方がアクセス可能なので、データ自体の独自性はないです。

2. 調査実施者が本当に知りたいデータが完全には揃っていない

調査実施者の知りたい形でデータが揃っていないことが多いです。

それゆえ、深い分析ができないケースもあります。

各データのメリット・デメリット（まとめ）

プライマリーデータ
(自ら集めたデータ)

セカンダリーデータ
(他人が既に集めたデータ)

メリット

- ・ 欲しいデータが集められる
- ・ 好きな調査方法で集められる

- ・ データがすぐに手に入り費用も安い
- ・ 専門家による分析もされている

デメリット

- ・ データ収集に時間と費用がかかる

- ・ 誰でもアクセスができる
- ・ 欲しいデータを十分に集められない

オブザベーション（顧客観察）

オブザベーション（顧客観察）とは

対象者（＝主に顧客）がどのような行動をしているか、どのような購買をしているかなどを、調査員が実際の現場に出向いて観察する調査方法。

一般的に、顧客へのアンケートやインタビューでは、あくまで顧客が意識している要望や不満の状況を知ることができるのですが、意識されていない不満や要望は出てきません。

意識していない顧客の要望や不満を見つけ出す手法として有効なのが、オブザベーションという調査手法となります。

まずはお店や施設に来店されるお客様をじっくりと観察することから始めてみましょう。

個人演習

個人演習①

オブザベーションの練習

あなたが働くお店または施設で顧客1人をじっくりと観察しましょう。その顧客はどんな行動をとっていますか。どんな表情でいますか。以下の質問に答える形で、気づいた点をまとめてみましょう。

観察ポイント①

顧客の基本情報（来店目的、性別、年齢、雰囲気、同行者有無）

観察ポイント②

顧客の行動（お店の中での顧客の行動を詳細に観察）

観察ポイント③

顧客の表情（顧客の表情や雰囲気を詳細に観察）

プライマリデータの活用方法②



アンケート調査について

【アンケート調査とは】

調査対象（特定商品の顧客や旅行者など）の意識や行動を把握するために、統計的な調査の場合は、一定のルールで調査の対象を選び（＝スクリーニング）、様々な調査方法で多数の人に回答を求め、特定の期間内で様式化した質問への回答をもとに、統計的処理を行う調査方法をアンケート調査といいます。

通常、アンケート調査の実施に際しては、アンケート調査票が事前に作られ、回答者はそれにこたえる形で回答します。



アンケート調査の設計ポイント

Why	なぜ実施？	・ 既存資料では欲しい情報が得られないから
	何の目的で？	・ 企業活動、行政、研究などに役立てるために
What	何を調査？	・ 調査対象の実態・意識・行動など
How (When/Where)	どのようにして 調べるか？	<ul style="list-style-type: none">・ 調査目的により、対象者の条件を定める・ 一定のルールで調査対象を選出（スクリーニング） → 同じ条件で実施すれば誰が行っても、また答える人が同じ調査対象に属する別の人であっても、同じ結果が得られる・ 統計的である必要がない場合は回答者を募集 → 統計的ではないアンケート調査は、同じ結果が得られる保証はない・ 多数の人に回答を求める・ 特定の期間内に調査する・ アンケート調査票を用い、様式化した質問で回答を求める・ 回答回収後、統計的処理を行う
Who	調査主体は？	・ 企業、官公庁、研究機関など

アンケート調査の種類

アンケートの依頼方法と回収方法により様々な方法があります。調査方法を決めるには、調査の「課題」、「目的」、「期間」、「経費」など、様々な事項を検討します。

【アンケート調査の実施概要】

①調査を依頼する場所

- ・対象者の居住場所など特定の場所、施設などで人が集まる場所で直接依頼する方法
- ・Web画面や新聞紙面などで告知する方法

②調査を依頼または募集メディア

- ・調査員、アンケートの募集員などの人力で実施
- ・郵便、電話、FAX、インターネットなどの情報通信メディアで実施
- ・新聞、雑誌などのマスメディアで実施
- ・チラシ、ポスターなどの広告メディアで実施 など

③アンケート票の記入者

- ・調査員が記入する他記式と、回答者自らが記入する自記式があります。

④アンケート票を回収する方法

- ・新聞、チラシなどで回答者を募集し、インターネットやはがきで回答を得るなど、依頼と回収の方法が異なる場合もあります。

【主な調査方法】

訪問面接調査、訪問留置き調査、街頭調査、来場者調査、郵送調査、電話調査、FAX調査、インターネットアンケート調査などがアンケートの調査方法には様々なやり方があります。

アンケート調査の種類

調査目的、調査内容、質問の数、対象者、
予算、実施期間などで最適な方法と種類を選びます

	調査を依頼する方法	調査方法の種類
直接依頼	調査員が訪問して依頼	<ul style="list-style-type: none">・ 訪問面接調査・ 訪問留置き調査
	街頭・施設等で調査員が依頼	<ul style="list-style-type: none">・ 街頭、来場者面接調査・ 街頭、来場者自記式調査・ 会場アンケート調査
	情報通信手段を利用して依頼	<ul style="list-style-type: none">・ 郵送調査・ 電話調査・ インターネットアンケート調査
公募	新聞や雑誌広告、チラシ、インターネット 広告、 店頭ポスター、パッケージ等で依頼	<ul style="list-style-type: none">・ インターネットアンケート調査・ ホームユーステスト・ FAX調査

アンケート調査の企画・設計

調査課題を設定し、予算、調査対象、調査方法、調査項目、分析方法などを決めます。企画・設計フェーズでは、どんな目的で、何を調べるために調査を実施するのか、という調査課題を明確に整理することがアンケート調査を成功させるために重要となります。

① 調査課題を設定する

PDCAサイクルに即して、調査課題の設定を行うと、何を課題とすべきかが分かりやすいです。

② 調査対象者を設定する

調査対象者の条件や属性などを定義し、対象者はどのようなリストから抽出するかを検討します。

③ 調査（実査）方法を検討する

対象者リストの有無、調査内容、地域、必要な調査精度、期間、費用などの条件を考慮して調査方法を決定します。

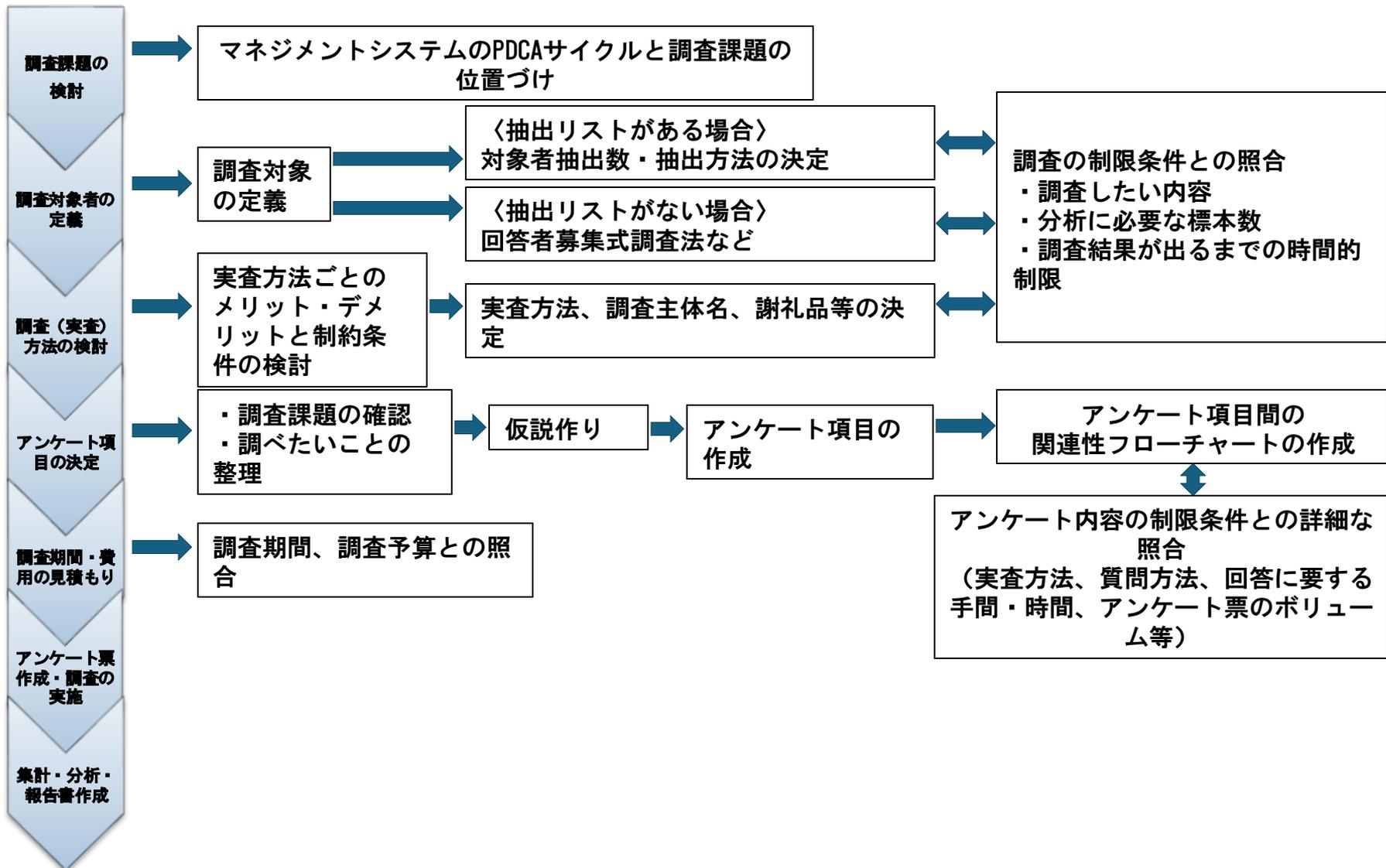
④ アンケート項目を決める

過去の調査結果や既存資料、社内検討会やグループインタビューなどで得た情報をもとにアンケートの調査項目を検討します。

⑤ 分析方法を決める

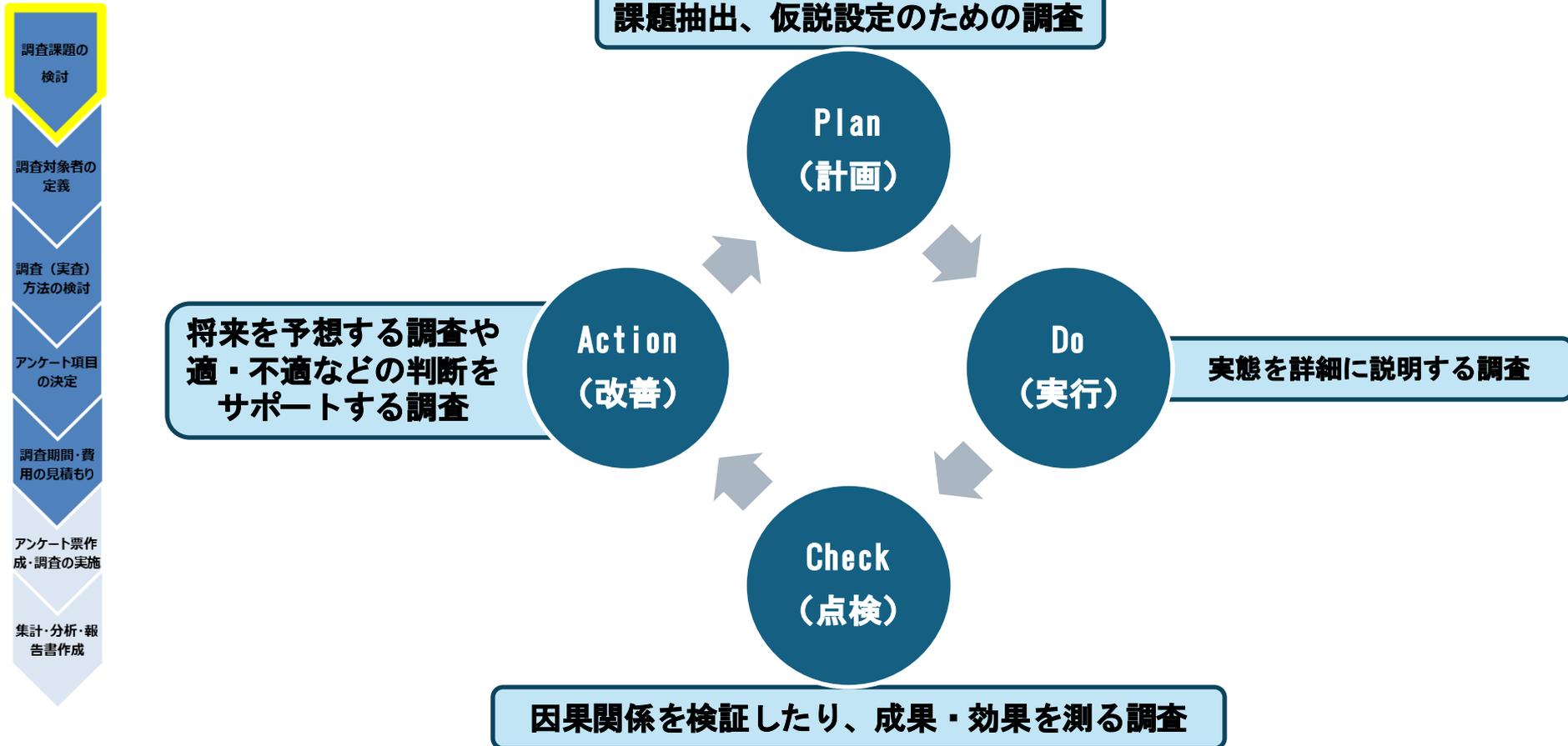
多変量解析などの高度な統計解析が必要な場合には、それに応じたアンケート票の設計と、標本数、期間、費用などをしっかりと決めることが必要です。

アンケート調査の企画・設計の進め方



調査課題を検討する

これから行おうとしている調査は、PDCAサイクルのどの段階に位置づけられるかを考え、調査課題を設定します。



調査対象者を定義する

調査対象の構成単位（個人か世帯かなど）、必要条件を時間的空間的に定義します。



① 調査対象と抽出リストを決める

マーケティングリサーチの場合、商品・サービスの利用者や見込み客の客層を対象とするのが一般的です。ターゲットが若者なら、調査対象は15～24歳の男女、小売業の見込客なら、施設を中心に半径5km圏内の住民などと定義します。

統計的な調査（標本調査）では、調査対象者のリストから、アンケートを依頼する人を確率的に抽出する方法である「サンプリング」を行い、実査の対象者を選びます。標本調査の対象（母集団）は、次のように定めます。

◆調査対象の構成単位・・・個人、世帯、会社、団体などの構成単位

◆調査目的に応じた必要条件・・・「性別、年齢、特定商品所有の有無」など個人の特性
「世帯の種類、上場か非上場か」など企業の

特性

◆空間的、時間的条件・・・「都道府県、市区町村、東京40km圏内」など地理的条件
「〇年〇月〇日現在で〇歳以上」などの時間的規定

② 抽出リストがない場合の対象者の決め方

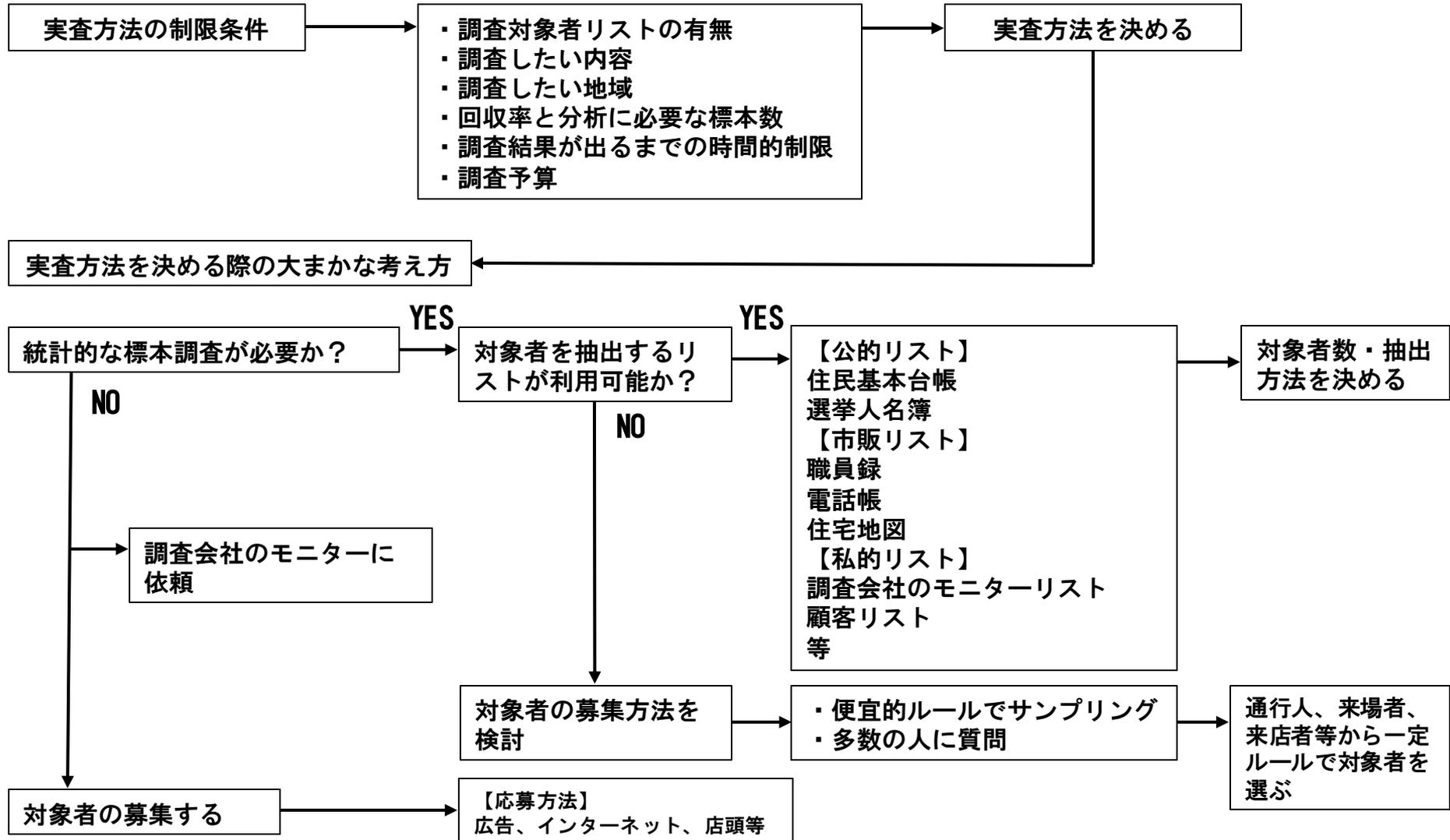
人が集まる場所での会場アンケートや、インターネットや広告メディアなどを通してアンケートを依頼することになります。

調査（実査）方法を検討する

調査対象者リストの有無、回収率、調査内容、調査地域、調査費用などを検討します。

- 
- ① **リストの有無を確認する**
調査対象者の住所などのリストがあれば、訪問・郵送調査などが可能です。
リストが無い場合は、調査対象者を募集する方法に限定されます。
 - ② **回収率を考慮する**
一般的に回収率が高いと、調査精度も高いと判断されます。
 - ③ **調査内容の量と質の両面を考える**
量では「回答に要する時間」、質では「対象者が回答しやすい方法」を検討します。
 - ④ **調査地域を考える**
調査目的や予算などを考慮しながら調査実施対象エリアを検討します。
 - ⑤ **調査期間を検討する**
いつまでに調査結果が必要かで、調査期間の許容範囲が決まります。
 - ⑤ **調査費用を検討する**
調査方法やサンプル数、対象エリアなどの条件で調査費用は大きく異なります。
例えば、訪問調査は、調査員の人件費など、費用が割高です。
調査目的と費用のバランスを考慮して調査方法を検討します。

調査（実査）方法を検討する手順



アンケート項目を決める

過去の資料なども参考にし、調べたい課題を大項目から小項目に細分化し、項目間の関連性を図化などして検討します。



① 過去の調査報告書類を参考にするときのチェックポイント

- ◆調査課題・・・PDCAサイクルのどの段階か判断します。
- ◆調査項目・・・目次構成やタイトルから、大項目や小項目を把握します。
- ◆アンケート項目・・・調査票の質問項目を参考にします。
- ◆詳細さの程度・・・どの項目をどの程度詳細に調べたかを参考にします。
- ◆因果関係の仮説検証・・・コメント部分を参考にします。
- ◆解析手法・・・統計的解析を行っている場合、統計解析技法に適合した質問項目とはどんなものか参考にします。

② 内部の検討会やグループインタビューなどの発言録から調べたい課題を抽出する

- ・キーワードを作る
- ・似たもの同士を集め、タイトルをつける
- ・行動と意識に分類し、因果関係を検討する

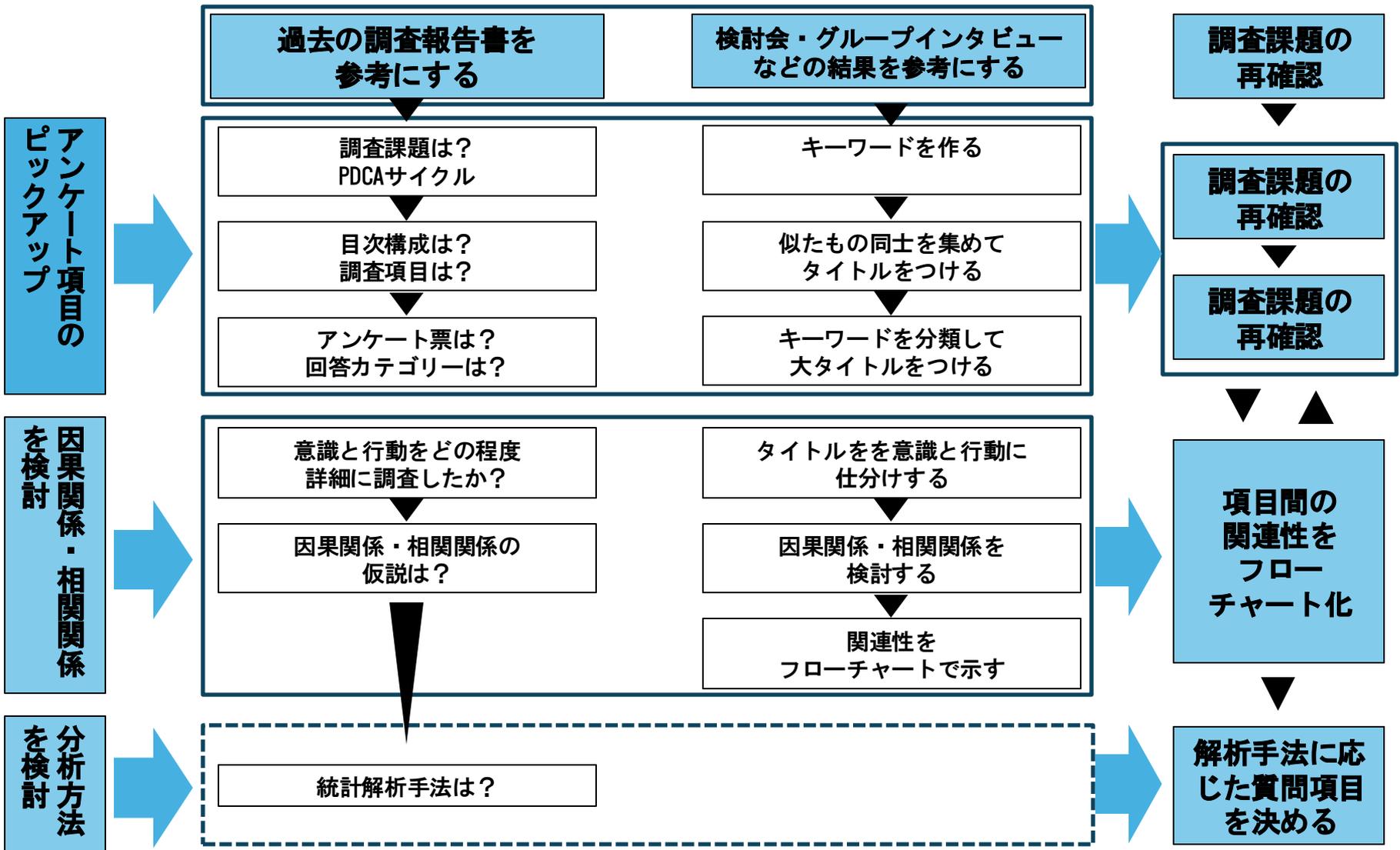
③ アンケート項目を考える

- ①・②の作業の後、アンケート項目を考えます。

④ 項目間の関連性フローチャートを作成する

項目間の関連性をフローチャートで示すと、集計分析の時に分析キーが理解しやすくなります。

アンケート項目を決める手順



調査に必要な期間を決める

調査期間は、目的、内容、地域、対象者、調査方法、標本抽出方法などを考慮して決めます。

調査期間を決める要素

調査目的	特定日に実施する必要がある場合は、実査期間が限定される
調査内容	回答に時間を要する質問は期間を加味
調査ボリューム	質問数が多い場合、休日を2回はさむように期間を設定
調査方法	調査方法により異なる
アンケート票の形式・印刷の有無	質問文と回答カテゴリ作成、印刷用レイアウト、印刷等の所要期間
対象地域	訪問面接調査等、調査員が現地に出向く場合は、地理的範囲や交通の便が所要期間に影響
対象者特性	対象者の年齢・職業により在宅率・在宅時間を考慮
対象者リストとサンプリング方法	多数の市町村に居住する対象者を、住民登録台帳、選挙人名簿から抽出する場合、抽出作業の所要期間が長い
調査員数	多数の調査員が必要な場合、調査員の募集、教育の期間も必要
データ処理	データ点検、入力の所要期間を見込む
自由回答の量	多くの自由回答を整理・集計する場合、集計所要期間が長くなる
解析の有無	複雑な統計解析（多変量分析）をする場合、期間がさらに必要

調査課題の
検討

調査対象者の
定義

調査（実査）
方法の検討

アンケート項目
の決定

調査期間・費用
の見積もり

アンケート票作
成・調査の実施

集計・分析・報
告書作成

アンケート票作成の手順

アンケート票は、協力をお願いするあいさつ部、質問本体部、対象者特性部の3部で構成します。



アンケート票の作成手順	アンケート票を作成する際のポイント
質問順の案作成	質問の流れをわかりやすく、答えやすい質問を先に、重要な質問はなるべく前に
各質問の質問方法、回答方法決定	質問内容に応じて、回答カテゴリの有無、回答カテゴリの種類（尺度）、答えの数、答えの形式等を決める
言葉遣いを決める	対象者、調査内容に応じた言葉遣い
各質問の回答カテゴリ案作成及び質問文作成	明瞭簡潔。誘導質問に注意 1 質問に複数質問が入らないように注意
回答方法の指示文案作成	回答方法に応じて指示文を統一 (当てはまるもの一つに○) (自由にお答えください) など
レイアウト案作成	調査ボリュームの調整 用紙サイズ、縦書き、横書き、罫線囲み、イラスト挿入 など
挨拶状(調査依頼状)作成	対象者の立場で
プリテスト	アンケート(案)を用いて、内部で実査を行い、問題点をチェック
見直し・問題点改善	
アンケート票のレイアウト決定	
印刷(またはコピー)	

アンケート調査の分析方法

その他、アンケート分析にはいくつか確立された手法があり、状況に応じて適切な手法を用いることで分析の効率を上げることができます。



クラスター分析

回答者を特定のルールで集団（クラスター）に分けて傾向を見る分析方法です。

クロス集計の新たな軸を作るときなどに活用することができます。

アソシエーション分析

数値同士の関連性（アソシエーション）を数値で算出する分析方法です。元データがあれば、ツールを使って自動的に算出を行うことが可能です。ビッグデータを扱うときなど、膨大な量のデータマイニングをする必要があるときなどに活用されています。

アンケート調査の分析方法

主成分分析

近い属性を集めて分析しやすくする方法を主成分分析といいます。

マーケティング戦略を作成する場面では、購買動機の分析やブランディング調査を行う際に使われています。

データをざっくりと見れるので、分析がしやすくなるメリットはありますが、細かな情報を捨てるデメリットもあります。

決定木分析

1つの原因から「もし〇〇だったら～」と仮説を繰り返し、複数の予測を立てていく分析方法です。

1つの原因を元に多くの仮説出しをすることができ、分析の過程で樹形図のような形ができあがっていく特徴があります。

プライマリーデータの活用方法③



個人演習①

アンケート調査票の作成

あなたが働くお店または施設における顧客向けアンケート調査票を作成します。アンケート票作成の手順を参考に、A4一枚片面のアンケート調査票を作成してみましょう。

【作成のポイント】

- ・ 調査課題を決める
- ・ 調査対象を決める
- ・ 調査目的を決める
- ・ 調査項目を決める（5～10問程度）



調査票の発表

グループで成果物を発表してください。

＜以下流れで発表&質疑応答＞

①グループの発表

②他のグループメンバーからの質問

③講師からの質問

→次の受講者へ

＜全員発表後＞

全体のまとめ（講師からのフィードバック）



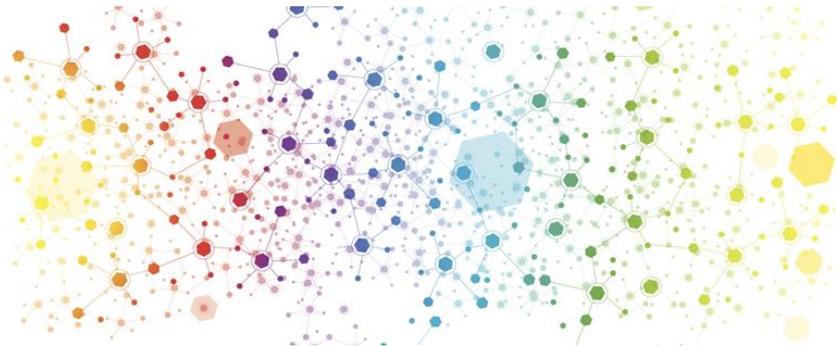
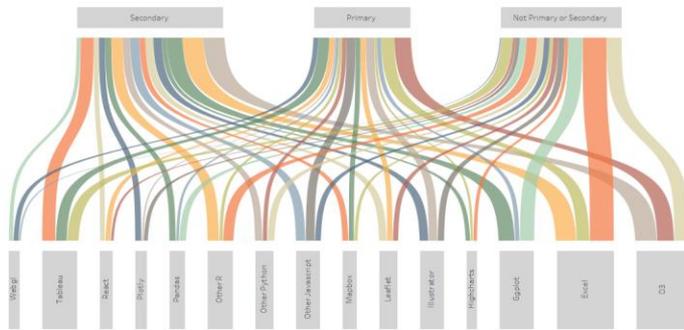
データ分析の流れ演習①



データの可視化

データの可視化

- データの可視化は近年「Data Visualization」としてひとつのジャンルを形成している。
- わかりやすい可視化はデータを用いたコミュニケーションを円滑にするツールとして効果が高い。



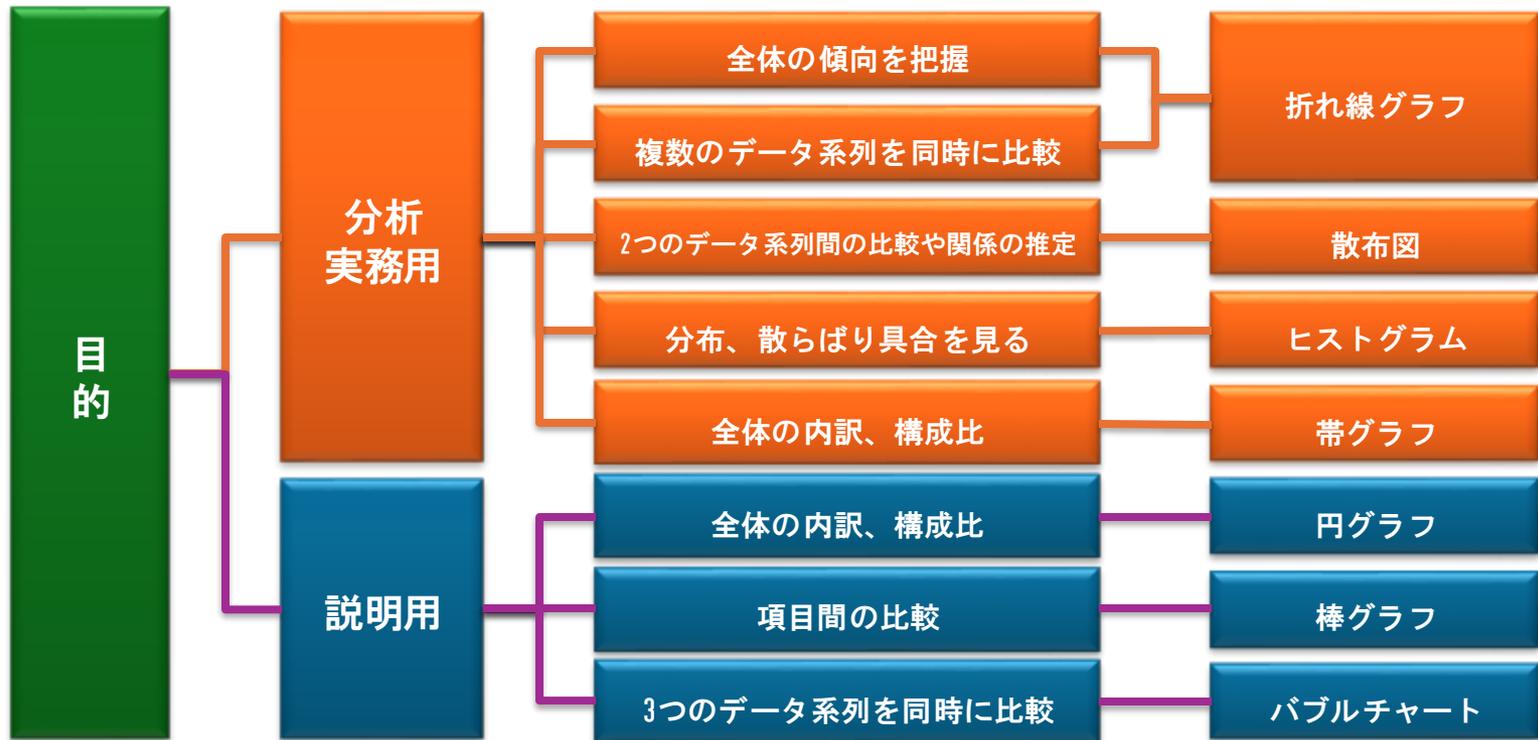
データの可視化のポイント

- データが表している内容を俯瞰して把握する。
 - そもそもデータが何を表しているかを理解する必要がある。
 - 俯瞰するためには情報の効率的な圧縮を考える。

- 適切なビジュアライゼーション手法を選択する。
 - 可視化手法には得意・不得意が存在する。

グラフの使い方

- グラフは適切な使い方があります。



グラフの特徴（1）

	折れ線グラフ	散布図	帯グラフ
図	<p>日本人の国内宿泊観光旅行の費用</p>	<p>サイト訪問回数と売上</p>	<p>アジア大洋州地域における主要国の国際会議開催件数</p>
目的	何らかの順序（時系列、量の多寡など）がある量の変化の傾向を見る。	2つのデータ系列を比較したり関係を推定するのに用いる。	構成比を比較する場合に用いる。
特徴	<ul style="list-style-type: none"> 汎用的に使える。 複数のデータ系列を一度にグラフにできる。 	<ul style="list-style-type: none"> 相関係数などの数値ではわからない情報が得られる。 	<ul style="list-style-type: none"> 構成比が見られるため、「質的な変化」を把握することができる。
注意点		<ul style="list-style-type: none"> 外れ値があるかを確認する。 	<ul style="list-style-type: none"> 比率が出てきたら帯グラフを用いる（円グラフではない）。

グラフの特徴（2）

	円グラフ	棒グラフ	バブルチャート																																				
図	<p>訪日外国人旅行者の内訳（2019年, 単位: 万人）</p> <table border="1"> <tr><th>地域</th><th>割合</th></tr> <tr><td>東アジア</td><td>12%</td></tr> <tr><td>東南アジア</td><td>6%</td></tr> <tr><td>北米</td><td>1%</td></tr> <tr><td>欧州主要5カ国</td><td>4%</td></tr> <tr><td>ロシア</td><td>2%</td></tr> <tr><td>オーストラリア</td><td>4%</td></tr> <tr><td>その他</td><td>71%</td></tr> </table>	地域	割合	東アジア	12%	東南アジア	6%	北米	1%	欧州主要5カ国	4%	ロシア	2%	オーストラリア	4%	その他	71%	<p>アジア大洋州地域における主要国の国際会議開催件数</p>	<p>国籍・地域別にみる訪日外国人旅行者1人あたり費目別旅行支出</p> <table border="1"> <tr><th>国籍・地域</th><th>支出額 (円)</th></tr> <tr><td>米国</td><td>30,506</td></tr> <tr><td>フランス</td><td>30,801</td></tr> <tr><td>ドイツ</td><td>23,464</td></tr> <tr><td>中国</td><td>108,788</td></tr> <tr><td>台湾</td><td>41,502</td></tr> <tr><td>香港</td><td>52,178</td></tr> <tr><td>韓国</td><td>17,939</td></tr> <tr><td>カナダ</td><td>22,970</td></tr> <tr><td>英国</td><td>23,218</td></tr> </table>	国籍・地域	支出額 (円)	米国	30,506	フランス	30,801	ドイツ	23,464	中国	108,788	台湾	41,502	香港	52,178	韓国	17,939	カナダ	22,970	英国	23,218
地域	割合																																						
東アジア	12%																																						
東南アジア	6%																																						
北米	1%																																						
欧州主要5カ国	4%																																						
ロシア	2%																																						
オーストラリア	4%																																						
その他	71%																																						
国籍・地域	支出額 (円)																																						
米国	30,506																																						
フランス	30,801																																						
ドイツ	23,464																																						
中国	108,788																																						
台湾	41,502																																						
香港	52,178																																						
韓国	17,939																																						
カナダ	22,970																																						
英国	23,218																																						
目的	メッセージをわかりやすく伝える。	複数の項目を比較する際に用いる。	3つ以上のデータ系列を一度に比べたい場合に用いる。																																				
特徴	<ul style="list-style-type: none"> 伝えたいメッセージが明確な場合の見た目の良さは群を抜いている。 分析実務の作業では円グラフを見て分かることは他のグラフにしても分かることが多い。 	<ul style="list-style-type: none"> 分析には折れ線グラフの方が使い勝手がよい。 積み上げをする際に有用。 	<ul style="list-style-type: none"> ポジショニング毎の特徴を見る際に有効（2軸でポジショニングを表し、バブルサイズで特徴を表現する）。 																																				
注意点	<ul style="list-style-type: none"> 構成にひ偏りがいないか 要素数が多すぎないか 	<ul style="list-style-type: none"> 横に並べる棒グラフはなるべく使わない。 	<ul style="list-style-type: none"> 情報量が増えるため、直感的な理解を妨げることがある。 																																				

ハンズオン

Three horizontal lines in blue, red, and green, stacked vertically, extending across the width of the page.

ハンズオン（1）

- ファイル名「演習データD3_1.xlsx」
- シート名「折れ線グラフ」



The screenshot shows the Microsoft Excel interface. The title bar indicates the file name is 'd31_3日目_セッション1'. The ribbon is set to 'ホーム' (Home). The spreadsheet data is as follows:

	A	B	C	D	E
1	日本人の国内宿泊観光旅行の費用と内訳（単位:円, 15歳以上）				
2					
3		宿泊費	交通費	その他	
4	2012年	17,700	13,580	14,290	
5	2013年	18,280	13,590	13,890	
6	2014年	19,350	14,230	14,700	
7	2015年	19,220	15,120	14,670	
8	2016年	19,770	14,540	14,520	
9	2017年	18,900	13,500	14,270	
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					

At the bottom of the screen, the chart menu is visible, with '折れ線グラフ' (Line Chart) highlighted by a red dashed circle. Other options include '散布図' (Scatter Plot), '帯グラフ' (Bar Chart), '円グラフ' (Pie Chart), '棒グラフ' (Bar Chart), and 'バブルチャート' (Bubble Chart).

ハンズオン (2)

自動保存 [オン] d31_3日目_セッション1

ファイル ホーム **挿入** 描画 ページレイアウト 数式 データ 校閲

ピボット テーブル テーブル

アドインを入手

個人用アドイン

おすすめ グラフ

A3

	A	B	C	D	E
1	日本人の国内宿泊観光旅行の費用と内訳 (単位:円, 15歳以上)				
2		宿泊費	交通費	その他	
4	2012年	17,700	13,580	14,290	
5	2013年	18,280	13,590	13,890	
6	2014年	19,350	14,230	14,700	
7	2015年	19,220	15,120	14,670	
8	2016年	19,770	14,540	14,520	
9	2017年	18,900	13,500	14,270	

領域選択

自動保存 [オン] d31_3日目_セッション1 検索

ファイル ホーム **挿入** 描画 ページレイアウト 数式 データ 校閲 表示 ヘルプ

ピボット テーブル テーブル

アドインを入手

個人用アドイン

おすすめ グラフ

A3

	A	B	C	D	E
1	日本人の国内宿泊観光旅行の費用と内訳 (単位:円, 15歳以上)				
2		宿泊費	交通費	その他	
4	2012年	17,700	13,580	14,290	
5	2013年	18,280	13,590	13,890	
6	2014年	19,350	14,230	14,700	
7	2015年	19,220	15,120	14,670	
8	2016年	19,770	14,540	14,520	
9	2017年	18,900	13,500	14,270	

2-D折れ線

折れ線

この種類のグラフの使用目的:

- 一定の時間 (年数、月数、日数) や項目にわたるデータの傾向を示します。

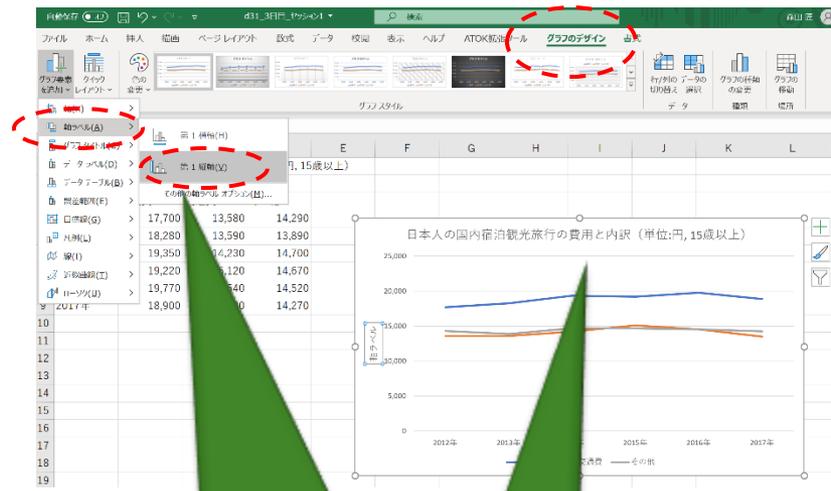
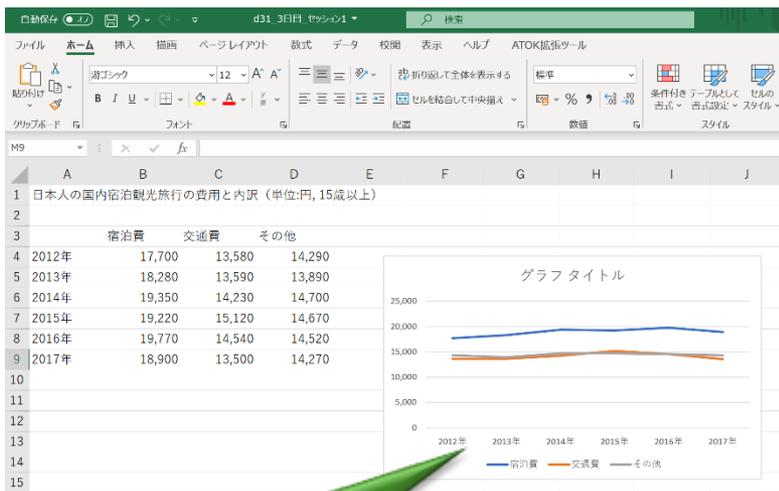
使用ケース:

- 項目の順序が重要な場合に使用します。
- データ要素の数が多い場合に使用します。

3-D 面

「挿入」タブのグラフ領域にある「2-D折れ線」から「折れ線グラフ」を選択する。

ハンズオン (3)

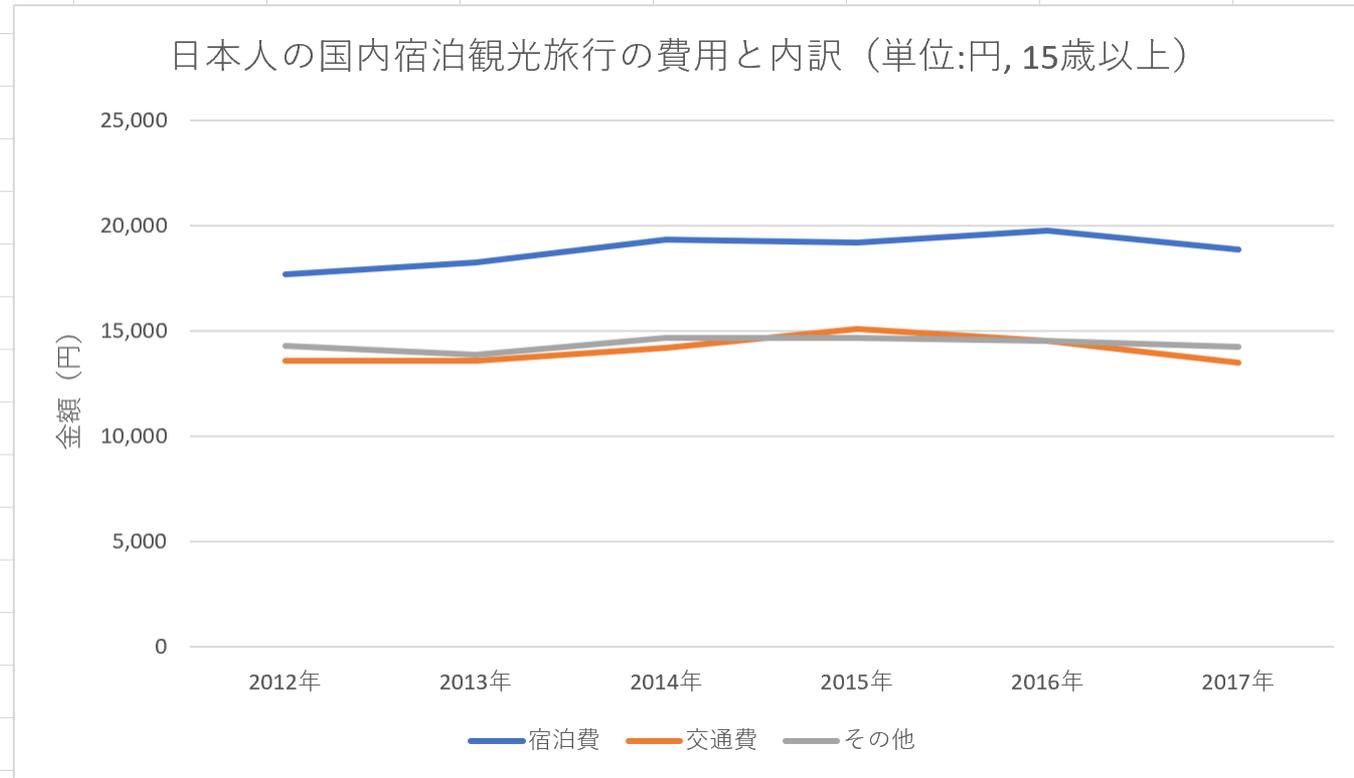


折れ線グラフが表示される。

軸ラベルの追加
「グラフのデザイン」タブの「グラフ要素を追加」→「軸ラベル」→「第1縦軸」で追加できる。

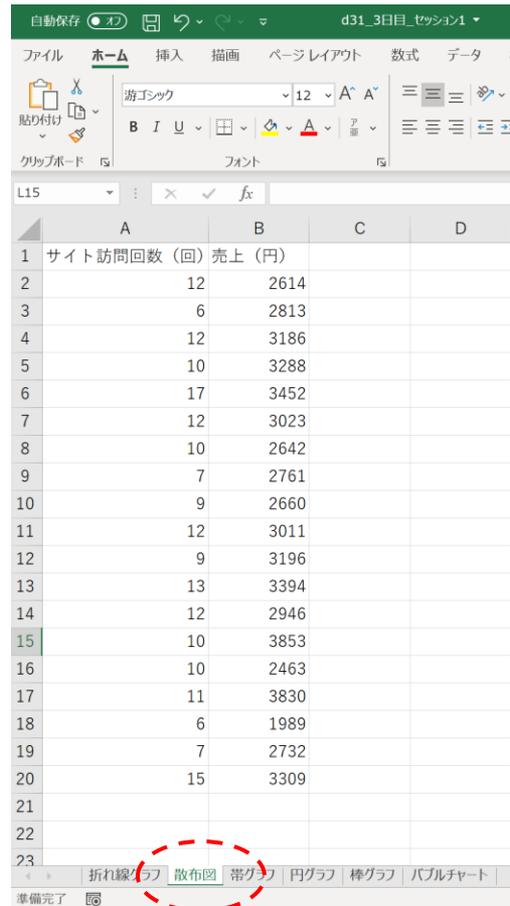
タイトルを追加

ハンズオン（４）



ハンズオン（5）

- ファイル名「演習データD3_1.xlsx」
- シート名「散布図」



The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D
1	サイト訪問回数 (回)	売上 (円)		
2	12	2614		
3	6	2813		
4	12	3186		
5	10	3288		
6	17	3452		
7	12	3023		
8	10	2642		
9	7	2761		
10	9	2660		
11	12	3011		
12	9	3196		
13	13	3394		
14	12	2946		
15	10	3853		
16	10	2463		
17	11	3830		
18	6	1989		
19	7	2732		
20	15	3309		
21				
22				
23				

The status bar at the bottom shows the following options: 折れ線グラフ, 散布図, 帯グラフ, 円グラフ, 棒グラフ, バブルチャート. The '散布図' option is circled in red.

ハンズオン (6)

	A	B	C	D
1	サイト訪問回数 (回)	売上 (円)		
2		12	2614	
3		6	2813	
4		12	3186	
5		10	3288	
6		17	3452	
7		12	3023	
8		10	2642	
9		7	2761	
10		9	2660	
11		12	3011	
12		9	3196	
13		13	3394	
14		12	2946	
15		10	3853	
16		10	2463	
17		11	3830	
18		6	1989	
19		7	2732	
20		15	3309	

「挿入」タブのグラフ領域にある「散布図」から「散布図」を選択する。

散布図

この種類のグラフの使用目的:

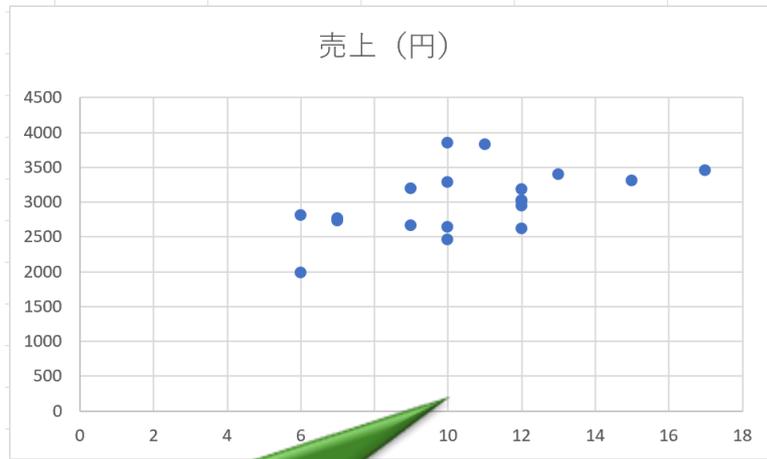
- 少なくとも 2 つのセットの値または 2 組のデータを比較します。
- 値のセットの関係を示します。

使用ケース:

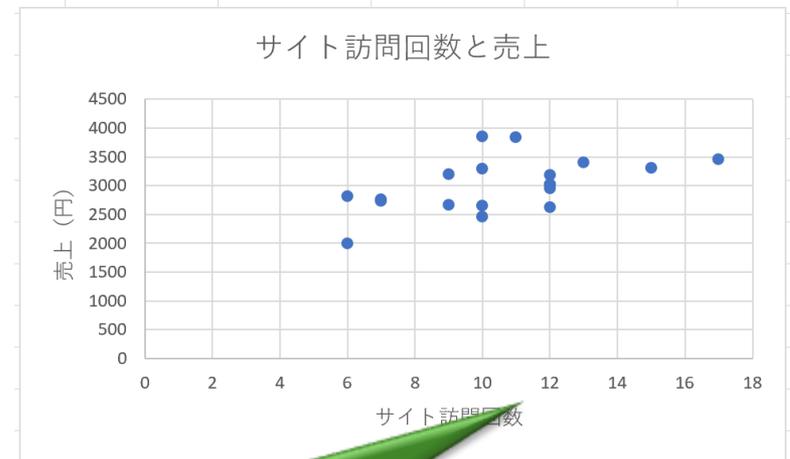
- データがばらばらしている場合に使用します。

その他の散布図(M)...

ハンズオン（7）



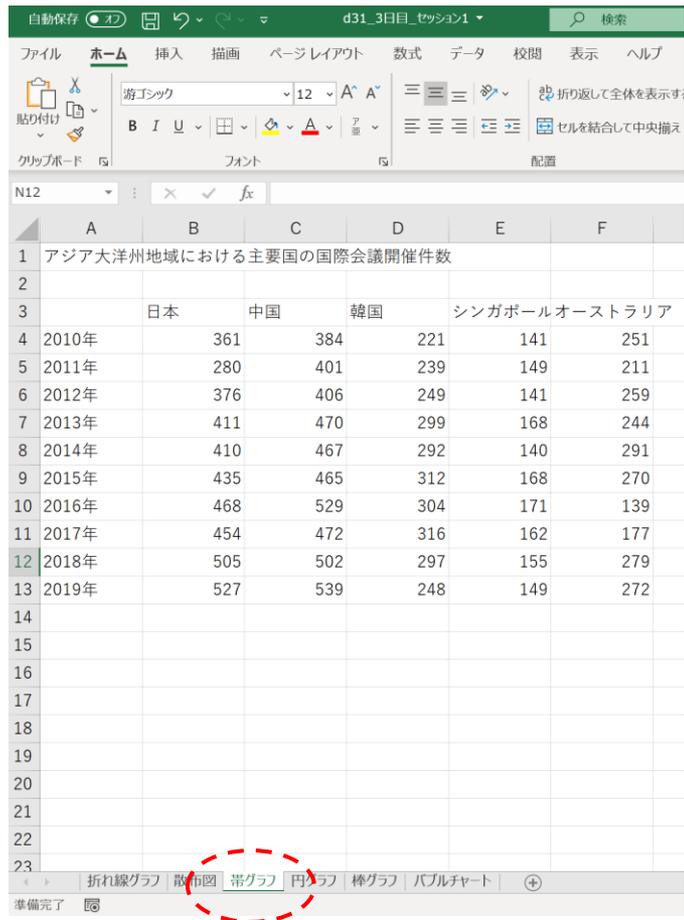
散布図が表示される。



タイトルと軸ラベルを追加

ハンズオン（8）

- ファイル名「演習データD3_1.xlsx」
- シート名「帯グラフ」



The screenshot shows the Microsoft Excel interface. The ribbon is set to '帯グラフ' (Bar Chart), which is highlighted with a red dashed circle. The data table is as follows:

	A	B	C	D	E	F
1	アジア大洋州地域における主要国の国際会議開催件数					
2						
3		日本	中国	韓国	シンガポール	オーストラリア
4	2010年	361	384	221	141	251
5	2011年	280	401	239	149	211
6	2012年	376	406	249	141	259
7	2013年	411	470	299	168	244
8	2014年	410	467	292	140	291
9	2015年	435	465	312	168	270
10	2016年	468	529	304	171	139
11	2017年	454	472	316	162	177
12	2018年	505	502	297	155	279
13	2019年	527	539	248	149	272
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						

ハンズオン（9）

	A	B	C	D	E	F
1	アジア大洋州地域における主要国の国際会議開催件数					
2						
3		日本	中国	韓国	シンガポール	オーストラリア
4	2010年	361	384	221	141	251
5	2011年	280	401	239	149	211
6	2012年	376	406	249	141	259
7	2013年	411	470	299	168	244
8	2014年	410	467	292	140	291
9	2015年	435	465	312	168	270
10	2016年	468	529	304	171	139
11	2017年	454	472	316	162	177
12	2018年	505	502	297	155	279
13	2019年	527	539	248	149	272

領域選択

「挿入」タブのグラフ領域にある「2-D横棒」から「100%積み上げ横棒」を選択する。

100% 積み上げ横棒

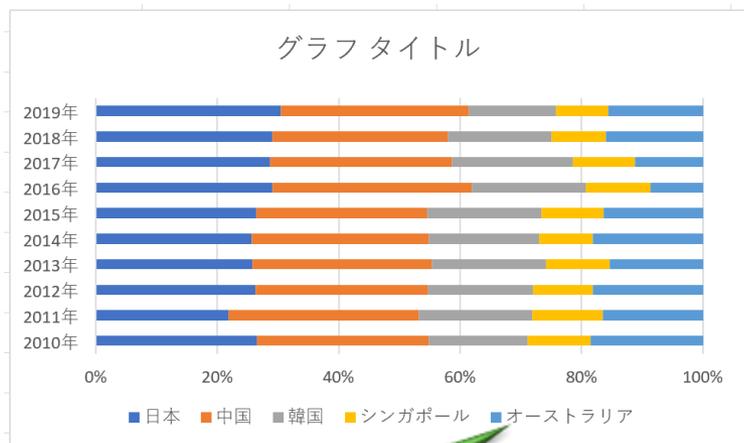
この種類のグラフの使用目的:

- 値の全体に対する割合を比較します。
- 値の全体に対する割合が時間とともにどのように変化するかを示します。

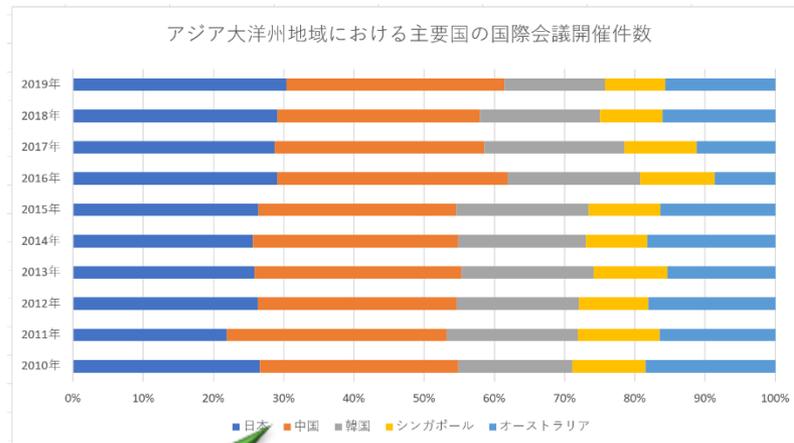
使用ケース:

- 項目の文字列が長い場合に使用します。

ハンズオン (10)

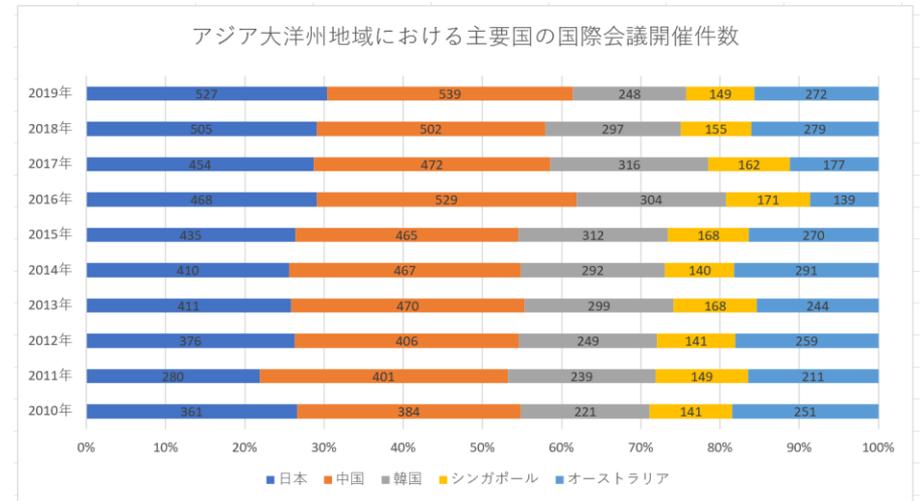
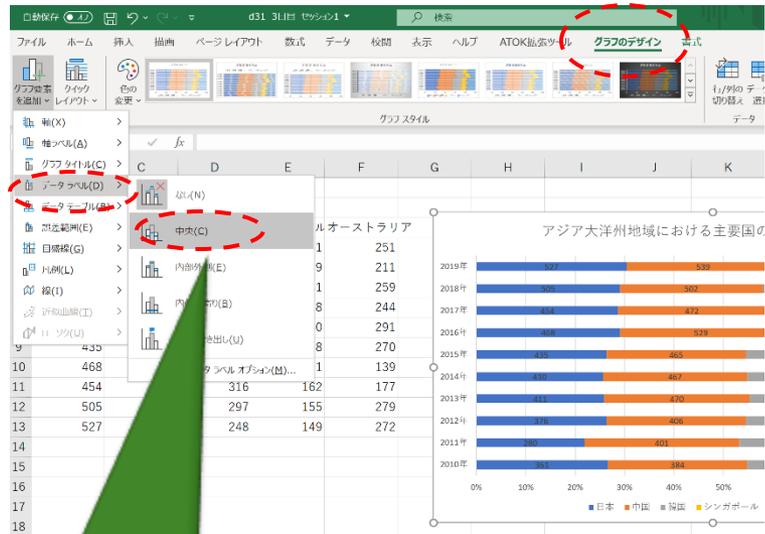


帯グラフが表示される。



タイトルを追加

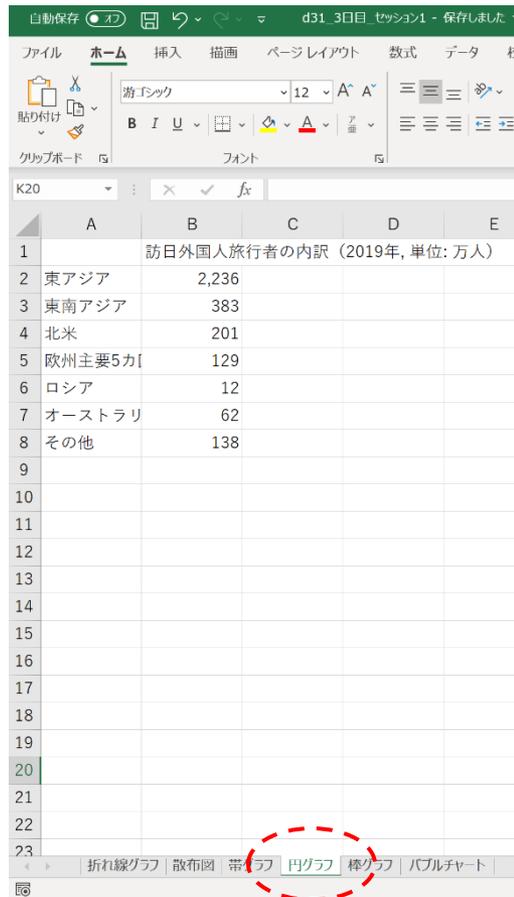
ハンズオン (11)



データラベルの追加
「グラフのデザイン」タブの「グラフ要素を追加」→「データラベル」→「中央」で追加できる。

ハンズオン (12)

- ファイル名「演習データD3_1.xlsx」
- シート名「円グラフ」



The screenshot shows the Microsoft Excel interface. The title bar indicates the file name is 'd31_3日目_セッション1 - 保存しました'. The ribbon is set to 'ホーム' (Home). The spreadsheet data is as follows:

	A	B	C	D	E
1		訪日外国人旅行者の内訳 (2019年, 単位: 万人)			
2	東アジア	2,236			
3	東南アジア	383			
4	北米	201			
5	欧州主要5カ国	129			
6	ロシア	12			
7	オーストラリア	62			
8	その他	138			
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					

At the bottom of the screen, the chart type menu is visible, with '円グラフ' (Pie Chart) highlighted by a red dashed circle.

ハンズオン (13)

自動保存 (オフ) d31_3日目_セッション1 - 保存しました

ファイル ホーム **挿入** 描画 ページレイアウト 数式 データ 校閲

ピボット テーブル

ピボット テーブル

	A	B	C	D	E
1		訪日外国人旅行者の内訳 (2019年, 単位: 万人)			
2	東アジア	2,236			
3	東南アジア	383			
4	北米	201			
5	欧州主要5カ国	129			
6	ロシア	12			
7	オーストラリア	62			
8	その他	138			
9					

領域選択

自動保存 (オフ) d31_3日目_セッション1 - 保存しました

ファイル ホーム **挿入** 描画 ページレイアウト 数式 データ 校閲 表示 ヘルプ ATOK

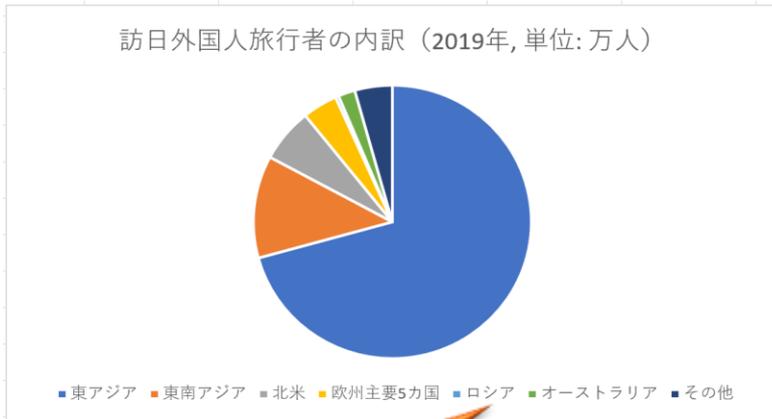
ピボット テーブル

ピボット テーブル

	A	B	C	D	E	F	G
1		訪日外国人旅行者の内訳 (2019年, 単位: 万人)					
2	東アジア	2,236					
3	東南アジア	383					
4	北米	201					
5	欧州主要5カ国	129					
6	ロシア	12					
7	オーストラリア	62					
8	その他	138					
9							

「挿入」タブのグラフ領域にある「2-D円」から「円」を選択する。

ハンズオン (14)



円グラフが表示される。

自動保存 (オン) d31_3日目_セッション1 検索

ファイル ホーム 挿入 描画 ページレイアウト 数式 データ 校閲 表示 ヘルプ

グラフ要素を追加 クイックレイアウト 色の変更

グラフスタイル

円グラフ!\$A\$2:\$A\$8,円グラフ!\$B\$2:\$B\$8,
E F
位: 万人)

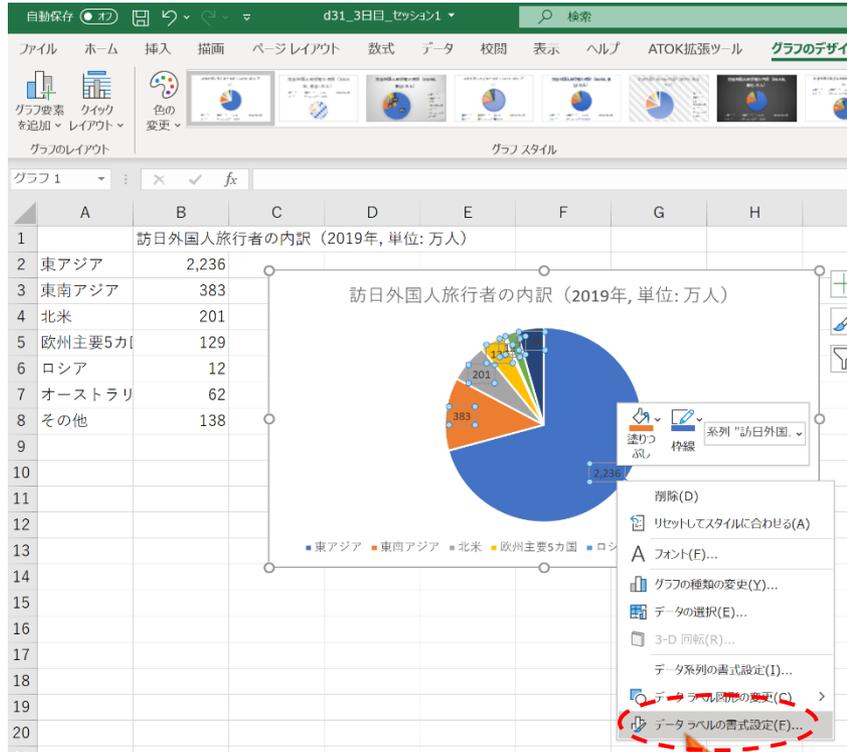
訪日外国人旅行者の内訳 (2019年, 単

円グラフ

2,236
383
201
129

データラベルの追加

ハンズオン (15)



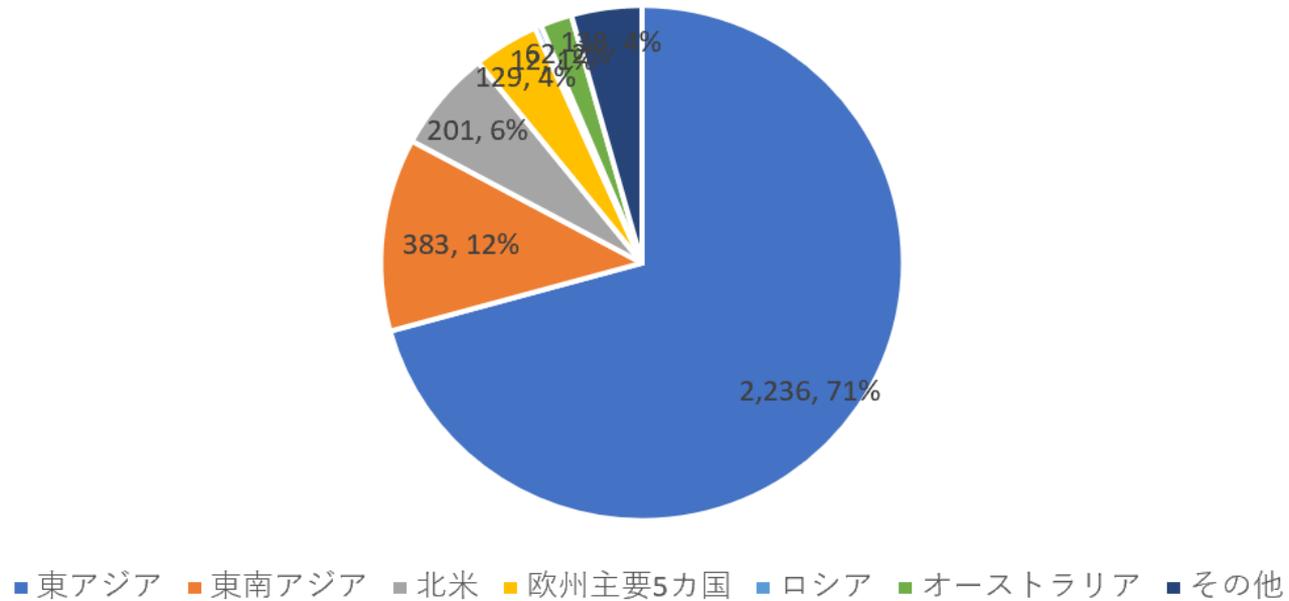
データラベルを選択した後に右クリック。「データラベルの書式設定」を選択。



パーセンテージにチェックを入れる。

ハンズオン (16)

訪日外国人旅行者の内訳 (2019年, 単位: 万人)



ハンズオン (17)

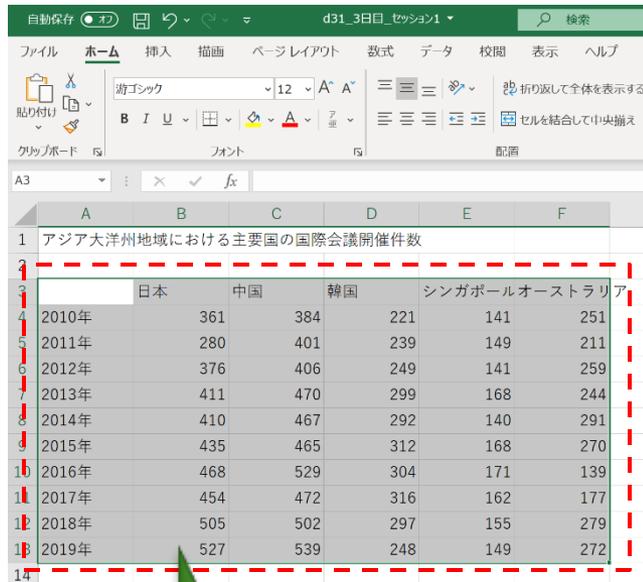
- ファイル名「演習データD3_1.xlsx」
- シート名「棒グラフ」

The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F	G
1	アジア大洋州地域における主要国の国際会議開催件数						
2							
3		日本	中国	韓国	シンガポール	オーストラリア	
4	2010年	361	384	221	141	251	
5	2011年	280	401	239	149	211	
6	2012年	376	406	249	141	259	
7	2013年	411	470	299	168	244	
8	2014年	410	467	292	140	291	
9	2015年	435	465	312	168	270	
10	2016年	468	529	304	171	139	
11	2017年	454	472	316	162	177	
12	2018年	505	502	297	155	279	
13	2019年	527	539	248	149	272	
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							

At the bottom of the screenshot, the chart type selector shows '棒グラフ' (Bar Chart) selected, indicated by a red dashed circle.

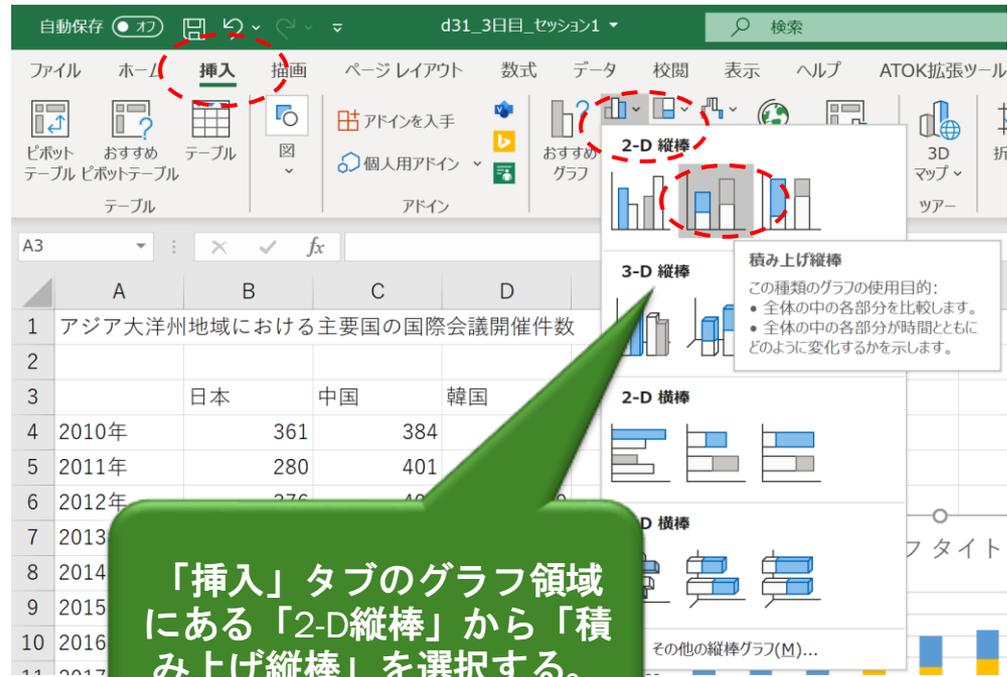
ハンズオン (18)



Excelの「挿入」タブの「グラフ」領域で「2-D縦棒」を選択する様子。赤い点線でデータ範囲が囲われており、緑色の吹き出しで「領域選択」と説明されている。

	日本	中国	韓国	シンガポール	オーストラリア
2010年	361	384	221	141	251
2011年	280	401	239	149	211
2012年	376	406	249	141	259
2013年	411	470	299	168	244
2014年	410	467	292	140	291
2015年	435	465	312	168	270
2016年	468	529	304	171	139
2017年	454	472	316	162	177
2018年	505	502	297	155	279
2019年	527	539	248	149	272

領域選択



Excelの「挿入」タブの「グラフ」領域で「2-D縦棒」を選択する様子。赤い点線で「挿入」タブと「2-D縦棒」が強調されており、緑色の吹き出しで「「挿入」タブのグラフ領域にある「2-D縦棒」から「積み上げ縦棒」を選択する。」と説明されている。

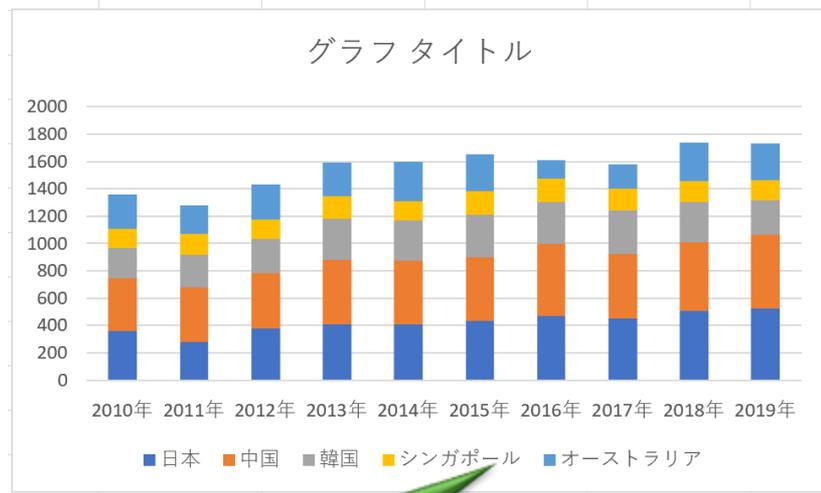
この種類のグラフの使用目的:

- 全体の中の各部分を比較します。
- 全体の中の各部分が時間とともにどのように変化するかを示します。

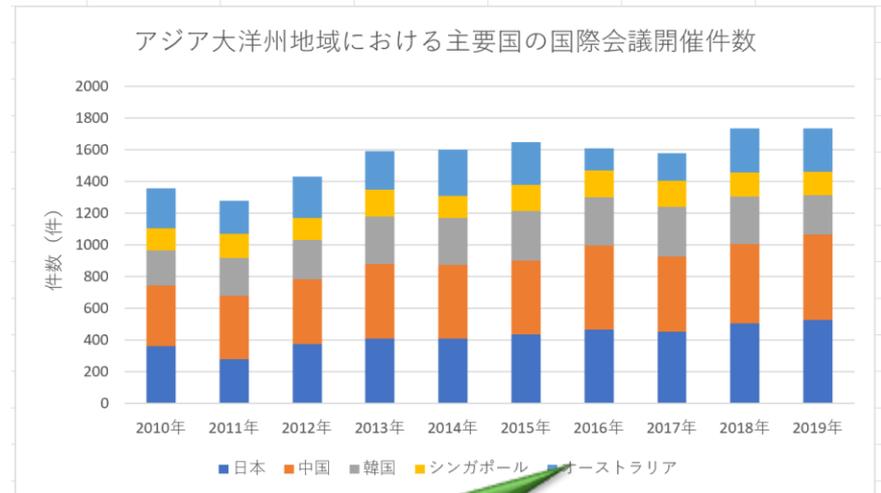
	日本	中国	韓国
2010年	361	384	
2011年	280	401	
2012年	376	406	
2013年	411	470	
2014年	410	467	
2015年	435	465	
2016年	468	529	
2017年	454	472	
2018年	505	502	
2019年	527	539	

「挿入」タブのグラフ領域にある「2-D縦棒」から「積み上げ縦棒」を選択する。

ハンズオン (19)

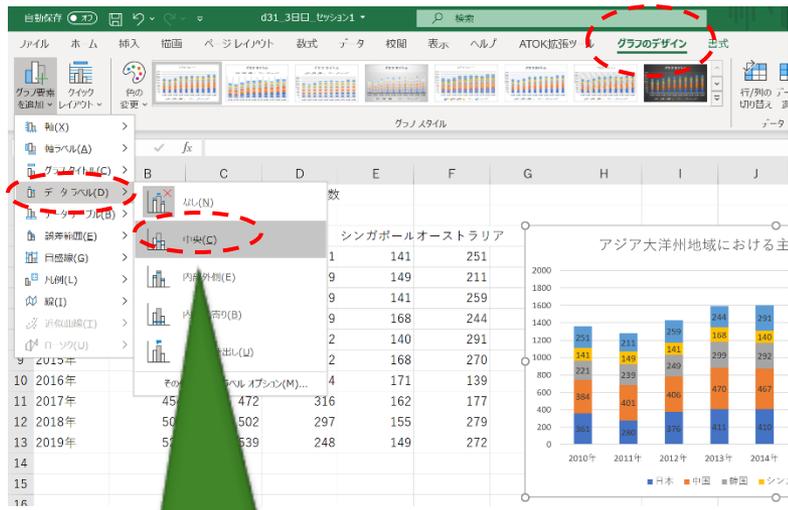


棒グラフが表示される。

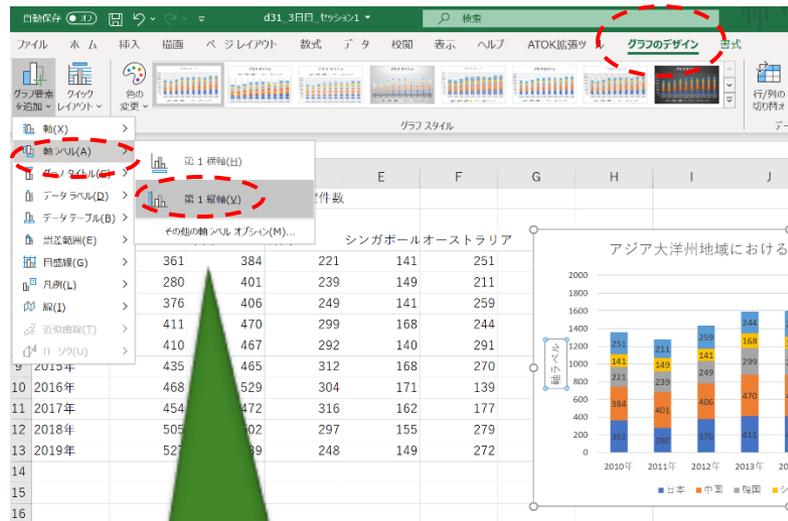


タイトルと軸ラベルを追加

ハンズオン (20)

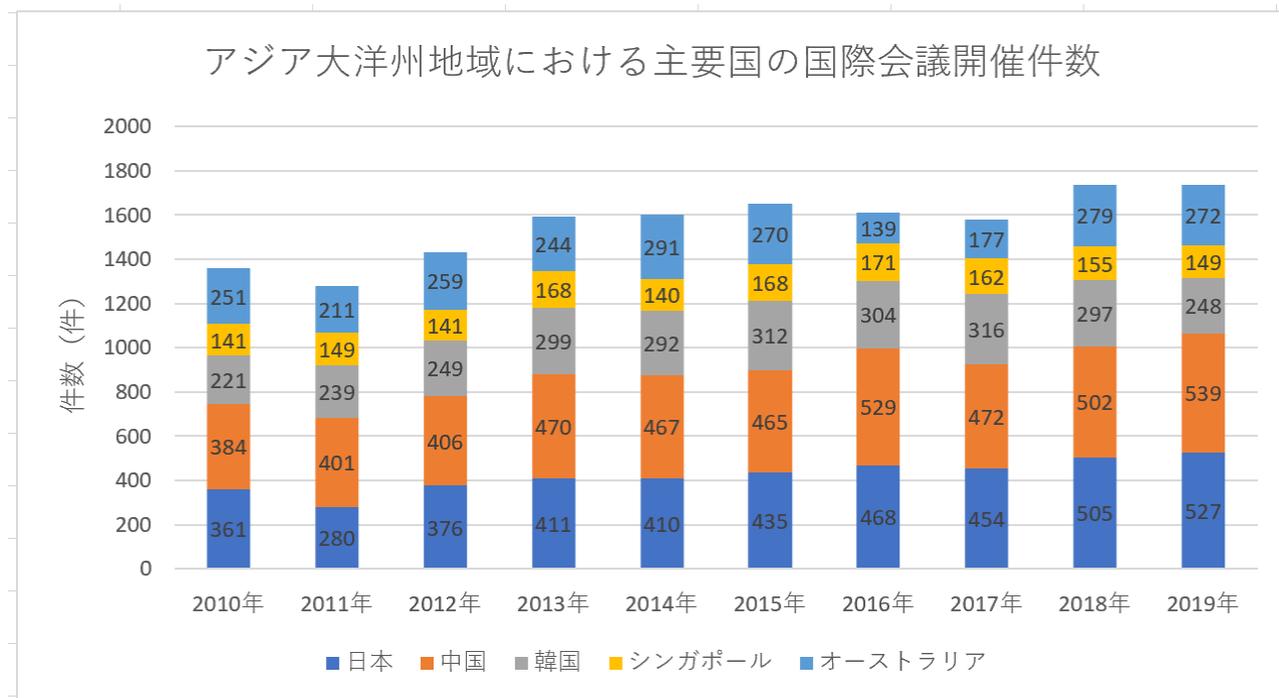


データラベルの追加



軸ラベルの追加

ハンズオン (21)



ハンズオン (22)

- ファイル名「演習データD3_1.xlsx」
- シート名「バブルチャート」

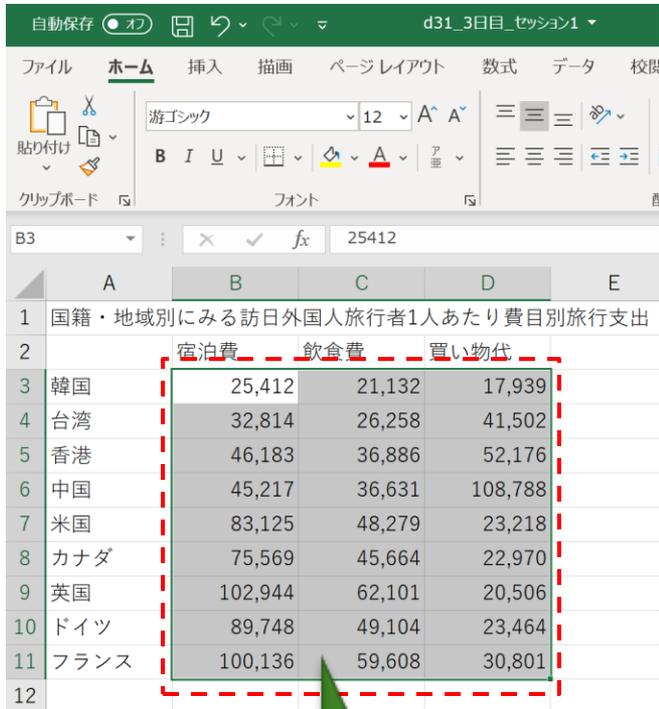


The screenshot shows the Microsoft Excel interface. The title bar indicates the file name is 'd31_3日目_セッション1'. The ribbon is set to 'ホーム' (Home). The active cell is K15. The data table is as follows:

	A	B	C	D	E
1	国籍・地域別にみる訪日外国人旅行者1人あたり費目別旅行支出				
2		宿泊費	飲食費	買い物代	
3	韓国	25,412	21,132	17,939	
4	台湾	32,814	26,258	41,502	
5	香港	46,183	36,886	52,176	
6	中国	45,217	36,631	108,788	
7	米国	83,125	48,279	23,218	
8	カナダ	75,569	45,664	22,970	
9	英国	102,944	62,101	20,506	
10	ドイツ	89,748	49,104	23,464	
11	フランス	100,136	59,608	30,801	
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					

At the bottom of the screenshot, the chart selection menu is visible, with 'バブルチャート' (Bubble Chart) highlighted by a red dashed circle.

ハンズオン (23)



自動保存 (オン) d31_3日目_セッション1

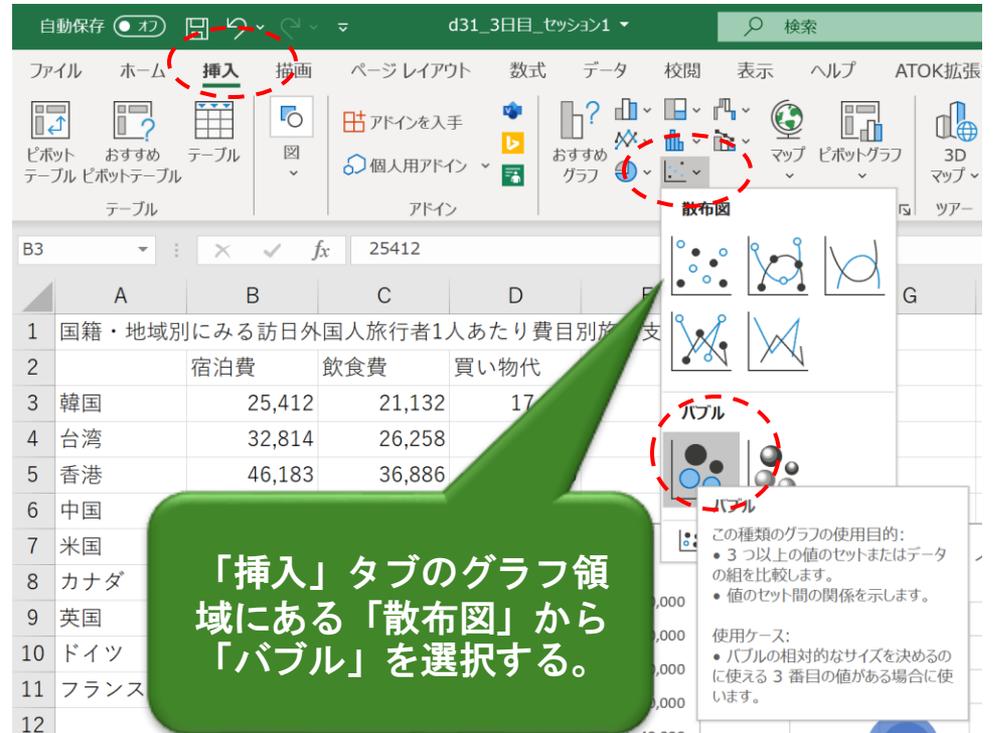
ファイル ホーム 挿入 描画 ページレイアウト 数式 データ 校閲

貼り付け 游ゴシック 12 A^ A^ B I U フォント

B3 25412

1	A	B	C	D	E
2		宿泊費	飲食費	買い物代	
3	韓国	25,412	21,132	17,939	
4	台湾	32,814	26,258	41,502	
5	香港	46,183	36,886	52,176	
6	中国	45,217	36,631	108,788	
7	米国	83,125	48,279	23,218	
8	カナダ	75,569	45,664	22,970	
9	英国	102,944	62,101	20,506	
10	ドイツ	89,748	49,104	23,464	
11	フランス	100,136	59,608	30,801	
12					

領域選択



自動保存 (オン) d31_3日目_セッション1 検索

ファイル ホーム 挿入 描画 ページレイアウト 数式 データ 校閲 表示 ヘルプ ATOK拡張

ピボット テーブル アドインを入手 アドインを入手 アドインを入手

おすすめ グラフ

散布図

バブル

この種類のグラフの使用目的:

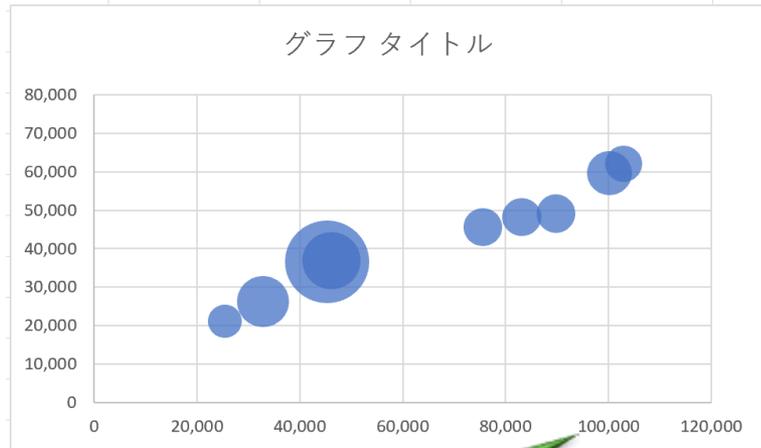
- 3 つ以上の値のセットまたはデータの組を比較します。
- 値のセット間の関係を示します。

使用ケース:

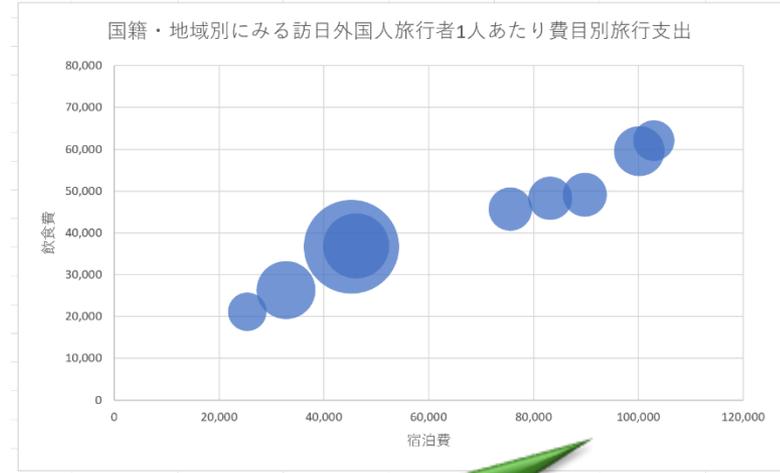
- バブルの相対的なサイズを決めるのに使える 3 番目の値がある場合に使います。

「挿入」タブのグラフ領域にある「散布図」から「バブル」を選択する。

ハンズオン (24)

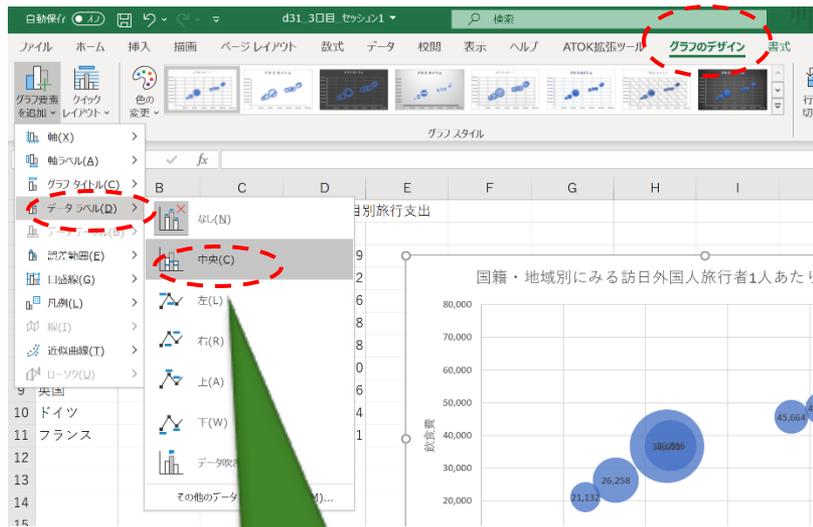


バブルチャートが表示される。

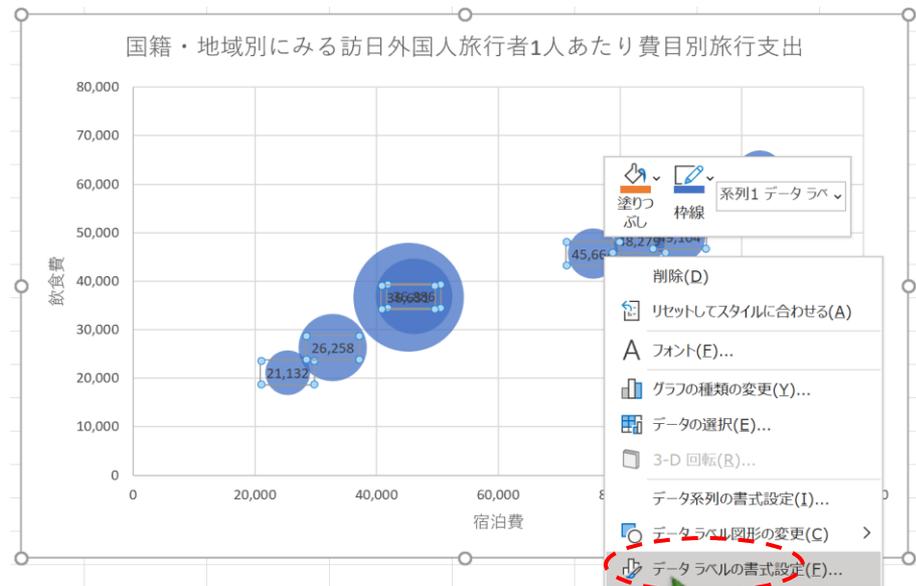


タイトルと軸ラベルを追加

ハンズオン (25)

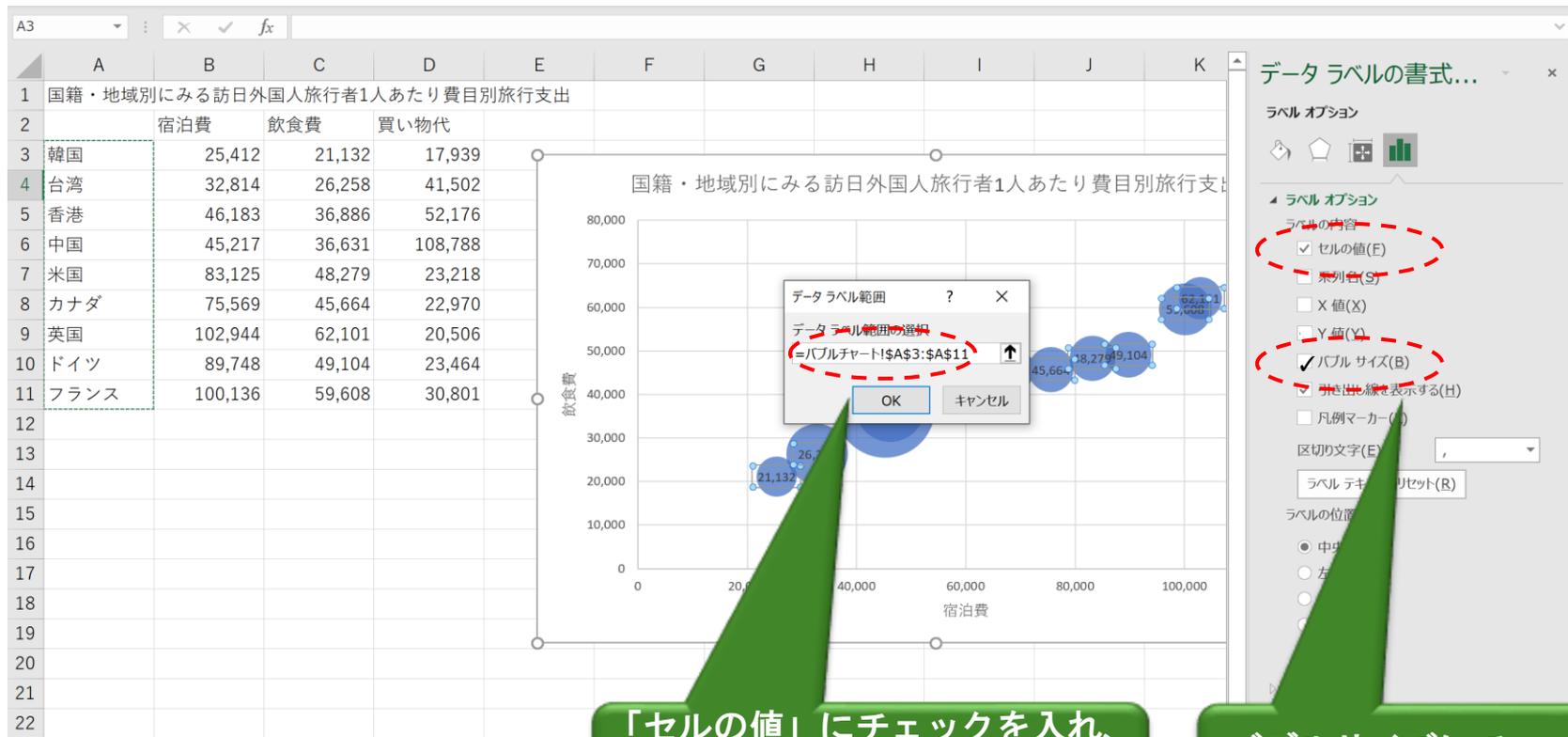


データラベルを追加

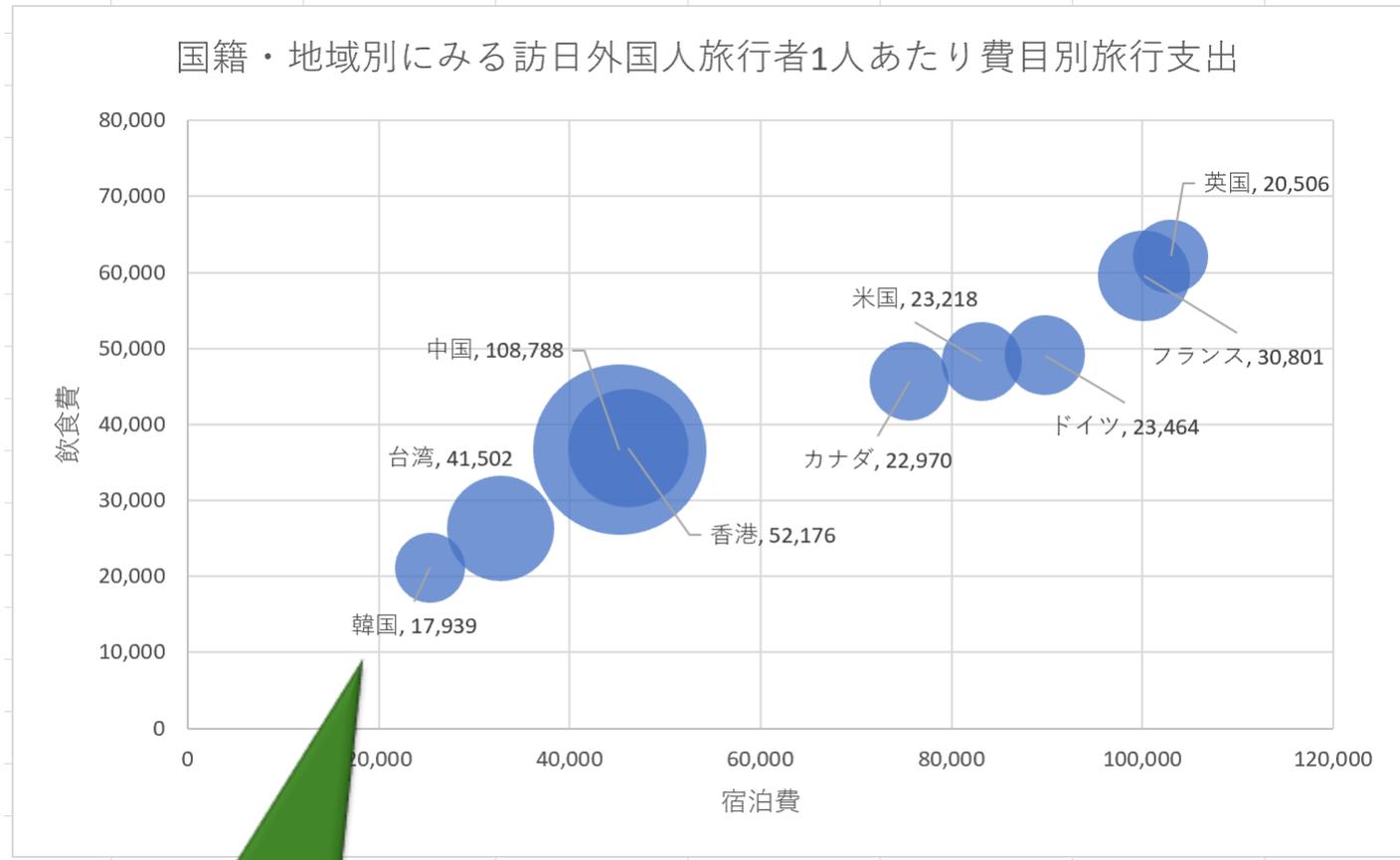


データラベルを選択した後に右クリック。「データラベルの書式設定」を選択。

ハンズオン (26)



ハンズオン (27)



データラベルをバランスよく配置する。

ワーク（1）

- 以下のワークシートを使ってハンズオンで実践したグラフから読み取れることを考えていく。

観光ビッグデータ	ワークシートg
データの読み取り	
1. 選択したグラフはなんですか？また、選んだ理由はなんですか？	
2. EXCELの表ではわかりずらかったものがグラフ化して見えてきたことはありますか？	
3. 選択したグラフで見える化ができそうなデータはどんなものがあるか できるだけたくさん書き出してみましょう。	
4. 自分の仕事において選択したグラフが役に立つ場面はどんなときですか？	

ワーク（2）

観光ビッグデータ

ワークシートg

データの読み取り

1. 選択したグラフはなんですか？また、選んだ理由はなんですか？

2. EXCELの表ではわかりずらかったものがグラフ化して見えてきたことはありますか？

3. 選択したグラフで見える化ができそうなデータはどんなものがあるか
できるだけたくさん書き出してみましょう。

4. 自分の仕事において選択したグラフが役に立つ場面はどんなときですか？

ハンズオンでグラフ化したグラフから
任意のグラフをひとつ選択して、選択
した理由を書き入れましょう。

ワーク (3)

観光ビッグデータ

ワークシート g

データの読み取り

1. 選択したグラフはなんですか？また、選んだ理由はなんですか？

2. EXCELの表ではわかりづらかったものがグラフ化して見えてきたことはありますか？

3. 選択したグラフで見える化ができそうなデータはどんなものがあるか
できるだけたくさん書き出してみましょう。

4. 自分の仕事において選択したグラフが役に立つ場面はどんなときですか？

データの表からは、わからなかったものが
グラフにすることによって見える化され、
わかるようになったことは何ですか？

ワーク（４）

観光ビッグデータ

ワークシートg

データの読み取り

1. 選択したグラフはなんですか？また、選んだ理由はなんですか？

2. EXCELの表ではわかりずらかったものがグラフ化して見えてきたことはありますか？

3. 選択したグラフで見える化ができそうなデータはどんなものがあるか
できるだけたくさん書き出してみましょう。

4. 自分の仕事において選択したグラフが役に立つ場面はどんなときですか？

選択したグラフは、他のどんなデータを扱うときに見える化ができそうかできるだけたくさん考えて見ましょう。絶対にできるという自信は必要ではありません。「使えるんじゃないかなあ」程度の予想で構いませんので、たくさん考えてみましょう。

ワーク（5）

観光ビッグデータ

ワークシートg

データの読み取り

1. 選択したグラフはなんですか？また、選んだ理由はなんですか？

2. EXCELの表ではわかりづらかったものがグラフ化して見えてきたことはありますか？

3. 選択したグラフで見える化ができそうなデータはどんなものがあるか
できるだけたくさん書き出してみましょう。

4. 自分の仕事において選択したグラフが役に立つ場面はどんなときですか？

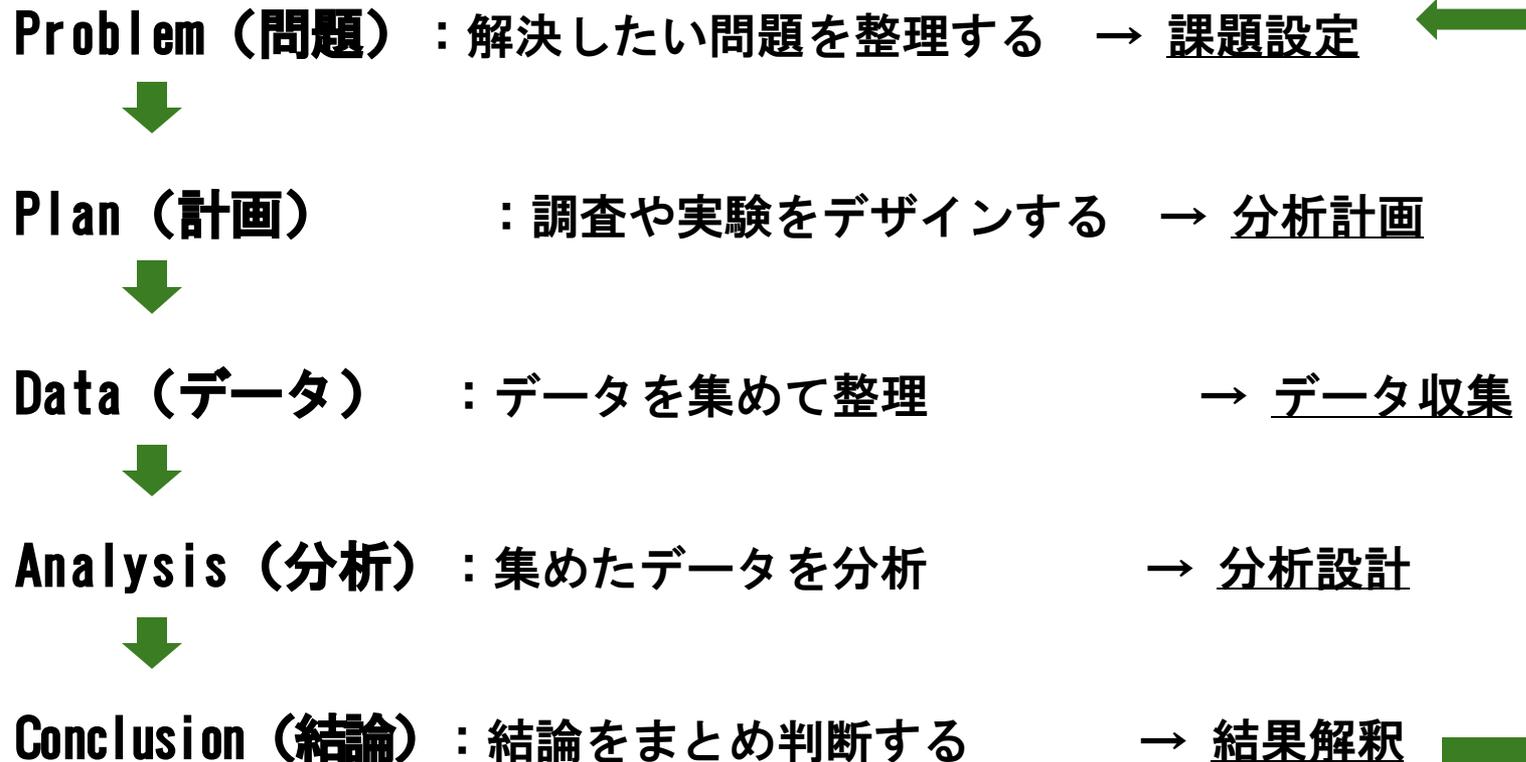
自分の仕事で、このグラフが役に立ちそうな場面を想像して、どう役に立っているのか想像してみましょう。

データ分析の流れ演習②



全体の流れ（１）

統計的な問題解決においては、次のPPDACによるサイクルで進めるのが一般的です。



全体の流れ（２）

PPDACサイクルでデータ分析を進める一例です。

Problem（問題）：大学で実施したTOEIC対策講座は有効だろうか？



Plan（計画）：受講学生とそれ以外の学生の点数を比べてみよう



Data（データ）：全学生の点数を入手し、受講の有無で2グループに分ける



Analysis（分析）：2グループの平均点を比較する（t検定）



Conclusion（結論）：あまり有意な差ではなかった…プログラムを見直そう

課題設定

身近な問題を明確にし、「何がしたいか（目的）」や「目的に関係することは何か」をはっきりさせるための段階

取り組む時のPoint

課題を設定する前の段階としては、少なくとも以下の点が考えられます。

実態を把握する→実際の状態を観察

目的を明確にする→理想の状態の定義

原因の仮説→実態を生んでいる要因の仮説立案

このあとの工程は、全てここで設定した目的から考えていく
→分析を通じて、何を実現したいのか最初に考えることが大切

いきなり分析を始めない！

課題をどうやって設定するか

データ分析の効用＝暗黙知からの脱却

過去の「経験と勘」を頼りに決めていたことを、根拠のある行動に変えていくことで成功確率を高める

→「何について、根拠のある行動に変えていきたいか」が課題設定のヒント！

「経験と勘」

以前は、こうだった気がする…
今回は、たぶんこうなる

データ中心の考え

分析結果からは、こうなる可能性がこれくらい高いといえる

分析計画

目的のために有効な調査・実験や分析をデザインする段階

取り組む時のPoint

分析を計画する際には、少なくとも以下のことを決める必要があります。

どのデータを使うのか

社内データ、オープンデータなどデータには様々な種類がある

どの分析手法を使うのか

様々な分析手法から、何を使うかは目的をもとに考えていく

どのような結果を期待するのか（仮説設定）

その期待をもとに分析の種類や高度さも考えることができる

一口に分析といっても様々です。

分析手法のいろいろ

より高度で柔軟で（抽象度が高い）ものが必ずしも優れているわけではない。

求める結果に応じて、適切な分析手法を選んでいくことが大切

より高度 & より柔軟

相関

重回帰分析

ニューラルネットワーク
(いわゆるAI)

分析の容易さ (コストの低さ)

ブラックボックス

データ収集

行ないたい分析のために必要なデータを収集し（分析で扱えるように）データを加工する（データ整理）する段階

取り組む時のPoint

データを取集し整理する段階では、少なくとも以下の作業が必要になります。

データを探して手に入れる

公的なデータを検索することもあれば、実験や調査で自ら手に入れることもある

データの有効性を検証する

集めたデータの信頼性や妥当性を改めて確認する

データを使えるように加工する（データクレンジング）

入力ミスや欠損によって使えなくなっている部分を加工して分析に備える

データはすんなりとは使えない

データクレンジング

収集したデータがそのまま数値的な分析にかけられるのは稀
データクレンジングという分析前の処理が必要になってくる

欠損値を埋める

入力ミスなどでデータが入っていないことはよくある
平均やデフォルトの値を入力するなどして、処理するのが一般的

外れ値処理

他の値から大きく外れた値のことをいう
測定ミス・記録ミス等に起因する異常値と実用上は区別できない

半角全角・大文字小文字などの表記ユレを統一

入力の段階で統一されていないと起こりうる
関数を使って一括変換できる場合もあるが、1つずつ置換していくしかない場合も多い

文字化け

OSの違いなどで発生 → 違う環境で処理する

ワーク



データ中心の考えに変えたいもの

身の回りで、「データ中心の考え」に変えたい分野を考える。

勘と経験に頼っていること

きちんとデータをとっていないこと

データはあるが、活用していないこと

など、

身の回りでデータ中心に変えていきたいものはないか考えてみます。

ワークの進め方（1）

Work Sheet

テーマ



個人ワーク



グループワーク

後で見分分かるように
テーマを書いておきましょう

個人の考えを書いておきましょう

データ中心に変えたい分野を考えます。
ここでは、思いつきでかまいません。
自分がデータ中心の考えに変えていきたい
分野を、どんどん書き出してみましよう。
自分の仕事関連でなくても、見たことがある
もの聞いたことがあるものでもかまいま
せん。
なるべく数多くだしていましよう。

ワークの進め方（3）

Work Sheet

テーマ



個人ワーク

グループで意見交換をしましょう。



グループワーク

個人書いたものを元に
グループ内で意見交換しましょう。

結論は出さなくて良いので
思いつくことをあげていきましょう。

ワークの進め方（3）

グループごとの意見を全体に共有しましょう。

どんな分野があがりましたか？

網羅的で良いので、他のグループに

伝えるためにも発表しましょう

※他のグループの説明から新しい視点を見つけてください。



分析設計

データからパターンを発見し、どのような要因が「目的」に影響を与えているかを分析していく段階

取り組む時のPoint

統計的な手段によって、分析できることは主に以下の通りです。

全体を把握する

統計（統べて計る）の言葉通り、情報の全体をとらえまとめて把握できる

複数の数値の関連性

別々のデータ（予測の対象と要因など）がどのように関わっているか（いなか）を

数学的に表現できる

その結果の精度

分析結果がどのくらい信頼できるのか（有意なのか）を数学的に算出できる

分かることと分からないことがある

分析による予測モデル

回帰分析などで予測モデルを構築すると説明変数から目的変数の予測が可能になります。

予測したい変数→目的変数
要因となる変数→説明変数

「何を予測したいか」から、目的変数が決まってきます。

x の函数 (Function) $y = f(x)$
(関数)

$$y = mx + b$$

従属変数
Dependent Variable

目的変数

独立変数
Independent Variable

説明変数

定数
Constant

中学数学の関数で扱った
 x と y に相当します

結果解釈

最初に設定した課題に対して結論を出し、新たな課題や今後の方針（追加分析）を検討・提示していく段階

取り組む時のPoint

分析結果が出たら、以下の作業が必要です。

結果の精度を把握する

仮説検定などの数学的な結果をもとに、その結果をどこまで信頼するのかを検討していく

→慣例はあっても数値だけで白黒つけられる場合ばかりではない

結果を評価する

事実と仮説（事前の期待）を比べ、結果から何が言えるのかを考察しまとめる

今後の方針を考える

分からなかったこと・もっと深く知るべきことを整理し次のステップを決めていく

分析は単発ではなくサイクル

統計的に有意とは

仮説検定という考え方で、算出した結果がどのくらい信頼できるのか数値的に判断できる。

算出の仕方や読み取り方の詳細は別単元で扱うが、ここでは「統計的に有意」の意味を確認する。

統計的に有意：

「実際に算出された結果」が偶然といえる可能性が極めて低いことを意味する。

例：新薬の治験の場合



②新薬と偽薬の効果に差がなくても①の差が偶然に発生してしまう確率を算出

③極めて低くければ（通常5%）、統計的に有意といえる

ワーク



分析を計画し、結果を読み取る

データを自由に分析して、結果を読み取ってみます。

ファイル「演習データE2.xlsx」を開きましょう。

そちらは、下記設定の架空のデータです。

大学生がオンラインで受けたクラス分けテストの結果

まずは、ここまで学習した方法を駆使して

データを様々な分析してみましよう。

ワークの進め方(1)

Work Sheet

テーマ



個人ワーク



グループワーク

後で見分分かるように
テーマを書いておきましょう

個人の考えを書いておきましょう

分析結果から読み取れることを書き込みます。

数値だけでなく
そこから推測できること（解釈）も
考えて書き加えていきましょう。

ワークの進め方(3)

Work Sheet

テーマ



グループで意見交換をしましょう。



個人での分析結果（算出した数値など）と解釈をグループ内で意見交換しましょう。結論は、出さなくても良いですが「全員が納得する解釈」はないか議論して探ってみましょう。

※分析手順に誤りがないかも見直しておきましょう。

ワークの進め方(4)

グループごとの分析結果を全体に共有しましょう。

- ①どんな分析をしたか
- ②どんな結果（数値）が出てきたか
- ③どんな解釈をしたか
- ④その理由はなにか

※他のグループの説明から新しい視点を見つけてください。



分析手法と解釈の例

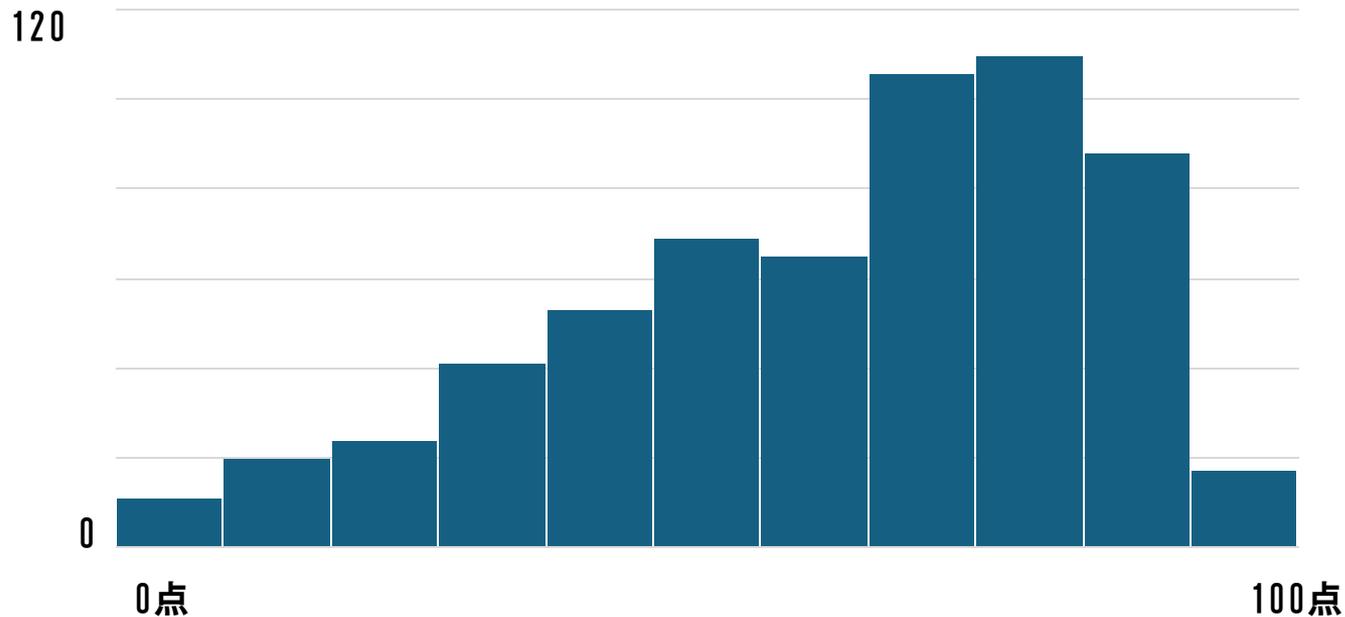
① 学科ごとの代表値と標準偏差を出す。

	全体	A学科	B学科	C学科	D学科
平均値	56	75	65	51	29
中央値	60	76	65	50	29
最頻値	71	67	71	50	42
標準偏差	20	7	8	11	13

解釈の視点（例）

- 平均値と標準偏差に関連する傾向はないか？
- 平均値と中央値の差と平均点の高さに関連する傾向はないか？

② ヒストグラムを作成し分布を可視化する

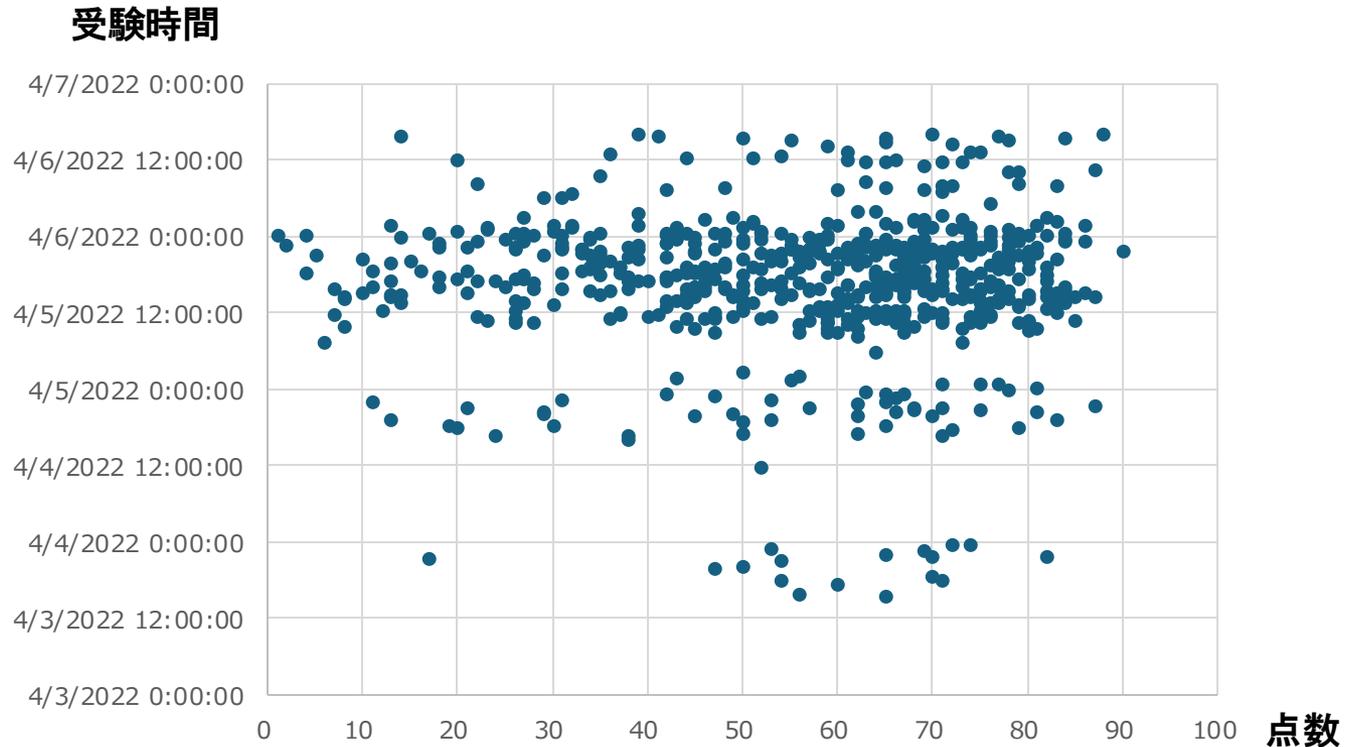


解釈の視点（例）

- どちらにどのように偏っているか？

分析手法と解釈の例

③ 受験時間とスコアに関連がないか散布図を作ってみる



解釈の視点（例）

● 受験時間とスコアに関連はありそうか？

データ分析の流れ演習③



データの種類と分類

統計学の目的

統計的な分析を行うことで、
経験者の暗黙知や目の前のできごとをベースとした施策を、
事実に基づく明確なコンセプトに基づいた戦略にすることが出来ます。

観光ビッグデータの精度は高くないので、
データの蓄積で精度を高める事と統計的な視点を持つことが重要です。

データの蓄積

条件や情報源により、
ミスリードを防ぐ
他のデータと組み合わせて
分析する
結果をフィードバックする
細分化による整理と
分析を行う

統計的な視点

分析の目的や動機を
明確化する
思い込みを無くし
客観的に分析する
データの整合性を確かめる
具体的な要素・視点を
取り入れる

データの分析の手法

「記述統計学」

目的：データを整理し、特徴を簡潔に表すこと。

方法：① 代表的な数値を用いて全体を簡潔化する。

例 … 平均値、中央値、偏差等

② 表やグラフを用いて傾向を表す。

例 … 度数分布表、折れ線グラフ等

「推測統計学」

目的：一部の標本データから全体を推測すること。

方法：① 推定：標本から母集団の値の範囲を推定する。

② 検定：標本から母集団について言えることを検証する。

「ベイズ統計学」

「多変量解析学」

「AI」

データの種類

データには「質的データ」と「量的データ」があります。
統計的な処理の仕方が変わるので違いを理解しましょう。

データ		性質	例	使用できる演算
データ	質的データ	名義尺度	区別のみ 順序はない	性別、居住地、職業区分、Yes/No 分類
		順序尺度	区別に加えて、 順序がある。 差は取れない。	5段階評価 分類、クラスの併合、累積
	量的データ	間隔尺度	差の大きさに意味がある。 比には意味がない。	摂氏温度、知能指数、偏差値、… 足し算や引き算 (合計、平均)
		比率尺度	差も比も意味がある。 (絶対ゼロがある)	身長、体重、得点、… 四則演算

代表値による分析

代表値とは①

データ全体を簡潔に説明する代表的な数値を代表値と呼びます。

代表値には、「データ数」「平均値」「中央値」「最小値・最大値」「四分位数（25%点、75%点）」「最頻値」等があります。

「データ数」

得られたサンプルの数。

利点：サンプルデータの信憑性を示す。（多いほど信頼しやすい。）

「平均値」

データを合計して、データの数で割ったもの。

利点：知名度が高く、分かりやすい。

欠点：データの歪みに大きく影響される。

代表値とは②

「中央値」

(データを大きさの順に並べて) データの中央に位置するデータ。

利点：データの歪みによる影響が少ない。

欠点：計算が面倒。

「最小値・最大値」

(データを大きさの順に並べて) データの中央に位置するデータ。

利点：データの歪みや外れ値があるかがわかる。

欠点：具体的な分析には使用できない。

「最頻値」

最も出現頻度が高いデータ。

利点：ボリュームゾーンがわかる。

欠点：全体の傾向はわからない。

代表値歪みの典型例

代表値が歪む典型例は「年収分布」。
平均値・中央値・最頻値が大きく異なる。
代表値は組み合わせて分析することが重要です。



散布度によるデータ分析

散布度とは①

データのばらつきの程度を示す値を散布度と言います。

散布度には、「偏差」「平均偏差」「分散」「標準偏差」等の数値的な表現方法や、「度数分布表」「パレート図」「ヒストグラム」等の視覚的な表現方法があります。

「偏差」

平均値との差を表します。（「データ」－「平均値」）

「平均偏差」

偏差の絶対値を合計して、データの数で割ったもの。

データ全体の平均値からのばらつきを表します。

※絶対値を取らずに偏差を合計すると必ず0となります。

散布度とは②

「分散」

偏差を二乗して合計し、データの数で割ったもの。

※絶対値の代わりに二乗をして合計しても0にならないようにしています。

「標準偏差」

分散の平方根を取ったもの。

平均偏差同様、データ全体の平均値からのばらつきを表します。

偏差値を求めるのにも使用します。

「度数分布表」

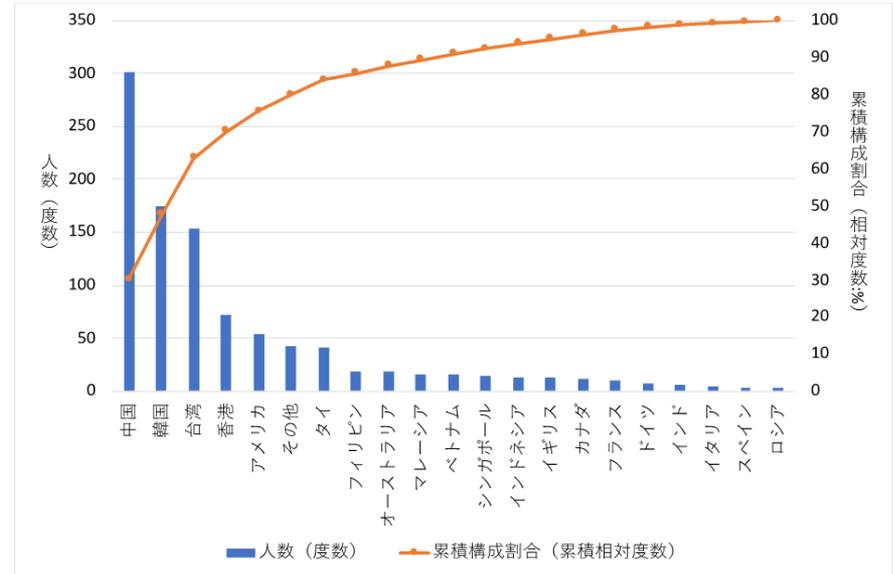
質的データならカテゴリ別に、
量的データなら「階級」を設定して、
データの頻度を集計したもの。

	人数 (度数)	累積度数	構成割合 (相対度数)	累積構成割合 (累積相対度数)
中国	301	301	30.1%	30.1%
韓国	175	476	17.5%	47.6%
台湾	153	629	15.3%	62.9%
香港	72	701	7.2%	70.1%
タイ	41	742	4.1%	74.2%
シンガポール	15	757	1.5%	75.7%
マレーシア	16	773	1.6%	77.3%
インドネシア	12	785	1.2%	78.5%

散布度とは③

「パレート図」 (質的データ)

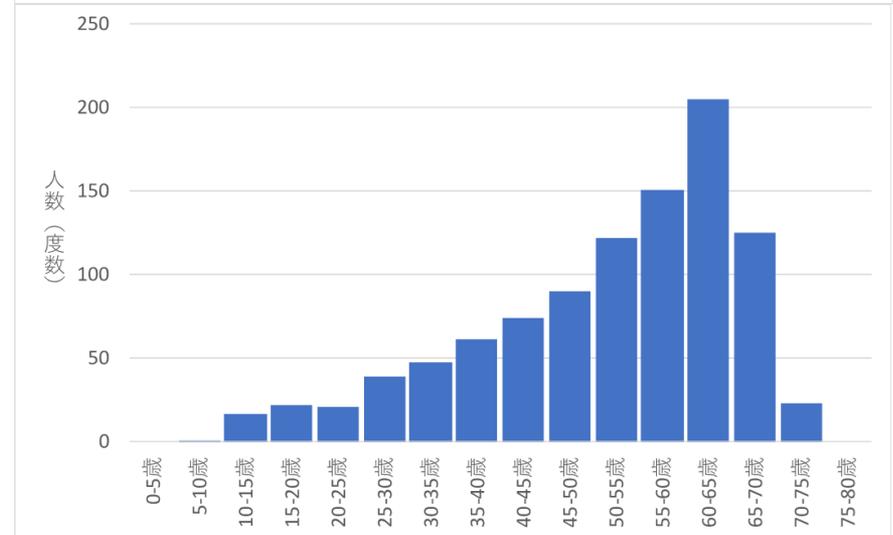
度数分布表をグラフ表示し、
頻度順に並び替え、
累積相対度数をプロットしたもの。
構成比の集中度合いがわかります。



「ヒストグラム」 (量的データ)

量的データの度数分布表を
グラフにしたもの。

山がいくつあるか、
左右への歪みがあるか、
上下の尖りや外れ値があるか等を
確かめられます。



データの可視化

データの可視化①

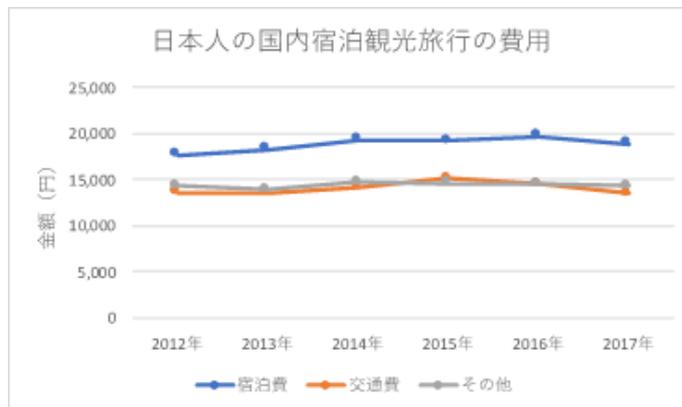
データを用いたコミュニケーションを円滑にするために適切な可視化をすることが重要です。

可視化の方法には、「折れ線グラフ」「散布図」「ヒストグラム」「帯グラフ」等のデータ分析用の方法や、「円グラフ」「棒グラフ」「バブルチャート」等のデータ説明用の方法があります。

「折れ線グラフ」

時系列や量の多寡等の順序がある量の変化を表します。

変化を表現するのに汎用的で、複数のデータ系列をまとめて表現できます。



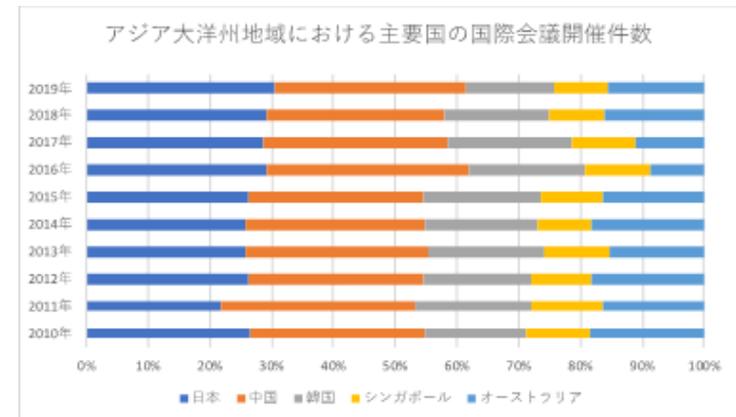
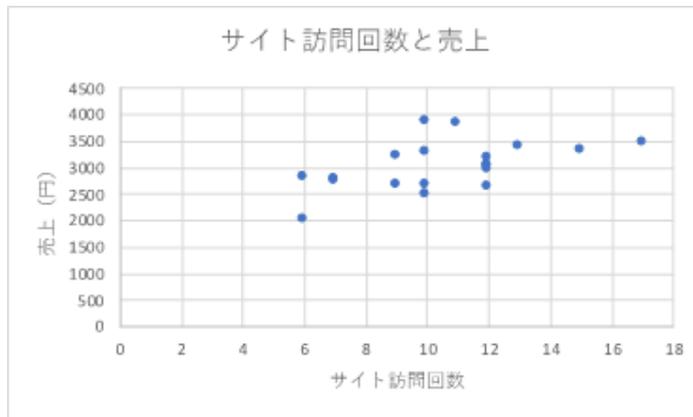
データの可視化②

「散布図」

2つのデータ系列を比較したり、関係を推定するのに用います。
外れ値など相関係数だけではわからない情報が得られます。

「帯グラフ」

構成比を比較する場合に用いられます。
比率で比べるため、質的な変化が把握できます。



データの可視化③

「円グラフ」

構成比に偏りがないかを表現する。

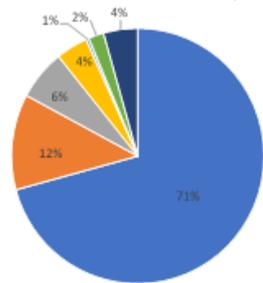
「棒グラフ」

折れ線グラフに、積み上げが必要な時に使用する。

「バブルチャート」

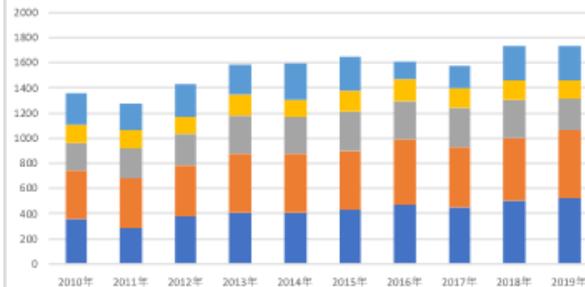
3つ以上のデータを一度に比べたい時に使用する。

訪日外国人旅行者の内訳 (2019年, 単位: 万人)



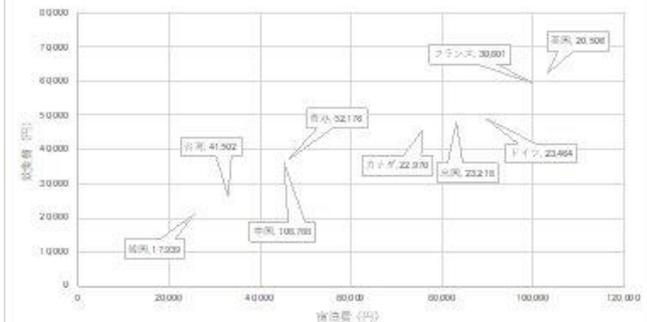
■ 東アジア ■ 東南アジア ■ 北米 ■ 欧州主要5カ国 ■ ロシア ■ オーストラリア ■ その他

アジア大洋州地域における主要国の国際会議開催件数



■ 日本 ■ 中国 ■ 韓国 ■ シンガポール ■ オーストラリア

国別・地域別にみる訪日外国人旅行者1人あたり費目別旅行支出



データ分析の流れ

データ分析の流れ①

統計的な問題解決においては、次のPPDACによるサイクルを進めるのが一般的です。

① Problem / 課題設定

身近な問題を明確にし、「何がしたいか（目的）」や「目的に関係することは何か」をはっきりさせるための段階。

いきなり分析を始めないことが重要です。

② Plan / 分析計画

目的のために有効な調査・実験や分析をデザインする段階。

使用するデータ、分析手法を適切に選択します。

データ分析の流れ②

③ Data / データ収集・クレンジング

行ないたい分析のために必要なデータを収集し（分析で扱えるように）データを加工する（データ整理）する段階。

データの信頼性・有効性を検証し、外れ値や欠損値を処理します。

④ Analysis / 分析設計

データからパターンを発見し、どのような要因が「目的」に影響を与えているかを分析していく段階。

全体を把握し、予測モデル等を用いて数学的に表現することを目標とします。

⑤ Conduction / 結果解釈

最初に設定した課題に対して結論を出し、新たな課題や今後の方針（追加分析）を検討・提示していく段階。

統計的に有意かを考え、結果を評価しまとめます。

ワーク



データを眺めて全体を把握

データを眺めて、歪みを予想してください。

ファイル「演習データE1.xlsx」を開きましょう。

データを”眺めて”ください。

どのような印象をもったか、グループで意見交換しましょう。

その後、実際に代表値を算出し、何が読み取れるか考えてみましょう。

ワークの進め方（1）

Work Sheet

テーマ



後で見分分かるように
テーマを書いておきましょう

データを眺めた印象を
グループで話し合いメモしましょう。

以下のようなことを予想してみましよう。

代表値は、それぞれのあたりでしょうか？
代表値に歪みは、ありそうでしょうか？

ワークの進め方（2）

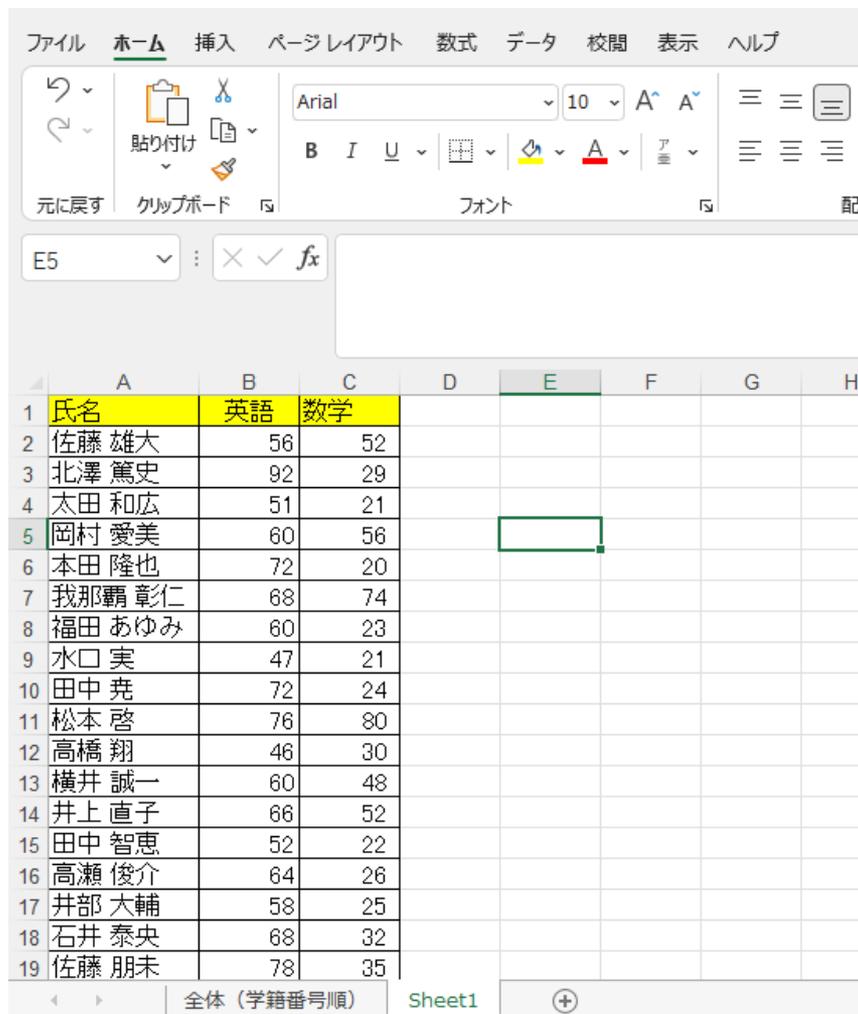
では、実際に代表値を算出してみましょう。

データ分析のアイデア

ハンズオン（答えあわせ）

ファイルを開く

- ファイル「演習データE1.xlsx」を開きましょう。



The screenshot shows the Microsoft Excel interface. The ribbon is set to 'Home' and the font settings are visible. The spreadsheet has columns for Name, English, and Math scores. The data is as follows:

	A	B	C	D	E	F	G	H
1	氏名	英語	数学					
2	佐藤 雄大	56	52					
3	北澤 篤史	92	29					
4	太田 和広	51	21					
5	岡村 愛美	60	56					
6	本田 隆也	72	20					
7	我那覇 彰仁	68	74					
8	福田 あゆみ	60	23					
9	水口 実	47	21					
10	田中 堯	72	24					
11	松本 啓	76	80					
12	高橋 翔	46	30					
13	横井 誠一	60	48					
14	井上 直子	66	52					
15	田中 智恵	52	22					
16	高瀬 俊介	64	26					
17	井部 大輔	58	25					
18	石井 泰央	68	32					
19	佐藤 朋未	78	35					

平均値を求める

The screenshot shows the Excel interface with the formula bar containing `=average(B2:B501)`. The spreadsheet data is as follows:

	A	B	C	D	E	F	G	H
488	伊藤 照久	61	27					
489	石川久美子	54	24					
490	新美 涼	42	48					
491	前田 智恵	88	34					
492	松田 恵子	56	30					
493	恵 泰輔	48	6					
494	木村 洋子	72	80					
495	高橋 芳明	58	52					
496	木山 直之	74	58					
497	大島 晶子	48	68					
498	粟屋 淳	58	52					
499	伊藤 陽子	82	22					
500	上原 俊樹	55	49					
501	山本 たかこ	66	52					
502	合計	29732	22482					
503	平均値	<code>B2:B501</code>						

The formula bar shows: `=average(B2:B501)`

Callout boxes indicate: 「平均値」を記入 (Enter the average) and `=average(B2:B501)`

The screenshot shows the same Excel spreadsheet after the formula has been calculated. The result is displayed in cell C496:

	A	B	C	D	E	F	G	H
488	伊藤 照久	61	27					
489	石川久美子	54	24					
490	新美 涼	42	48					
491	前田 智恵	88	34					
492	松田 恵子	56	30					
493	恵 泰輔	48	6					
494	木村 洋子	72	80					
495	高橋 芳明	58	52					
496	木山 直之	74	58					
497	大島 晶子	48	68					
498	粟屋 淳	58	52					
499	伊藤 陽子	82	22					
500	上原 俊樹	55	49					
501	山本 たかこ	66	52					
502	合計	29732	22482					
503	平均値	59.464						

The formula bar shows: `58`

ワークの続き

ワークの進め方（2）

Work Sheet

テーマ



後で見分分かるように
テーマを書いておきましょう

個人の考えを書いておきましょう



算出した代表値から
どのようなことが読み取れますか？
個人の考えを整理しましょう。

ワークの進め方（3）

Work Sheet

テーマ



グループで意見交換をしましょう。



個人で読み取ったことを
グループ内で意見交換しましょう。
結論は、出さなくても良いですが
「全員が納得する読み取り」はないか
議論して探ってみましょう。

ワークの進め方（4）

グループごとの意見を全体に共有しましょう。

グループで出た意見の中で

良かった・興味深かった

変わっている（たぶん他では出てないだろう）

自分が気に入っている（元々自分の意見でもOK）

特に理由はないけど、皆に聞いて欲しい

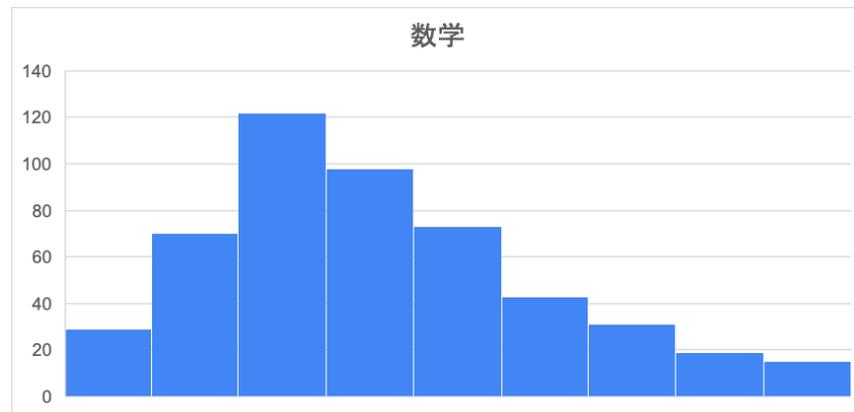
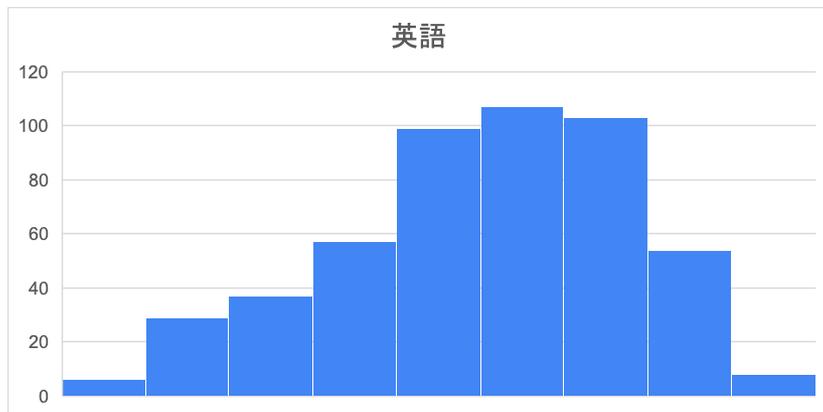
というものを、共有してください。

※他のグループの説明から新しい視点を見つけてください。



代表値から見えること

二つのデータの分布の外観は、下記のようにになります。



3つの代表値からだけでも、例えば

「数学より英語の方が、高得点に偏っている」
ということくらいは分かります。

しかし、代表値からだけでは分からないことも多いので
他にもいろいろな統計値が必要になってくるのです。

データ分析に求められる着眼点①



事実と意見の違い

事実：真偽が証明できる
⇒誰もが同意する

人間には、水が必要だ

意見：真偽が証明できない
⇒その人の主観による

水を飲むと元気が湧く

事実と意見の区別

円安が大幅に加速し、1ドル140円台にまで進行しました。

1ドル140円台に → 真偽が確認できるので「事実」

大幅に加速し → 「大幅」か否かは個人の主観なので「意見」

事実と意見の区別

(例)

4月30日の国内新型コロナウイルスの感染者は、
2万5182人で依然としてパンデミックは継続傾向にある。

感染者は2万5182人 → 真偽が確認できるので「事実」

継続傾向 → データを元にはいるが、解釈なので「意見」

ワーク

話し合いのなかの事実と意見

これから、グループでディスカッションをしてもらい、その中での意見と事実の使い方を振り返ります。

クラスを大きく2つ（Aサイド・Bサイド）に分けます。

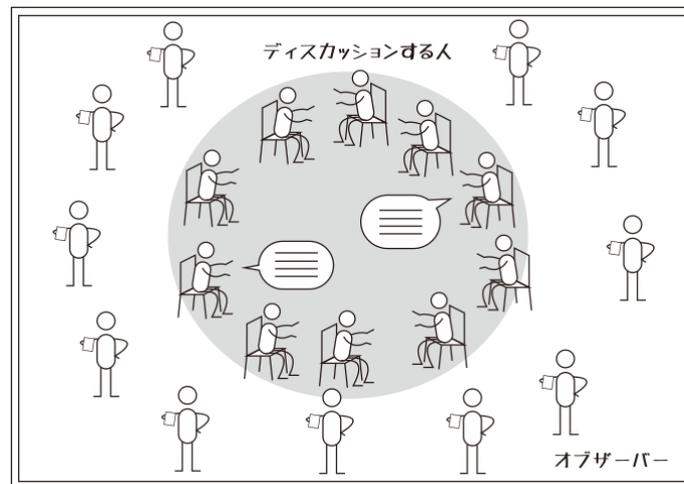
→そのなかでさらに4-5人のグループになってもらいます。

Aサイドのグループがディスカッションをしている間、Bサイドのグループは観察します。

Bサイドのグループがディスカッションをしている間、Aサイドのグループは観察します。

観察の中では、下記を行ってください。

- ① 発言内容を記録する（特に主張の根拠）
- ② その発言が事実なのか意見なのかを分類する。



話し合いのなかの事実と意見

Work Sheet

テーマ



個人ワーク



グループワーク

後で見て分かるように
テーマを書いておきましょう

観察対象グループの発言を記録
※書き取れる範囲で

記録しながら、発言を分類します。

事実＝客観的で真偽判定可能

意見＝主観的で審議判定が不可能

分類例

事実を四角、意見を丸で囲む

事実には☆マーク、意見には○をつける

Aサイド用テーマ 1

ライフラインの整備されていない無人島で1年間サバイバル生活をしなくてはならなくなりました。島には、命を脅かす猛獣や毒を持つ危険な生物はいません。きれいな水が流れている川もあります。

家から何かアイテムを3つ持って行っていいことになっているときに、何を持っていきますか？

Bサイド用テーマ 1

もはや日本を代表するメニューのカレーライスとラーメン。アフリカのテレビ局で、カレーライスかラーメンを紹介するテレビ番組の企画を担当することになったチームのメンバーとして、カレーライスかラーメンのどちらを選ぶか討論し、決定してください。

ただし、アフリカのどこの国で放送されるかは、まだ知らされていません。

話し合いのなかの事実と意見

Work Sheet

テーマ



グループで意見交換をしましょう。



個人での記録と分類をグループ内で共有しグループとしての観察結果を考えてください。

なるべく結論がでることを目指してください。

話し合いのなかの事実と意見

グループごとの観察結果を全体に共有しましょう。

共有の観点

- ①どんな発言があったか
- ②事実と意見の比率は、どれくらいか
- ③どちらの方が根拠として効果的に作用していたと思うか



※他のグループの説明から新しい視点を見つけてください。

なぜ、事実と意見を混同するのか？

○原因①：話し手自身が事実と意見を混同している

「富士山は、世界で一番美しい山だ」

「今の政治はダメだ」

⇒ 発言者は事実だと思っており、悪意はない

○原因②：あえて事実と意見を混同するように書いている

「カップメンを食べた人の死亡率は100%です」

→ 100年後・200年後であれば、死亡して当然である

「弊社の8割の製品に最新の環境基準が適応されています。」

→ 実は残りの2割が主力消費で全体の生産高の8割を担っている

⇒ 悪意有り（印象操作により受け手が誤解するように仕向けている）

事実と意見の区別

自分の発言や報告で、無意識に事実と意見を混同しないようにしましょう。

事実と意見を意識的に混同させようとしている発言や文章に気を付けましょう。

事実と意見を分けて考える

以前に分析した下記結果を元に

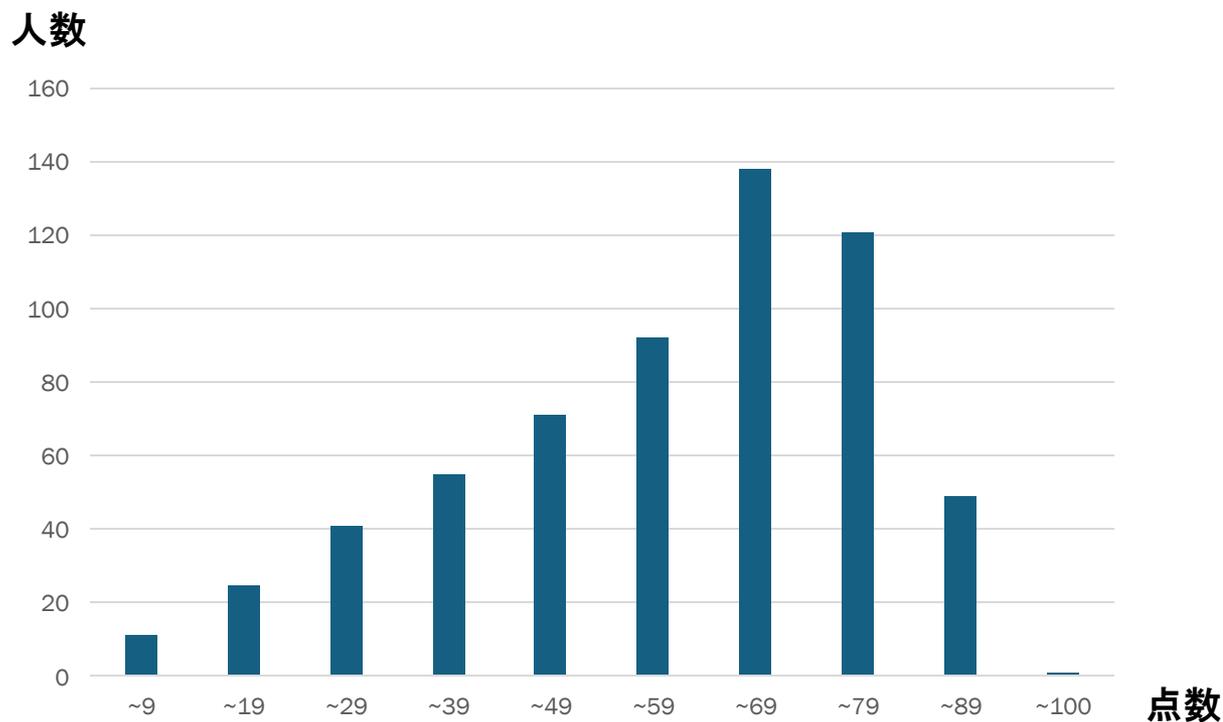
事実と意見を分けて考える練習をしましょう。

	全体	A学科	B学科	C学科	D学科
平均値	56	75	65	51	29
中央値	60	76	65	50	29
最頻値	71	67	71	50	42
標準偏差	20	7	8	11	13
平均値と中央値の差 (絶対値)	4.4	1.1	0.2	0.8	0.0

- ①結果から読み取れる事実をあげていってください。
- ②そこから導き出せる意見（解釈）をあげてください。

事実と意見を分けて考える

下記結果を元に事実と意見を分けて考える練習をしましょう。



- ① グラフから読み取れる事実をあげていってください。
- ② そこから導き出せる意見（解釈）をあげてください。

ワークの進め方（1）

Work Sheet

テーマ



後で見えて分かるように
テーマを書いておきましょう

分析結果から読み取れることを
整理していきましょう。

記入しながら事実と意見（解釈）を分類します。

分類例

事実を四角、意見を丸で囲む

事実には☆マーク、意見には○をつける

ワークの進め方（2）

Work Sheet

テーマ



グループで意見交換をしましょう。



個人での読み取りと分類をグループ内で共有し、グループとしての解釈を考えてください。

結論は、出さなくても良いですが「全員が納得する読み取り・解釈」はないか議論して探ってみましょう。

ワークの進め方（3）

グループごとの意見を全体に共有しましょう。

共有の観点

- ①どんな事実が見つかったか
- ②どのような解釈が考えられるか
- ③その根拠はなにか

※他のグループの説明から新しい視点を見つけてください。



解答例

	全体	A学科	B学科	C学科	D学科
平均値	56	75	65	51	29
中央値	60	76	65	50	29
最頻値	71	67	71	50	42
標準偏差	20	7	8	11	13
平均値と中央値の差 (絶対値)	4.4	1.1	0.2	0.8	0.0

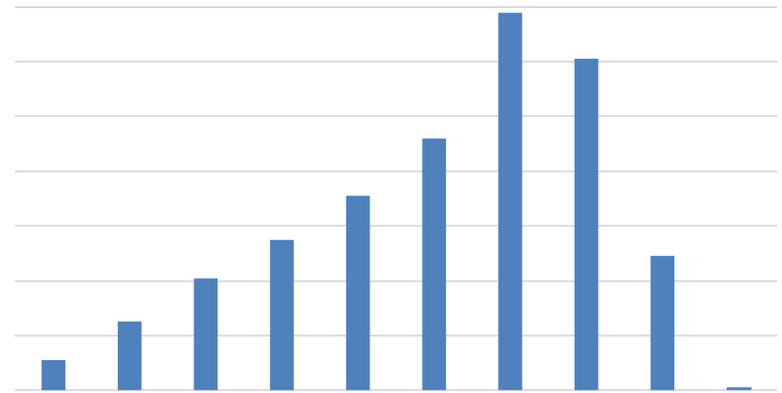
①結果から読み取れる事実をあげていってください。

- 1) 平均点は、A学科>B学科>C学科>D学科となっているが、標準偏差は逆の順序になる
- 2) 平均値と中央値の差はA学科が一番大きく、D学科が一番小さい

②そこから導き出せる意見（解釈）をあげてください。

- 1) より、平均点の高い学科ほど、学生の差が大きいと考えられる（標準偏差が大きい）
- 2) より、平均点の高い学科は一部の学生が全体の平均を押し上げている可能性がある

解答例



①結果から読み取れる事実をあげていってください。

1) 60～69点が最も多い

2) 90～100点より、0～9点の方が多くいる

②そこから導き出せる意見（解釈）をあげてください。

1)2)より、全体的に高得点（50点より上）に偏っている

データ分析に求められる着眼点②



統計とはなにか？

はじめに（１）

以下の数値において、統計としての数値に当てはまるかどうか考えてみましょう。また、そのように判断した理由を考えましょう。

- 中学生10人の平均身長
- 2022年、生産人口世代の年収の中央値
- 世界大学ランキングの順位
- 各政党の支持率

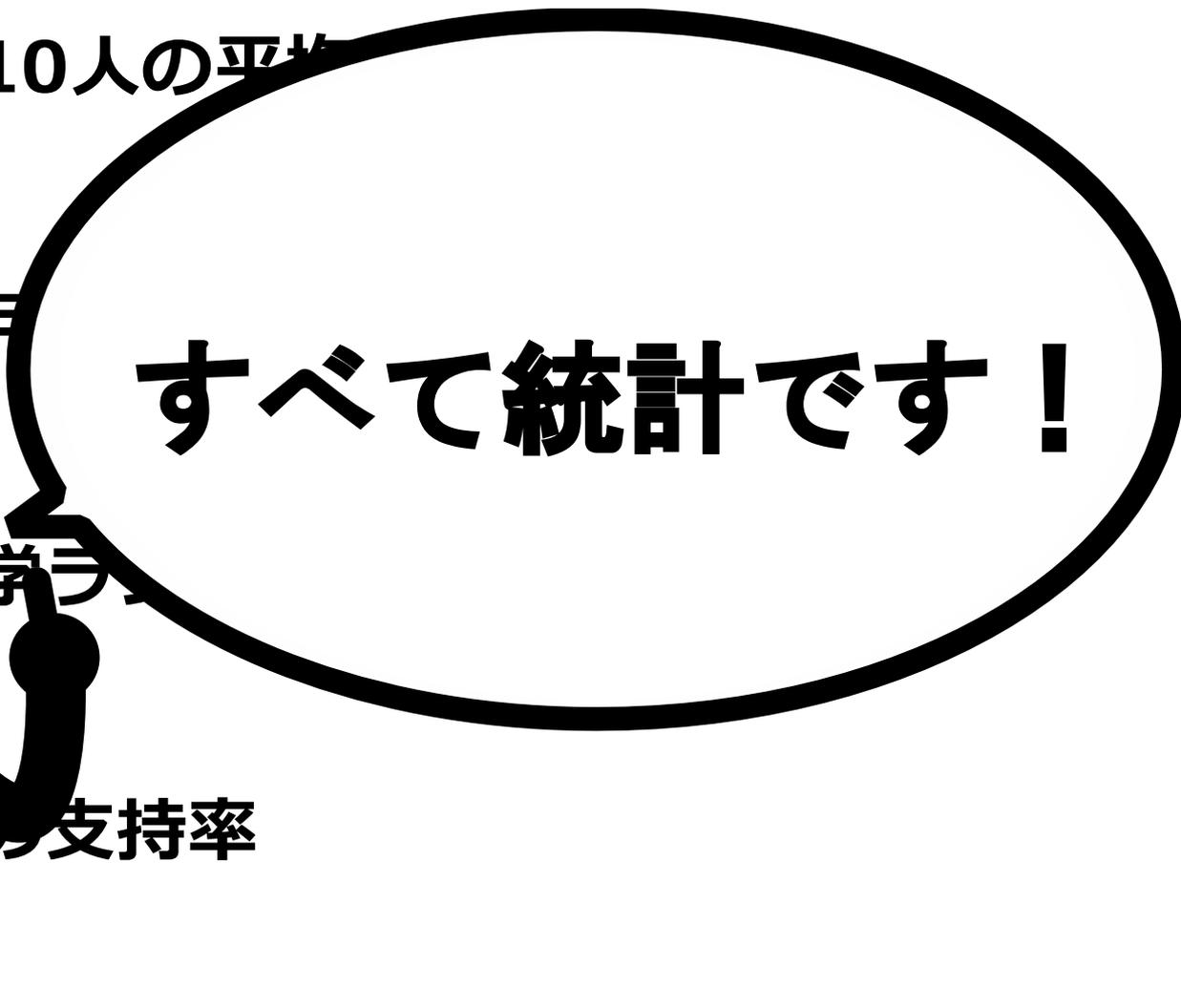
はじめに (2)

- ・ 中学生10人の平均

- ・ 2022年

- ・ 大学ラ

- ・ 大学の支持率



すべて統計です！

統計とは何か？（1）



多くのものを
1つにまとめる



ものの数値を調べる
(重さや長さなど)



全てをあつめて計算する

統計とは何か？（2）

統計には大きく分けて3種類あります。

古典
統計学



新しい
統計学



記述統計学

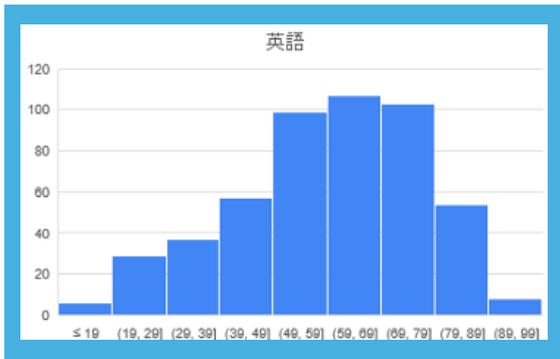
推測統計学

ベイズ統計学

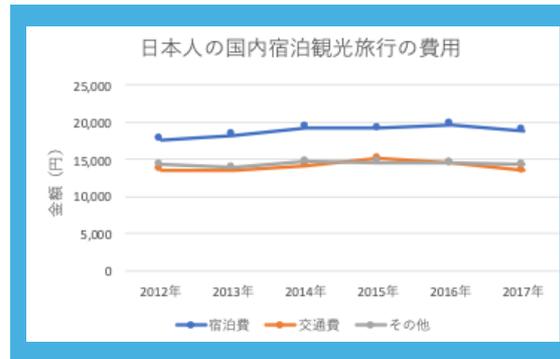
記述統計学とは（1）

データを集約して、表やグラフなど見やすい形に
整え、データの特徴を捉えること

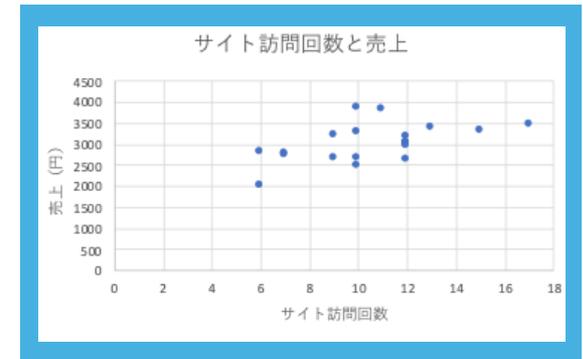
→人に何かデータを伝えるときに使用する



ヒストグラム



折れ線グラフ



散布図

記述統計学とは（2）

例：どちらの数値の方が大きいと感じますか？

A {2, 5, 10, 25, 50, 100}

B {5, 30, 40, 50, 60, 70}

Aには最大値の100があるため、大きい？

Aには最小値2があるため、小さい？

Bは各要素の合計値が大きいから大きい？

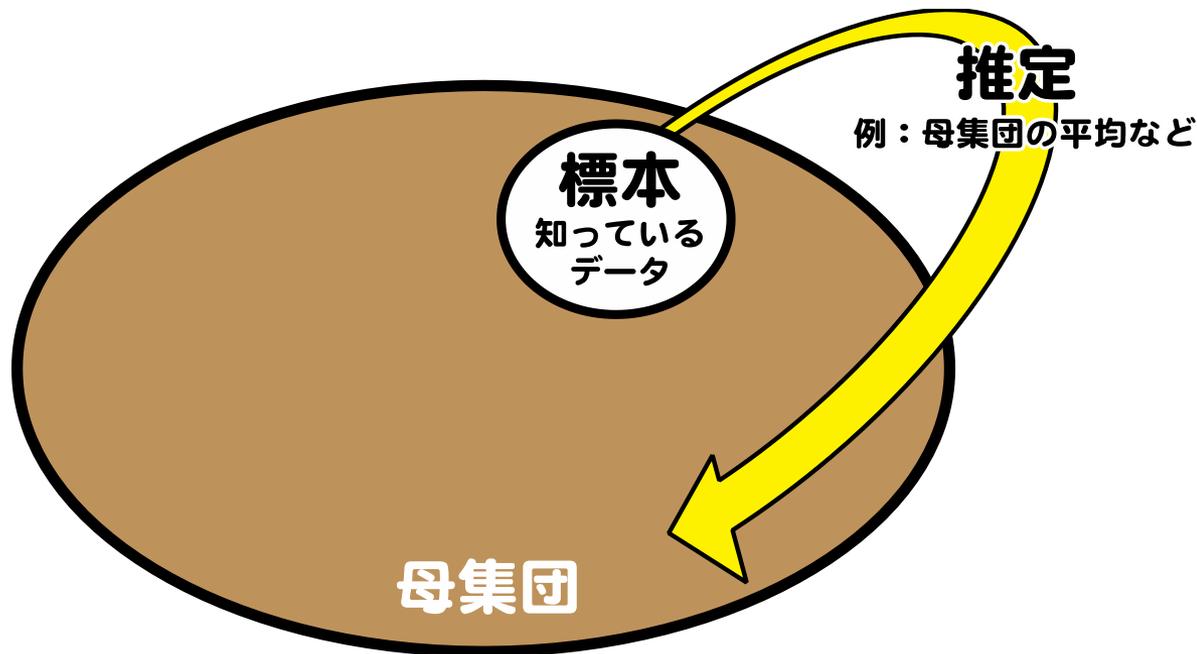
Bには3桁の数値がないため、小さい？

今回は数値が少ないのでイメージしやすいけども、これが1万データあったときにどっちが大きいかわかるか判断できますか？
それを表やグラフで可視化するのが記述統計学です。

推測統計学とは（1）

必要なデータを全部持ってなかったとしても、
持っているデータから全体の傾向を推測すること

→ビジネスシーンでよく使用される



推測統計学とは（2）

例：あるPC企業は

- ・ 80cmの落下耐久があります
- ・ キーボードが10万回押しても壊れません
- ・ 70度の気温下で使用しても壊れません

としてPCを販売していますが、
販売するPCの全てをチェックすることは可能でしょうか？

→ 無作為に5千台程度に1台抽出してチェックを行い、
3つの試験に合格しているため、このように記載が可能になります。

今あるデータのみから母集団を考える統計学

→ 推測統計よりも主観的にデータを扱いながら
統計を進められる

通常の統計学

母集団／真の確率は
決まっている
例：コインを投げる

実際に投げて観察する

表が何回出たという事象

ベイズ統計

特定の事象がある
例：コインを投げた時に
表が何回出たという事象

主観的に推測する

コインを投げると表が
この確率で出るだろう



ワーク



統計ではないもの

身の回りの数値や、データで統計ではないものにはどんなものがあるか考えてみましょう。また、それはどうして統計ではないと考えた理由をグループで共有してみましょう。

Work Sheet

テーマ


個人ワーク


グループワーク

これは違うかも？
と思ったものでも、どんどん書き出していきましょう。
書き出した後に、統計かそうでないかを考えていきましょう。

考えたものをグループで共有していきましょう。
なぜ、統計ではないと思ったのかをグループメンバーに説明していきましょう。

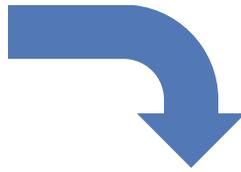
統計でわかること・わからないこと

統計が得意とすること

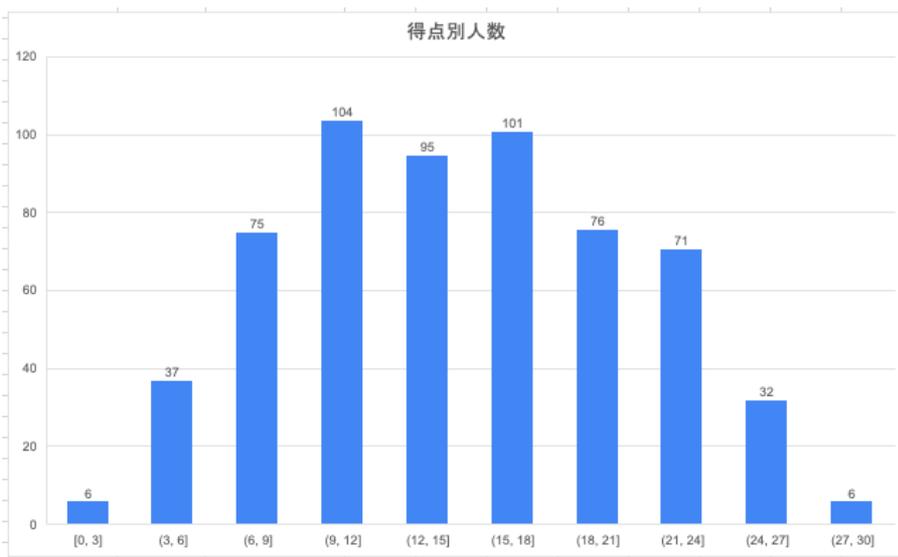
大量かつ複雑なデータを人間が見て理解しやすくする

Excelの一覧表

1	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
タイムスタンプ	スコア	学部	学級番号	姓(漢字)	姓(全角カタカナ)	名(漢字)	名(全角カタカナ)	問1	問2	問3	問4	問5	問6	問7	問8	問9	
4/4/2021 13:49:27	16/30	人間工学部	BA211082	川青	カワサキ	みゆ	ミユ	選択2	選択2	選択3	選択3	選択4	選択4	選択1	選択1	選択2	選択2
4/4/2021 13:56:07	15/30	情報デザイン学部	BA211105	井上	イノウエ	雅和	マサカズ	選択1	選択1	選択5	選択5	選択4	選択1	選択1	選択2	選択2	選択1
4/4/2021 13:58:10	6/30	情報デザイン学部	BA211089	伊藤	イトウ	真代	マヨ	選択2	選択2	選択2	選択1	選択1	選択4	選択2	選択1	選択2	選択1
4/4/2021 13:59:57	11/30	情報デザイン学部	BA211029	澤本	サエモト	芳彦	ヨシヒコ	選択1	選択3	選択5	選択2	選択1	選択4	選択1	選択5	選択5	選択5
4/4/2021 14:05:15	16/30	法学部	AA211097	入江	イリエ	紗織	サオリ	選択5	選択3	選択5	選択2	選択4	選択4	選択1	選択3	選択3	選択5
4/4/2021 14:08:49	14/30	工学部	BA411093	阿部	アベ	康二	コウジ	選択5	選択5	選択3	選択3	選択4	選択1	選択5	選択5	選択5	選択5
4/4/2021 14:16:43	7/30	法学部	AA211154	西川	ニシカワ	恭	ケイ	選択2	選択4	選択3	選択1	選択1	選択5	選択2	選択4	選択5	選択5
4/4/2021 14:22:27	11/30	工学部	AA211105	尾永	オノナガ	智絵	チエ	選択5	選択3	選択5	選択4	選択1	選択4	選択1	選択2	選択3	選択3
4/4/2021 14:24:07	14/30	工学部	BA411110	吉原	ヨシハラ	マイ	マイ	選択5	選択3	選択5	選択3	選択4	選択4	選択3	選択4	選択3	選択3
4/4/2021 14:41:30	11/30	人間工学部	BA211114	森尾	モリオ	和弘	カズヒロ	選択2	選択3	選択5	選択2	選択4	選択4	選択2	選択3	選択3	選択3
4/4/2021 14:43:46	5/30	法学部	AA211094	安部	アベ	雅	トヨル	選択2	選択5	選択3	選択3	選択1	選択3	選択2	選択2	選択3	選択3
4/4/2021 14:44:23	13/30	情報デザイン学部	BA211063	山下	ヤマシタ	孝明	コウメイ	選択5	選択3	選択3	選択2	選択1	選択4	選択4	選択3	選択5	選択4
4/4/2021 14:45:15	12/30	法学部	AA211158	石原	イシハラ	希美	ノゾミ	選択1	選択3	選択5	選択4	選択4	選択1	選択2	選択2	選択5	選択5
4/4/2021 14:48:37	10/30	情報デザイン学部	BA211112	宮家	ミヤカ	隆善	タカユキ	選択2	選択4	選択5	選択2	選択1	選択1	選択4	選択5	選択2	選択2
4/4/2021 15:01:12	21/30	人間工学部	BA311054	塩谷	シオヤ	千景子	チエコ	選択5	選択3	選択5	選択3	選択4	選択1	選択2	選択5	選択3	選択3
4/4/2021 15:01:12	25/30	法学部	AA211041	奥谷	オクニ	英雄	ヒデオ	選択2	選択3	選択5	選択3	選択4	選択1	選択2	選択5	選択5	選択5
4/4/2021 15:01:51	12/30	情報デザイン学部	BA211113	丸山	マルヤマ	健介	ケンスケ	選択5	選択1	選択4	選択2	選択4	選択3	選択1	選択5	選択3	選択3
4/4/2021 15:03:47	7/30	工学部	BA411030	栗川	クリガワ	由香里	ユカ	選択1	選択5	選択5	選択2	選択4	選択4	選択3	選択2	選択2	選択2
4/4/2021 15:04:45	5/30	法学部	AA211094	津谷	ツルヤ	隆	タカシ	選択2	選択3	選択5	選択3	選択1	選択3	選択3	選択2	選択3	選択3
4/4/2021 15:06:49	11/30	情報デザイン学部	BA211080	奥田	オクダ	大貴	タイキ	選択1	選択1	選択5	選択2	選択2	選択5	選択2	選択3	選択1	選択1
4/4/2021 15:08:58	5/30	工学部	BA411022	田島	タナベ	尚	ナオ	選択2	選択3	選択5	選択4	選択5	選択2	選択2	選択3	選択1	選択1
4/4/2021 15:12:19	20/30	法学部	AA211093	石塚	イシヅカ	龍之	リュウキ	選択5	選択3	選択5	選択3	選択1	選択1	選択5	選択5	選択5	選択5
4/4/2021 15:12:48	19/30	法学部	AA211153	太田	オオタ	和美	カズミ	選択5	選択3	選択5	選択4	選択4	選択1	選択2	選択5	選択5	選択5
4/4/2021 15:13:09	18/30	人間工学部	BA311073	野口	ノグチ	孝二	コウジ	選択5	選択3	選択5	選択3	選択4	選択3	選択1	選択3	選択3	選択2
4/4/2021 15:15:29	11/30	工学部	BA411018	加藤	カノウ	洋介	ヨウスケ	選択5	選択3	選択5	選択2	選択4	選択3	選択4	選択2	選択3	選択3
4/4/2021 15:20:38	8/30	工学部	BA411072	吉田	ヨシダ	希	ノゾミ	選択5	選択3	選択5	選択5	選択2	選択2	選択4	選択3	選択3	選択3
4/4/2021 15:25:26	19/30	人間工学部	BA311110	西	ニシ	真緒	マオ	選択3	選択3	選択5	選択3	選択1	選択1	選択2	選択5	選択5	選択5
4/4/2021 15:29:03	25/30	工学部	BA411052	川原	カワハラ	光希	ミツノブ	選択5	選択3	選択3	選択1	選択4	選択1	選択1	選択5	選択5	選択5
4/4/2021 15:34:14	12/30	人間工学部	BA311038	竹中	タケナカ	優子	ユウコ	選択2	選択3	選択5	選択5	選択4	選択4	選択3	選択3	選択4	選択4
4/4/2021 15:35:13	15/30	情報デザイン学部	BA211119	上村	ウエムラ	舞	マイ	選択5	選択3	選択5	選択3	選択4	選択4	選択1	選択1	選択5	選択1
4/4/2021 15:35:53	17/30	法学部	AA211110	渡辺	ワタナベ	清貴	サヤカ	選択2	選択2	選択5	選択4	選択1	選択1	選択1	選択1	選択5	選択1
4/4/2021 15:36:40	14/30	法学部	AA211057	梶子	カシコ	洋介	ヨウスケ	選択5	選択3	選択5	選択4	選択1	選択3	選択3	選択5	選択5	選択5
4/4/2021 15:39:32	13/30	工学部	BA411027	大田	オオタ	由里	ユリ	選択5	選択5	選択1	選択3	選択5	選択1	選択5	選択5	選択5	選択5
4/4/2021 15:41:18	9/30	法学部	AA211114	橋	タチバナ	絵	エリ	選択2	選択1	選択5	選択2	選択4	選択5	選択2	選択5	選択2	選択2
4/4/2021 15:41:52	13/30	人間工学部	BA311044	西澤	ニシザワ	拓	タクミ	選択3	選択4	選択5	選択3	選択3	選択4	選択5	選択3	選択3	選択2
4/4/2021 15:42:12	15/30	情報デザイン学部	BA211101	佐藤	サトウ	タクマ	タクマ	選択5	選択3	選択5	選択3	選択4	選択3	選択3	選択3	選択3	選択3
4/4/2021 15:43:45	14/30	人間工学部	BA311052	長	ナガシ	健	ケン	選択5	選択1	選択5	選択3	選択4	選択1	選択1	選択2	選択5	選択5
4/4/2021 15:48:00	24/30	人間工学部	BA311088	塩原	シオハラ	慶子	ケイコ	選択1	選択3	選択5	選択3	選択4	選択1	選択1	選択1	選択5	選択5
4/4/2021 15:48:17	17/30	法学部	AA211078	山崎	ヤマザキ	惠子	ケイコ	選択5	選択3	選択5	選択2	選択4	選択4	選択5	選択3	選択3	選択5
4/4/2021 15:50:11	16/30	工学部	BA411054	白鳥	シラトリ	由美	ユミ	選択1	選択3	選択5	選択3	選択4	選択2	選択4	選択2	選択1	選択1
4/4/2021 15:50:32	7/30	情報デザイン学部	BA211005	大塚	オオツカ	朋子	カオル	選択2	選択4	選択5	選択2	選択3	選択4	選択2	選択5	選択5	選択5
4/4/2021 15:51:59	21/30	情報デザイン学部	BA211009	佐藤	サトウ	明	アカキ	選択5	選択3	選択5	選択3	選択4	選択1	選択1	選択1	選択5	選択3
4/4/2021 15:52:27	12/30	工学部	BA411112	永高	ナガタ	ナガタ	ナガタ	選択5	選択3	選択5	選択3	選択4	選択3	選択3	選択2	選択5	選択5

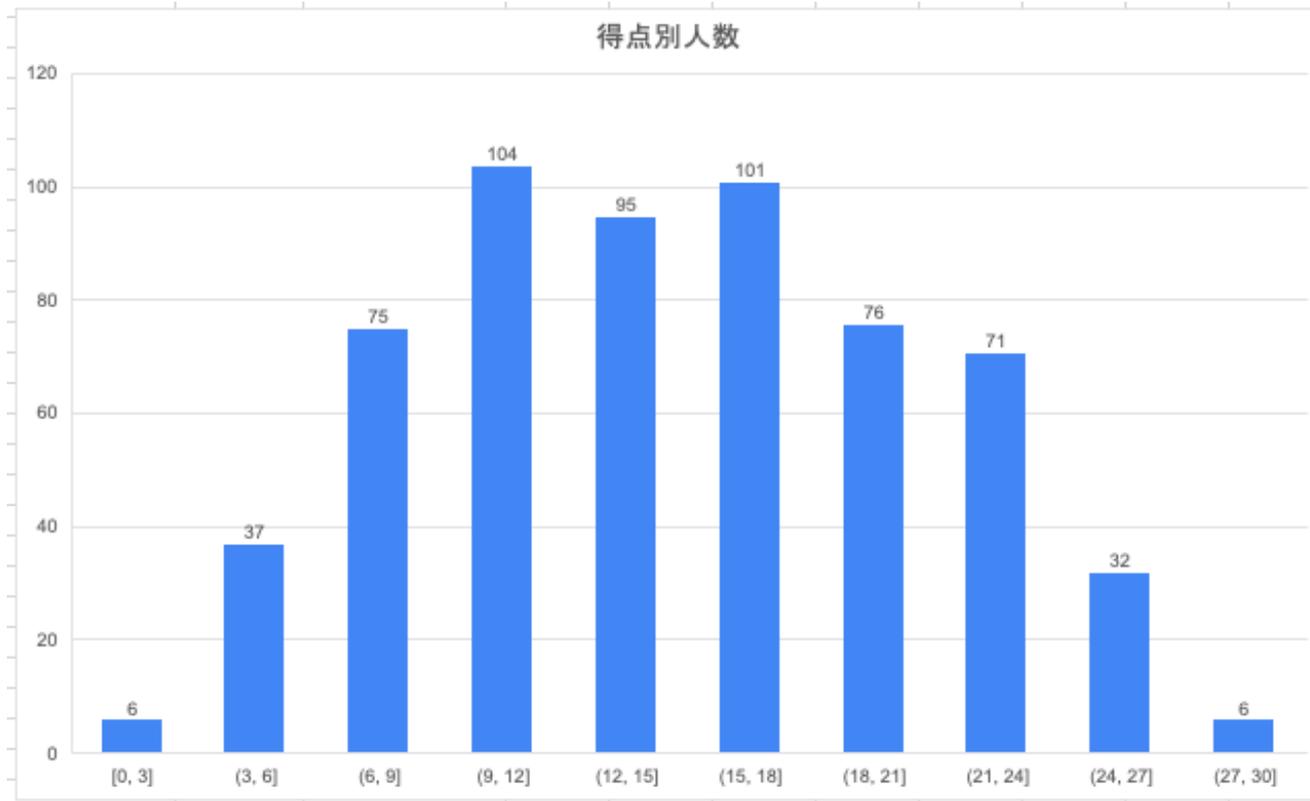


分析したいデータをヒストグラム化



統計ではできないこと

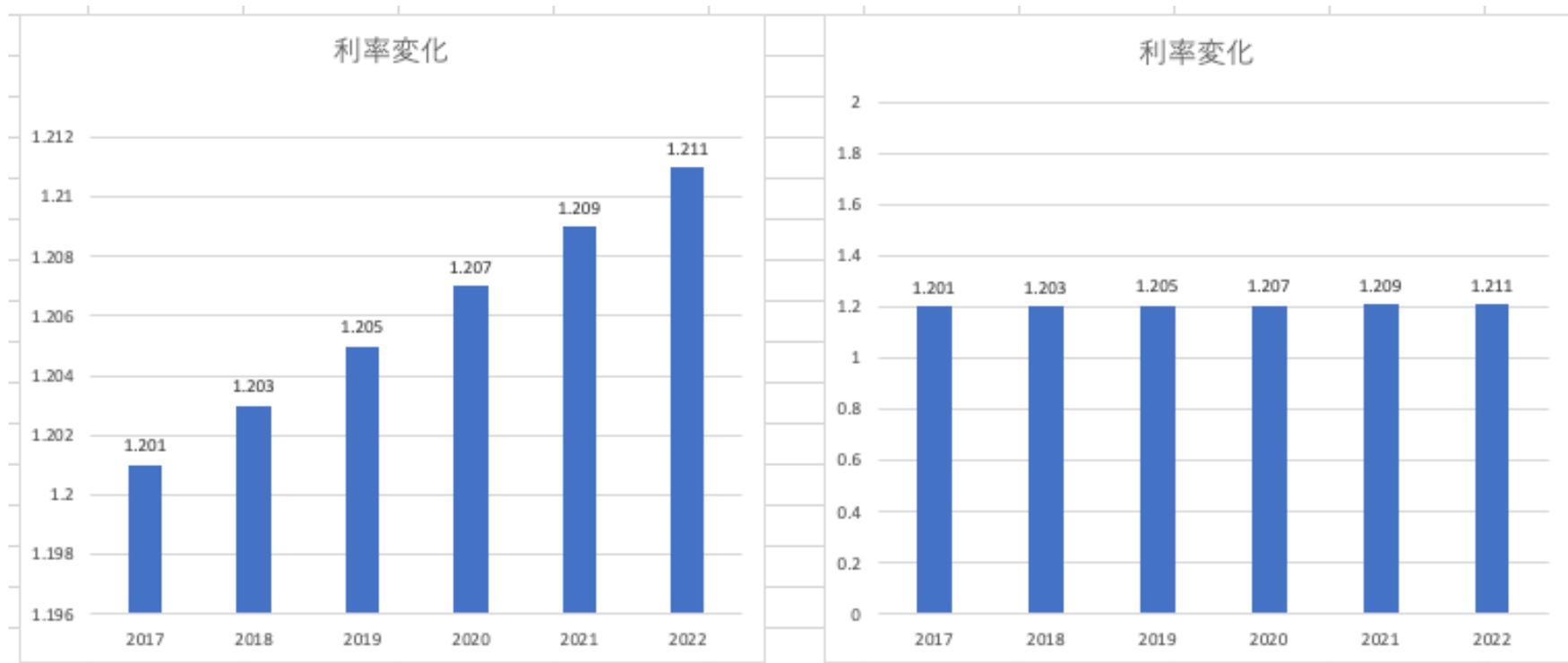
データ (結果) を見て考察するのは自分



このグラフからわかることはなんだろうか？

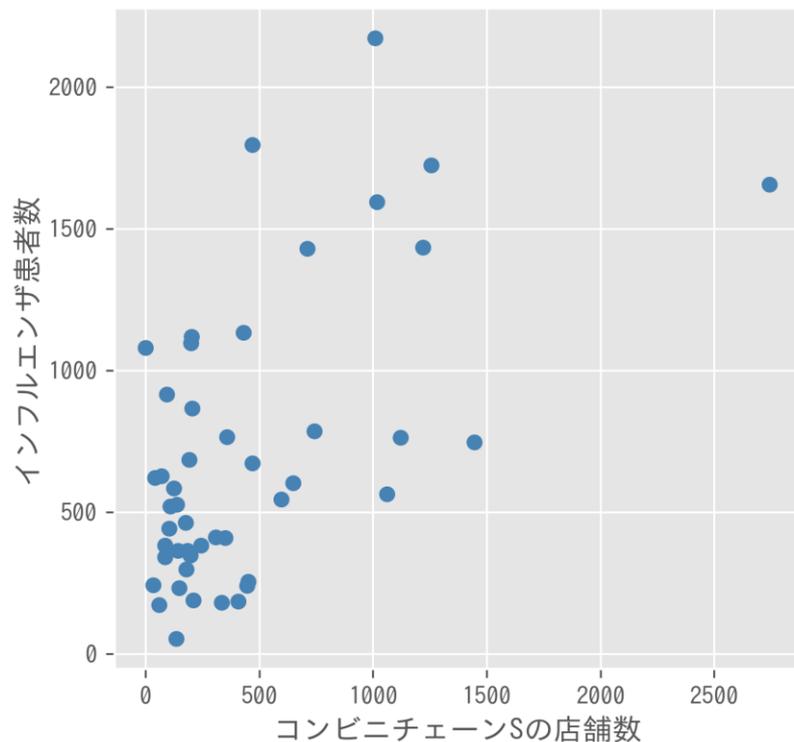
統計を行う上での注意点（１）

見た目に騙されないようにする



統計を行う上での注意点（2）

結果から間違った方向性に向かないようにする



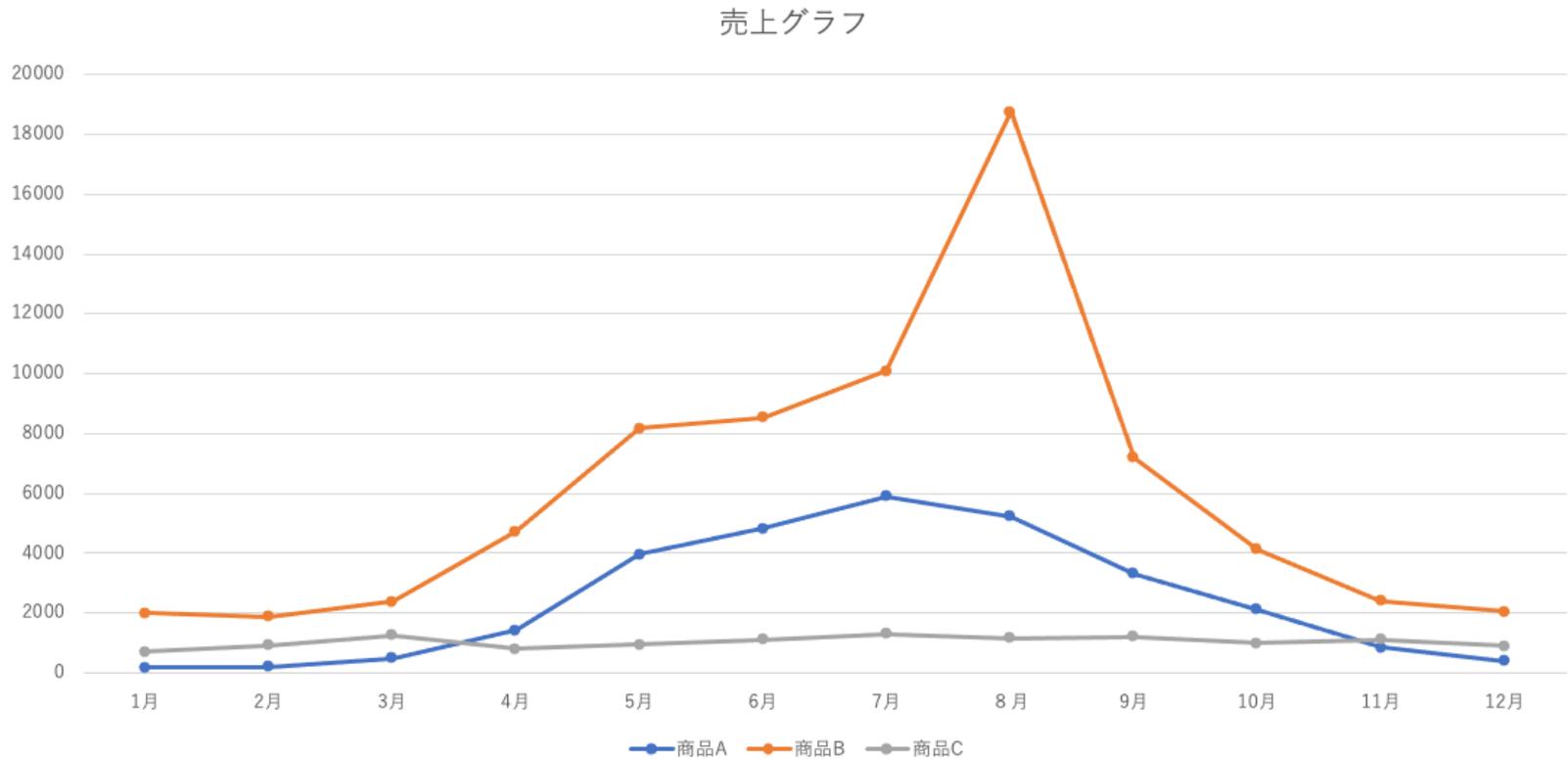
左の図はインフルエンザ患者数とコンビニチェーン店の店舗数の散布図です。コンビニチェーン店が多いとインフルエンザの患者数が多くなるようなので、コンビニの数を減らせばインフルエンザの患者数は減る…？

ワーク



ワーク①

以下のグラフからわかることを書き取りなさい。



ワーク②

Work Sheet

テーマ



テーマを書き込む

グラフを見て、直感的に分かることから
どんどん書き出していきましょう。

何を読み取れたかをグループで話し
あいましょうグループメンバーの意
見で、いいと思ったことは、どんど
ん取り入れていきましょう。

ワーク③

グループごとの意見を全体に共有しましょう。

どんなことを読み取れましたか？

※他のグループの説明から新しい視点を見つけてください。



データ分析に求められる着眼点③



確率とはなにか？

確率とは何か？

方程式や関数、図形の定理など、中学校・高校で学んできた数学には、いろいろな分野・単元がありました。そのなかで、特殊な職業に就かれています方以外で、最も日常的に聞く言葉が「確率」ではないでしょうか？

サイコロを振ったときに、1の目が出る確率は？
明日の降水確率は？
このプロジェクトの成功の確率は？

確率とは、ある事柄や行動の結果が特定の結果となる割合

つまり、「そうなるだろう」という目安

どうして、こんなことを考えるのか？

世界は、偶然でできている！

数学的確率

起きうる結果が何通りあるか計算して求める確率

統計的確率

実際に起こったことの結果を集計して求める確率

サイコロをふったとき、1が出る確率は？

サイコロには1～6までの目があります。ふったときには、1～6の目のうち、どれかひとつに決定します。つまり、起きうる結果は1～6までの6通りで、結果は、1～6のうちのどれか一つです。
したがって、

決定する結果

$$1 \div 6 = \frac{1}{6}$$



あり得る結果の
パターンの数

確率

今日の午後雨が降る確率は30%です。

今日の午後の天気図や、その他の気象条件を過去の類似した状況に照らし合わせて、過去にどれくらいの割合で雨が降っていたかの数値を確率として出したものです。

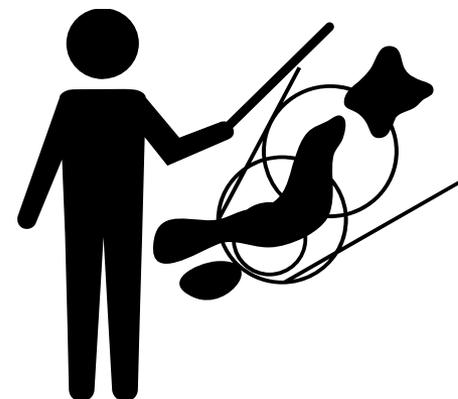
したがって、計算の意味は

$$30 \div 100 = \frac{3}{10}$$

過去の同様な
気象条件の時に、
雨が降った回数

調べた過去の同様な
気象条件の件数

確率



確率の求め方

場合の数とあたりの数

$$\text{確率} = \frac{\text{当てはまる数}}{\text{場合の数}}$$

ここの算出の仕方の例を見ていきましょう。

サイコロをふって、1が出る確率

$$1 \div 6 = \frac{1}{6}$$

サイコロをふって、偶数が出る確率

$$3 \div 6 = \frac{3}{6} = \frac{1}{2}$$

2, 4, 6のどれかが
出ればいから、
当てはまる数は3

確率の計算では、場合の数を正しく算出することが非常に重要!

大数の法則

たとえば、今ここでサイコロを6回振ったとしましょう。はたして、1の目は何回でるのでしょうか？

数学的確率から考えていけば、1回は1の目が出るように思われますが、実際はそうはならないケースの方が多いのではないのでしょうか。これは、サイコロの出目は、偶然によって偏るという事から起こります。

しかし、サイコロを振っていく回数を増やしていくと、どんどん1の目が出る確率は $\frac{1}{6}$ に近づいていきます。60回振った場合では、8回であったり13回であったものが、600回、6000回、60000回と続けていくことにより、どんどん $\frac{1}{6}$ に近づいていきます。

このように、試行を繰り返していくことで、統計的確率が数学的確率に近づいていくことを、「**大数の法則**」といいます。

確率の豆知識 ①

年末ジャンボ宝くじで1等に当選する確率と 自動車事故で死亡する確率

年末ジャンボの1等にあたる確率を考えてみましょう。

年末ジャンボの販売枚数は4億4千枚万です。1等7億円は22本あります。1等の当る確率は次の計算式で求められます。

数学的確率

$$22 \div 440,000,000 = \frac{1}{20,000,000}$$

2000万分の1！

今度は、自動車事故で死亡する確率を考えてみましょう。

日本の人口は1億2580万人です。警察庁の発表によると令和元年の交通事故による死者数は3215人です。（1年間の確率です。）

統計的確率

$$3,215 \div 125,000,000 = \frac{1}{40,000}$$

4万分の1！

練習問題

2つのサイコロを投げた時の確率について

大小2つのサイコロを振った時、次の確率を求めなさい。

1. 2つのサイコロの出目が同じ確率は？
2. 大きいサイコロの出目が、小さいサイコロの出目よりも大きい確率は？
3. 小さいサイコロと大きいサイコロの出目の差が3である確率は？
4. 2つのサイコロの出目の和が、7になる場合と7になる場合では、どちらの方が確率が高いか？

2つのサイコロを投げた時の確率について（解1）

組み合わせ表を使って、確率を考える。組み合わせ表は、すべての場合の数を考えることができるので、確率を考える時に有効な手段のひとつです。

1. 2つのサイコロの出目が同じ確率は？

小 \ 大	1	2	3	4	5	6
1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

すべての組み合わせのパターンは36通りあります。その中で、出目が同じパターンは6通りです。したがって確率は、

$$\frac{6}{36} = \frac{1}{6}$$

当てはまる数

場合の数

2つのサイコロを投げた時の確率について（解2）

2. 大きいサイコロの出目が、小さいサイコロの出目よりも大きい確率は？

小 \ 大	1	2	3	4	5	6
1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

オレンジでマークした部分が辺りのパターンとなります。

$$\frac{15}{36} = \frac{5}{12}$$

当てはまる数

場合の数

2つのサイコロを投げた時の確率について（解3）

3. 小さいサイコロと大きいサイコロの出目の差が3である確率は？

小 \ 大	1	2	3	4	5	6
1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

グリーンでマークした部分が辺りのパターンとなります。

$$\frac{6}{36} = \frac{1}{6}$$

当てはまる数

場合の数

2つのサイコロを投げた時の確率について（解4）

4. 2つのサイコロの出目の和が、6になる場合と7になる場合では、どちらの方が確率が高いか？

小 \ 大	1	2	3	4	5	6
1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

和が6になる確率 $\frac{5}{36}$

和が7になる確率 $\frac{6}{36} = \frac{1}{6}$

よって、和が7になる確率の方が高い

3つのサイコロを投げた時の確率について

3つのサイコロを振った時、出目の和が9になる場合と10になる場合では、
どちらの確率が高いか考えてみましょう。

3つのサイコロを投げた時の確率について（解）

サイコロを3つ振った時に目目の和が9になるパターンは、(1, 2, 6), (1, 3, 5), (1, 4, 4), (2, 2, 5), (2, 3, 4), (3, 3, 3)の6パターン

サイコロを3つ振った時に目目の和が10になるパターンは、(1, 3, 6), (1, 4, 5), (2, 2, 6), (2, 3, 5), (2, 4, 4), (3, 3, 4)の6パターン

どちらも、6パターンなので、同じ確率のように考えてしまいそうですが、大小2つのサイコロの時のように組み合わせ表を考えてみましょう。大きな表になるので、現実的に書いてみるのは面倒ですが、大中小の3つのサイコロの組み合わせを考えていくと次のように考えられます。

(1, 2, 6)のパターン

(1, 2, 6), (1, 6, 2), (2, 1, 6), (2, 6, 1), (6, 1, 2), (6, 2, 1)

このように、和が9になるパターンと、10になるパターンのそれぞれのサイコロの組み合わせを順番をつけて考えていくと、和が9になるパターンは25通り、和が10になるパターンは27通りあります。

したがって、3つのサイコロを振った時に和が10になる確率の方が和が9になる確率より高いと言えます。

場合の数・当りの数の求め方

場合の数を考える

ドリンクバーでのソフトドリンクの選択と、好きなコーヒーの選択

8種類のドリンクがあるドリンクバーで、1回好きなものを飲める。（混ぜるのはなし！）

このときの場合の数は、8ですよね。

それでは、好きなコーヒーを選んだ場合の場合の数を次の条件から考えてみましょう。

選択出来るチョイス

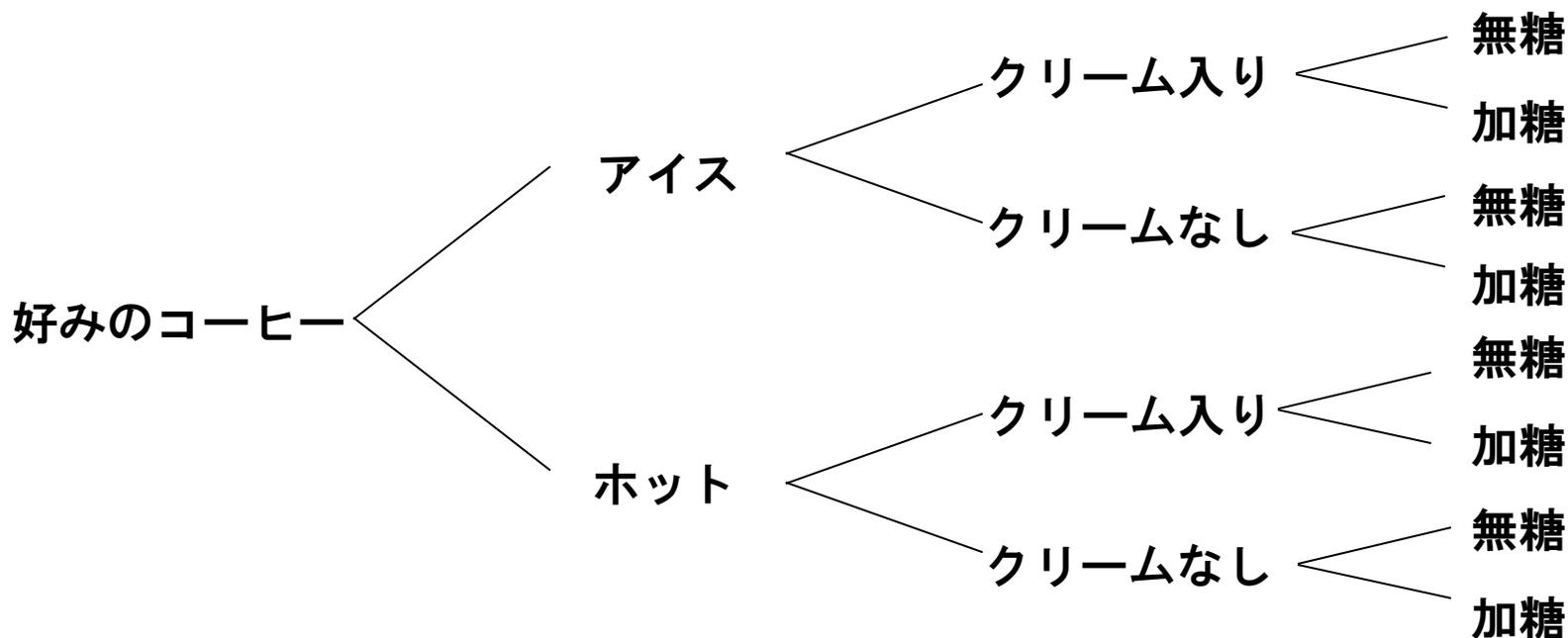
1. アイスかホット
2. クリームは入れるか入れないか
3. 砂糖は入れるか入れないか



樹形図で考える

樹形図で整理してみましよう。

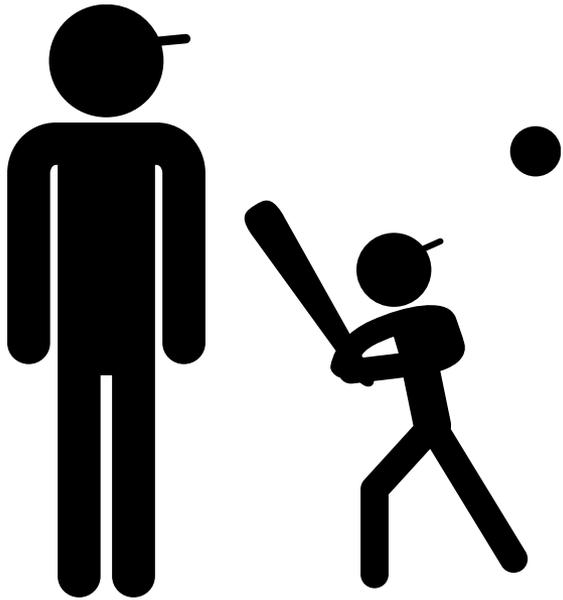
樹形図を書くと、分岐とパターンがわかりやすくなり、場合の数を求めるときに、間違いが起こりにくくなります。



2つの選択肢で3回分岐する、つまり $2^3 = 8$ の場合の数となります。

階乗を使って計算する 1

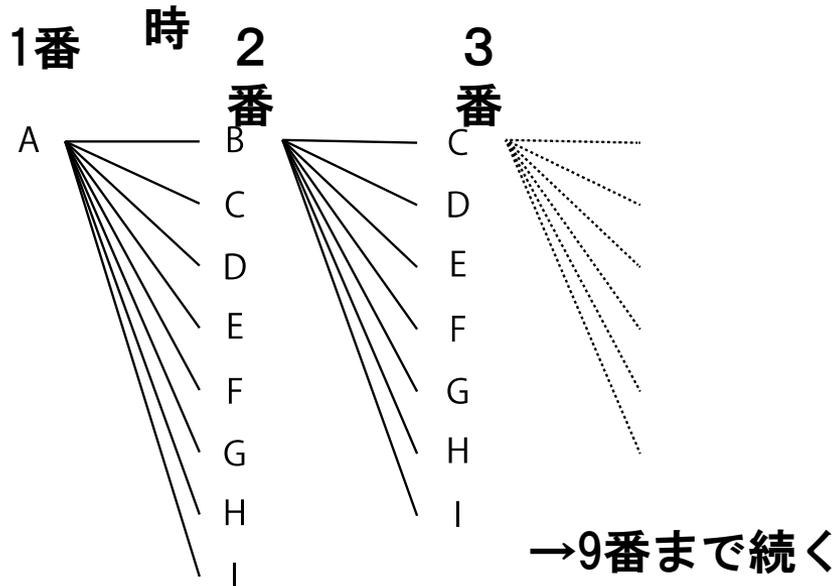
少年野球のチームでやっとメンバーが9人揃った。各ポジションはコーチと選手の話し合いで決まったが、打順が決まらない。そこで、コーチは、打順1番から9番まで、すべてのパターンを試してみしてから決めることを提案した。打順は何パターンあるのだろうか？



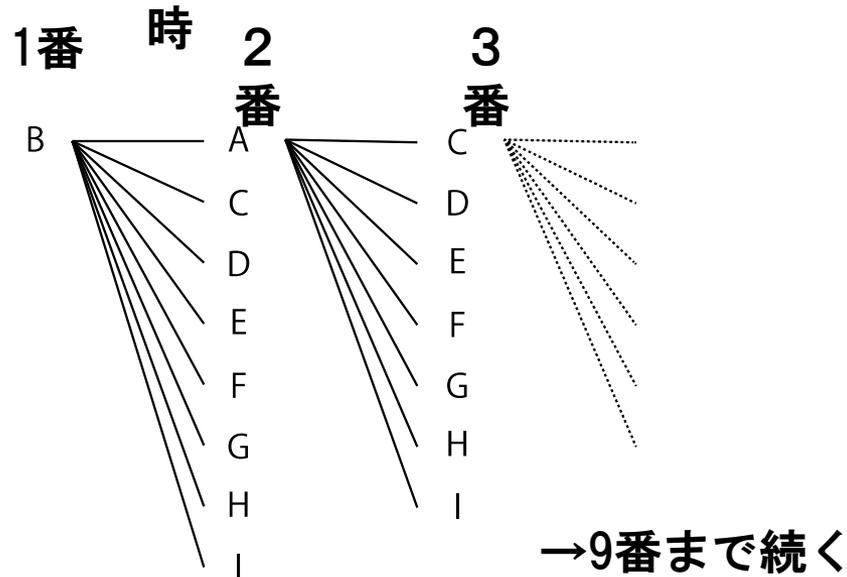
階乗を使って計算する2

AからIまでの9人がいる。打順1番に入れる可能性は9人。2番に入れるのは、1番に入らなかった8人、3番に入れるのは1番2番に入らなかった7人と、言うカタチで順番を考えていくことができる。

Aを1番とした



Bを1番とした



階乗を使って計算する3

1番に入ることの出来る選手は9人。2番に入ることの出来る選手は1番に入らなかった8人。3番に入ることのできる選手は、1番と2番に入らなかった7人、というようにかけ算で求めることができます。

$$9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 362,880$$

この式を階乗記号「！」を使って表すことができます。

$$9! = 362,880$$

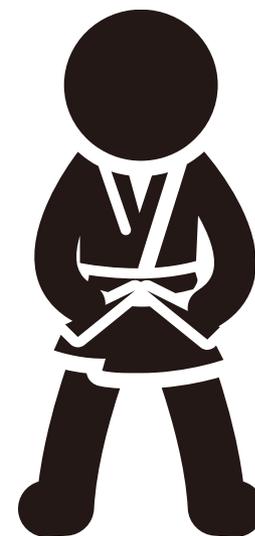
仮に1日1パターンの打順を試してみて、毎日試合をしたとして全部のパターンを試すまでに約1000年かかります。あまり現実的な提案とは言えません。

順列を考える 1

今度は、少年野球ではなくある中学の柔道部の場合を考えてみましょう。

柔道部員が8名いる。大会に出るために5人の選手を決めなければならない。

先鋒、次鋒、中堅、副将、大将のパターンは何通り考えられるか？



順列を考える2

野球の打順の時のように、順番が関係ある並び方のパターンを考えますが、今回は、5人選んだ後の3人の順番はパターンには関係ありません。ですので、次のような式で計算することができます。

$$8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 6,720$$

ここの順番は関係ないので計算に含めない。

参考（順列の公式）

異なる n 個のものから r 個選んだときの順列 $= {}_n P_r = \frac{n!}{(n-r)!}$

$$\text{この問題の場合、} {}_8 P_5 = \frac{8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{3 \times 2 \times 1} = 6,720$$

組み合わせを考える 1

次は、ある高校のバスケットボール部の場合を考えてみましょう。

バスケットボール部員が8名いる。レギュラー5人を選んだ時のパターンは何通りあるか？ただし、このバスケットボールチームにはポジションの考えがないものとする。



組み合わせを考える2

野球や柔道の時と違って、レギュラー5人選考したときの順番は関係ありません。その時には、計算方法はどのように変わのでしょうか。選手の8名をAからHとします。順列の時のように5人選考したときのことを考えてみましょう。

仮に、A、B、C、D、Eの5人が選ばれたとします。

このときに、順列では別のパターンとして考えていた、B、D、A、E、Cや、E、B、D、E、Aなどは、同じメンバーで構成されているので、重複していると考えられます。ですから、順列で求めたパターンから重複している部分を取り除いて行けばいいわけです。

たとえばこの場合ならば、8人から選んだ5人の順列のパターン数を重複する5人の順列（同じメンバーで順番が違う）で割ってあげればいいわけです。

$$(8 \times 7 \times 6 \times 5 \times 4) \div (5 \times 4 \times 3 \times 2 \times 1) = 56$$

参考（組み合わせの公式）

異なる n 個のものから r 個選んだときの組み合わせ $= {}_n C_r = \frac{{}_n P_r}{r!}$

$$\text{この問題の場合、} {}_8 C_5 = \frac{8 \times 7 \times 6 \times 5 \times 4}{5 \times 4 \times 3 \times 2 \times 1} = 56$$

順列と組み合わせのまとめ

順列 選んだものの順番や順序が関係ある場合。

異なる n 個のものから r 個選んだときの順列 $= {}_n P_r = \frac{n!}{(n-r)!}$

- ・ 10個の中から、4つ選んだ場合

$${}_{10}P_4 = \frac{10!}{(10-4)!} = \frac{10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{6 \times 5 \times 4 \times 3 \times 2 \times 1} = 5,040$$

組み合わせ 選んだものの順番や順序が関係がない場合。

異なる n 個のものから r 個選んだときの組み合わせ $= {}_n C_r = \frac{{}_n P_r}{r!}$

- ・ 10個の中から、4つ選んだ場合

$${}_{10}C_4 = \frac{{}_n P_r}{r!} = \frac{(10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1) \div (6 \times 5 \times 4 \times 3 \times 2 \times 1)}{4 \times 3 \times 2 \times 1} = 210$$

練習問題

順列と組み合わせの練習問題

1. 5人の陸上部員がいる。100メートル走の選考会で、1位になった選手は奨学金と大学への推薦がもらえる。2位になった選手は、奨学金はもらえないが大学への推薦がもらえる。このとき、1位と2位になった選手の組み合わせは何通り考えられるか。

2. 5人の陸上部員がいる。100メートル走の選考会で、1位と2位になった選手はオリンピック候補生として強化合宿に招待される。このとき、1位と2位になった選手の組み合わせは何通り考えられるか。



順列と組み合わせの練習問題（解）

1. 5人から2人を選んだ順列となる。

したがって、 ${}_5P_2$ となり、
計算は $(5 \times 4 \times 3 \times 2 \times 1) \div (3 \times 2 \times 1) = 20$

20通りのパターンが考えられる。

2. 5人から2人を選んだ組み合わせとなる。

したがって、 ${}_5C_2$ となり、
計算は $(5 \times 4) \div (2 \times 1) = 10$

10通りのパターンが考えられる。

データ分析に求められる着眼点④



乗法定理と加法定理

乗法定理と加法定理

確率の計算をする上で、乗法定理と加法定理の意味をきちんと理解しておくことは重要です。乗法定理と加法定理は一般的に、積の法則、和の法則とも呼ばれることがあります。

乗法定理

お互いに影響を与えない、独立した事象が連続した時の確率を求める場合は、それぞれの事象の確率の積を求める。

加法定理

ある事象で、どちらか一方が起きるともう片方が起きえない（同時には起こりえない事象）で、どちらかが起きる確率を求める場合は、それぞれの事象の和を求める。

言葉の説明を聞いても、イメージしにくいので、具体例を交えて見ていきましょう。

乗法定理

サイコロを2つふった時、2つとも1の目が出る確率は？

2つのサイコロをサイコロAとサイコロBとしましょう。この2つのサイコロの出る目は、お互いに影響しあいません。ですから、サイコロAをふった時に1が出る確率とサイコロBをふった時に1が出る確率は、それぞれ独立した事象の連続と考えることができるので、それぞれの確率の積で求めることができます。

サイコロAで1が出る確率

$$1 \div 6 = \frac{1}{6}$$

×

サイコロBで1が出る確率

$$1 \div 6 = \frac{1}{6}$$

かけ算で求める！

$$\frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

加法定理

サイコロを2つふった時、どちらか一方が1になる確率は？

どちらか一方が1になるということは、サイコロAをふった時に1が出た場合には、サイコロBでは1でない目が出ています。逆に、サイコロBで1が出た場合には、サイコロAでは1でない目が出ています。つまり、どちらか一方が1の場合に、もう片方が1でない事象の計算をする場合は、たし算を使います。

サイコロAで1が出て、
サイコロBで1以外が出る確率

サイコロA

$$1 \div 6 = \frac{1}{6}$$

サイコロB

$$5 \div 6 = \frac{5}{6}$$

$$\frac{1}{6} \times \frac{5}{6} = \frac{5}{36}$$

+

サイコロAで1以外が出て、
サイコロBで1が出る確率

サイコロA

$$5 \div 6 = \frac{5}{6}$$

サイコロB

$$1 \div 6 = \frac{1}{6}$$

$$\frac{5}{6} \times \frac{1}{6} = \frac{5}{36}$$

たし算で求める！

$$\frac{5}{36} + \frac{5}{36} = \frac{10}{36} = \frac{5}{18}$$

少なくとも・・・の場合

サイコロを2つふった時、少なくともどちらか一方が1になる確率は？

乗法定理と加法定理の基本を理解した上で、次の事象の確率を考えてみましょう。

サイコロAとサイコロBの2つのサイコロをふった時に、少なくともどちらか一方が1になる確率を求めてみましょう。

ポイント

少なくとも一方が1になるという事象を考えてみましょう。この事象は、2つのサイコロのどちらかが1になった時だけでなく、両方とも1の場合でも当てはまります。また、両方とも1になる事象と、どちらか一方が1になる事象は同時には起こりえません。ですので、求めたい確率は、両方とも1になる確率と、どちらか一方が1になる確率の和で求められます。

サイコロA、サイコロBの
両方で1が出る確率

$$\frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

+

サイコロA、サイコロBの
どちらか一方で1が出る確率

$$\frac{5}{36} + \frac{5}{36} = \frac{10}{36} = \frac{5}{18}$$

たし算で求める！ $\frac{1}{36} + \frac{5}{18} = \frac{11}{36}$

乗法定理の簡単な例

トランプを使った確率を考えてみましょう。

ジョーカーを抜いたトランプが1セット、52枚あります。ここから、数字の7を引く確率は $\frac{1}{13}$ です。

それでは、このトランプの山から2枚引いたときに、2枚とも7である確率はいくつか？



乗法定理の簡単な例（解）

1枚引いたときに $\frac{1}{13}$ というのは、次の計算からでてきますね。

7の枚数

$$\frac{4}{52} = \frac{1}{13}$$

トランプ全体の枚数

2枚引いたというのは、1枚引いた後に、戻さずにもう1枚ひくのと 同じなので、2枚目が7である確率は

残りの7の枚数

$$\frac{3}{51} = \frac{1}{17}$$

トランプ全体の
残りの枚数

したがって、2枚とも7である確率は $\frac{1}{13} \times \frac{1}{17} = \frac{1}{221}$

加法定理の簡単な例

トランプを使った確率を考えてみましょう。

ジョーカーを抜いたトランプが1セット、52枚あります。ここから、数字の7を引く確率は $\frac{1}{13}$ です。

それでは、このトランプの山から2枚引いたときに、どちらか1枚が7である確率はいくつか？



加法定理の簡単な例（解）

1枚が7で、2枚目が7以外の確率

7の枚数

$$\frac{4}{52}$$

×

$$\frac{48}{51} = \frac{16}{221}$$

7以外の枚数

トランプ全体の枚数

1枚引いた後の
残りの枚数

1枚が7以外で、2枚目が7の確率

7以外の枚数

$$\frac{48}{52}$$

×

$$\frac{4}{51} = \frac{16}{221}$$

7の枚数

トランプ全体の枚数

1枚引いた後の
残りの枚数

この2つの現象は同時に起こりえないので、
求める確率は

$$\frac{16}{221} + \frac{16}{221} = \frac{32}{221}$$

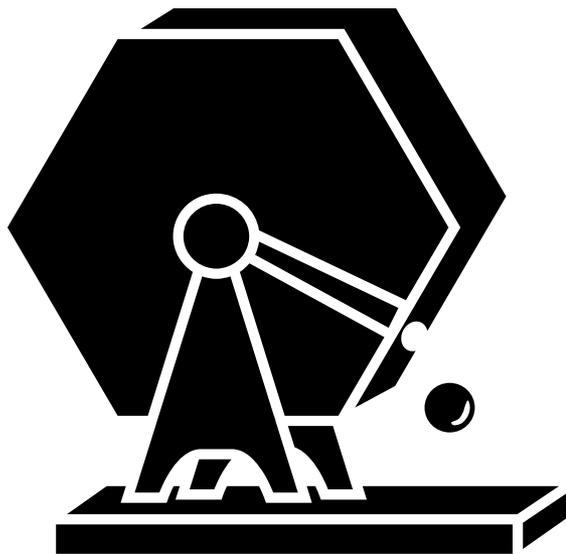
くじを戻した場合と、戻さない場合

ガラポン（下の絵のような福引き）を考えてみます。

玉が100個入っています。当りは1つで残りの99個はハズレです。以下の2つのケースを考えてみましょう。

1. 1回引くたびにハズレの玉を戻して、毎回100個でくじ引きをする。
2. くじを引いても、玉を戻さずにくじ引きを続ける。

1のケースと2のケースで13番目にくじを引いた時に当りを引く確率を考えてみましょう。



くじを戻した場合と、戻さない場合（解）

1と2とどちらもケースでも、13番目にくじを引いた人が当りを引く確率は $\frac{1}{100}$ です。

1のケース：1回引くたびにハズレの玉を戻して、毎回100個でくじ引きをする。

くじを引く事象は、毎回独立した事象と考えられます。前のくじ引きと次のくじ引きの間には、なんら関係はありません。ですから、毎回毎回100個の玉のなかから1個の当りを狙う $\frac{1}{100}$ の確率でのくじ引きが続いていきます。

2のケース：くじを引いても、玉を戻さずにくじ引きを続ける。

くじを引く事象は、連続した事象と考えられます。それでは13番目に当りを引く確率を計算してみましょう。13番目にくじを引くということは、前の12人目まではハズレが続いていると考えると、

$$\frac{99}{100} \times \frac{98}{99} \times \frac{97}{98} \times \frac{96}{97} \times \frac{95}{96} \times \frac{94}{95} \times \frac{93}{94} \times \frac{92}{93} \times \frac{91}{92} \times \frac{90}{91} \times \frac{89}{90} \times \frac{88}{89} \times \frac{1}{88} = \frac{1}{100}$$

1回目のハズレ

6回目のハズレ

約分で1になる

13回目で当り

練習問題

袋の中の玉

袋のなかに赤い玉が3個，白い玉が3個，青い玉が4個が入っている。同時に2個を取り出したとき，同じ色になる確率はいくらか。

袋の中の玉（解）

$$\text{2つとも赤い玉が出る確率は、} \frac{3}{10} \times \frac{2}{9} = \frac{1}{15}$$

$$\text{2つとも白い玉が出る確率は、} \frac{3}{10} \times \frac{2}{9} = \frac{1}{15}$$

$$\text{2つとも青い玉が出る確率は、} \frac{4}{10} \times \frac{3}{9} = \frac{2}{15}$$

それぞれの事象は同時には起きえないので、たし算で確率を求めることができる。

$$\frac{1}{15} + \frac{1}{15} + \frac{2}{15} = \frac{4}{15}$$

先に3勝する確率は？

AとBの2人でチェスをしている。AがBに勝つ確率が $\frac{1}{3}$ であるとき、
Aが先に3勝する確率はいくらか？

先に3勝する確率は？

AがBに3勝するパターンは10通りある。それぞれの確率を求めて和を出せばよい。

それぞれの試合は、独立した事象なので毎回のAの勝率は $\frac{1}{3}$ です。

① Aの3勝0敗が1パターン (AAA)

$$\frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} = \frac{1}{27}$$

② Aの3勝1敗が3パターン (AABA、ABAA、BAAA)

$$\frac{1}{3} \times \frac{1}{3} \times \frac{2}{3} \times \frac{1}{3} = \frac{2}{81} \quad (\text{かけ算中の } \frac{2}{3} \text{ の位置は負けたパターンにより変わります。})$$

$$3 \text{ パターンあるので } \frac{2}{81} + \frac{2}{81} + \frac{2}{81} = \frac{6}{81} = \frac{2}{27}$$

③ Aの3勝2敗が6パターン (AABBA、ABABA、ABBAA、BAABA、BABAA、BBAAA)

$$\frac{1}{3} \times \frac{1}{3} \times \frac{2}{3} \times \frac{2}{3} \times \frac{1}{3} = \frac{4}{243} \quad (\text{かけ算中の } \frac{2}{3} \text{ の位置は負けたパターンにより変わります。})$$

$$6 \text{ パターンあるので } \frac{4}{243} + \frac{4}{243} + \frac{4}{243} + \frac{4}{243} + \frac{4}{243} + \frac{4}{243} = \frac{24}{243} = \frac{8}{81}$$

$$\text{いずれも同時には起こらないので } \frac{1}{27} + \frac{2}{27} + \frac{8}{81} = \frac{17}{81}$$

直感と計算

計算の大切さ（１）

次の3つを確率が高い順に並べてください。

※まずは計算をしないで、直感で順番を考えてから、計算をして確認してみましよう。

- ① 赤玉50個、白玉50個入った袋から、1回で赤玉を取り出す。
- ② 赤玉90個、白玉10個入った袋から、7回連続で赤玉を取り出す。
但し取り出した玉は、毎回元の袋に戻します。
- ③ 赤玉10個、白玉90個入った袋から、7回連続で球を取り出し、
1回以上が赤玉である。但し取り出した玉は、毎回元の袋に戻します。

計算の大切さ（２）

実際の確率計算に反して、多くの人は②→①→③の順に確率が高いと感じるそうです。

- ① 赤玉50個、白玉50個入った袋から、1回で赤玉を取り出す。

→ **50%**

- ② 赤玉90個、白玉10個入った袋から、7回連続で赤玉を取り出す。

→ **48%**

$$0.9 \times 0.9 \times 0.9 \times 0.9 \times 0.9 \times 0.9 \times 0.9 = 0.478$$

- ③ 赤玉10個、白玉90個入った袋から、7回連続で球を取り出し、1回以上が赤玉である。

→ **52%**

$$1 - (0.9 \times 0.9 \times 0.9 \times 0.9 \times 0.9 \times 0.9 \times 0.9) = 0.521$$

参考：ダニエル・カーネマン「ファスト&スロー」

直感の不確定さ（1）

人間の直感がそれぞれの人生において非常に重要なモノであることは間違いないと思われませんが、データ分析や統計・確率を学んでいく中で直感に頼って考察に飛びついてしまうのは非常に危険です。
ここでは、直感と計算結果が乖離している例をいくつか考えていきましょう。

簡単な例から考えていきましょう。

6面のサイコロがあります。このサイコロは数字が書かれているのではなく、色分けされています。3面が赤、2面が青、1面が黄色で塗られています。このサイコロを2個ふった時に、色の組み合わせとして一番高い確率ででてくるものは何色と何色の組み合わせでしょうか？



直感の不確定さ（２）

簡単な計算で検証することができます。

直感的には、赤で塗られている面が多いので、直感的に両方とも赤になる確率が高そうに感じられますが、それぞれの確率を計算してみましょう。

両方とも赤になる確率

$$\frac{3}{6} \times \frac{3}{6} = \frac{9}{36} = \frac{1}{4}$$

赤と青の組み合わせになる確率

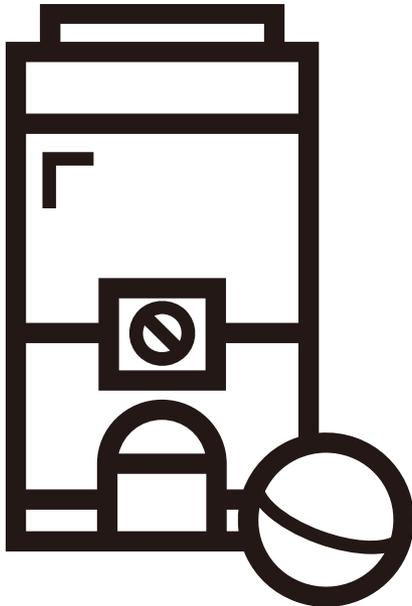
$$\frac{3}{6} \times \frac{2}{6} + \frac{2}{6} \times \frac{3}{6} = \frac{12}{36} = \frac{1}{3}$$

計算をしてみると、赤と青の組み合わせになる方が高いということが分かります。

直感と余事象（1）

5種類の景品の内、欲しいモノは1種類だけ

1回200円で引けるガチャに、5種類のフィギュアが入っている。すでに、4種類持っていて残りはあと1種。その時たまたまネット通販で、最後の1種のフィギュアが送料込み1000円ででていた。どうしてもコンプリートしたい。ガチャをやるべきか通販で買ったほうがいいか。



直感と余事象（2）

5種類の景品の内、欲しいモノは1種類だけ

欲しいフィギュアが5種類のうちの1種類。つまり確率としては $\frac{1}{5}$ となります。

この余事象を考えていくとわかりやすいと思います。

$\frac{1}{5}$ であたるもの = $\frac{4}{5}$ ではずれる、ということですね。

通販の1000円分を使ったとして、5回連続ではずれる確率を考えてみましょう。

$$\frac{4}{5} \times \frac{4}{5} \times \frac{4}{5} \times \frac{4}{5} \times \frac{4}{5} = \frac{1024}{3125}$$

$\frac{1}{5}$ の確率と聞いた直後の直感で、5回に1回は当りが引けるように感じてしま

いがちですが、余事象を考え計算すると、おおよそ $\frac{1}{3}$ の確率で1000円使っても

ガチャでゲットすることはできません。

もちろん運次第では、200円で手にすることもあります。

誕生日のパラドックス（１）

出会った人が自分と同じ誕生日である確率は $\frac{1}{365}$ です。40人のクラスの中に同じ誕生日のペアがいる確率はどれくらいだろうか？



誕生日のパラドックス（2）

40人のクラスの任意の1人を選びます。任意に2人目を選んだ時に、1人目と2人目の誕生日が異なる確率は $\frac{364}{365}$ です。3人目を選んだ時に、3人目の誕生日が前の2人と異なる誕生日である確率は $\frac{363}{365}$ となります。このように4人目、5人目と全員が異なる誕生日の確率を計算していくと次のような式で計算できます。

$$\frac{364}{365} \times \frac{363}{365} \times \frac{362}{365} \times \frac{361}{365} \times \frac{360}{365} \times \frac{359}{365} \times \frac{358}{365} \times \frac{357}{365} \cdots \times \frac{326}{365} \doteq 0.108768$$

2人目

3人目

40人目

少なくとも1組の同じ誕生日のペアがいるということは、全員が異なる誕生日の余事象と考えられるので、

$$1 - 0.108768 = 0.891231$$

約89%の確率で40名のクラスには同じ誕生日のペアがいることがわかります。

ちなみに40人のクラスで、自分と同じ誕生日のクラスメートがいる確率は

$$1 - \left(\frac{364}{365}\right)^{40} \doteq 0.103932 \text{ 約10\%となります。}$$

データ分析に求められる着眼点⑤



期待値（１）

確率を活用していく上で「期待値」の計算は判断材料として非常に有効です。

期待値とは、ある事象で得られる値の平均値を表し、得られる可能性があるすべての値と、それぞれが起こる確率の和との積を求めます。

文章で読んでもイメージしづらいので、具体例を考えてみましょう。

100円で1回サイコロをふることができます。出た目に応じて下の表のように景品が貰えます。サイコロの目はどの目も $\frac{1}{6}$ の確率で出ます。期待値は、どうやって求めたらいいのでしょうか。

サイコロの目	景品
1	賞品A（1000円相当）
2	ハズレ
3	賞品B（300円相当）
4	ハズレ
5	賞品C（10円相当）
6	ハズレ

期待値（２）

「期待値」は、それぞれが起こりえる確率と、その場合にえられる値の積の和で求めます。この場合は出たサイコロの目でもらえる景品の値と、その目が出る確率の積がその場合の値になり、すべての場合の和が期待値となります。

サイコロの目	景品	確率 $\frac{1}{6}$ をかけた金額
1	賞品A（1000円相当）	約166.7円
2	ハズレ	0円
3	賞品B（300円相当）	50円
4	ハズレ	0円
5	賞品C（10円相当）	約1.7円
6	ハズレ	0円
合計		218.4円

期待値（3）

このサイコロゲームの期待値は218.4円とでました。この218.4円とは、100円払ってこのゲームに参加して得られる値の平均値となり、投資額100円に対して、218.4円というリターンが平均的に得られるということになり、かなり率のよい投資と考えることができます。

参考（期待値の公式）

期待値 = （受け取る価値 × 確率）の合計

一般的には、 $\frac{\text{（受け取る価値} \times \text{確率）の合計}}{\text{投入価値}} \times 100$ の式を利用して、百分率で表すことが多い。

期待値（４）

期待値の計算方法がわかると、次のようなケースを考える際に参考となる数値を導き出すことができるようになる。

それぞれの期待値を算出して比較してみましょう。

期待値の高いプロジェクトはどれだろうか？

1. 実行予算が1000万円のプロジェクトで、20%の確率で成功する。
成功報酬は1200万円。
2. 実行予算が500万円のプロジェクトで、70%の確率で成功する。
成功報酬は800万円。
3. 実行予算が200万円のプロジェクトで、5%の確率で成功する。
成功報酬は5000万円。

期待値（5）

1. $1200万 \times 20\% = 240万$
実行予算が1000万円なので、1万円あたりの投資額に対する期待値は、
 $240万 \div 1000 = 2400$
投資1万円に対する期待額は2400円
2. $800万 \times 70\% = 560万$
実行予算が500万円なので、1万円あたりの投資額に対する期待値は、
 $560万 \div 500 = 11200$
投資1万円に対する期待額は1万1200円
3. $5000万 \times 5\% = 250万$
実行予算が200万円なので、1万円あたりの投資額に対する期待値は、
 $250万 \div 200 = 12500$
投資1万円に対する期待額は1万2500円

3番のプロジェクトの期待値が一番高く算出されます。

条件付きの確率

条件付きの確率（1）

現実社会での確率は、今まで学んできたような単純な条件で整理できる事象は少数派です。現実には、モノゴトを確率で考えていこうとすると、事象と起こりえる結果には様々な条件が加わっているケースがほとんどです。

条件が加わった事象の確率を考えてみましょう。

ジョーカーを抜いたトランプが1セット52枚あります。この中から無作為に1枚引いた時のカードはハートでした。このカードがAである確率はいくつか？



条件付きの確率（2）

確率を求めたい事象に対してある条件が加わった時の計算方法を考えてみます。

この場合、A選ぶという確率を求めるうえで、ハートのカードのなかから選ぶという条件が加わっていると考えることができます。よって確率の基本的な求め方の考えに戻って整理して考えてみると、当りは「ハートのAが選ばれる」すなわち52枚のカードの1枚を選ぶことになり、確率は $\frac{1}{52}$ の数を考えてみましょう。この事象はハートのカードが選ばれた時にだけおきることなので、全体の $\frac{1}{4}$ からハートのカードが選ばれる確率が $\frac{1}{13}$ となります。確率は $\frac{1}{13}$ です。

したがって、求める確率は $\frac{1}{13}$ となります。

$$\frac{\frac{1}{52}}{\frac{1}{4}} = \frac{1}{52} \div \frac{1}{4} = \frac{1}{13}$$

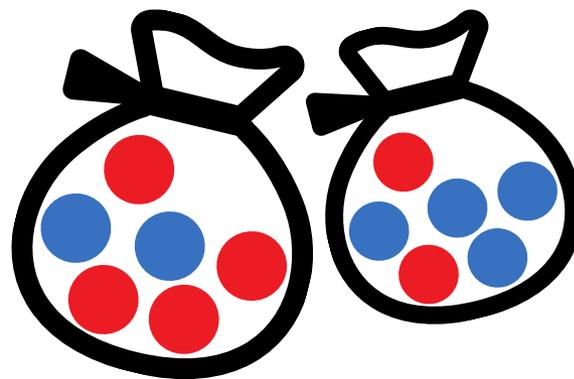
ハートのAが
選ばれる確率

ハートが
選ばれる確率

条件付きの確率（3）

次の事象の確率を考えてみましょう。

AとBの2つの袋があります。それぞれ6つずつボールが入っています。Aの袋には赤いボールが4つと青いボールが2つ、Bの袋には青いボールが4つと赤いボールが2つ入っていることが分かっています。今、無作為にどちらかの袋からボールをひとつとり出した時に、そのボールは赤でした。このボールがAの袋から取り出された確率を考えてみましょう。



条件付きの確率（４）

この場合、赤いボールを選ぶという確率を求めるうえで、Aの袋から選ぶという条件が加わっていると考えることができます。よって確率の基本的な求め方の考えに戻って整理して考えてみると、当りは

「Aの袋から選ばれる」かつ「赤いボールが選ばれる」

となります。次に場合の数を考えてみましょう。この事象は赤いボールが選ばれた時にだけおきることなので、全体の場合から**赤いボールが選ばれる確率**が場合の数となります。したがって、求める式は

$$\frac{\text{「Aの袋から選ばれる確率」} \times \text{「赤いボールが選ばれる確率」}}{\text{「赤いボールが選ばれる確率」}} = \text{Aの袋から赤いボールを選んだ確率}$$

となり、求める確率は $\frac{2}{3}$ となります。

Aの袋が
選ばれる確率

赤いボールが
選ばれる確率

Aの袋で
赤いボールが選ばれる確率

$$\frac{\frac{1}{2} \times \frac{4}{6}}{\left(\frac{1}{2} \times \frac{4}{6}\right) + \left(\frac{1}{2} \times \frac{2}{6}\right)} = \frac{1}{3} \div \frac{1}{2} = \frac{2}{3}$$

加法定理

Bの袋で
赤いボールが選ばれる確率

演習



演習（1）

あるところにA、Bという2つの商店があります。商店Aでは、すべてのお買い物に対して5%引きというキャンペーンをやっています。一方商店Bでは、それに対抗してお買い上げ20名あたりに1人無料にするというキャンペーンを始めました。どちらがお得か考えてみましょう。



それぞれの期待値を計算して比較してみよう。

演習（解）

仮に10万円の商品を購入したとして期待値の計算をしてみます。

商店A

$$10\text{万円} \times \frac{5}{100} \times \frac{100}{100} = 2500\text{円}$$

5%の割引

全員が当る

商店B

$$10\text{万円} \times \frac{100}{100} \times \frac{1}{20} + 10\text{万円} \times 0 \times \frac{19}{20} = 2500\text{円}$$

100%の割引

20人に1人が
当る

割引なし

20人に19人が
はずれる

計算をしてみると、10万円の買い物に対しての期待値は同額です。お店側にとっての負担は、全く同じと考えられます。

ですが、当たった時の額が商店Bの方が大きいので、話題性やお得感などは商店Bの方が上手であると思われます。

実数と割合

実数と割合（１）

次の表を見てみましょう。この表は、A市、B町、C村の人口と面積とコンビニの店舗数をまとめたものです。

	A市	B町	C村
人口	500,000	50,000	5,000
面積	75km ²	25km ²	50km ²
コンビニの店舗数	50店	3店	6店

1. どの自治体にコンビニがたくさんあるのか考えてみましょう。
2. もし自分がコンビニのオーナーになるならば、どこで新規営業を始めるか考えてみましょう。

実数と割合（1）

次の表を見てみましょう。この表は、A市、B町、C村の人口と面積とコンビニの店舗数をまとめたものです。

	A市	B町	C村
人口	500,000	50,000	5,000
面積	75km ²	25km ²	50km ²
コンビニの店舗数	50店	3店	6店

1. どの自治体にコンビニがたくさんあるのか考えてみましょう。
店舗数で比較するとA市のコンビニが一番多いですが、
人口に対する割合や面積に対する割合だと多い都市が変わります。
2. もし自分がコンビニのオーナーになるならば、どこで新規営業を始めるか考えてみましょう。
店舗からのある距離以内に住んでいる人口の多さで選ぶ方法も、
周りに店舗が少ない場所を選ぶ方法もあります。

標本誤差

平均点が60点のテストがあります。

そのテストで自分が100点をとった時の影響について考えてみましょう。

1. テストの受験者が10人だった場合。

平均点は $(60 \times 10 + 100) / 11 = 63.6$ 点で、3点以上上がります。

2. テストの受験者が10000人だった場合。

平均点は $(60 \times 10 + 10000) / 10001 = 60$ 点で、変わりません。

サンプルが少ない時は、一つ一つのデータの影響が大きくなります。

反対に、サンプルが多い時は、特異なデータの影響が小さくなります。

サンプルから得られた値が、母集団とどれくらい一致しているか（標本誤差）を考えることが重要です。

条件付き割合

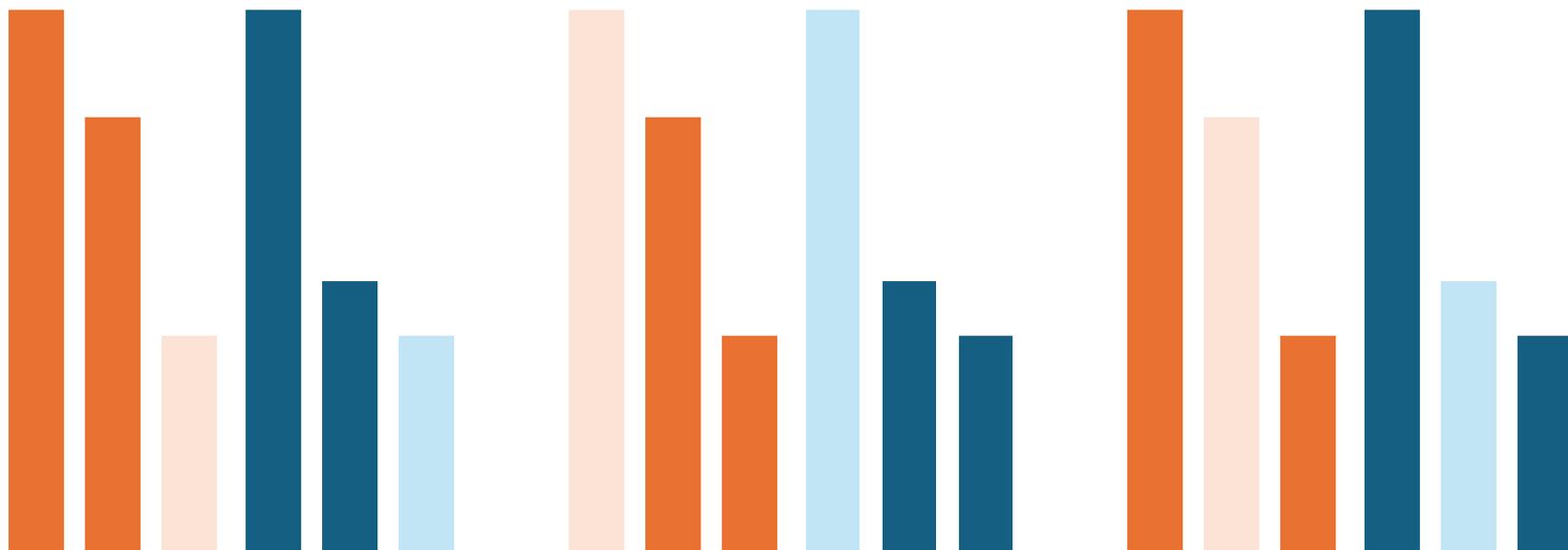
ある新規スポーツ用品ブランドの認知度を、競技場でアンケート調査を行った時、1000人のサンプルの80%が「知っている」と答え、「知っている」と答えたサンプルの50%が「良いイメージを持っている」と答えた。

ある新規スポーツ用品ブランドの認知度を、郊外のショッピングモールでアンケート調査を行った時、10000人のサンプルの50%が「知っている」と答え、知っていると答えたサンプルの80%が「良いイメージを持っている」と答えた。

どちらも、あるブランドを「知っていて良いイメージを持っている」人のサンプル全体に対する割合は等しく40%ですが、、

条件付き割合

割合は、基準をとる場所（条件）により変化します。



サンプル全体を基準に
知っている割合を考える

知っている人を基準に
好印象の割合を考える

サンプル全体を基準に
好印象の割合を考える

※赤い棒 … 競技場でのアンケート調査

青い棒 … ショッピングモールでのアンケート調査

一番左をサンプル全体として1にした時の、「知っている割合（中央）」と
「知っていて良いイメージを持っている割合（右）」

実数分析と比率分析

実数分析

- 実際の数値を用いて分析をします。
- 直感的にわかりやすく、「量」の比較を行えます。
- 具体的には…
 - 東京駅からの所要時間で交通の便を比較する。
 - SNSの投稿数から話題性を分析する。

比率分析

- 数値の比率を求めて分析をします。
- 効率性や有効性などの、「質」の比較を行えます。
- 具体的には…
 - 所要時間に対する費用で交通手段を比較する。
 - 日本円と外貨の為替レートを見て景気を分析する。

ワーク



ワーク（1）

人口100万人の地域で、90%の住民がある病気の予防接種を受けました。

その病気に罹った4万人のうち、予防接種を受けた人は3万人、予防接種を受けていない人は1万人でした。

予防接種を受けた効果はあるのでしょうか？

条件付き割合で整理してみよう！



考えたことを、グループで共有しましょう。

Work Sheet

テーマ



個人ワーク

個人で考えてみましょう。



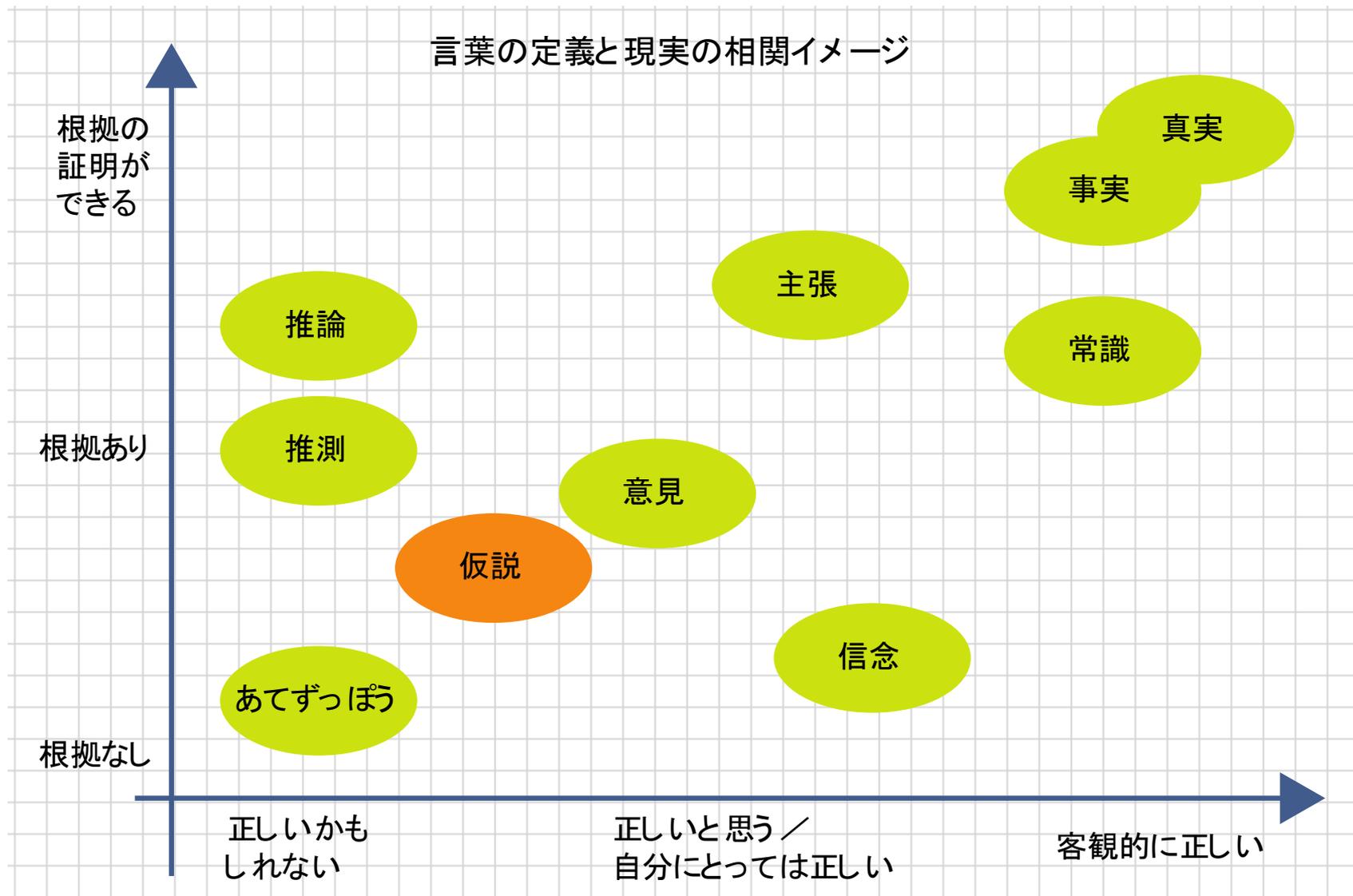
グループワーク

データの読み方演習①



仮説とは

仮説とは



仮説の定義

現象や法則を説明するために仮に立てる説。
実験・観察などによる検証を通じて、真偽を確認していく。
本来の意味（科学的定義）での仮説は、実験によってどのような結果が現れたら正しいと言えるのか、または間違っていると言えるのかが、あらかじめ明確になっていることが必要である。

データ分析における仮説

データ分析の世界では「分析結果（数値やグラフ）がそうなっている理由を説明する仮の理由」くらいの理解で問題ない。

その理由が正しいかどうか、検証できることが理想であるが、データが手に入らない・存在しないなど必ずしも検証が可能とは限らないことは覚えておく必要がある。

大雑把であっても、仮の理由を考えることで次の調査・分析の動機なり、実際の行動（分析）へとつながっていく。

ワーク



仮説作りの練習

近隣の観光名所について仮説を立ててみます。

① 近隣の観光名所をあげます。

② その観光名所に関する仮説を立てます。

「その観光名所でありそうなこと・おきてそうなことの予想」
くらいで大丈夫です。

ただし、可能な限り数値的な予想を行ってください。

〇時ごろは閑散としている< 〇時ごろは〇人くらいしかいない

③ その仮説の真偽を確かめる方法を考えます。

ワークの進め方(1)

Work Sheet

テーマ



後で見ても分かるように
テーマを書いておきましょう

個人の考えを書いておきましょう

- ①近所の観光名所
- ②その観光名所に関する仮説
- ③その仮説を確かめる方法

ワークの進め方(2)

Work Sheet

テーマ



グループで意見交換をしましょう。



仮説と検証方法をグループ内で共有します。
結論を1つにする必要はないので
様々な名所の様々な仮説を考えてください。

ワークの進め方(3)

グループごとの成果を全体に共有しましょう。

名所と仮説は、思いついたものを網羅的に伝えてください。

※他のグループの説明から新しい視点を見つけてください。



データの読み取りと次の調査

結果の読み取りのポイント

分析の結果（グラフなど）を見たら、最低限以下の点を考察する。

①何が読み取れるか

※事実と解釈・意見を混同しないように注意

②その原因は何と推測できるか（仮説）

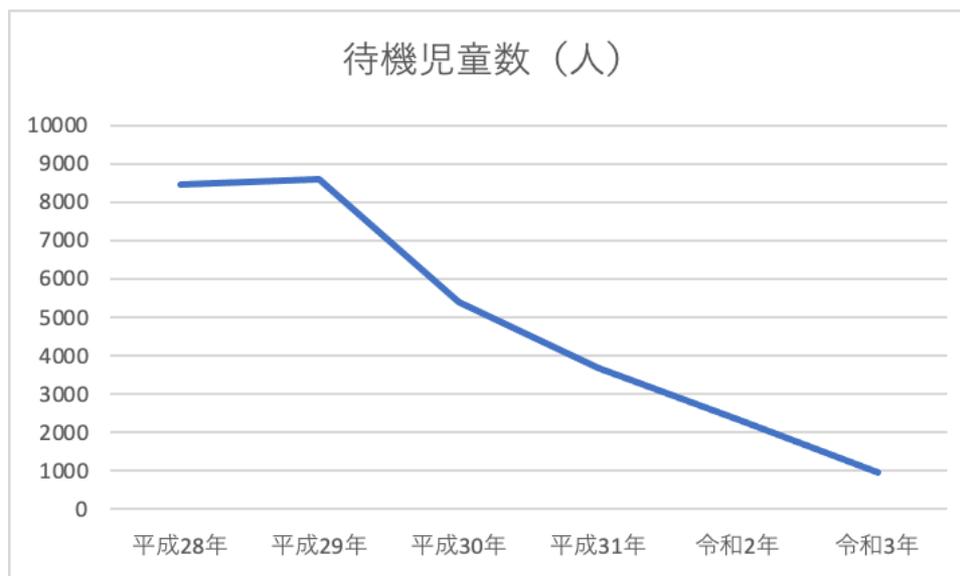
③仮説を確認するためには、何が必要か（今後の方針）

⇒どのようなデータで、どのような分析をするべきか

結果の読み取りの例

下記の「待機児童数」の場合

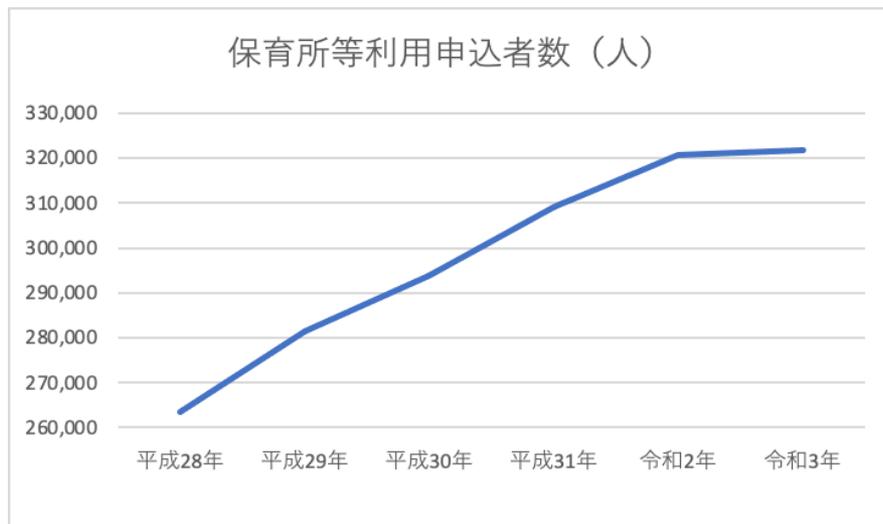
- ① 事実→平成29年から令和3年は減少している
 解釈→8000超から1000未満なので、“激減”だ
- ② 保育サービスの充実か？あるいは、少子化の影響か？
- ③ 保育サービス利用数や就学前児童数を調べる



出典：都内の保育サービスの状況について

実際に調べてみる①

保育サービス利用数を調べる⇒

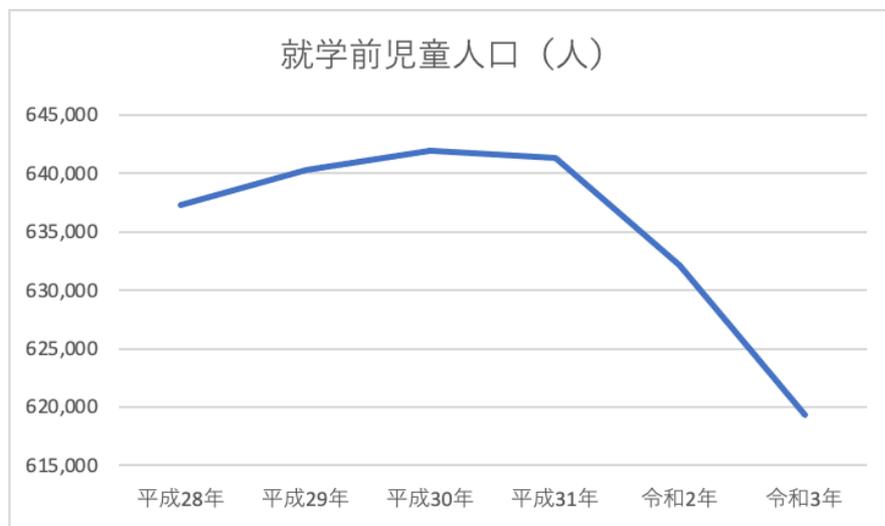


出典：都内の保育サービスの状況について

- サービス利用数は見つからなかった
- 利用申込者数があったので代用
- 増加傾向にある
- 申込者が増えて待機児童が減っているならサービス利用数が増えていると推測可能

実際に調べてみる②

就学前児童数を調べる⇒



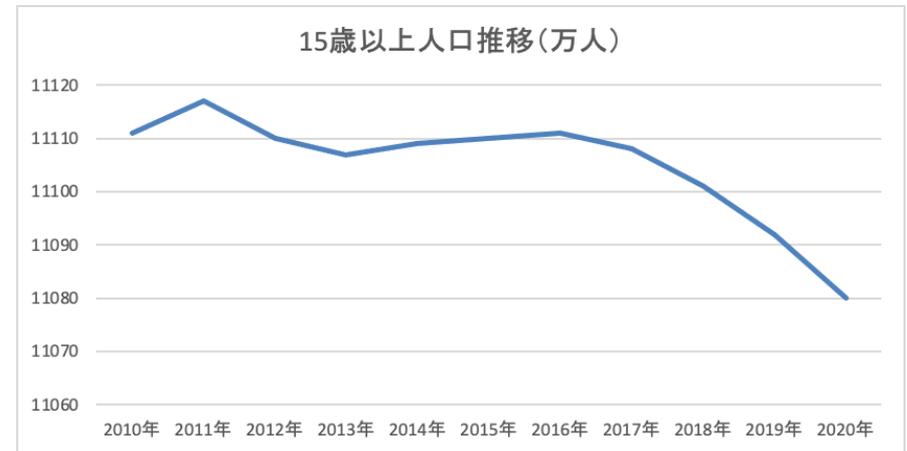
出典：都内の保育サービスの状況について

就学前児童数は減少傾向
→仮説と一致する

ワーク（読み取り練習）

データの読み取りと次の調査

読み取り練習



出典：労働力調査（基本集計） 2021年（令和3年）7月分結果

- ①何が読み取れるか
- ②その原因は何と推測できるか（仮説）
- ③その仮説を確認するためには、どんなデータが必要か

ワークの進め方（1）

Work Sheet

テーマ



個人ワーク



グループワーク

後で見ても分かるように
テーマを書いておきましょう

個人の考えを書いておきましょう

- ①何が読み取れるか
- ②その原因は何と推測できるか（仮説）
- ③その仮説を確認するためには、
どんなデータが必要か

ワークの進め方（3）

Work Sheet

テーマ



グループで意見交換をしましょう。



個人での読み取り・仮説・次の方針を
グループ内で意見交換しましょう。
グループで議論して、全員が合意する
結論を目指しましょう。

ワークの進め方（4）

グループごとの意見を全体に共有しましょう。

- ①何が読み取れたか
- ②その原因の仮説はなにか
- ③その仮説を確認するためには、どんなデータが必要か

※他のグループの説明から新しい視点を見つけてください。



データの読み取りと次の調査

※回答の一例です。



①何が読み取れるか

事実

2012年以降、就業者は増加しているが2020年には減少している

解釈

一時的な減少か、減少傾向の始まりかはまだ分からない

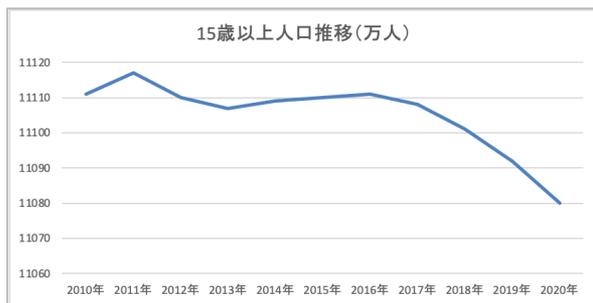
②その原因は何と推測できるか

15歳以上の人口が減少しているのに就業者が増えているのは、なぜ？（違和感）

→今まで働いていなかった世代が働き出しているのか？（仮説）

③その仮説を確認するためには、どんなデータが必要か

年齢別の就業者数、特に65歳以上の就業者数の推移を確認する



データの読み方演習②



ワーク(分析結果から次の調査を考える)

分析結果から次の調査を考える

データを自由に分析して、結果を読み取ってみます。

ファイル「演習データE3.xlsx」を開きましょう。

まずは、ここまで学習した方法を駆使して

データを様々な分析してみましよう。

ワークの進め方(1)

Work Sheet

テーマ



個人ワーク



グループワーク

後で見分分かるように
テーマを書いておきましょう

個人の考えを書いておきましょう

- ①何が読み取れるか
- ②その原因は何と推測できるか（仮説）
- ③その仮説を検証するためには、
どんなデータが必要か

ワークの進め方(3)

Work Sheet

テーマ



グループで意見交換をしましょう。



個人での読み取り・仮説・次の方針を
グループ内で意見交換しましょう。
グループで議論して、全員が合意する
結論を目指しましょう。

ワークの進め方(4)

グループごとの意見を全体に共有しましょう。

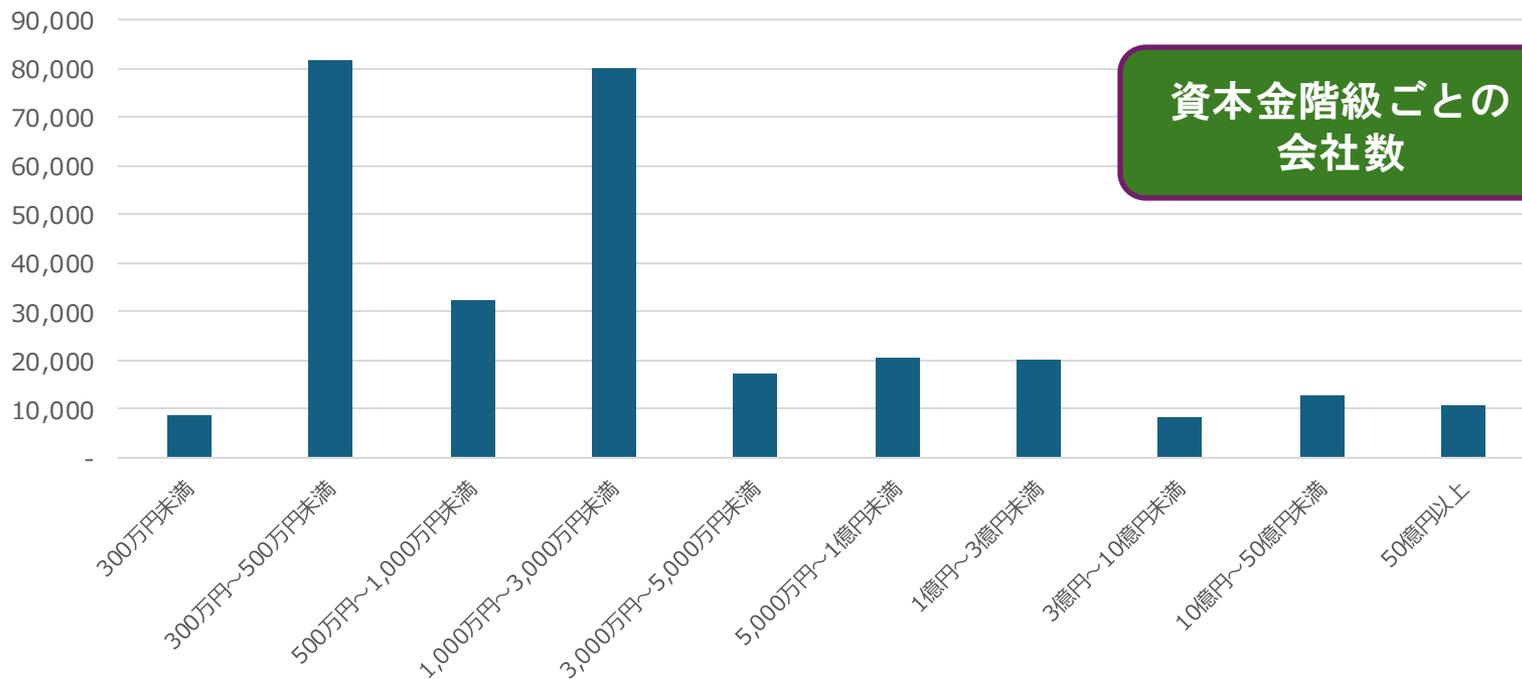
- ①何が読み取れたか
- ②その原因の仮説はなにか
- ③その仮説を検証するためには、どんなデータが必要か

※他のグループの説明から新しい視点を見つけてください。



分析結果から次の調査を考える

※分析・解釈・仮説の一例です。



①何が読み取れるか

事実

500万円～1,000万円未満が
その両隣より少なくなっている

解釈

この資本金帯だけ、他と比べて不自然に少なく見える

②その原因は何と推測できるか（仮説）

500万円～1,000万円未満という資本金帯は
家族経営・一人社長の会社には大き過ぎて
通常の観光系企業にとっては、小さすぎる

③その仮説を検証するためには

観光業の企業における家族経営・一人社長の
会社の割合を調べる

基本的な統計・データ分析スキル①



相関とクロス表・データの関係性

戦略／関係性のパターン

- データの種類は大きく分けると2種類
 - 量的変数
 - 質的変数
- 2種類の組み合わせで関係性のパターンが作られる。
 - パターンに応じて適切な方法を選択する。

関係性のパターン		データ2	
		量的変数	質的変数
データ1	量的変数	相関係数	相関比 (今回は対象外)
	質的変数	相関比 (今回は対象外)	クロス表

関係性のパターン (1)

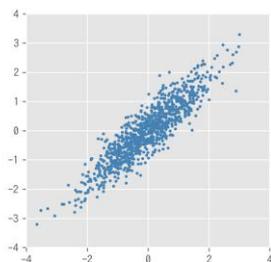
相関係数



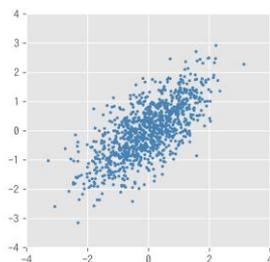
相関係数の基本

- 比較する2つの量で散布図を作り分布の形を調べる。
- 関係性には「大きさ」と「向き」が存在する。

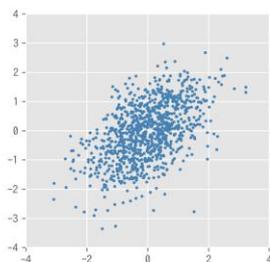
0.9



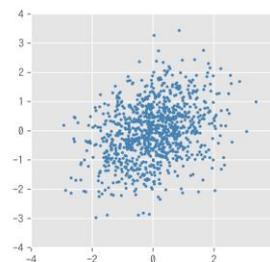
0.7



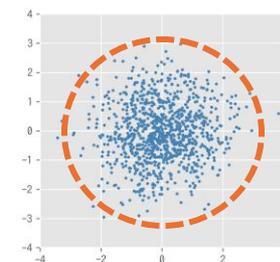
0.5



0.3



相関係数がゼロ
の場合は「独立」「無相関」と呼ぶ。
散布図は円のような分布になる。

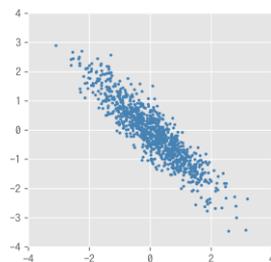


0.0

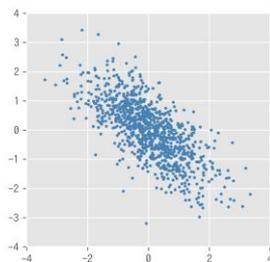
強い相関がある

弱い相関がある

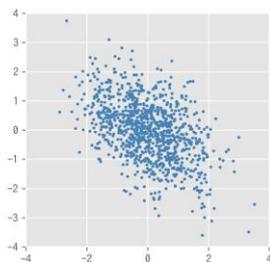
-0.9



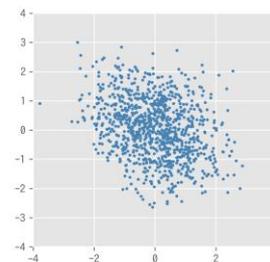
-0.7



-0.5

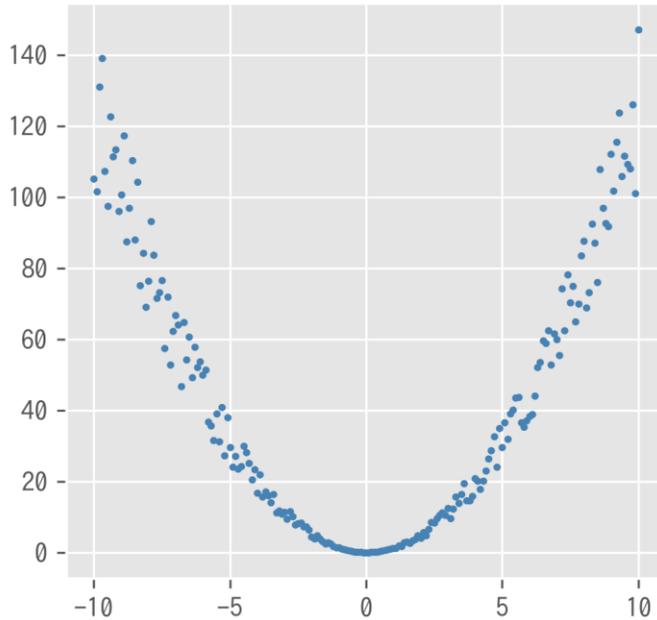


-0.3



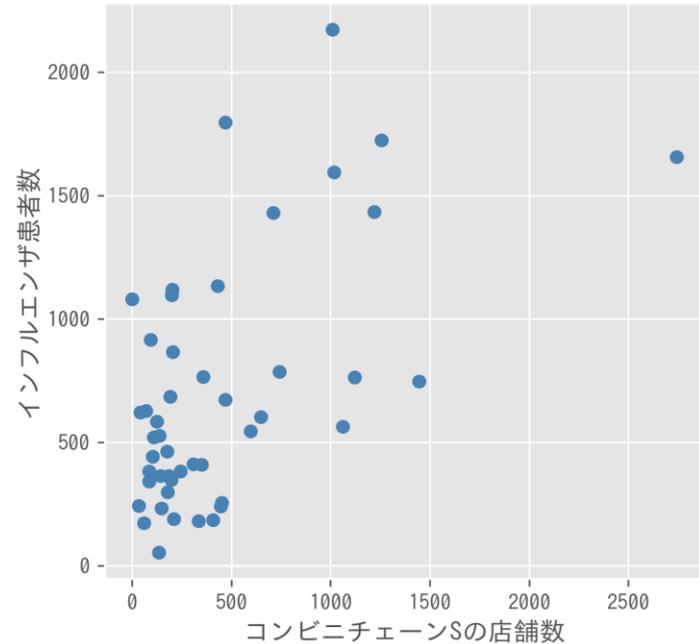
相関係数の注意点

● 「直線の関係」のみ有効。



相関係数: -0.004

● 関係性と「因果関係」は別問題。

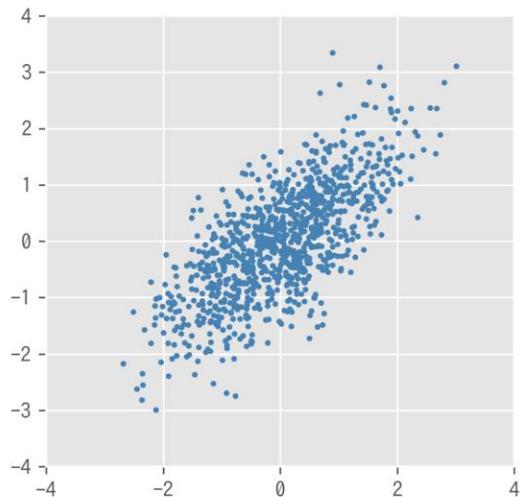


相関係数: 0.573

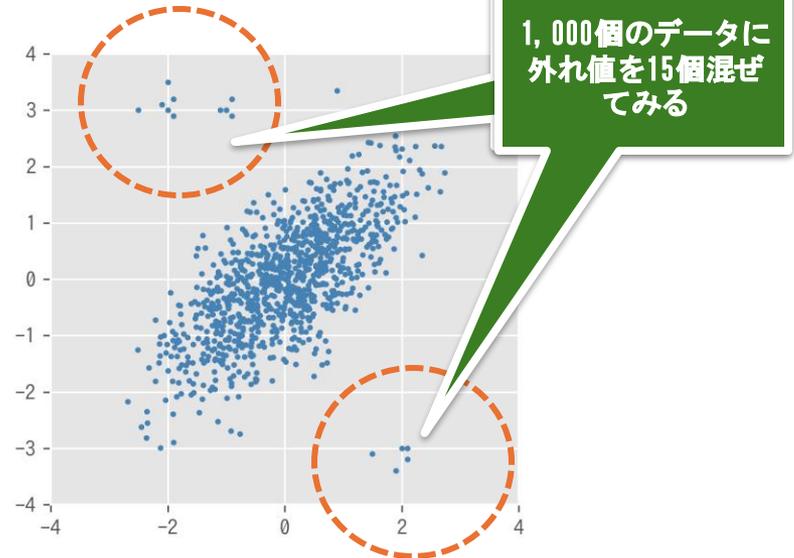
※コンビニチェーンの店舗数とインフルエンザ患者数には弱い相関がある。コンビニチェーンを減らせばインフルエンザ患者数が減るのではないか？

相関係数アドバンス

- ここまで取り扱った相関係数: Pearsonの積率相関
 - これは平均値と同じくデータの歪みに弱い。
- 中央値と同じように歪みに強い相関係数: Spearmanの順位相関
 - 数値のデータをランキングデータに変換して相関係数を算出する。
- 2種類の相関係数を計算し、結果が大きく異なるようであれば外れ値などが含まれている可能性がある。



Pearsonの積率相関: 0.709



Pearsonの積率相関: 0.572
Spearmanの順位相関: 0.622

実 習

相関係数



実習 (1)

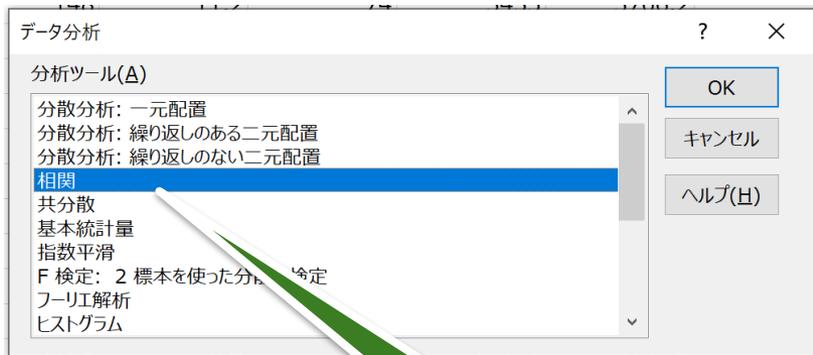
- ファイル「演習データD4_1.xlsx」シート「相関係数」

The screenshot shows an Excel spreadsheet with the following data:

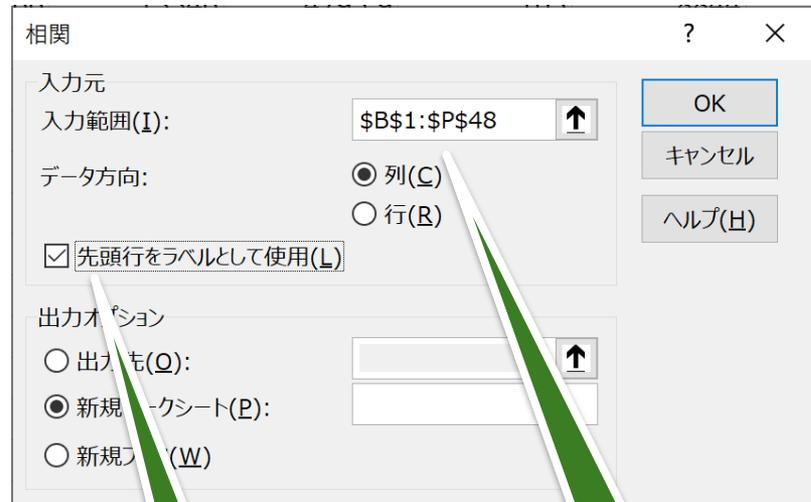
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
1		人口 (人)	面積 (離島などインフラ整備不全)	72797											
2	北海道	5506419	83424	2171	1010	9.3	66	28160	4102.3	374	26330	11285	38190	59432	酒類
3	青森県	1373339	9645	914	93	11	76	4608	3586.9	85	5723	2335	9126	17112	
4	岩手県	1330147	15275	231	148	11.2	74	3435	3706.2	77	5398	2003	9231	15969	
5	宮城県	2348165	6859	184	407	13.5	68	14929	4452.4	138	11985	4162	18949	22214	
6	秋田県	1085997	11638	521	108	12.5	74	2428	3628.5	55	3865	1485	6177	14824	
7	山形県	1168924	6652	363	183	12.7	71	3975	3759.9	48	4741	1675	8159	15029	
8	福島県	2029064	13784	254	450	14.2	69	11127	4101.5	105	9069	3246	14546	23611	
9	茨城県	2969770	6097	602	648	14.8	74	24809	4862.9	178	14323	5047	22358	30368	
10	栃木県	2007683	6408	241	445	14.8	68	12767	4684.1	126	10152	3626	15588	20591	
11	群馬県	2008068	6362	672	470	15.6	63	13105	4553.8	123	9031	3511	14732	21661	
12	埼玉県	7194556	3768	1433	1219	15.9	63	63383	4768.9	323	36279	13138	57470	60264	
13	千葉県	6216289	5083	763	1120	16.8	67	52974	4877.1	356	31375	11290	48343	53603	
14	東京都	13159388	2106	1656	2744	16.4	69	125251	6235.4	750	88067	24855	109986	110507	
15	神奈川県	9048331	2416	747	1445	16.9	70	53628	5442.6	429	49769	16587	74320	72970	
16	新潟県	2374450	10364	1133	429	14.5	74	12757	4026.8	113	9965	3276	17066	28383	
17	富山県	1093247	2046	53	134	15.2	72	5330	4350.3	52	4739	1562	7722	12547	
18	石川県	1169788	4186	527	138	15.7	70	5393	4363.9	73	5340	1812	9449	12223	
19	福井県	806314	4190	626	69	15.6	72	3231	4271.6	47	3744	1181	6461	8764	
20	山梨県	863075	4201	189	210	15.7	63	4617	4533.2	81	3961	1473	6198	9441	
21	長野県	2152449	13104	1795	468	13.1	73	9535	4458	183	9933	3420	16326	24303	
22	岐阜県	2080773	9769	348	197	16.9	65	14897	4376.2	118	9492	3227	16000	21518	
23	静岡県	3765007	7254	786	741	17.6	69	20869	4654.9	227	18463	6732	30260	38393	
24	愛知県	7410719	5123	563	1061	17	65	65511	5406.3	344	42302	13077	66825	62395	
25	三重県	1854724	5759	462	175	16.9	66	13346	4795.9	105	8844	3281	14514	19690	
26	滋賀県	1410777	3767	383	242	15.8	73	8737	4766	73	7465	2329	13015	12233	
27	京都府	2636092	4612	765	357	17.1	66	18603	4873.9	167	12746	4581	20106	25332	

実習（2）

- 分析ツールを使って計算する。



相関を選択



先頭にラベルがある場合はチェック

計算するセルの範囲を指定

実習 (3)

- 相関行列が出力される。

Excel screenshot showing a correlation matrix. The ribbon is set to 'データ' (Data). The matrix is displayed in the following table:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1		人口 (人)	島など除く)	レエンザ患者イ	レブン店舗	平均気温(°C)	平均湿度(%)	数 (認知件)	マン平均年	き家数(千戸)	婚姻件数(件)	離婚件数(件)	出生数(人)	死亡数(人)	消費量(キロ)	喫茶店件数(件)
2	人口 (人)	1														
3	面積 (離島)	0.082813	1													
4	インフルエ	0.593126	0.409491	1												
5	セブンイレ	0.969526	0.102835	0.573348	1											
6	年平均気温(0.023967	-0.53115	-0.00728	-0.03537	1										
7	年平均湿度(-0.42373	-0.11999	-0.1608	-0.37945	0.020315	1									
8	刑法犯総数	0.965738	-0.03312	0.568421	0.924129	0.091996	-0.43538	1								
9	サラリーマ	0.782581	-0.16918	0.22686	0.777851	0.089873	0.785145	-0.53993	1							
10	空き家数(千	0.975876	0.140095	0.652743	0.951414	0.012415	-0.42873	0.965409	0.768178	1						
11	婚姻件数(件	0.990102	0.033063	0.564211	0.973543	0.057698	-0.38084	0.965205	0.789561	0.966193	1					
12	離婚件数(件	0.996218	0.102572	0.619094	0.96336	0.050889	-0.41798	0.968868	0.76273	0.98298	0.987788	1				
13	出生数(人)	0.995811	0.031598	0.57906	0.962216	0.085472	-0.40689	0.966432	0.784423	0.966845	0.991997	0.993132	1			
14	死亡数(人)	0.992524	0.159007	0.639921	0.965544	-0.021	-0.42292	0.957768	0.752437	0.985239	0.976079	0.990845	0.982	1		
15	酒類の総消	0.969868	0.078435	0.603849	0.962352	0.015246	-0.35681	0.962582	0.745798	0.968822	0.986003	0.972308	0.965867	0.99	1	
16	喫茶店件数(0.936473	0.008582	0.524722	0.901909	0.042919	-0.4368	0.940206	0.818856	0.926992	0.943009	0.928168	0.94601	0.9	0.925606	1

出力は下半分のみ出る。(リーグ戦と同じなので)

実習 (4)

- 見やすい形に整形する。

Excelのスクリーンショット。相関係数マトリックスが表示されており、色（オレンジ系）で相関の強さを示しています。オレンジが濃いほど相関が強いことを示しています。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1																
2	人口 (人)	1.000														
3	面積 (離島など除く) (km2)	0.083	1.000													
4	インフルエンザ患者数(人)	0.593	0.409	1.000												
5	セブンイレブン店舗数(件)	0.970	0.103	0.573	1.000											
6	年平均気温(℃)	0.024	-0.531	-0.007	-0.035	1.000										
7	年平均湿度(%)	-0.424	-0.120	-0.161	-0.379	0.020	1.000									
8	刑法犯総数 (認知件数) (件)	0.966	-0.033	0.568	0.924	0.092	-0.435	1.000								
9	サラリーマン平均年収(千円)	0.783	-0.169	0.227	0.778	0.090	-0.540	0.785	1.000							
10	空き家数(千戸)	0.976	0.140	0.653	0.951	0.012	-0.429	0.965	0.768	1.000						
11	婚姻件数(件)	0.990	0.033	0.564	0.974	0.058	-0.381	0.965	0.790	0.966	1.000					
12	離婚件数(件)	0.996	0.103	0.619	0.963	0.051	-0.418	0.969	0.763	0.983	0.988	1.000				
13	出生数(人)	0.996	0.032	0.579	0.962	0.085	-0.407	0.966	0.784	0.967	0.992	0.993	1.000			
14	死亡数(人)	0.993	0.159	0.640	0.966	-0.021	-0.423	0.958	0.752	0.985	0.976	0.991	0.983	1.000		
15	酒類の総消費量(キロリットル)	0.970	0.078	0.604	0.962	0.015	-0.357	0.963	0.746	0.969	0.986	0.972	0.966	0.967	1.000	
16	喫茶店件数(件)	0.936	0.009	0.525	0.902	0.043	-0.437	0.940	0.819	0.927	0.943	0.928	0.946	0.930	0.926	1.000

相関の強さに応じて網掛けの色を変えると全体の傾向がわかりやすくなる。

実習（5）

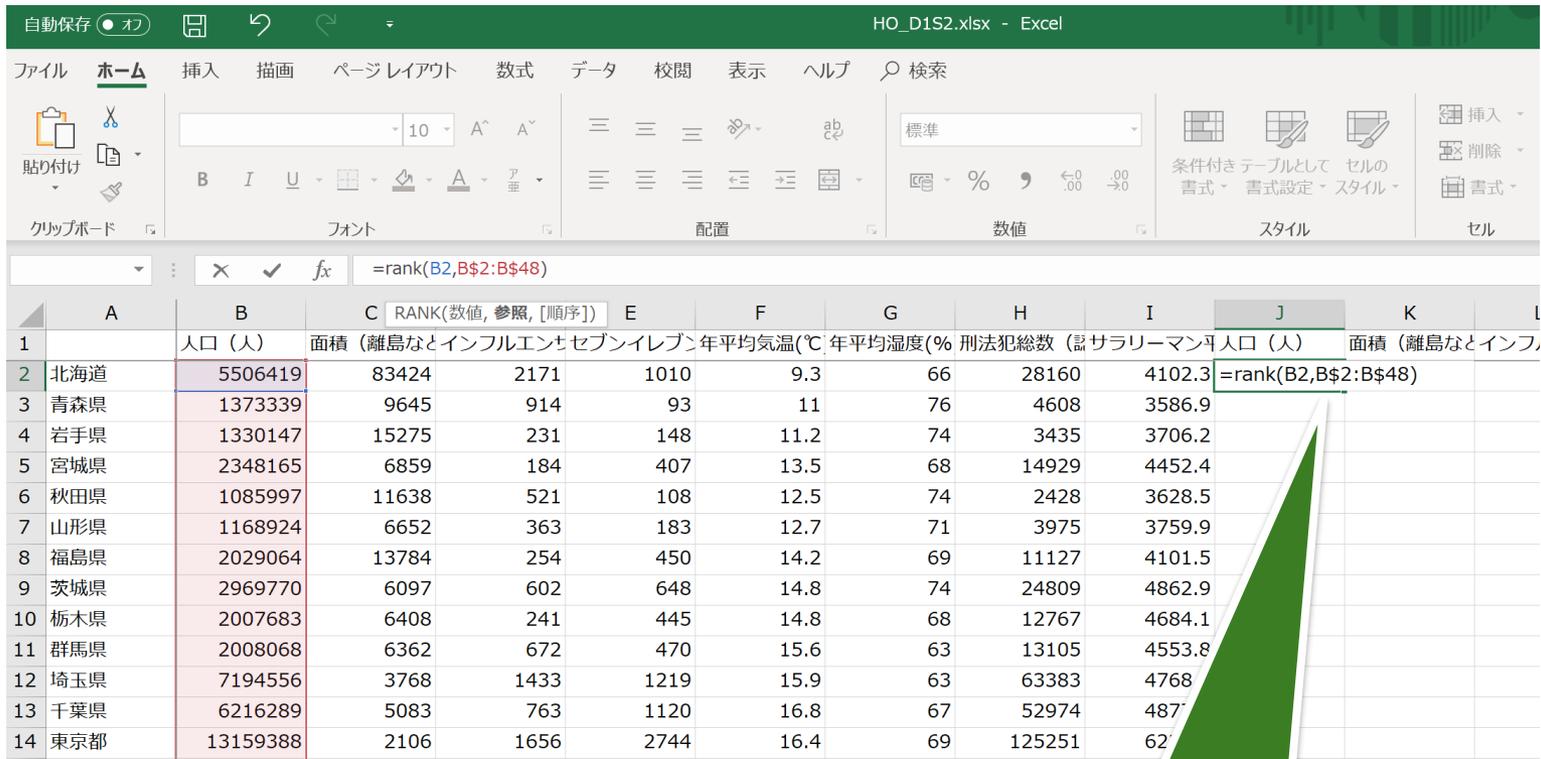
- スピアマンの順位相関係数を算出する。シート「順位相関」

The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		人口 (人)	面積 (離島などインフルエンタセブンイレブ			年平均気温(°C)	年平均湿度(%)	刑法犯総数 (罰サラリーマン	人口 (人)		面積 (離島などインフルエンタセブンイレブ			年平均気温(°C)
2	北海道	5506419	83424	2171	1010	9.3	66	28160	4102.3					
3	青森県	1373339	9645	914	93	11	76	4608	3586.9					
4	岩手県	1330147	15275	231	148	11.2	74	3435	3706.2					
5	宮城県	2348165	6859	184	407	13.5	68	14929	4452.4					
6	秋田県	1085997	11638	521	108	12.5	74	2428	3628.5					
7	山形県	1168924	6652	363	183	12.7	71	3975	3759.9					
8	福島県	2029064	13784	254	450	14.2	69	11127	4101.5					
9	茨城県	2969770	6097	602	648	14.8	74	24809	4862.9					
10	栃木県	2007683	6408	241	445	14.8	68	12767	4684.1					
11	群馬県	2008068	6362	672	470	15.6	63	13105	4553.8					
12	埼玉県	7194556	3768	1433	1219	15.9	63	63383	4768.9					
13	千葉県	6216289	5083	763	1120	16.8	67	52974	4877.1					
14	東京都	13159388	2106	1656	2744	16.4	69	125251	6235.4					
15	神奈川県	9048331	2416	747	1445	16.9	70	53628	5442.6					
16	新潟県	2374450	10364	1133	429	14.5	74	12757	4026.8					
17	富山県	1093247	2046	53	134	15.2	72	5330	4350.3					
18	石川県	1169788	4186	527	138	15.7	70	5393	4363.9					
19	福井県	806314	4190	626	69	15.6	72	3231	4271.6					
20	山梨県	863075	4201	189	210	15.7	63	4617	4533.2					
21	長野県	2152449	13104	1795	468	13.1	73	9535	4458					
22	岐阜県	2080773	9769	348	197	16.9	65	14897	4376.2					
23	静岡県	3765007	7254	786	741	17.6	69	20869	4654.9					
24	愛知県	7410719	5123	563	1061	17	65	65511	5406.3					
25	三重県	1854724	5759	462	175	16.9	66	13346	4795.9					
26	滋賀県	1410777	3767	383	242	15.8	73	8737	4766					
27	京都府	2636002	4612	765	357	17.1	66	18603	4873.9					

実習（6）

- Rank関数を使って数字を順位データに変える。



Excelの画面キャプチャ。ワークシートには人口データが並び、J2セルに順位計算の式が入力されている。

	A	B	C	D	E	F	G	H	I	J	K
1		人口（人）	面積（離島などインフラ整備セブナイレブ			年平均気温(℃	年平均湿度(%	刑法犯総数（誌	サラリーマン	人口（人）	面積（離島などインフ
2	北海道	5506419	83424	2171	1010	9.3	66	28160	4102.3	=rank(B2,B\$2:B\$48)	
3	青森県	1373339	9645	914	93	11	76	4608	3586.9		
4	岩手県	1330147	15275	231	148	11.2	74	3435	3706.2		
5	宮城県	2348165	6859	184	407	13.5	68	14929	4452.4		
6	秋田県	1085997	11638	521	108	12.5	74	2428	3628.5		
7	山形県	1168924	6652	363	183	12.7	71	3975	3759.9		
8	福島県	2029064	13784	254	450	14.2	69	11127	4101.5		
9	茨城県	2969770	6097	602	648	14.8	74	24809	4862.9		
10	栃木県	2007683	6408	241	445	14.8	68	12767	4684.1		
11	群馬県	2008068	6362	672	470	15.6	63	13105	4553.8		
12	埼玉県	7194556	3768	1433	1219	15.9	63	63383	4768		
13	千葉県	6216289	5083	763	1120	16.8	67	52974	487		
14	東京都	13159388	2106	1656	2744	16.4	69	125251	62		

順位を求める関数
=rank(B2, B\$2:B\$48)
絶対参照を使って列の範囲を固定する。

実習 (7)

- 数式をコピーする。その後は分析ツールの「相関」で同様に計算できる。

自動保存 (オ) | HO_D1S2.xlsx - Excel | 森田 匠

ファイル ホーム 挿入 描画 ページレイアウト 数式 データ 校閲 表示 ヘルプ 検索

Meiryoo 10 A A

B I U 標準

条件付き書式 テーブルとしてセルの書式設定 スタイル

挿入 削除 書式

Σ 並べ替えとフィルター 検索と選択 アイデア

共有 コメント

Q26 =RANK(I26,I\$2:I\$48)

	A	E	F	G	H	I	J	K	L	M	N	O	P	Q
1		セブンイレブン	年平均気温(℃)	年平均湿度(%)	刑法犯総数 (※サラリーマン)	人口 (人)		面積 (離島など)	インフルエンザ	セブンイレブン	年平均気温(℃)	年平均湿度(%)	刑法犯総数 (※サラリーマン)	年平均
2	北海道	1010	9.3	66	28160	4102.3	8	1	1	8	47	37	9	30
3	青森県	93	11	76	4608	3586.7	31	8	12	40	46	4	37	46
4	岩手県	148	11.2	74	3435	3777.2	32	2	42	33	45	10	43	43
5	宮城県	407	13.5	68	14929	32.4	15	16	44	18	41	33	14	20
6	秋田県	108	12.5	74	2428	4628.5	38	5	27	38	44	10	47	45
7	山形県	183	12.7	71	3975	3759.9	35	19	35	30	43	23	40	41
8	福島県	450	14.2	69	1111	4101.5	18	3	39	15	40	29	20	31
9	茨城県	648	14.8	74	1111	4862.9	11	23	22	11	37	10	10	7
10	栃木県	445	14.8	68	1111	4684.1	20	20	41	16	37	33	18	12
11	群馬県	470	15.6	63	1111	4553.8	19	21	19	13	34	45	17	16
12	埼玉県	1219	15.9	67	1111	4768.9	5	37	6	4	29	45	4	10
13	千葉県	1120	16.8	68	1111	4877.1	6	28	16	5	23	36	6	5
14	東京都	2744	16.4	67	1111	6235.4	1	44	4	1	26	29	1	1
15	神奈川県	1445	16.9	67	1111	5442.6	2	42	17	2	20	26	5	2
16	新潟県	429	14.7	74	1111	4026.8	14	6	8	17	39	10	19	34
17	富山県	134	14.7	72	1111	4350.3	37	45	47	36	36	19	33	25
18	石川県	138	14.7	70	1111	4363.9	34	34	26	35	32	26	32	24
19	福井県	69	14.7	72	1111	4271.6	43	33	20	43	34	19	44	29
20	山梨県	210	14.7	63	1111	4533.2	41	32	43	24	32	45	36	18
21	長野県	199	14.7	73	1111	4458	16	4	2	14	42	15	22	19
22	岐阜県	199	14.7	65	1111	4376.2	17	7	36	28	20	40	15	23
				69	1111	4654.9	10	13	14	9	12	29	11	13
				65	1111	5406.3	4	26	24	6	19	40	3	3
				66	1111	4795.9	22	24	28	32	20	37	16	9
				73	1111	4766	28	38	32	23	31	15	25	11
				66	1111	4873.9	13	31	15	19	18	37	12	6

対象とする列に同じように順位を算出。

関係性のパターン (2)

クロス表



クロス表

- クロス表とは、2つ以上の属性を行・列として集計したもの
 - Excelでは「ピボットテーブル」と呼ばれる
- データ分析では、このクロス表をできるだけたくさん作り、確認する事が重要になる。
 - このプロセスを通して、データに関する感覚が身につき、分析のあたりをつける事ができる。
- クロス表を作ったら、行と列で割合を計算する。
 - 割合にすることで、異なる行・列を比較することができる。

クロス表データの前処理

- クロス表は「質的変数」x「質的変数」の軸を用意する。
- 刻み幅が細かいと複雑になるため、情報を減らす。

カテゴリ	年齢	金額
物販	53	2820
物販	46	3150
旅行	51	95330
物販	34	5110
レッスン(学び)	43	21100
物販	68	2040
旅行	53	83220
旅行	55	90420
旅行	42	79980
物販	43	3330
物販	54	3010
物販	45	2760
物販	45	3380
レジャー	31	23220
グルメ(飲食)	42	5700
物販	24	1050
物販	50	3500
ビューティー(美容)	49	8420
物販	40	3280
グルメ(飲食)	54	4990
物販	61	1810
レジャー	56	28390
ビューティー(美容)	55	9320



カテゴリ	年代	金額
物販	50代	1
物販	40代	1
旅行	50代	1
物販	30代	1
レッスン(学び)	40代	1
物販	60代	1
旅行	50代	1
旅行	50代	1
旅行	40代	1
物販	40代	1
物販	50代	1
物販	40代	1
物販	40代	1
レジャー	30代	1
グルメ(飲食)	40代	1
物販	20代	1
物販	50代	1
ビューティー(美容)	40代	1
物販	40代	1
グルメ(飲食)	50代	1
物販	60代	1
レジャー	50代	1
ビューティー(美容)	50代	1

クロス表の結果（１）

- クロス表を作成したら、「何が言えるか」を常に検討する。

	グルメ (飲食)	ビューティー (美容)	レジャー	レッスン (学び)	医療 (美容・歯科)	物販	旅行	合計
20代	271	299	139	86	9	588	193	1,585
30代	1,344	1,202	806	424	92	2,810	1,099	7,777
40代	3,596	2,292	2,135	770	136	6,716	2,862	18,507
50代	4,718	2,244	2,618	773	194	8,047	4,202	22,796
60代	2,679	820	1,525	223	95	4,409	2,729	12,480
70代	500	92	301	17	18	789	569	2,286
合計	13,108	6,949	7,524	2,293	544	23,359	11,654	65,431

- ポイントとなる箇所を強調する。

	グルメ (飲食)	ビューティー (美容)	レジャー	レッスン (学び)	医療 (美容・歯科)	物販	旅行	合計
20代	271	299	139	86	9	588	193	1,585
30代	1,344	1,202	806	424	92	2,810	1,099	7,777
40代	3,596	2,292	2,135	770	136	6,716	2,862	18,507
50代	4,718	2,244	2,618	773	194	8,047	4,202	22,796
60代	2,679	820	1,525	223	95	4,409	2,729	12,480
70代	500	92	301	17	18	789	569	2,286
合計	13,108	6,949	7,524	2,293	544	23,359	11,654	65,431

クロス表の結果（２）

- クロス表から比率を計算する。

	グルメ (飲食)	ビューティー (美容)	レジャー	レッスン (学び)	医療 (美容・歯科)	物販	旅行	合計
20代	A 271	299	139	86	9	588	193	B 1,585
30代	1,344	1,202	806	424	92	2,810	1,099	7,777
40代	3,596	2,292	2,135	770	136	6,716	2,862	18,507
50代	4,718	2,244	2,618	773	194	8,047	4,202	22,796
60代	2,679	820	1,525	223	95	4,409	2,729	12,480
70代	500	92	301	17	18	789	569	2,286
合計	13,108	6,949	7,524	2,293	544	23,359	11,654	65,431

- 各世代における相対的な比率が求まる。

合計部分は100%になる

	グルメ (飲食)	ビューティー (美容)	レジャー	レッスン (学び)	医療 (美容・歯科)	物販	旅行	合計
20代	17.1%	18.9%	8.8%	5.4%	0.6%	37.1%	12.2%	100.0%
30代	17.3%	15.5%	10.4%	5.5%	1.2%	36.1%	14.1%	100.0%
40代	19.4%	12.4%	11.5%	4.2%	0.7%	36.3%	15.5%	100.0%
50代	20.7%	9.8%	11.5%	3.4%	0.9%	35.3%	18.4%	100.0%
60代	21.5%	6.6%	12.2%	1.8%	0.8%	35.3%	21.9%	100.0%
70代	21.9%	4.0%	13.2%	0.7%	0.8%	34.5%	24.9%	100.0%
合計	20.0%	10.6%	11.5%	3.5%	0.8%	35.7%	17.8%	100.0%

$A \div B$

合計部分は100%にならない

クロス表の結果（3）

- クロス表から比率を計算する。

	グルメ (飲食)	ビューティー (美容)	レジャー	レッスン (学び)	医療 (美容・歯科)	物販	旅行	合計
20代	A 271	299	139	86	9	588	193	1,585
30代	1,344	1,202	806	424	92	2,810	1,099	7,777
40代	3,596	2,292	2,135	770	136	6,716	2,862	18,507
50代	4,718	2,244	2,618	773	194	8,047	4,202	22,796
60代	2,679	820	1,525	223	95	4,409	2,729	12,480
70代	500	92	301	17	18	789	569	2,286
合計	C 13,108	6,949	7,524	2,293	544	23,359	11,654	65,431

- 各カテゴリにおける相対的な比率が求まる。

合計部分は100%にならない

	グルメ (飲食)	ビューティー (美容)	レジャー	レッスン (学び)	医療 (美容・歯科)	物販	旅行	合計
20代	2.1%	4.3%	1.8%	3.8%	1.7%	2.5%	1.7%	2.4%
30代	10.3%	17.3%	10.7%	18.5%	16.9%	12.0%	9.4%	11.9%
40代	27.4%	33.0%	28.4%	33.6%	25.0%	28.8%	24.6%	28.3%
50代	36.0%	32.3%	34.8%	33.7%	35.7%	34.4%	36.1%	34.8%
60代	20.4%	11.8%	20.3%	9.7%	17.5%	18.9%	23.4%	19.1%
70代	3.8%	1.3%	4.0%	0.7%	3.3%	3.4%	4.9%	3.5%
合計	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

$A \div C$

合計部分は100%になる

リフト値

- 先ほど計算した比率と反対の方向の比率を計算する。
- 計算結果は一致し、その結果を「リフト値」と呼ぶ。
- リフト値は、全体を1.0とした場合の各セグメントの起こりやすさを表す。

	グルメ (飲食)	ビューティー (美容)	レジャー	レッスン (学び)	医療 (美容・歯科)	物販	旅行	合計
20代	17.1%	18.9%	8.8%	5.4%	0.6%	37.1%	12.2%	100.0%
30代	17.3%	15.5%	10.4%	5.5%	1.2%	36.1%	14.1%	100.0%
40代	18.7%	12.4%	11.5%	4.2%	0.7%	36.3%	15.5%	100.0%
50代	20.7%	9.8%	11.5%	3.4%	0.9%	35.3%	18.4%	100.0%
60代	21.5%	6.6%	12.2%	1.8%	0.8%	35.3%	21.9%	100.0%
70代	21.9%	4.0%	13.2%	0.7%	0.8%	34.5%	24.9%	100.0%
合計	20.0%	10.6%	11.5%	3.5%	0.8%	35.7%	17.8%	100.0%

	グルメ (飲食)	ビューティー (美容)	レジャー	レッスン (学び)	医療 (美容・歯科)	物販	旅行	合計
20代	2.1%	4.3%	1.8%	3.8%	1.7%	2.5%	1.7%	2.4%
30代	10.3%	17.3%	10.7%	18.5%	16.9%	12.0%	9.4%	11.9%
40代	27.6%	33.0%	28.4%	33.6%	25.0%	28.8%	24.6%	28.3%
50代	36.0%	32.3%	34.8%	33.7%	35.7%	34.4%	36.1%	34.8%
60代	20.4%	11.8%	20.3%	9.7%	17.5%	18.9%	23.4%	19.1%
70代	3.8%	1.3%	4.0%	0.7%	3.3%	3.4%	4.9%	3.5%
合計	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

$$A \div B \div C$$

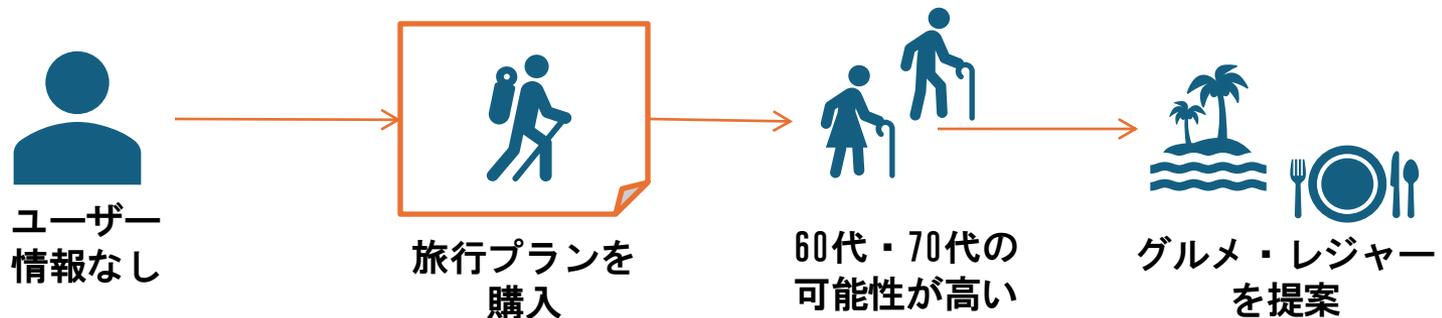
$$A \div C \div B$$

	グルメ (飲食)	ビューティー (美容)	レジャー	レッスン (学び)	医療 (美容・歯科)	物販	旅行	合計
20代		1.78	0.76	1.55	0.68	1.04	0.68	1.00
30代	0.86	1.46	0.90	1.56	1.42	1.01	0.79	1.00
40代	0.97	1.17	1.00	1.19	0.88	1.02	0.87	1.00
50代	1.03	0.93	1.00	0.97	1.02	0.99	1.03	1.00
60代	1.07	0.62	1.06	0.51	0.92	0.99	1.23	1.00
70代	1.09	0.38	1.15	0.21	0.95	0.97	1.40	1.00
合計	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

「20代はビューティー（美容）の購入可能性が1.78倍」→「ユーザーが20代と判明したらビューティー（美容）の広告を出す」

リフト値の活用事例

- リフト値は行と列の一方の情報が出た場合に他方の可能性がどの程度高いかを定量的に比較することが可能となる。
- 全体のデータが増えることでリフト値の精度が高まり、個人のデータが増えることでより属性予測が正確になる。



実習

クロス表



実習（1）

- ファイル「演習データD4_2.xlsx」シート「クロス表」

自動保存 ● オフ HO_D1S2.xlsx - Excel

ファイル ホーム 挿入 描画 ページレイアウト 数式 データ 校閲 表示 ヘルプ 検索

貼り付け Meiryo 10 A A 標準 条件付きテーブルとして セルの書式 書式設定 スタイル

Q22 fx

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	カテゴリ	年齢	金額										
2	レジャー	53	38920										
3	ビューティー	21	5210										
4	レジャー	44	23070										
5	物販	27	1210										
6	物販	39	6400										
7	物販	55	2800										
8	旅行	61	65350										
9	グルメ (飲食)	59	7960										
10	物販	57	3510										
11	グルメ (飲食)	36	4330										
12	グルメ (飲食)	51	3330										
13	物販	31	4140										
14	物販	21	640										
15	レジャー	54	6390										
16	旅行	26	18720										
17	物販	56	4270										
18	グルメ (飲食)	43	5470										
19	レッスン (習)	37	20470										
20	グルメ (飲食)	35	3090										
21	物販	61	2580										
22	ビューティー	50	8470										
23	物販	41	2390										
24	グルメ (飲食)	37	2880										
25	ビューティー	31	13180										
26	グルメ (飲食)	39	2900										
27	レッスン (習)	61	21540										

相関係数 順位相関 外れ値処理 クロス表

実習（2）

- 情報を加工する
 - クロス表は刻み幅が細かいと非常に見にくい表になってしまう。

Excel screenshot showing a formula bar with the formula `=floor.math(B2,10)&"代"`. The table below shows the data being processed:

	A	B	C	D	E	F	G
1	カテゴリ	年齢	金額	年代	購入有無		
2	レジャー	53	38920	<code>=floor.math(B2,10)&"代"</code>			
3	ビューティー	21	5210				
4	レジャー	44	23070				
5	物販	27	1210				
6	物販	39	6400				
7	物販	55	2800				
8	旅行	61	65350				
9	グルメ（飲	59	7960				
10	物販	57	3510				
11	グルメ（飲	36	4330				
12	グルメ（飲	51	3330				
13	物販	31	4140				

年齢を年代に変換
`=floor.math(B2,10)&"代"`

Excel screenshot showing a formula bar with the formula `=floor.math(B2,10)&"代"`. The table below shows the data being processed:

	A	B	C	D	E	F	G
1	カテゴリ	年齢	金額	年代	購入有無		
2	レジャー	53	38920	50代	1		
3	ビューティー	21	5210	20代	1		
4	レジャー	44	23070	40代	1		
5	物販	27	1210	20代	1		
6	物販	39	6400	30代	1		
7	物販	55	2800	50代	1		
8	旅行	61	65350	60代	1		
9	グルメ（飲	59	7960	50代	1		
10	物販	57	3510	50代	1		
11	グルメ（飲	36	4330	30代	1		
12	グルメ（飲	51	3330	50代	1		
13	物販	31	4140	30代	1		

1を入れる

実習（3）

- カテゴリのセルをコピーして挿入する。

The screenshot shows the Microsoft Excel interface. The ribbon is set to 'ホーム' (Home). The '貼り付け' (Paste) icon is selected. The font settings are 'Meiryo', size 10. The table below has the following data:

	A	B	C	D	E	F	G
1	カテゴリ	年齢	金額	カテゴリ	年代	購入有無	
2	レジャー	53	38920	レジャー	50代	1	
3	ビューティ	21	5210	ビューティ	20代	1	
4	レジャー	44	23070	レジャー	40代	1	
5	物販	27	1210	物販	20代	1	
6	物販	39	6400	物販	30代	1	
7	物販	55	2800	物販	50代	1	
8	旅行	61	65350	旅行	60代	1	
9	グルメ（飲	59	7960	グルメ（飲	50代	1	
10	物販	57	3510	物販	50代	1	
11	グルメ（飲	36	4330	グルメ（飲	30代	1	
12	グルメ（飲	51	3330	グルメ（飲	50代	1	
13	物販	31	4140	物販	30代	1	

実習（４）

- ピボットテーブルを選択して必要な情報を入れる

自動保存 (丸) 保存 復元

HO_D1S2.xlsx - Excel

ファイル ホーム **挿入** 描画 ページレイアウト 数式 データ 校閲 表示 ヘルプ 検索

ピボット おすすめ テーブル
テーブル ピボットテーブル

アドインを入手
個人用アドイン

おすすめ グラフ

マップ

ピボット

テーブル

アドイン

グラフ

K4

	A	B	C	D	E	F	G	H	I	J	K	L
1	カテゴリ	年齢	金額	カテゴリ	年代	購入有						
2	レジャー	53	38920	レジャー	50代							
3	ビューティー	21	5210	ビューティー	20代							
4	レジャー	44	23070	レジャー	40代							
5	物販	27	1210	物販	20代							
6	物販	39	6400	物販	30代							
7	物販	55	2800	物販	50代							
8	旅行	61	65350	旅行	60代							
9	グルメ (飲食)	59	7960	グルメ (飲食)	50代							
10	物販	57	3510	物販	50代							
11	グルメ (飲食)	36	4330	グルメ (飲食)	30代							
12	グルメ (飲食)	51	3330	グルメ (飲食)	50代							
13	物販	31	4140	物販	30代							
14	物販	21	640	物販	20代							
15	レジャー											
16	旅行											
17	物販											
18	グルメ (飲食)											

ピボットテーブルの作成

分析するデータを選択してください。

テーブルまたは範囲を選択(S)

テーブル/範囲(I): クロス表!\$D\$1:\$F\$65432

外部データソースを使用(U)

接続の選択(C)...

接続名:

このブックのデータ モデルを使用する(D)

ピボットテーブル レポートを配置する場所を選択してください。

新規ワークシート(N)

既存のワークシート(E)

場所(L): クロス表!\$K\$4

複数のテーブルを分析するかどうかを選択

このデータをデータ モデルに追加する(M)

OK キャンセル

集計する列を指定

新規ワークシートの方がその後の加工がしやすくなる。

実習（5）

- ピボットテーブルの設定画面が出る。

The screenshot displays the Microsoft Excel interface with the PivotTable Tools ribbon selected. The ribbon includes tabs for '分析' (Analysis) and 'デザイン' (Design). The '分析' tab is active, showing options for 'アクティブなフィールド' (Active Fields), 'グループ' (Groups), 'フィルター' (Filters), 'データ' (Data), '計算方法' (Calculation Methods), and 'ツール' (Tools). The 'デザイン' tab shows options for 'フィールド/アイテム/セット' (Fields/Items/Sets), 'OLAP ツール' (OLAP Tools), and 'リレーションシップ' (Relationships). The 'ツール' tab shows options for 'フィールドリスト' (Field List), '+/- ボタン' (Plus/Minus Buttons), and 'フィールドの見出し' (Field Headings).

The main workspace shows a PivotTable named 'ピボットテーブル1' (PivotTable1) in cell B8. The PivotTable Fields task pane is open on the right side of the screen, titled 'ピボットテーブルのフィールド' (PivotTable Fields). The task pane contains a search box and a list of fields to be added to the report. The fields listed are 'カテゴリ' (Category), '年代' (Year), and '購入有無' (Purchase Status). The task pane also includes a section for '次のボックス間でフィールドをドラッグしてください:' (Drag fields between the following boxes:), with a 'フィルター' (Filter) box and a '列' (Column) box. At the bottom of the task pane, there are checkboxes for '行' (Row) and 'Σ 値' (Sum of Values).

The status bar at the bottom of the Excel window shows '関係係数 | 順位相関 | 外れ値処理 | Sheet8 | クロス表'.

実習（6）

- 集計したい情報を3種類選択する。

自動保存 ● オフ | HO_D1S2.xlsx - Excel | ピボットテーブル ツール | 森田 匠

ファイル ホーム 挿入 描画 ページレイアウト 数式 データ 校閲 表示 ヘルプ **分析** デザイン 検索

アクティブなフィールド: カテゴリ | ドブルックダウン | ドブルックアップ

グループ: → グループの選択 | 目録グループ解除 | フィールドのグループ化

フィルター: スライサーの挿入 | タイムラインの挿入 | フィルターの接続

データ: 更新 | データソースの変更 | アクション

計算方法: フィールド/アイテム/セット | OLAP ツール | リレーションシップ

ツール: ピボットグラフ | おすすめピボットテーブル

表示: フィールドリスト | +/- ボタン | フィールドの見出し

行ラベル	合計 / 購入有無
グルメ (飲食)	13108
20代	271
30代	1344
40代	3596
50代	4718
60代	2679
70代	500
ビューティー (美容)	6949
20代	299
30代	1202
40代	2292
50代	2244
60代	820
70代	92
レジャー	7524
20代	139
30代	806
40代	2135
50代	2618
60代	1525
70代	301
レッスン (学び)	2293
20代	86
30代	474

ピボットテーブルのフィールド

レポートに追加するフィールドを選択してください:

検索

カテゴリ
 年代
 購入有無
その他のテーブル...

次のボックス間でフィールドをドラッグしてください:

フィルター	列

行: カテゴリ | Σ 値: 合計 / 購入有無
年代

レイアウトの更新を保留する | 更新

100%

実習（7）

- 行と列の項目名を決める。

自動保存 (オフ) HO_D152.xlsx - Excel

ファイル ホーム 挿入 描画 ページレイアウト 数式 データ 校閲 表示 ヘルプ 分析 デザイン 検索

アクティブなフィールド: 合計 / 購入有無

アクティブなフィールド: ドリルダウン ドリルアップ

グループ: → グループの選択, 解除 グループ解除, フィールドのグループ化

フィルター: スライサーの挿入, タイムラインの挿入, フィルターの接続

データ: 更新, データソースの変更, アクション

計算方法: フィールド/アイテム/セット, OLAP ツール, リレーションシップ

ツール: ピボットグラフ, おすすめピボットテーブル

表示: フィールドリスト, +/- ボタン, フィールドの見出し

A3 合計 / 購入有無

行ラベル	グルメ (飲食)	ビューティー (美容)	レジャー レッスン (学び)	医療 (美容・歯科)	物販	旅行	総計	
20代	271	299	139	86	9	588	193	1585
30代	1344	1202	806	424	92	2810	1099	7777
40代	3596	2292	2135	770	136	6716	2862	18507
50代	4718	2244	2618	773	194	8047	4202	22796
60代	2679	820	1525	223	95	4409	2729	12480
70代	500	92	301	17	18	789	569	2286
総計	13108	6949	7524	2293	544	23359	11654	65431

ピボットテーブルのフィールド

レポートに追加するフィールドを選択してください:

検索

カテゴリ
 年代
 購入有無
その他のテーブル...

次のボックス間でフィールドをドラッグしてください:

▼ フィルター

■ 列
カテゴリ

■ 行
年代

Σ 値
合計 / 購入有無

レポートの更新を保留する 更新

100%

相関係数 順位相関 外れ値処理 Sheet8 クロス表

カテゴリを「列」に移す

合計だけでなく平均値なども求められる。

実習（8）

- ピボットテーブルの結果を別のシートにコピーする（値貼り付けする）

自動保存 ● オフ

HO_D1S2.xlsx - Excel

ファイル ホーム 挿入 描画 ページレイアウト 数式 データ 校閲 表示 ヘルプ 検索

Meiryo 10 A⁺ A⁻

標準

条件付き書式 テーブルとしてセルの書式設定 スタイル

R18

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1															
2			グルメ (飲食)	ビューティ (レジャー)	レッスン (医療 (美容))	物販	旅行	総計							
3		20代	271	299	139	86	9	588	193	1585					
4		30代	1344	1202	806	424	92	2810	1099	7777					
5		40代	3596	2292	2135	770	136	6716	2862	18507					
6		50代	4718	2244	2618	773	194	8047	4202	22796					
7		60代	2679	820	1525	223	95	4409	2729	12480					
8		70代	500	92	301	17	18	789	569	2286					
9		総計	13108	6949	7524	2293	544	23359	11654	65431					
10															
11															
12															
13															
14															
15															
16															
17															
18															
19															
20															
21															
22															

ピボットテーブルの出力結果は加工に向かないので、データだけを別シートに貼り付けて作業する。

実習（9）

- リフト値計算用の枠を作る

The screenshot shows the Microsoft Excel interface with a spreadsheet containing data. The ribbon is set to 'ホーム' (Home). The spreadsheet has columns A through U and rows 1 through 24. The data is organized into two main sections, each with a header row (row 2) and a list of age groups (rows 3-8) followed by a total row (row 9).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1																					
2			グルメ (飲食ビュティ-レジャー)	レッスン (兮医療 (美容・物販	旅行	総計							グルメ (飲食ビュティ-レジャー)	レッスン (兮医療 (美容・物販	旅行	総計					
3		20代	271	299	139	86	9	588	193	1585		20代									
4		30代	1344	1202	806	424	92	2810	1099	7777		30代									
5		40代	3596	2292	2135	770	136	6716	2862	18507		40代									
6		50代	4718	2244	2618	773	194	8047	4202	22796		50代									
7		60代	2679	820	1525	223	95	4409	2729	12480		60代									
8		70代	500	92	301	17	18	789	569	2286		70代									
9		総計	13108	6949	7524	2293	544	23359	11654	65431		総計									
10																					
11																					
12			グルメ (飲食ビュティ-レジャー)	レッスン (兮医療 (美容・物販	旅行	総計							グルメ (飲食ビュティ-レジャー)	レッスン (兮医療 (美容・物販	旅行	総計					
13		20代										20代									
14		30代										30代									
15		40代										40代									
16		50代										50代									
17		60代										60代									
18		70代										70代									
19		総計										総計									
20																					
21																					
22																					
23																					
24																					

Three green callout boxes with white text and arrows pointing to the data tables in the spreadsheet:

- Top right callout: 左上の表の枠のみをコピーする。
- Bottom left callout: 左上の表の枠のみをコピーする。
- Bottom right callout: 左上の表の枠のみをコピーする。

実習 (10)

- 行方向の比率を計算する。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T			
1																							
2			グルメ (飲食・ビューティー・レジャー)				レッスン (美容・医療・物販)				旅行	総計		グルメ (飲食・ビューティー・レジャー)				レッスン (美容・医療・物販)				旅行	総計
3		20代	271	299	139	86	9	588	193	1585		20代	=C3/\$J3										
4		30代	1344	1202	806	424	92	2810	1099	7777		30代											
5		40代	3596	2292	2135	770	136	6716	2862	18507		40代											
6		50代	4718	2244	2618	773	194	8047	4202	22796		50代											
7		60代	2679	820	1525	223	95	4409	2729	12480		60代											
8		70代	500	92	301	17	18	789	569	2286		70代											
9		総計	13108	6949	7524	2293	544	23359	11654	65431		総計											

=C3/\$J3
絶対参照を使うとそのまま数式を
コピーできる

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T			
1																							
2			グルメ (飲食・ビューティー・レジャー)				レッスン (美容・医療・物販)				旅行	総計		グルメ (飲食・ビューティー・レジャー)				レッスン (美容・医療・物販)				旅行	総計
3		20代	271	299	139	86	9	588	193	1585		20代	0.170978										
4		30代	1344	1202	806	424	92	2810	1099	7777		30代											
5		40代	3596	2292	2135	770	136	6716	2862	18507		40代											
6		50代	4718	2244	2618	773	194	8047	4202	22796		50代											
7		60代	2679	820	1525	223	95	4409	2729	12480		60代											
8		70代	500	92	301	17	18	789	569	2286		70代											
9		総計	13108	6949	7524	2293	544	23359	11654	65431		総計											

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T			
1																							
2			グルメ (飲食・ビューティー・レジャー)				レッスン (美容・医療・物販)				旅行	総計		グルメ (飲食・ビューティー・レジャー)				レッスン (美容・医療・物販)				旅行	総計
3		20代	271	299	139	86	9	588	193	1585		20代	0.170978	0.188644	0.087697	0.054259	0.005678	0.370978	0.121767	1			
4		30代	1344	1202	806	424	92	2810	1099	7777		30代	0.172817	0.154558	0.103639	0.05452	0.01183	0.361322	0.141314	1			
5		40代	3596	2292	2135	770	136	6716	2862	18507		40代	0.194305	0.123000	0.115362	0.041606	0.007349	0.36289	0.154644	1			
6		50代	4718	2244	2618	773	194	8047	4202	22796		50代	0.206966	0.098438	0.033909	0.00851	0.353001	0.184331	1				
7		60代	2679	820	1525	223	95	4409	2729	12480		60代	0.214663	0.065705	0.120000	0.007612	0.353285	0.21867	1				
8		70代	500	92	301	17	18	789	569	2286		70代	0.218723	0.040249	0.000000	0.000000	0.000000	0.000000	0.000000	1			
9		総計	13108	6949	7524	2293	544	23359	11654	65431		総計	0.200333	0.106200	0.000000	0.000000	0.000000	0.000000	0.000000	1			

M3をコピーして全体に貼り付けると
数式が反映される。

基本的な統計・データ分析スキル②



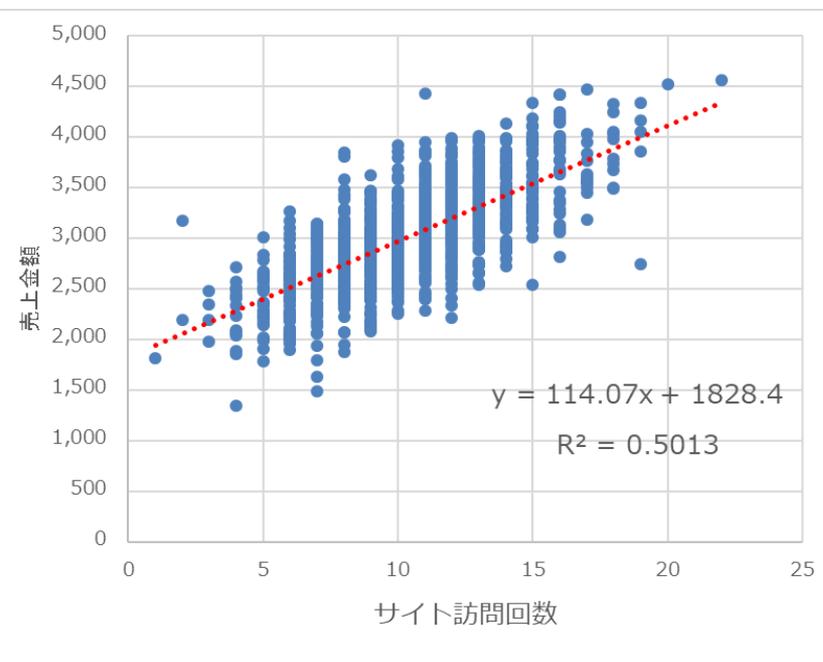
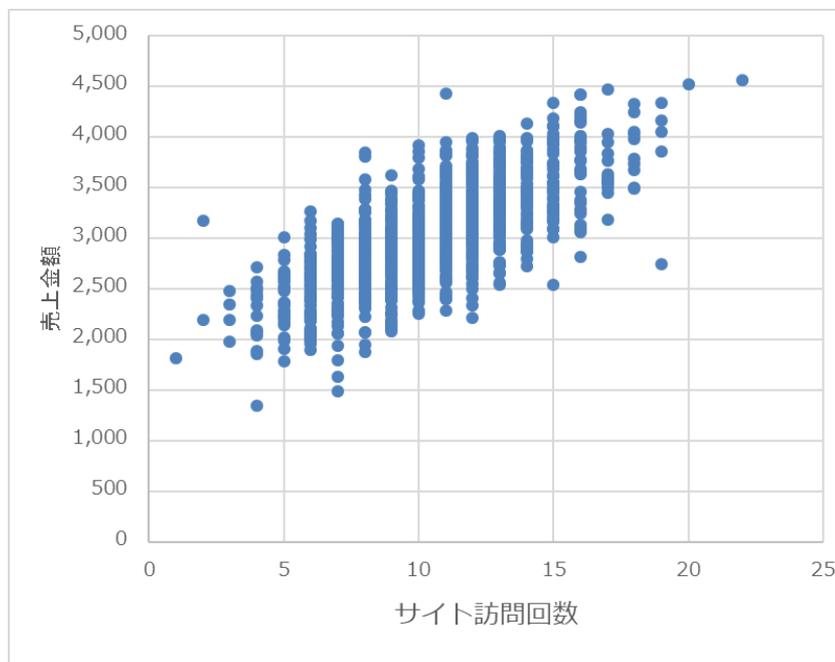
データの掛け合わせ

回帰分析とは？

- 予測したい変数（目的変数）をいくつかの要因となる変数（説明変数）の影響度で関係を説明する方法。
 - 説明変数が1つの場合：単回帰分析
 - 説明変数が複数の場合：重回帰分析

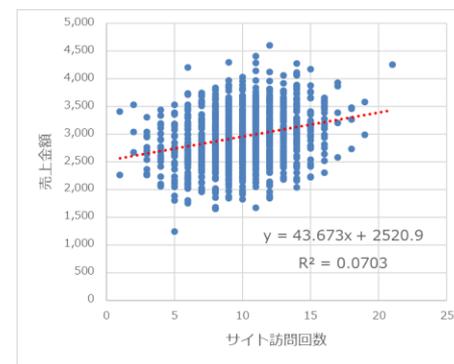
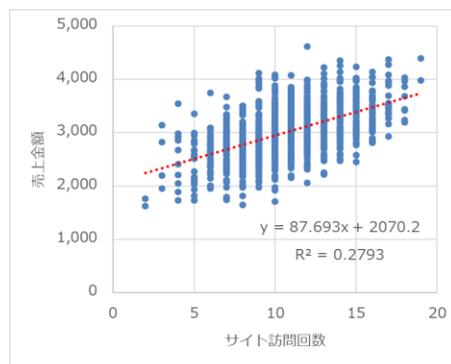
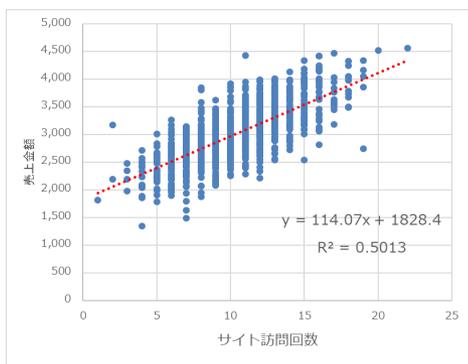
単回帰分析 (1)

- 説明変数が1つの場合の回帰分析を「単回帰分析」という。
 - 単回帰分析は「散布図」から算出することができる。
 - 単回帰分析のモデル式は、「 $y = ax + b$ 」となる。



単回帰分析（2）

- 単回帰分析の結果のイメージをきちんと作ることがその先に進むために重要。
 - 決定係数がモデルの当てはまりの良さを表す。
 - 決定係数が0.5以上あれば当てはまりがよく、0.2未満であればモデルを見直す。



モデル式

$$y = 114.07x + 1828.4$$

$$y = 87.693x + 2070.2$$

$$y = 43.673x + 2520.9$$

相関係数

$$R = 0.708$$

$$R = 0.5285$$

$$R = 0.2652$$

決定係数

$$R^2 = 0.5013$$

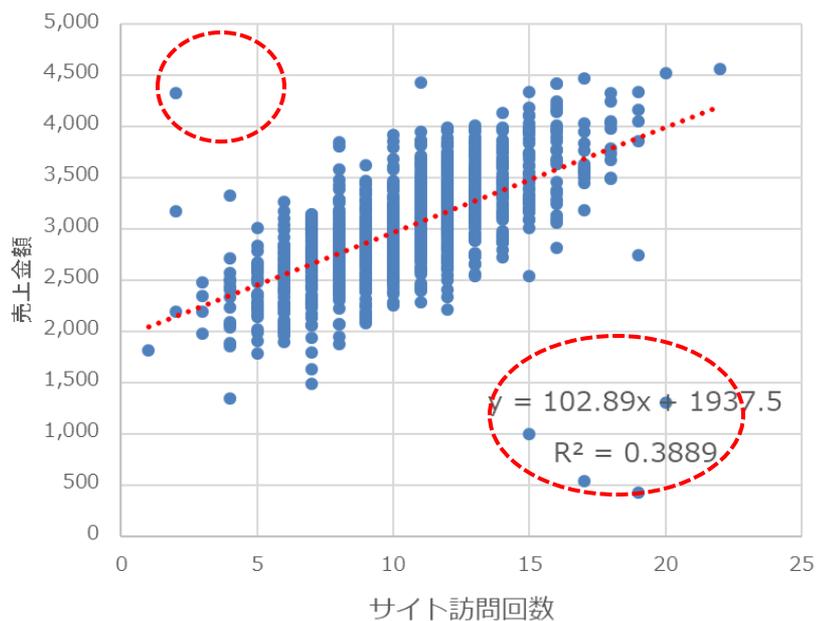
$$R^2 = 0.2793$$

$$R^2 = 0.0703$$

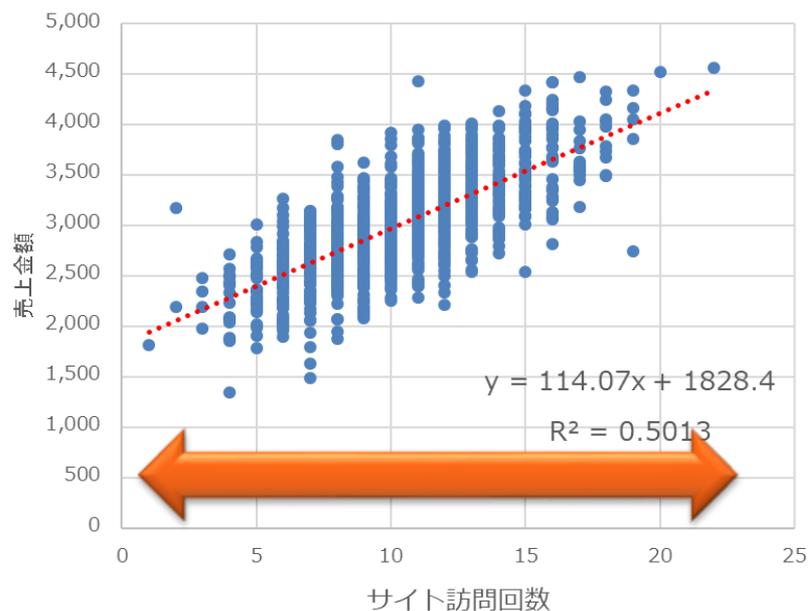
単回帰分析 (3)

- 回帰分析の注意点

外れ値に弱い



モデルを構築した範囲外のことは何も言えない



単回帰分析

実習



実習 (2)

- データを選択し、散布図を作成する。

The screenshot shows the Microsoft Excel interface with the 'Insert' ribbon selected. The 'Charts' group is expanded to show the 'Scatter' (散布図) options. A tooltip is displayed over the 'Scatter' options, providing information about the different types of scatter plots available.

散布図

この種類のグラフの使用目的:

- 少なくとも 2 つのセットの値または 2 組のデータを比較します。
- 値のセットの関係を示します。

バブル

使用ケース:

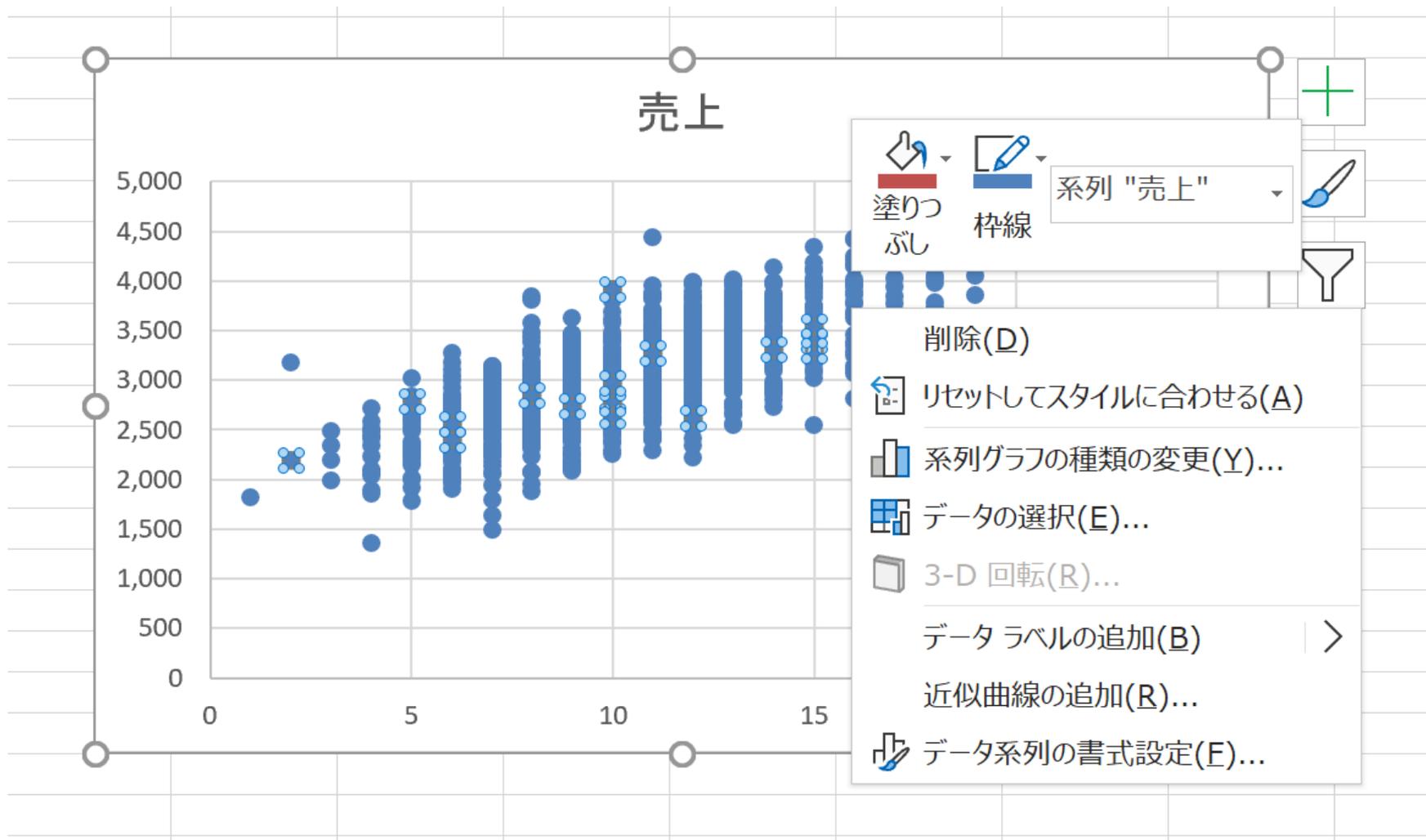
- データがばらついている場合に使用します。

その他の散布図(M)...

	A	B	C	D	E	F
1	サイト訪問回数	売上				
2	12	2,614				
3	6	2,813				
4	12	3,186				
5	10	3,288				
6	17	3,452				
7	12	3,023				
8	10	2,642				
9	7	2,761				
10	9	2,660				
11	12	3,011				
12	9	3,196				
13	13	3,394				
14	12	2,946				
15	10	3,853				
16	10	2,463				
17	11	3,830				
18	6	1,989				
19	7	2,732				
20	15	3,309				
21	9	3,278				
22	11	2,847				

実習 (3)

- 散布図のデータ点を右クリックし、「近似曲線の追加」を選択



実習 (4)

- 近似曲線の書式設定で、「線形近似」「グラフに数式を表示する」「グラフにR-2乗値を表示する」を選択する。

近似曲線の書式設定

近似曲線のオプション

近似曲線のオプション

- 目盛線 (Y)
- 前傾線 (L)
- 対数近似 (Q)
- 多項式近似 (P) 次数(D) 2
- 累乗近似 (W)
- 移動平均 (M) 区間(E) 2

近似曲線名

- 自動(A) 線形 (売上)
- ユーザー設定(C)

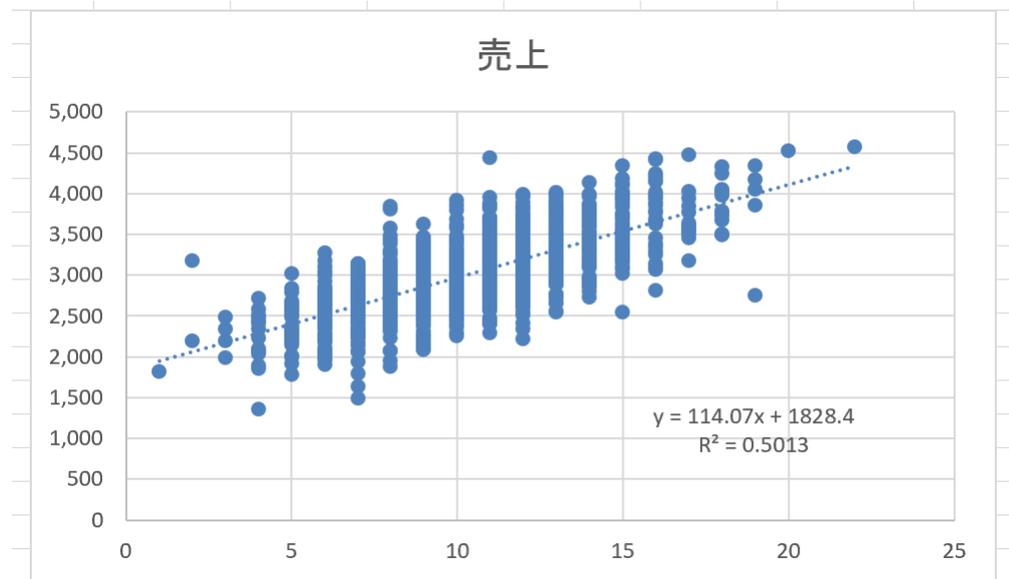
予測

前方補外(E) 0.0

後方補外(B) 0.0

切片(S) 0.0

- グラフに数式を表示する(E)
- グラフに R-2 乗値を表示する(R)



「身近なデータで掛け合わせ」の計画

ワークの進め方

- 以下のワークシートを使って分析を進め、発表に備えます。

観光ビッグデータ	ワークシート p データ分析実習準備
1. 普通の業務で使っているデータは何ですか？	
2. 普通の業務で求められていることは何ですか？	
3. 1のデータの中で回帰分析をして2に役立ちそうなものはありますか？	
4. 1のデータと何を掛け合わせる必要がありますか？	
5. 分析をしたら、どんな結果になると思いますか？（仮説）	
6. 4のデータがすぐに手に入りますか？ Yes→実際に分析して結果を書き込みましょう。 No →5の仮説通りになったら、どう活かせるか考えましょう。	

ワークの進め方

観光ビッグデータ

ワークシート p

データ分析実習準備

1. 普段の業務で使っているデータは何ですか？

2. 普段の業務で求められていることは何ですか？

3. 1のデータの中で回帰分析をして2に役立ちそうなものはありますか？

4. 1のデータと何を掛け合わせる必要がありますか？

5. 分析をしたら、どんな結果になると思いますか？（仮説）

6. 4のデータがすぐに手に入りますか？

Yes→実際に分析して結果を書き込みましょう。

No →5の仮説通りになったら、どう活かせるか考えましょう。

まずは、普段の業務で扱ってるデータを
挙げていってください。
”データ”であれば、なんでもOKです。

<データであれば何でもOKの例>

出席回数・出席率

課題の提出数／未提出数

テストの点数

グループワークで自分が担当している
作業数・タスク数

ワークの進め方

観光ビッグデータ

ワークシート p

データ分析実習準備

1. 普段の業務で使っているデータは何ですか？

.....

2. 普段の業務で求められていることは何ですか？

.....

3. 1のデータの中で回帰分析をして2に役立ちそうなものはありますか？

.....

4. 1のデータと何を掛け合わせる必要がありますか？

.....

5. 分析をしたら、どんな結果になると思いますか？（仮説）

.....

6. 4のデータがすぐに手に入りますか？

Yes→実際に分析して結果を書き込みましょう。

No →5の仮説通りになったら、どう活かせるか考えましょう。

あなたが” 普段の業務” で求めら得ていること（ミッション）を書いてください。提出するシートではないので” 自分の言葉” で書いていただいて大丈夫です。

ワークの進め方

観光ビッグデータ

ワークシート p

データ分析実習準備

1. 普段の業務で使っているデータは何ですか？

.....

2. 普段の業務で求められていることは何ですか？

.....

3. 1のデータの中で回帰分析をして2に役立ちそうなものはありますか？

.....

4. 1のデータと何を掛け合わせる必要がありますか？

.....

5. 分析をしたら、どんな結果になると思いますか？（仮説）

.....

6. 4のデータがすぐに手に入りますか？

Yes→実際に分析して結果を書き込みましょう。

No→5の仮説通りになったら、どう活かせるか考えましょう。

回帰分析をしたら、自分の業務・ミッションに役立ちそうなものはないか考えてみましょう。
（無くて無理やり捻り出してください）

<役に立つの例>

売上UP（客単価が上がる・顧客数が増える）

経費が削減される

業務効率があがる

仕事が楽しくなる

企業の社会的価値が上がる

ワークの進め方

観光ビッグデータ

ワークシート p

データ分析実習準備

1. 普段の業務で使っているデータは何ですか？

.....

2. 普段の業務で求められていることは何ですか？

.....

3. 1のデータの中で回帰分析をして2に役立ちそうなものはありますか？

.....

4. 1のデータと何を掛け合わせる必要がありますか？

.....

5. 分析をしたら、どんな結果になると思いますか？（仮説）

.....

6. 4のデータがすぐに手に入りますか？

Yes→実際に分析して結果を書き込みましょう。

No →5の仮説通りになったら、どう活かせるか考えましょう。

回帰分析に必要な
「もう一つの軸」は、何でしょうか

ワークの進め方

観光ビッグデータ

ワークシート p

データ分析実習準備

1. 普段の業務で使っているデータは何ですか？

.....

2. 普段の業務で求められていることは何ですか？

.....

3. 1のデータの中で回帰分析をして2に役立ちそうなものはありますか？

.....

4. 1のデータと何を掛け合わせる必要がありますか？

.....

5. 分析をしたら、どんな結果になるとおもいますか？（仮説）

.....

6. 4のデータがすぐに手に入りますか？

Yes→実際に分析して結果を書き込みましょう。

No →5の仮説通りになったら、どう活かせるか考えましょう。

分析結果を予想し仮説を立てましょう。
仮説は間違っても良いので
「どんな分析結果が出そうか」を記述します。

ワークの進め方

＜最低限書くべき結果＞

モデル式 ($y = \circ x + \Delta$)

決定係数

見直しの必要はあるか (決定係数は0.2以上か)

もう一つの軸が手に入れば、早速分析してみましょう。
手に入らない場合は「仮説通りだった」先の
計画を立ててください。

公的データ分析演習①



講義の目的

分析と提案の流れを実際に体験し身につけてもらうために
公的なデータを使って分析を行います。

分析と提案のテーマ

地域の現状・課題について、公的データを使って分析し
解決策となるようなアイデアを提案してください。

公的データの分析と提案

公的データの分析と提案

分析と提案のテーマ

地域の現状・課題について、公的データを使って分析し
解決策となるようなアイデアを提案してください。
下記を満たしていれば、どのような提案でも構いません。

データを活用して

- ①分析テーマや対象地域の概要を把握できている
- ②対象地域を全国や他の地域と比較できている
- ③対象地域の課題を説明できている

※この課題は、内閣府地方創生推進室主催「地方創生☆政策アイデアコンテスト 2022」を参考に作成しています。

公的データの分析と提案

公的データとは

公的機関、組織によって提供されているデータであり、自前の独自データと掛合わせることで分析を深めることができる。

※過去に学習した内容も活用してください。

公的データの例

●経済・統計データ

人口推計、消費者物価指数など

●位置データ・地図データ

位置情報ビッグデータ、国土数値情報など

●気象データ

気象庁公開の各種データ

公的データ入手先の例

地域経済分析システム「RESAS」

Regional Economy (and) Society Analyzing System

地域経済に関する様々な官民のビッグデータを可視化

【RESAS サイト】 <https://resas.go.jp/>

e-Stat（政府統計の総合窓口）

各府省が公表する統計データを一つにまとめたポータルサイト

統計データの検索や提供など、さまざまな機能を備えている

各種統計情報のワンストップ提供を目指し、平成20年4月から運用

【e-Stat サイト】 <https://www.e-stat.go.jp/>

グループで分析計画を立ててください。

計画立案のポイント

①どこを「地域」とするか

②解決したい課題は何か

その地域にある課題は、どのようなものが予想されるか（仮説）

③どのようなデータが必要か

課題設定に必要なデータは何か、提案に必要なデータは何か



公的データの分析と提案

グループで必要なデータの収集を行ってください。

データ活用のポイント

- ①分析テーマや対象地域の概要を把握する
- ②対象地域を全国や他の地域と比較する
- ③対象地域の課題を説明する
- ④伝えたいことに応じて適切なデータを選択する
- ⑤複数のデータを活用し多角的・多面的に分析する
- ⑥対象地域の先行する取り組みを調査・分析する
- ⑦先行する他地域の類似事例を調査・分析する



公的データ分析演習②



講義の目的

分析と提案の流れを実際に体験し身につけてもらうために
公的なデータを使って分析を行います。

分析と提案のテーマ

地域の現状・課題について、公的データを使って分析し
解決策となるようなアイデアを提案してください。

公的データの分析と提案

公的データの分析と提案

グループでデータ分析を行い提案を作成してください。

提案資料作成の際のポイント

- ①分析テーマや対象地域の概要を把握している
- ②対象地域を全国や他の地域と比較している
- ③対象地域の課題を説明できている
- ④分析したデータや地域特性を踏まえた提案になっている
- ④伝えたいことに応じて適切なデータを選択している
- ⑥対象地域・他地域の先行する取り組み・類似事例を調査・分析している
- ⑦データや引用における出所を明記している
- ⑧聞き手を意識した理解しやすい提案になっている



グループで提案を発表してください。

＜以下流れで発表&質疑応答＞

- ①グループの提案の発表
- ②他のグループメンバーからの質問
- ③講師からの質問

→次の受講者へ

＜全員発表後＞

全体のまとめ（講師からのフィードバック）



実業務データ分析実習①



講義の目的

分析と提案の流れを確認し、身近なデータで分析を行います。
次回の発表（データ分析分野での最終発表）に備えます。

データ分析実習

以下のワークシートを使って分析を進め、発表に備えます。

観光ビッグデータ	ワークシート a
データ分析実習 (1)	
1. 自分の係わっている業務のデータで使えそうなデータはありますか？	
2. そのデータをここまで学んだ方法で分析したらどんな結果になると思いますか？ (仮説)	
3. 仮説通りの結果になったらどんな場面で活用できそうですか？	
4. 実際に分析した結果を書き込みましょう。	
- MEMO -	

観光ビッグデータ	ワークシート b
データ分析実習 (2)	
1. 分析結果から、どのような内容が読み取れますか？ (考察)	
2. 今後の業務での改善ポイントはなんでしょうか？ (提案)	
- MEMO -	

観光ビッグデータ

ワークシート a

データ分析実習 (1)

1. 自分の関わっている業務のデータで使えそうなデータはありますか？

.....

2. そのデータをここまで学んだ方法で分析したらどんな結果になると思いますか？ (仮説)

.....

3. 仮説通りの結果になったらどんな場面で活用できそうですか？

.....

4. 実際に分析した結果を書き込みましょう。

.....

-MEMO-

まずは、普段の業務で扱ってるデータで分析に使えそうなものがないか考えてみましょう。

もし、業務内に見当たらなければ「身近な」ものから探していただいてOKです。

データ分析実習

観光ビッグデータ

ワークシート a

データ分析実習 (1)

1. 自分の関わっている業務のデータで使えそうなデータはありますか？

.....

2. そのデータをここまで学んだ方法で分析したらどんな結果になると思いますか？ (仮説)

.....

3. 仮説通りの結果になったらどんな場面で活用できそうですか？

.....

4. 実際に分析した結果を書き込みましょう。

.....

- MEMO -

分析結果を予想し仮説を立てましょう。
仮説は間違っても良いので
「どんな分析結果が出そうか」を
可能な限り詳細に記述してください。

データ分析実習

観光ビッグデータ

ワークシート a

データ分析実習 (1)

1. 自分の関わっている業務のデータで使えそうなデータはありますか？

.....

2. そのデータをここまで学んだ方法で分析したらどんな結果になると思いますか？ (仮説)

.....

3. 仮説通りの結果になったらどんな場面で活用できそうですか？

.....

4. 実際に分析した結果を書き込みましょう。

.....

-MEMO-

続けて、もし立てた仮説が正しければ
その結果をどんな形で
活用できるかを考えてください。

ここまでで分析前の準備は完了です。

データ分析実習

データ分析前の準備は、以上です。

分析のために必要なデータを用意してください。

データ分析実習

データ分析実習

観光ビッグデータ

ワークシート a

データ分析実習 (1)

1. 自分の関わっている業務のデータで使えそうなデータはありますか？

.....

2. そのデータをここまで学んだ方法で分析したらどんな結果になると思いますか？ (仮説)

.....

3. 仮説通りの結果になったらどんな場面で活用できそうですか？

.....

4. 実際に分析した結果を書き込みましょう。

.....

- MEMO -

分析のためのデータ準備を進めてください

分析結果を後ほど
ここに書き込みます。

ペアを組んで、ここまでの見直しを行います。

説明役：分析計画と使用予定データについて説明します。

聞き役：分析計画のブラッシュアップを手伝います。

<聞き役のタスク>

- ①良かった点を伝える
- ②改善点（もっとこうだったら良い）を伝える
- ③質問（疑問点を確認する）
- ④追加のアイデアを提供する（あれば）



実業務データ分析実習②



講義の目的

分析と提案の流れを確認し、身近なデータで分析を行います。
次回の発表（データ分析分野最終発表）に備えます。

データ分析実習

データ分析実習

観光ビッグデータ

ワークシート a

データ分析実習 (1)

1. 自分の関わっている業務のデータで使えそうなデータはありますか？

2. そのデータをここまで学んだ方法で分析したらどんな結果になると思いますか？ (仮説)

3. 仮説通りの結果になったらどんな場面で活用できそうですか？

4. 実際に分析した結果を書き込みましょう。

- MEMO -

実際にデータ分析をしてください。
ここまで学んだ手法の何を使ってもOKです。

その結果をここに書き込みましょう。
ただし、書くのは「結果」のみです。

<結果と考察の違い>

アイスの売り上げと気温の相関は-0.7だ→結果
暑い方がアイスは売れる傾向がある →考察

ペアを組んで、ここまでの見直しを行います。

説明役：分析結果の読み取りを説明します。

聞き役：読み取りに問題がないか確認します。

<聞き役の注意点>

この段階では「結果のみ」の読み取りです。

結果の読み取りが間違っていなければ、指摘は不要です。

※解釈に違いがあっても、今は問題なし



データ分析実習

観光ビッグデータ

ワークシートb

データ分析実習 (2)

1. 分析結果から、どのような内容が読み取れますか？ (考察)

2. 今後の業務での改善ポイントはなんですか？ (提案)

- MEMO -

データ分析の結果から、何が読み取れるかを記述してください。

推測になってしまう部分もあるかもしれませんがこの段階ではOKです。

観光ビッグデータ

ワークシートb

データ分析実習 (2)

1. 分析結果から、どのような内容が読み取れますか？ (考察)

考察の結果を踏まえての次のアクションを
考えて記述してください。
考察の段階で「予測」になってしまったものがあれば
それは新たな「仮説」になります。
その仮説を検証する次の分析を企画してください。

2. 今後の業務での改善ポイントはなんですか？ (提案)

- MEMO -

<次の分析・調査の企画>

最低限必要なもの

- ①分析するデータ
→どんなデータが欲しいか
- ②有効と思われる分析手法
→どの方法で分析するか
- ③仮説の検証方法
→どんな結果が出たら仮説は正しいと言えるか

ペアを組んで、ここまでの見直しを行います。

説明役：考察・提案について説明します。

聞き役：分析のブラッシュアップを手伝います。

<聞き役のタスク>

①改善点（もっとこうだったら良い）を伝える

→特に考察で説明役が気がついていないことを指摘

②質問（疑問点を確認する）

③追加のアイデアを提供する（特に提案について）



データ分析実習

データ分析実習は、以上です。

次回の発表に備えて、説明の準備をしておいてください。

※ペアワークで得たアドバイスを反映させましょう。



企画書・事業計画書作成演習①



学習目標

1. これまで学習してきたデータマーケティングの知識を活用して、企画書・事業戦略書にまとめる
2. その内容を発表して、評価をしてもらう
3. 他者の発表を聞いて、よい点をフィードバックする

1. 4～5名のグループを作成する
2. グループ毎に取り組むテーマを決める
3. 役割分担を決める
4. 今回を含めて3回の授業の中で、企画書・事業計画書を完成させる
5. 最終回で発表する

まとめ

今日、決定したことと、次回取り組むことをまとめましょう。

【決定事項】

【次回取り組むこと】

企画書・事業計画書作成演習②



前回決めたグループに分かれて、自分の担当を再確認したうえで、ワークを始めましょう。

授業終了15分前に、今回決まったことをまとめてみましょう。また、次回の授業で企画書・事業計画書のまとめが終わるようにしましょう。

まとめ

今日、決定したことと、次回取り組むことをまとめましょう。

【決定事項】

【次回取り組むこと】

企画書・事業計画書作成演習③



進め方

グループワークの結果として、企画書・事業計画書にまとめるところまでを完成させましょう。

次回はグループ発表となるので、発表についての役割分担も決めましょう。

まとめ

次回の発表に向けて、決定したことを記録しておきましょう。

【役割分担】 自分の役割

【決定事項】

企画書・事業計画書作成演習④



各グループで取り組んだことを発表します。
他のグループの発表については、よいと思った点を意識しながら聞きましょう。

作成したマーケティング戦略書を発表します。