

下山 紗代子

一般社団法人リンクデータ  
代表理事

2020.2.6 (Thu.)

令和元年度  
政策評価に関する統一研修

行政プロセスに  
データ分析を  
取り入れるために  
知っておきたい  
知識と事例

# 本資料の電子ファイル入手方法

---

- 本資料では参考資料のURLを入れていたり、カラーでのビジュアルライズを多用しているため、電子ファイルでの閲覧をおすすめいたします
- 本資料は総務省の「政策評価に関する統一研修」の公式ページからも後日公開予定ですが、すぐに閲覧したい方は以下のURLからご覧頂けます

**<https://bit.ly/3602N28>**





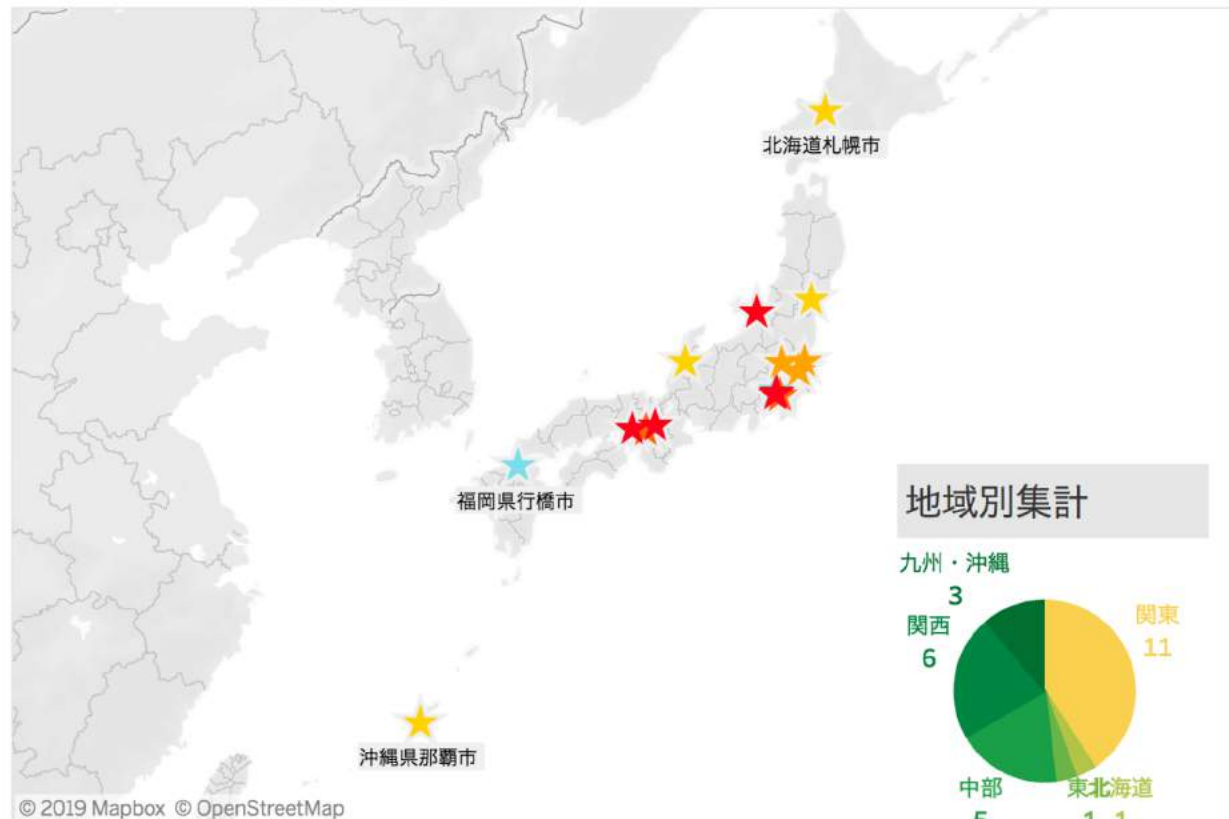
# Sayoko Shimoyama

# 2018年度の地方自治体におけるデータ活用支援

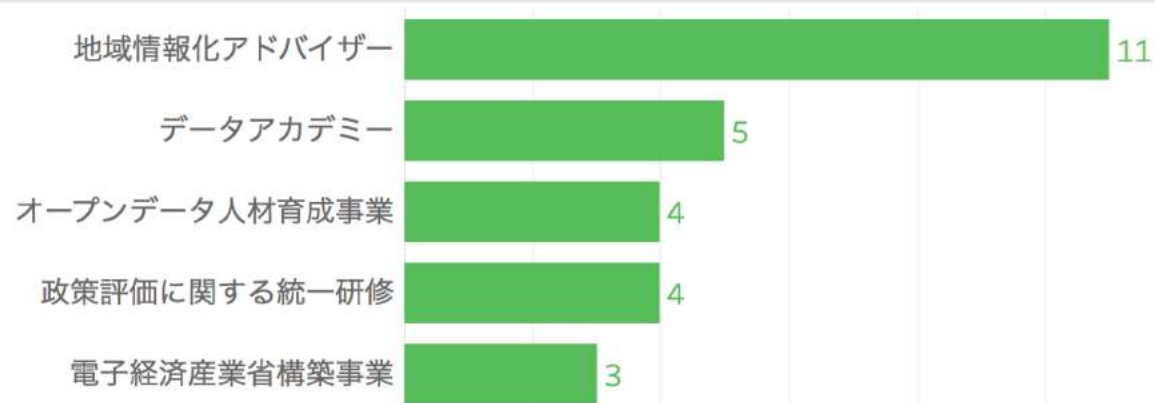
トータル

# 27回

※1自治体での1日の対応につき1回とカウント



## 事業別集計



## 月別集計



- 事業名
- 地域情報化アドバイザー
  - データアカデミー
  - オープンデータ人材育成事業
  - 政策評価に関する統一研修
  - 電子経済産業省構築事業

# 目次

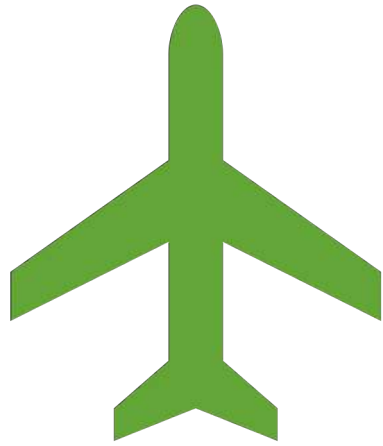
---

1. なぜ、データを使う必要があるのか
2. 公共データの共有がもたらす社会的インパクト
3. 政策立案のためのデータ分析の流れ
4. データ活用におけるよくある勘違い
5. お薦めの文献・資料

# 1. なぜ、 データを使う 必要があるのか

どちらの事故の方が怖いと感じますか？

---

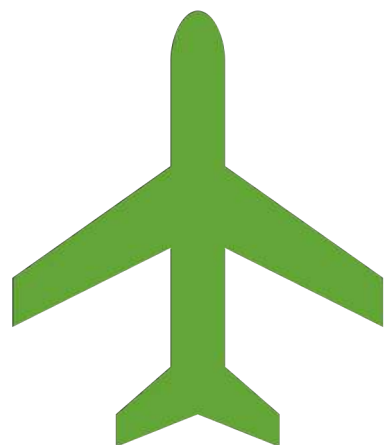


飛行機事故



自動車事故

# 実際に死亡事故に遭う確率を調べてみると...



約**252**万件に1回  
起こる確率

飛行機事故

**0.00004%**

2018年に世界で発生した飛行機事故件数：15 / 年間総フライト数：37,800,000  
出典元：Aviation Safety Network



約**3.5**万件に1回  
起こる確率

自動車事故

**0.00279%**

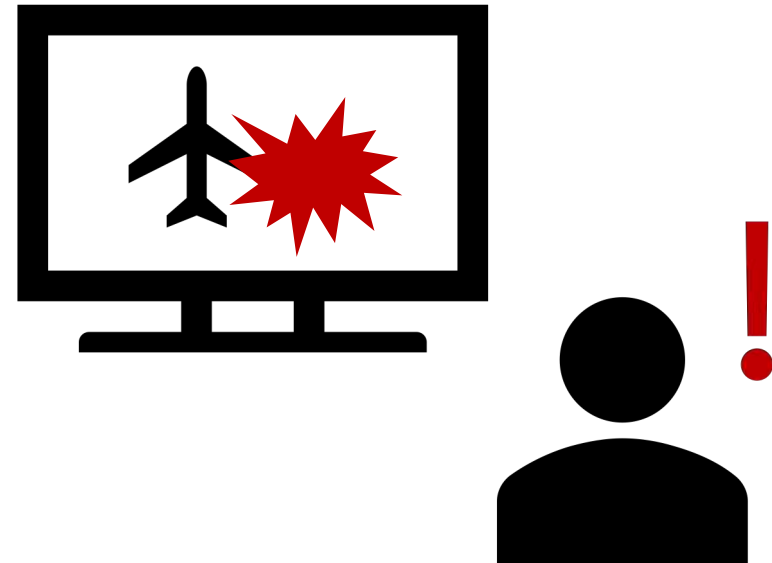
2018年の国内の自動車交通事故による死亡者数：3,532 / 総人口：126,435,000  
出典元：警視庁, 総務省統計局

飛行機の  
約**70倍**



# 「可用性ヒューリスティック」と呼ばれる判断の傾向が原因

- 人は、ある事例を思い浮かべやすければ、起こりやすいと判断しやすい傾向がある
- 認識、理解、決定の際に、**思い出しやすい情報だけに基づいて判断する傾向を**  
「可用性ヒューリスティック (Availability heuristic)」と呼ぶ
- 人が陥りやすい「認知バイアス」のうちの1つ



飛行機事故は大きく報道されることが多く、印象に残りやすい

# 認知バイアス =無意識な思考の偏り・誤り



**認知バイアス**（cognitive bias）とは、認知心理学や社会心理学での様々な観察者効果の一種であり、非常に基本的な統計学的な誤り、社会的帰属の誤り、記憶の誤り（虚偽記憶）など**人間が犯しやすい問題**である。転じて認知バイアスは、事例証拠や法的証拠の**信頼性を大きく歪める**。

出典：

<https://ja.wikipedia.org/wiki/%E8%AA%8D%E7%9F%A5%E3%83%90%E3%82%A4%E3%82%A2%E3%82%B9>

# 認知バイアスの例①： アンカリング効果

## ○ アンカリング効果

人間は商品を相対的な比較でしか判断することができない

アンカリング効果なし



価格: ¥ 12,000

ん～高いなあ.....  
買うのどうしようかな  
他も検討するか...



アンカリング効果あり



参考価格: ~~¥ 56,000~~

価格: ¥ 12,000

お！普通より安く  
なっている！  
買おう！



Twitter:@never\_be\_a\_pm  
<https://dividable.net>

# 認知バイアスの例②： サンクコスト（埋没費用）

## ○ 埋没費用

回収できない投資額を回収しようとしてしまう認知バイアス

埋没費用なし



チケット価格: ¥0

この映画つまらない  
から見るのやめよう



埋没費用あり



チケット価格: ¥1000

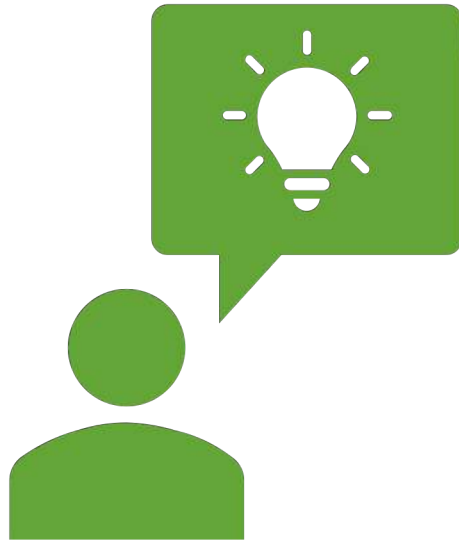
1000円払ってしまったし、もったいない  
から見つけよう...



Twitter:@never\_be\_a\_pm  
<https://dividable.net>

# どうすれば認知バイアスを外せるか？

---



どんな認知バイアスがあるか  
知っておけば気を付けられる...？

# 名前の付いている認知バイアスは190以上

<http://lelang.sites-hosting.com/naklang/method.html>

認知バイアス一覧で社会心理学入門  
～社会科学の知の蓄積を活用した社会教育の実現に向けて～

曉 美焰 (Xiao Meiyao) 社会学研究家, 2019.7.18 祝2.6版完成!

疑似科学を生み出すのは人間の思考が本来持っている誤りやすい傾向である。それ故に「科学と疑似科学の境界」を判断するためには、社会科学の知識は避けて通れない。ここでは人間の誤りやすい傾向について、人類の英知である「社会心理学」の偉大なる成果である認知バイアス一覧を英語版に基づいて紹介する。ここに紹介する知識は知っておくだけで人生に役立つ知識である事に間違いない。これらの概念を紹介する日本語のウェブサイトを探すのが難しい事自体が、日本人が論理的思考をしていない事の証明であろう。これまで社会科学とは縁の無かった科学技術系の人達が、少しでも社会科学に興味を持っていただく事を祈る。

## 1. 基本用語

- 帰属** 出来事や他人の行動や自分の行動の原因を説明する心的過程 (...のせいにする)。簡単にいえば、人柄のせいにするのが内的帰属であり、事情のせいにするのが外的帰属である。
- 動機づけ** 行動を始発させ、目標に向かって維持・調整する過程・機能。好奇心や関心によってもたらされ、賞罰に依存しない行動が内発的動機づけであり、義務、賞罰、強制などによってもたらされる行動が外発的動機づけである。

以下の認知バイアスの日本語訳は、できるだけ既に翻訳されている用語を用いている。しかし、翻訳が見つからなかった場合には勝手に日本語を造語し

## 2. 信念形成の際の認知バイアス一覧

- 後知恵バイアス (Hindsight bias)**  
物事が起きてからそれが予測可能だったと考える傾向。
- 可用性カスケード (Availability cascade)**  
主張を何度も聞いているうちに、真理であると確信する傾向。
- 可用性ヒューリスティック (Availability heuristic)**  
認識、理解、決定の際に、思い出しやすい情報だけに基づいて判断する傾向。
- 観察者期待効果 (Observer-expectancy effect)**  
観察者が期待する効果を観察する観点を解析し、無意識のうちにデータを誤って解釈する傾向。観察者効果の一種。
- 感情移入ギャップ (Empathy gap)**  
怒ったり恋愛したりしている時に、その感情を持たない視点で考える事ができない傾向。

常に全てを  
意識するのは  
無理！！

# 認知バイアスは避けられない

---

自分も、他人も、認識や判断には

常に認知バイアスが  
かかっている可能性がある

という前提でいる必要がある



# 認知バイアス対策

---

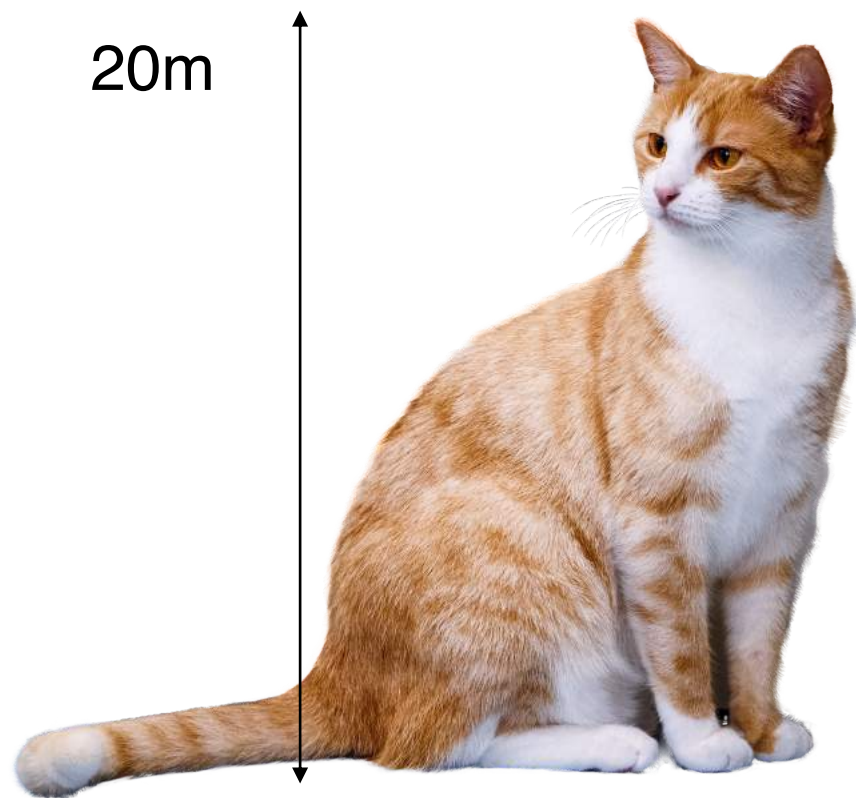
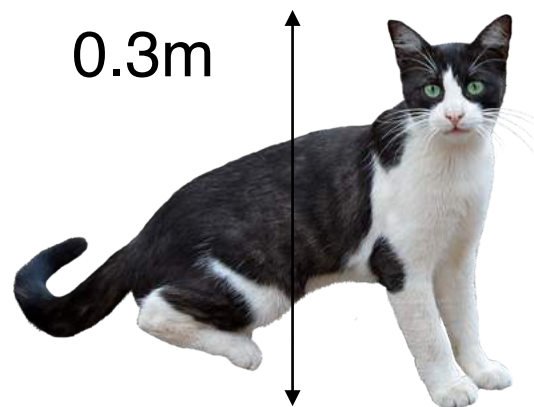
どうすればバイアスがかかっていることに気付き、  
正しい判断に近付くことが出来るのか？



主観的な情報だけでなく、客観的な情報を使う



# 客観的 = 誰が見ても大体同じ解釈ができる



客観的

Q1.  
どちらの猫が大きい？

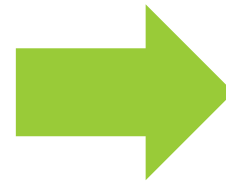
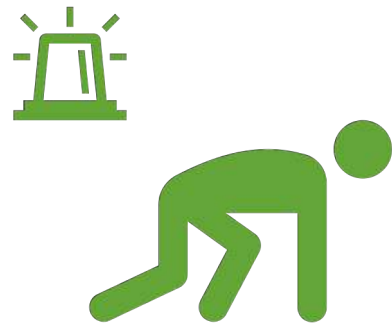
主観的

Q2.  
どちらの猫が可愛い？

# 数値を使うと解釈の齟齬を減らせる

例) ある市における救急隊員が抱えていた課題

配属先が変わって  
業務負担が増えた



しかし上司の理解が得られない



# 数値を使うと解釈の齟齬を減らせる

市内の各隊の隊員1人あたりの  
出動件数\*を算出して比較

※使用データ：消防年報の年間救急出動件数÷各隊の配属人数



最大で **7.75倍** の差

- 数字は言葉よりも解釈の幅が狭い
- 解釈の齟齬が言葉に比べて起こりにくい



出典元：Code for Japan データアカデミーにおける某自治体のチームの当日発表資料

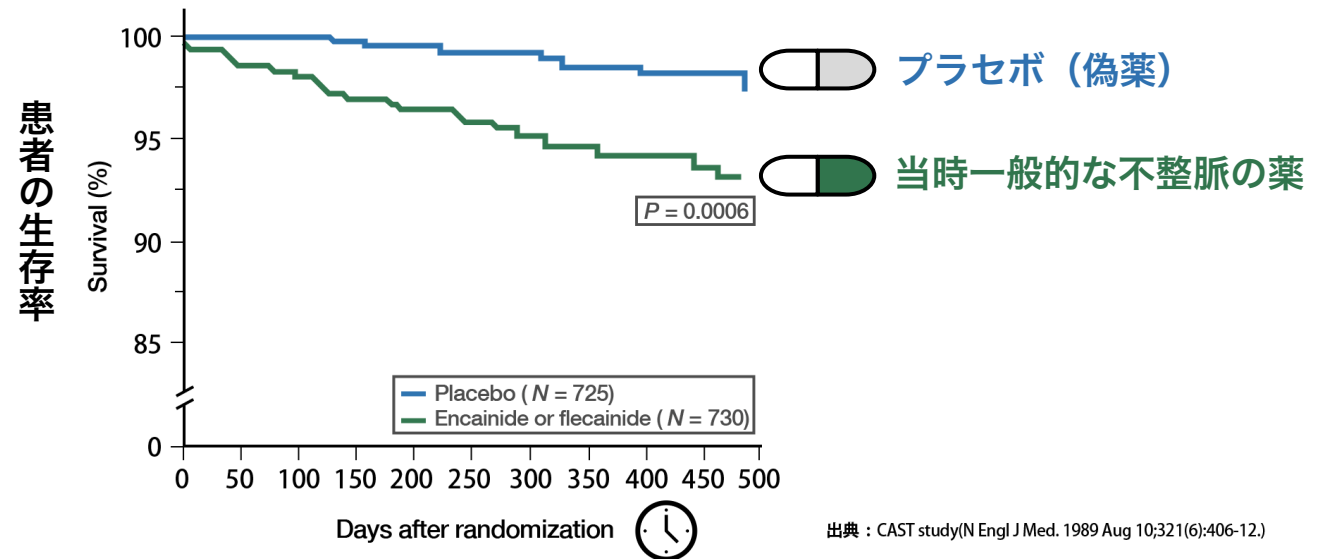
# 経験やカンだけではなく、それを裏付ける客観的なデータが必要

## ■ EBPM : Evidence Based Policy Making

...エビデンス (科学的根拠) に基づいた政策立案

## ■ EBPMは医学 (Evidence Based Medicine) から派生した考え方

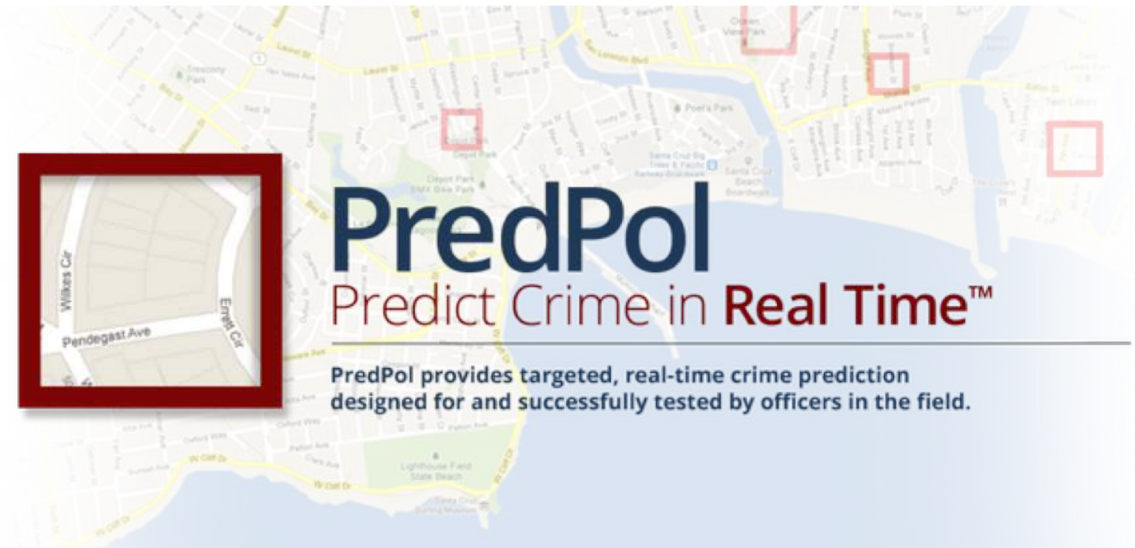
- 1989年、当時一般的に使用されていた不整脈の薬の効果をデータを取って検証したところ、服用によって死亡率が高まることが判明



経験的知識のみで判断することの危険性

# データ分析結果が経験やカンを越えた例： PredPol（米国の犯罪発生予測システム）

- PredPol はカリフォルニア大学の研究グループが開発した犯罪予測システム
- 過去の膨大な犯罪情報のデータベースをもとに、将来「いつ、どこで」犯罪が起きるのか、機械学習をベースに予測
- ロサンゼルス市警のプロの犯罪分析官とPredPolの予測を比較したところ、犯罪分析官に比べ**2倍以上**PredPolの方が正確に予測できた



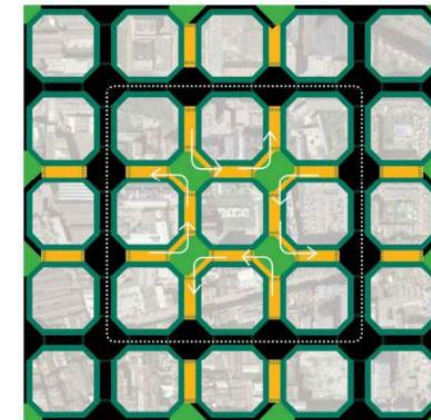
<https://www.predpol.com/>

従来の経験やカン頼りの方法ではなく、  
**データ**に基づく客観的な根拠から判断する方法へシフト

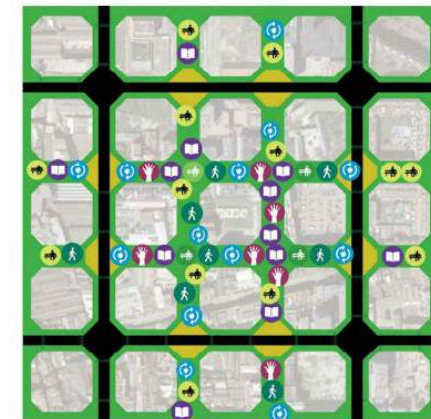
# データで根拠を示すことで合意形成を進めた例： スーパーストリック構想（バルセロナ市）

- バルセロナでは自動車の交通量が多く、汚染物質やCO2、騒音などによる公害が課題
- 「市民中心都市」をコンセプトにかかげ、段階的に車が通れる道を制限する計画を実施
- 交通量や環境に関するセンシングデータを用いて分析を行い、計画が進んだ場合にどのくらい問題が改善するのか、具体的な数値で示すことで**合意形成を推進**
  - 交通量が**21%**減
  - **94%**の市民が危険なレベルの粒子状物質に晒されることはなくなる
  - **73.5%**の市民は「65デシベル以上の騒音」を経験せずにすむ

<https://citiesofthefuture.eu/superblocks-barcelona-answer-to-car-centric-city/>



Phase 1



Phase 2

# 2. 公共データの 共有がもたらす 社会的インパクト

# サンフランシスコ市 × Yelp から見える データがもたらす社会変革



- 世界最大級のローカルビジネスのレビューサイト
- 日本でいうところの「食べログ」的なサービス

- サンフランシスコ市では、飲食店に対して実施した保健衛生検査の結果をデータとして公開している
- Yelpはそれをサービスに取り込み、“Health Score”として100点満点のスコア化して各飲食店のレビューページに表示





# サンフランシスコ市 × Yelp から見える データがもたらす社会変革

**yelp** Find Restaurants Near San Francisco, CA, United States

Home Services Restaurants Auto Services More Write a Review

## Fog Harbor Fish House

Claimed

4.5 stars 4550 reviews

Write a Review Add Photo Share Save

\$\$ Seafood, Bars

Pier 39 Ste A-202 San Francisco, CA 94133

Get Directions (415) 421-2442 fogharbor.com

Skip the line, Join our waitlist via your phone! Join Now

Make a Reservation Thursday, September 13, 2018 7:00 pm 2 people Find a Table

Today 11:00 am - 10:00 pm Closed now Full menu

Price range \$11-30

Health Score 87 out of 100

衛生スコア：  
87点

# サンフランシスコ市 × Yelp から見える データがもたらす社会変革

## Fog Harbor Fish House

November 20, 2017 – Routine Inspection

### Violations

- Improper storage use or identification of toxic substances
- High risk food holding temperature [ date violation corrected: 12/4/2017 ]
- Improper food storage
- Wiping cloths not clean or properly stored or inadequate sanitizer [ date violation corrected: 11/20/2017 ]

### Inspections

Date	Inspection Type	Violations	Score
November 20, 2017	Routine	4	87
July 17, 2017	Routine	2	86
April 25, 2016	Routine	3	92

直近の保健衛生  
検査において違反  
があった項目

Health Score

87

out of 100

### About Health Scores

We collect public inspection data directly from your local health department. Due to the local health department's inspection schedule as well as the time it takes to pass that information on to us, it is possible that we may not display the most recent inspection data.

Please report data inaccuracies via one of the methods below:

- [Email](#)
- [Website](#)

Yelpはユーザーにより安心・安全を届けられる  
サービスを提供可能に

# サンフランシスコ市 × Yelp から見える データがもたらす社会変革



もしサンフランシスコ市が  
衛生調査結果をサイトで公表している  
だけだったら？

一部の人しか見に来ない

強み：市民の安全を守るために実施  
した調査結果として、公平なデータ  
を提供できる



もしYelpが  
独自に飲食店を評価する  
スコアを付けていたら？

公平性が担保されない

強み：店舗の営業データや口コミ  
データを集めて、多くのユーザが  
集まる場を持っている

データが共有されることで  
「組織を超えて得意分野を分担できる社会」になる

# 行政のデータは誰にでも使えるように公開されている必要がある



もしサンフランシスコ市が  
Yelpだけに衛生調査結果データを  
渡していたら？



不公平に対し非難が集中

- 衛生調査は市の税金を使って運用されているので、納税者は等しくそのデータを使う権利を持っている
- 誰でも自由に使えるデータである必要がある

オープンデータ

# オープンデータの定義

国、地方公共団体及び事業者が保有する官民データのうち、国民誰もがインターネット等を通じて容易に利用（加工、編集、再配布等）できるよう、次のいずれの項目にも該当する形で公開されたデータをオープンデータと定義する

【出典】「オープンデータ基本指針」平成29年5月30日  
高度情報通信ネットワーク社会推進戦略本部・官民データ活用推進戦略会議決定

1. 営利目的、非営利目的を問わず二次利用可能なルールが適用されたもの
2. 機械判読に適したもの
3. 無償で利用できるもの

# “Open By Default”

2013年 G8 「オープンデータ憲章」による国際的な合意



「税金を使って作られた  
データは全て公共財として  
公開するべきである」  
という考え方を  
原則として採用

“オープンデータ”とは、  
元々皆さんのものだったデータを  
返すことだと考えています。

- Miquel Mateu, Open Data Manager of Barcelona City



14-18 June 2016, World Data Viz Challenge 2016,  
Barcelona

# 2016年：官民データ活用推進基本法施行

→国及び自治体はオープンデータに取り組むことが義務付けられた

## オープンデータ基本指針 概要

### 本基本指針の位置づけ

平成28年12月14日に公布・施行された「官民データ活用推進基本法」において、国、地方公共団体、事業者が保有する官民データの容易な利用等について規定された。本文書は、これまでの取組を踏まえ、オープンデータ・バイ・デザイン<sup>(注)</sup>の考えに基づき、国、地方公共団体、事業者が公共データの公開及び活用に取り組む上での基本方針をまとめたものである。

### 1. オープンデータの意義

- (1) 国民参加・官民協働の推進を通じた諸課題の解決、経済活性化
- (2) 行政の高度化・効率化
- (3) 透明性・信頼の向上

### 2. オープンデータの定義

- ① 営利目的、非営利目的を問わず  
二次利用可能なルールが適用されたもの
- ② 機械判読に適したもの
- ③ 無償で利用できるもの

### 3. オープンデータに関する基本的ルール

- (1) 公開するデータの範囲・・・各府省庁が保有するデータは、原則オープンデータとして公開。公開することが適当でない公共データは、公開できない理由を原則公開するとともに、限定的な関係者間での共有を図る「限定公開」といった手法も積極的に活用。
- (2) 公開データの二次利用に関するルール・・・原則、政府標準利用規約を適用。
- (3) 公開環境・・・特にニーズが高いと想定されるデータは、一括ダウンロードを可能とする仕組みの導入や、APIを通じた提供を推進。
- (4) 公開データの形式等・・・機械判読に適した構造及びデータ形式で掲載することを原則。法人情報を含むデータは、法人番号を併記。
- (5) 公開済みデータの更新・・・可能な限り迅速に公開するとともに適時適切な更新。

### 4. オープンデータの公開・活用を促す仕組み

- (1) オープンデータ・バイ・デザインの推進・・・行政手続き及び情報システムの企画・設計段階から必要な措置
- (2) 利用者ニーズの反映・・・各府省庁の保有データとその公開状況を整理したリストを公開→利用者ニーズを把握の上、ニーズに即した形での公開

### 5. 推進体制

- (1) 相談窓口の設置・・・総合的な相談窓口（内閣官房IT総合戦略室）・相談窓口（各府省庁）の設置
- (2) 推進体制・・・内閣官房IT総合戦略室は、政府全体のオープンデータに関する企画立案・総合調整、各施策のレビュー、フォローアップを実施等

### 6. 地方公共団体、独法、事業者における取組

地方公共団体・・・官民データ法の趣旨及び本基本指針を踏まえて推進。  
独立行政法人・・・国費によって運営されていること又は実施している事業や研究があることに鑑み、基本指針に準拠して取組を推進することが望ましい。  
公益事業分野の事業者・・・その公益性に鑑み、本基本指針及び利用者ニーズを踏まえて推進することが望ましい。

(注)公共データについて、オープンデータを前提として情報システムや業務プロセス全体の企画、整備及び運用を行うこと。



# オープンデータの意義

---

1. 国民参加・官民協働の推進を通じた諸課題の解決、経済活性化
2. 行政の高度化・効率化
3. 透明性・信頼の向上

【出典】 「オープンデータ基本指針」 平成29年5月30日  
高度情報通信ネットワーク社会推進戦略本部・官民データ活用推進戦略会議決定

# オープンデータの意義

---

1. 国民参加・官民協働の推進を通じた  
諸課題の解決、経済活性化

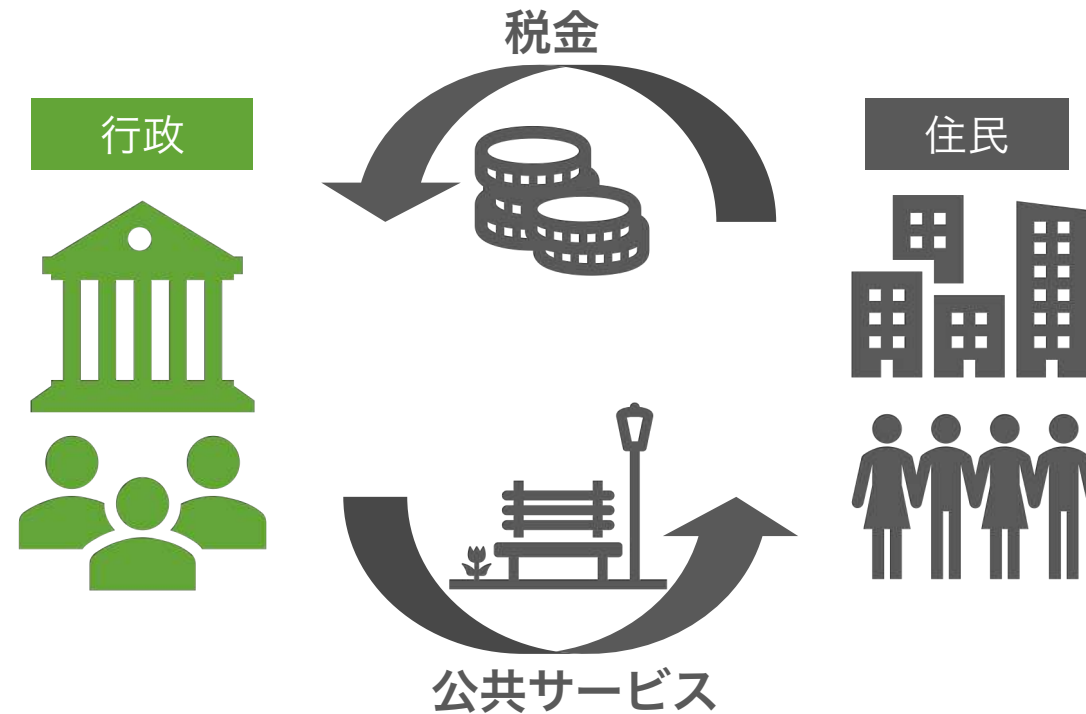
2. 行政の高度化・効率化

3. 透明性・信頼の向上

【出典】 「オープンデータ基本指針」平成29年5月30日  
高度情報通信ネットワーク社会推進戦略本部・官民データ活用推進戦略会議決定

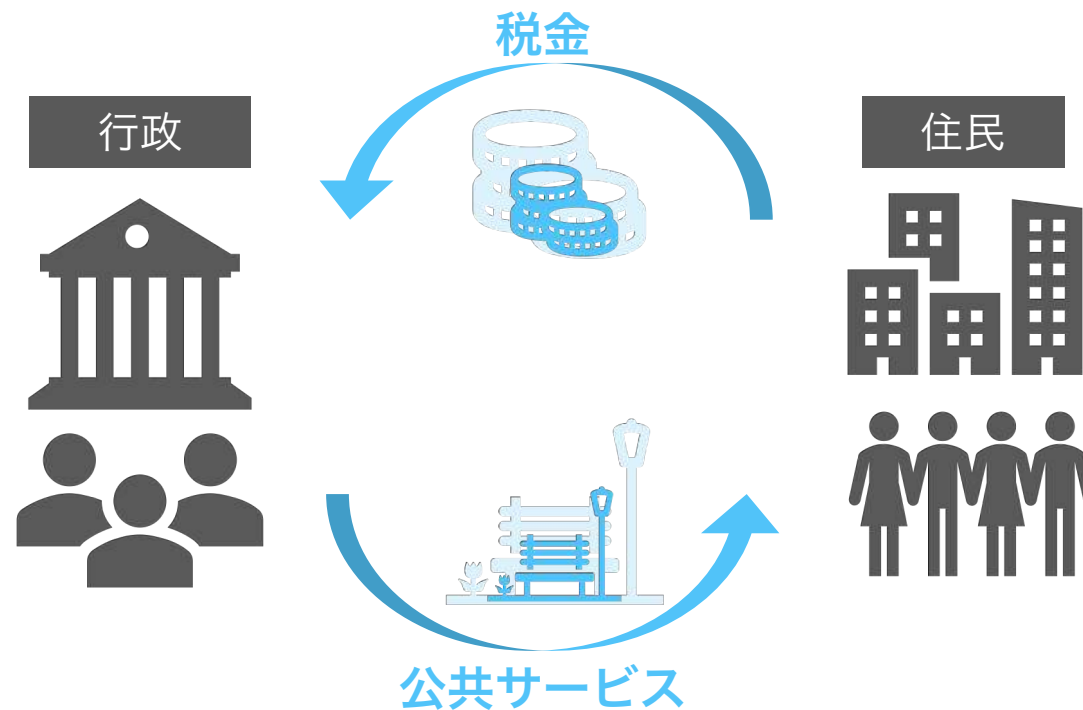
# なぜ官民協働が必要なのか？

これまでは行政が公共サービス運営の主体だった

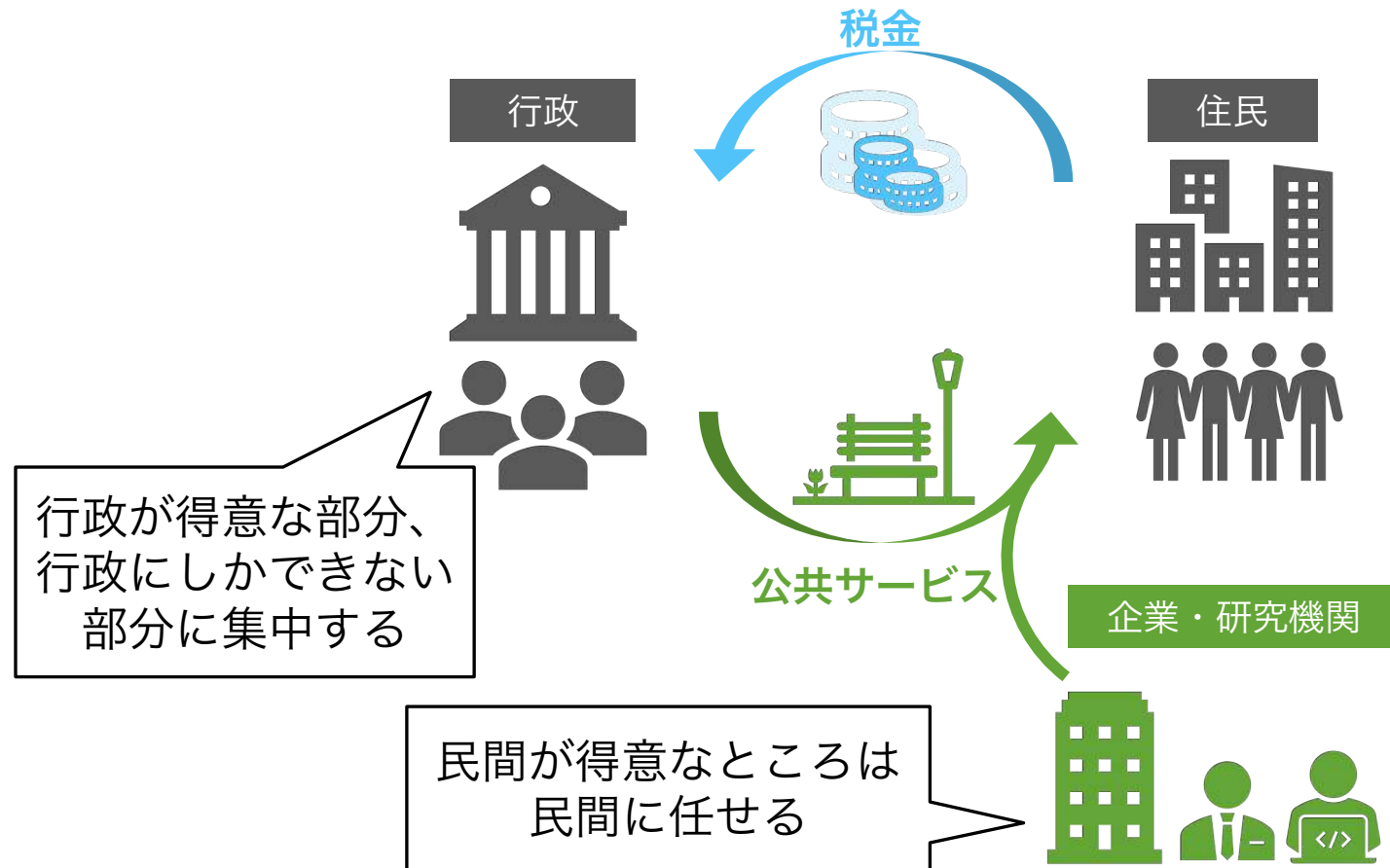


# なぜ官民協働が必要なのか？

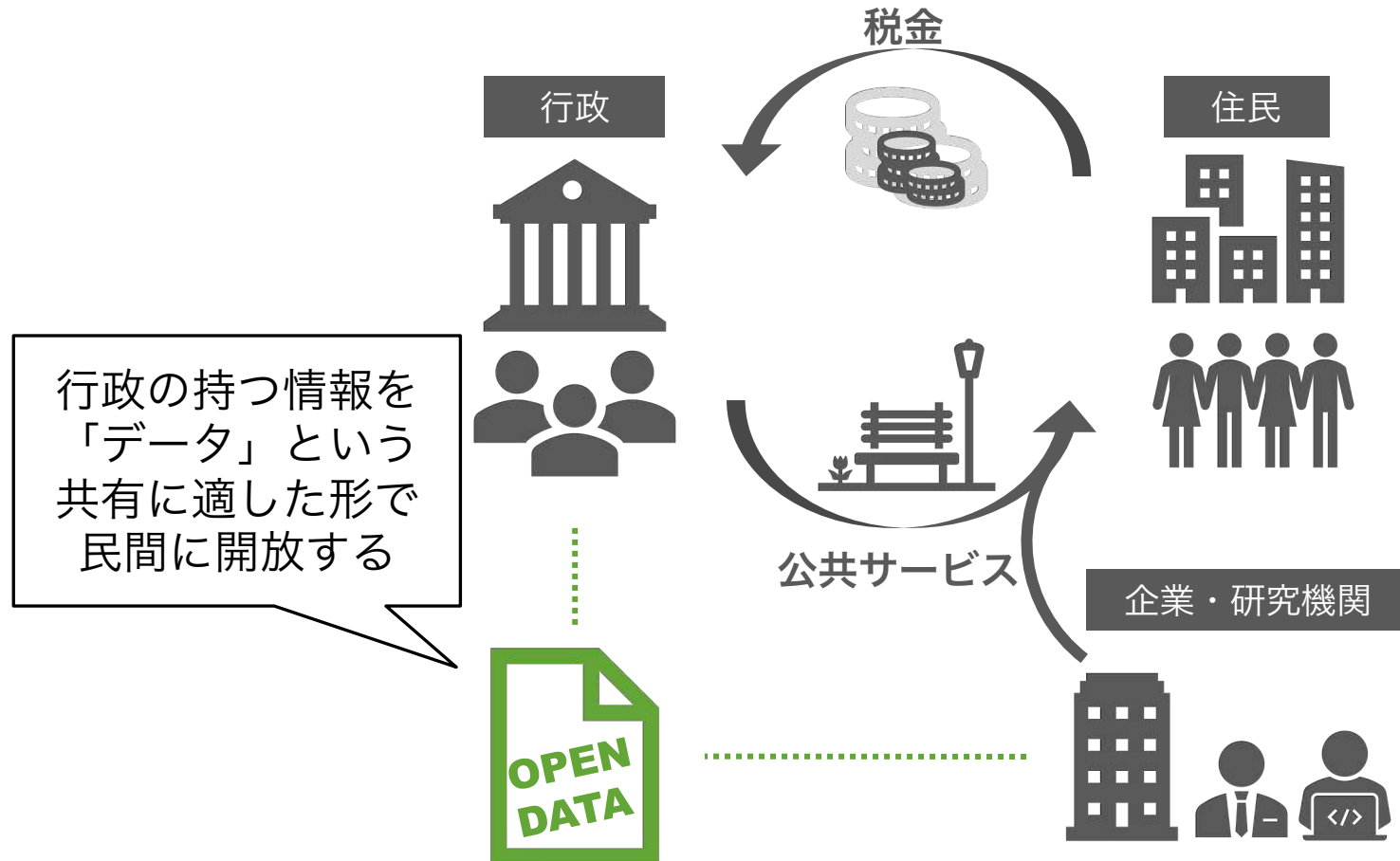
税収が減ると、これまで通りの公共サービスは維持できなくなる



# 官民協働によって地域を持続可能にする



# オープンデータによって、 組織を超えて得意分野を分担できる社会 が実現する



# 事例：「しずみちInfo」 (静岡市 × トヨタIT開発センター)

## 静岡市「しずみちInfo」

- 道路交通情報をオープンデータ化
- 通行規制や災害発生などのデータをリアルタイムで配信



## トヨタIT開発センター

- 共同実験で「しずみちInfo」のデータを取得しカーナビで情報配信

(データ公開元)  
規制関連データ：静岡市建設局道路部道路保全課  
防災関連データ：静岡市総務局危機管理総室



# 事例：「しずみちInfo」

## 従来は必要な情報を必要な人に届けられなかった

BEFORE

行政

チラシや公式Webサイト等で  
情報発信



通行規制  
情報



ドライバー

必要なタイミングで  
必要な情報が得られない



# 事例：「しずみちInfo」 オープンデータでリアルタイムに情報が届 けられるように

AFTER

行政

オープンデータとして  
通行規制情報を公開



通行規制  
オープン  
データ

カーナビ

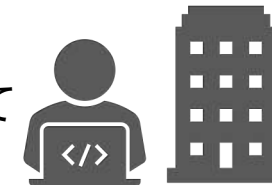
ドライバー

必要なタイミングで  
最新情報が得られる



民間企業

開発している製品に  
オープンデータを載せて  
サービス提供



※静岡市ではカーナビなどの機械から  
オープンデータにアクセスできるように  
するために、API (Application  
Programming Interface) という仕組み  
でデータが公開されています

# オープンデータの意義

---

1. 国民参加・官民協働の推進を通じた  
諸課題の解決、経済活性化

2. 行政の高度化・**効率化**

3. 透明性・信頼の向上

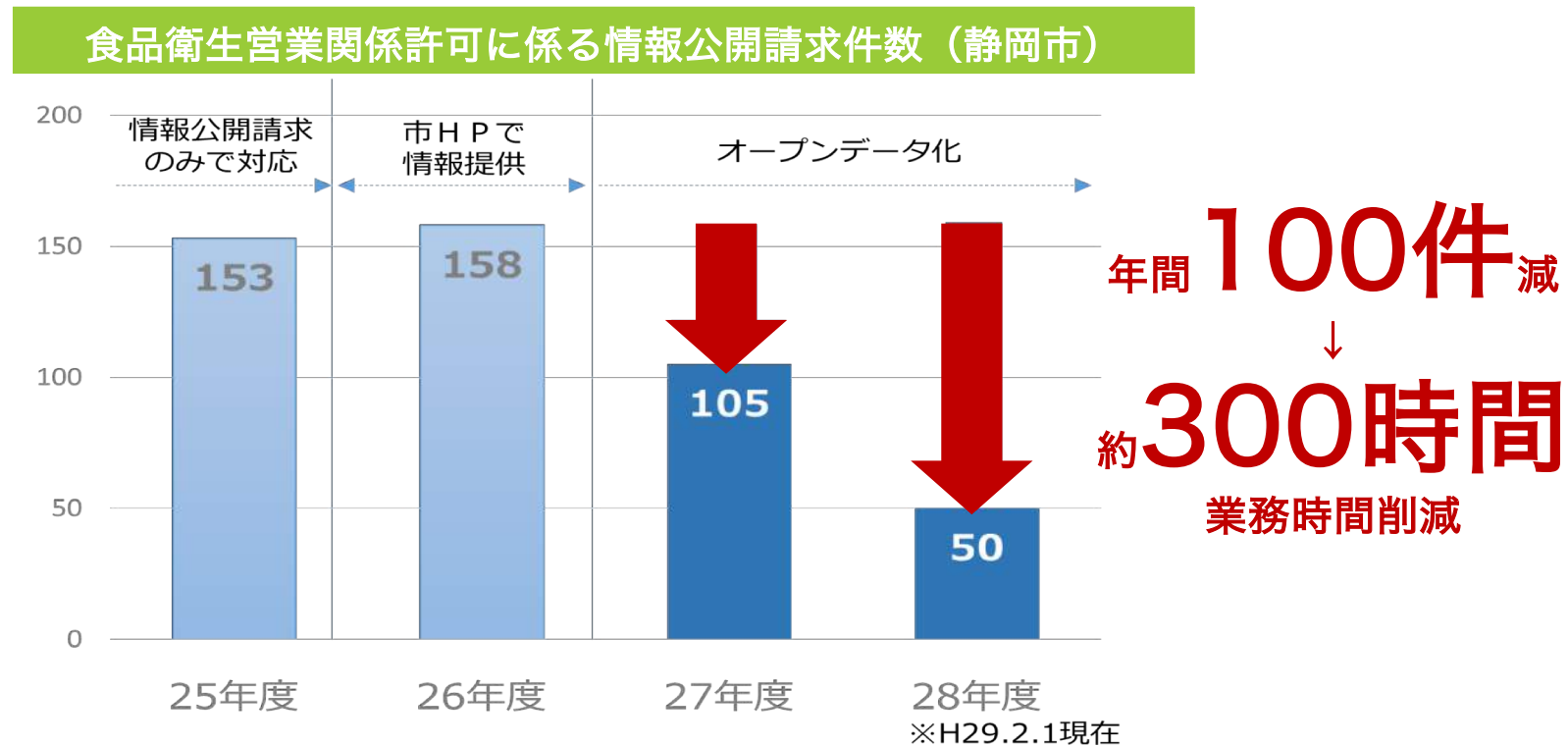
オープンデータ化  
する作業の分、  
仕事が増えるのでは…？



【出典】 「オープンデータ基本指針」平成29年5月30日  
高度情報通信ネットワーク社会推進戦略本部・官民データ活用推進戦略会議決定

# 事例： 情報公開請求の軽減（静岡市）

静岡市では、情報公開請求件数の多かった食品衛生許可に関するデータをオープンデータとして公開することにより、業務負担の軽減に成功



出典：VLED「第4回データ運用検討分科会（2017年2月13日）」静岡市プレゼン資料をもとに作成

# トータルコストが高いのはどちら？ 情報公開請求 VS オープンデータ

静岡市のデータとヒアリング結果※を参考に、費用をシミュレーション

## A. オープンデータ化しない場合

請求件数（1年目）： 150件  
請求件数（2年目）： 150件  
請求件数（3年目）： 150件

請求1件あたりの対応時間： 3時間  
職員の時給： 2,300円

## B. オープンデータ化した場合

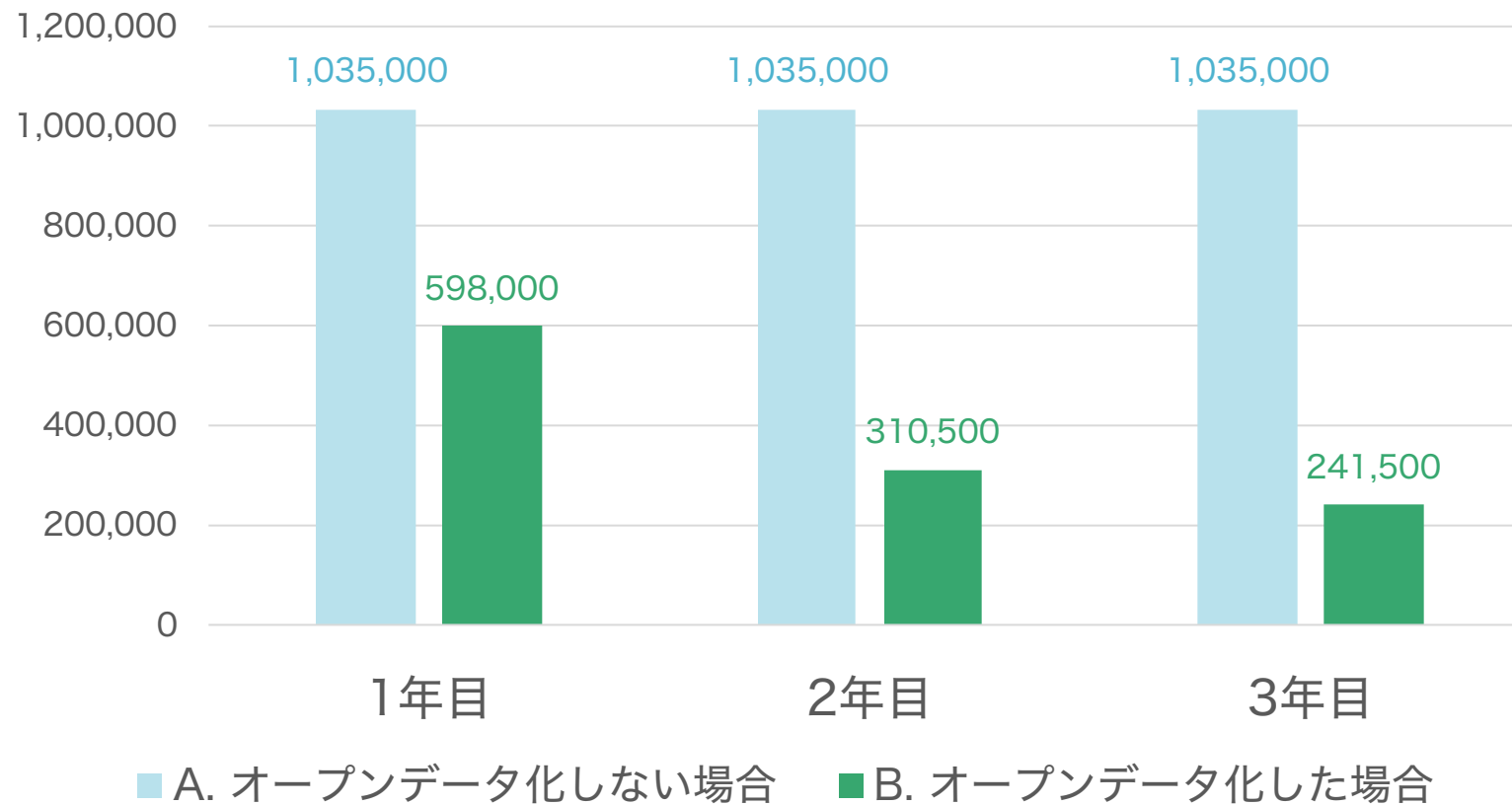
データ作成対応（初期コスト）： 50時間  
データ更新対応（毎月）： 5時間

請求件数（1年目）： 100件  
請求件数（2年目）： 50件  
請求件数（3年目）： 30件

請求1件あたりの対応時間： 1.5時間  
職員の時給： 2,300円

※主に対処時間に関して、静岡市職員/地域情報化アドバイザーの新庄大輔氏に情報提供頂きました

# シミュレーション結果： オープンデータ化することで対応コストを大幅に削減できる

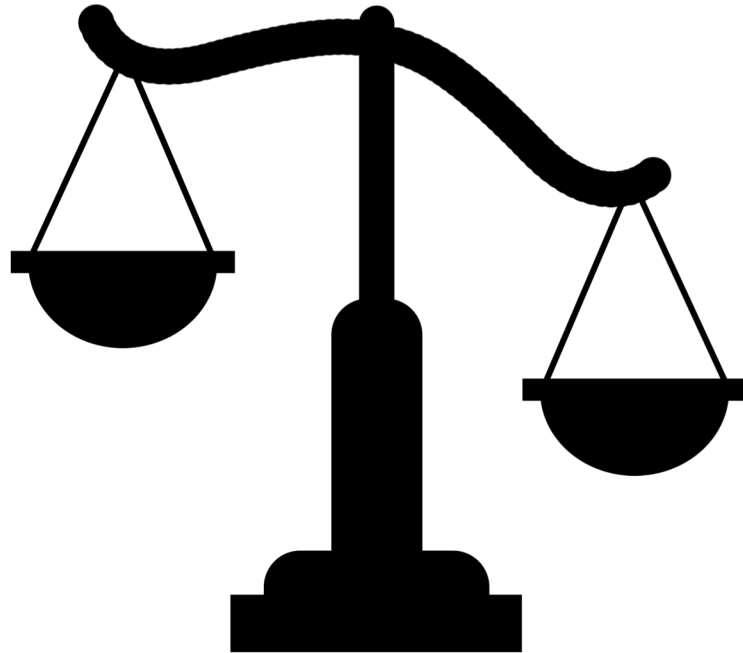


# オープンデータの意義

---

1. 国民参加・官民協働の推進を通じた諸課題の解決、経済活性化
2. 行政の高度化・効率化
3. 透明性・信頼の向上

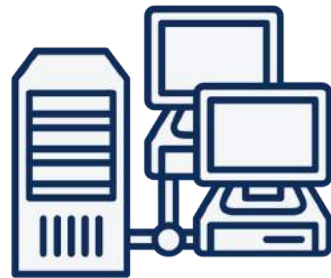
【出典】 「オープンデータ基本指針」 平成29年5月30日  
高度情報通信ネットワーク社会推進戦略本部・官民データ活用推進戦略会議決定



## 透明性・信頼の向上は “共創”を進めるために必要

---

- 共創 = Co-creation
- 共創を進めるためには、行政と民間が対等な立場で話ができる必要がある
- そのためには、行政と民間の「情報の非対称性」を解消する必要がある
- 客観的なデータを共有して、同じものを見ながら話ができるのがベスト



オープンデータは  
庁内でデータを適切に管理できて  
いるかを示す役割もある

---

- そもそも庁内でデータをきちんと管理できていなかったら、公開はできない
- 不正の抑止力にもなることが期待できる



# オープンデータの意義

---

1. 国民参加・官民協働の推進を通じた  
諸課題の解決、経済活性化
2. 行政の高度化・効率化
3. 透明性・信頼の向上

【出典】 「オープンデータ基本指針」平成29年5月30日  
高度情報通信ネットワーク社会推進戦略本部・官民データ活用推進戦略会議決定

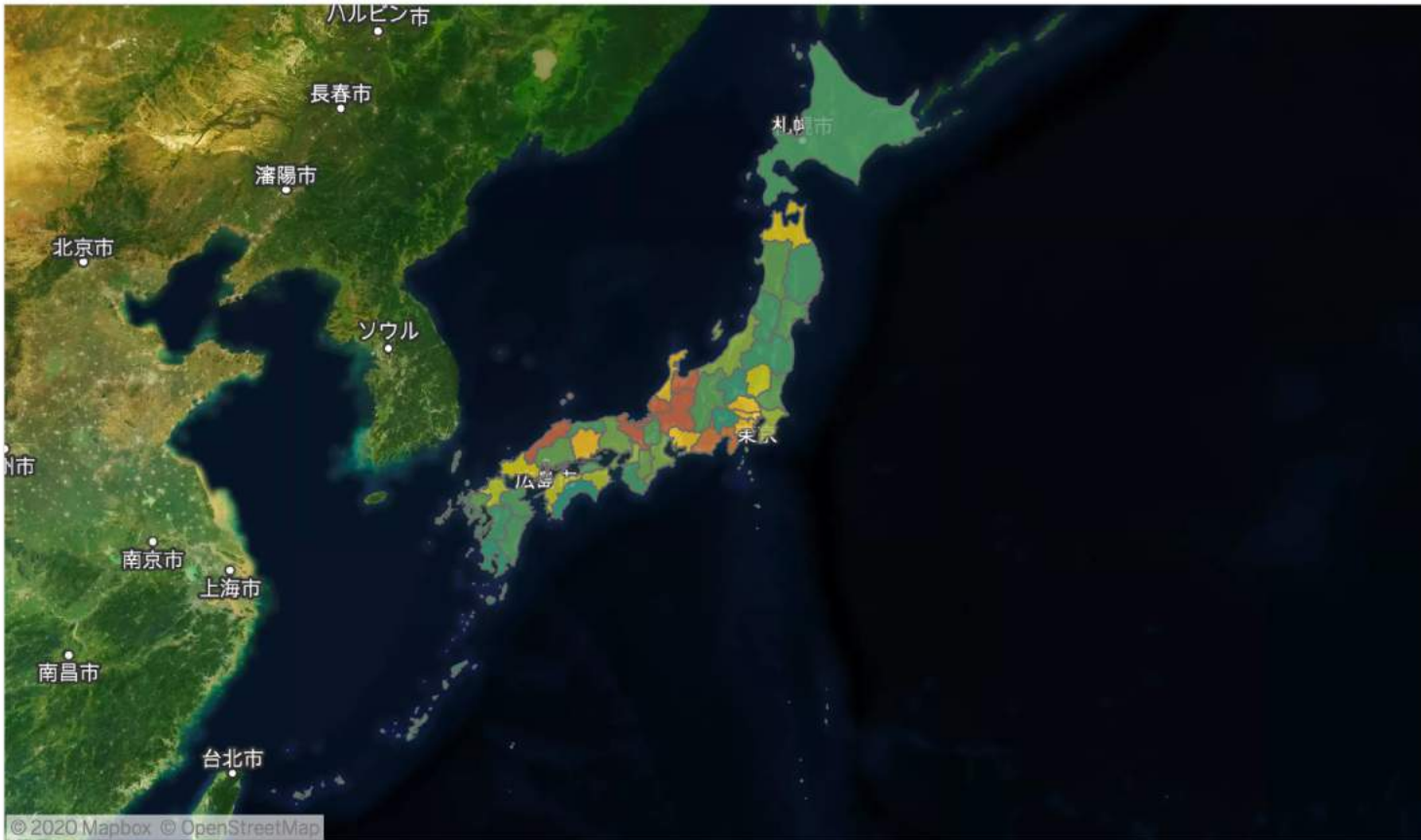
# 日本のオープンデータ取組自治体

データ出典元：CC BY 4.0 政府CIOポータル>オープンデータ>オープンデータ取組自治体一覧（2019年12月16日時点）

<https://cio.go.jp/policy-opendata/#jichitaisuu>

## 622 / 1,742 自治体 (36%)

地図上の都道府県、または棒グラフをクリックすると、選択した都道府県の詳細が表示されます↓



## 都道府県別オープンデータ取組率の比較

1	岐阜県	100%
2	京都府	100%
3	島根県	100%
4	富山県	100%
5	福井県	100%
6	静岡県	94%
7	岡山県	74%
8	愛知県	70%
9	東京都	69%
10	石川県	68%
11	埼玉県	68%
12	神奈川県	67%
13	青森県	60%
14	山口県	53%

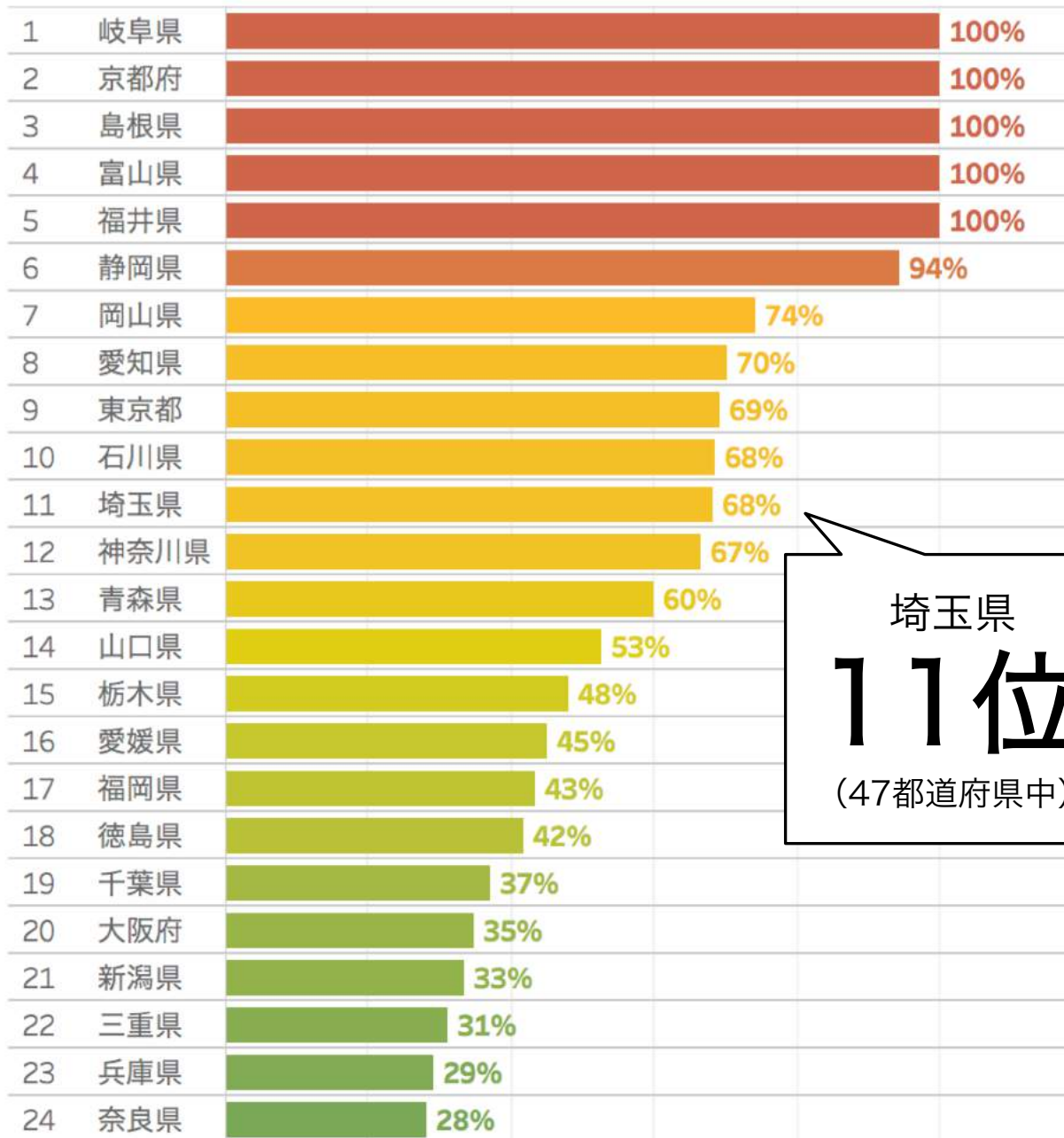
## 都道府県別オープンデータ取組自治体数の比較

(灰色のバーは自治体数総数)

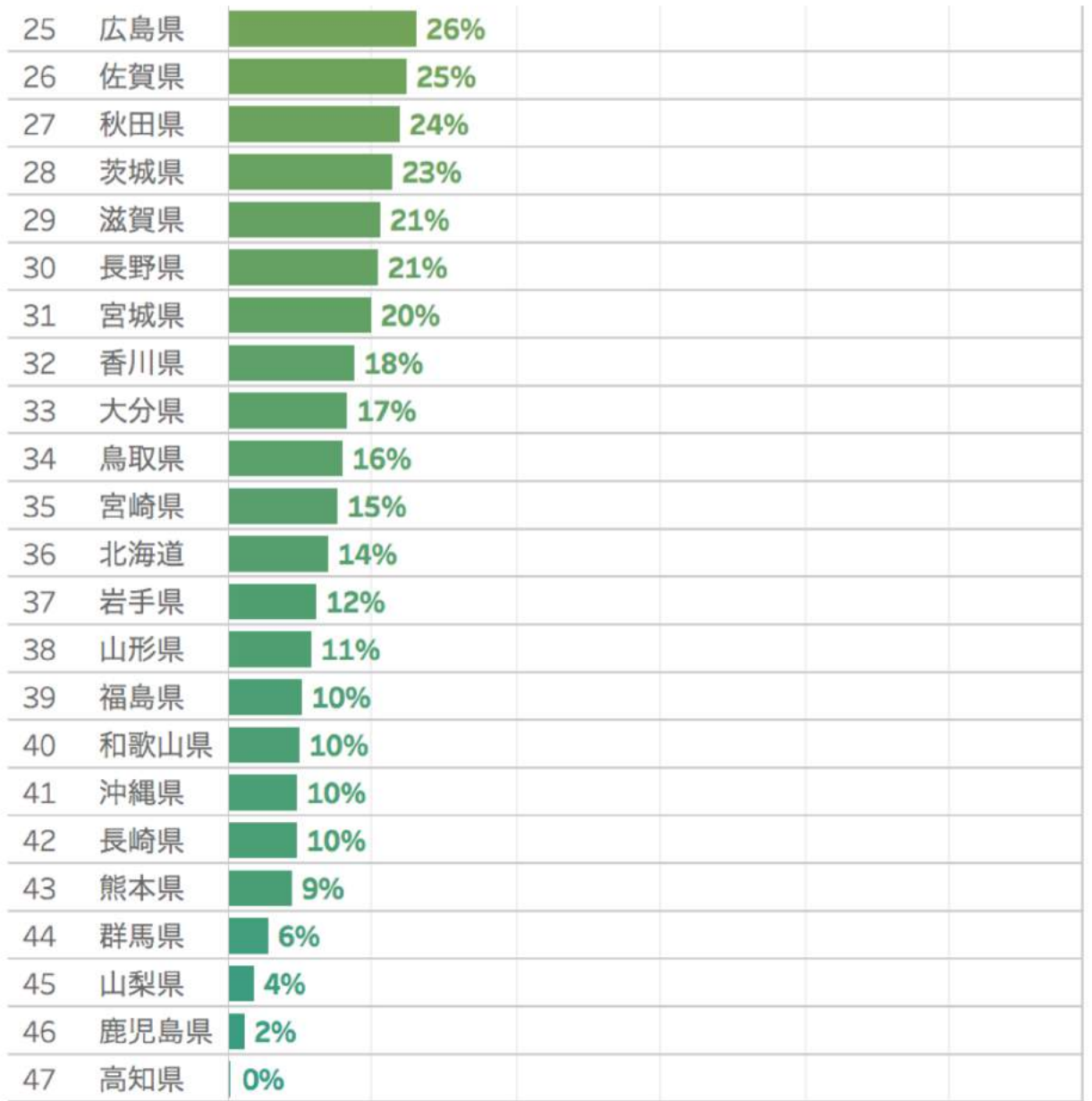
1	埼玉県	43	63
2	東京都	43	62
3	岐阜県	42	
4	愛知県	38	54
5	静岡県	33	
6	京都府	26	
7	福岡県	26	60
8	北海道	25	179
9	青森県	24	40
10	神奈川県	22	33
11	岡山県	20	27
12	千葉県	20	54



<https://public.tableau.com/profile/sayoko.shimoyama.linkdata#!/vizhome/5946/sheet0>



埼玉県  
**11位**  
 (47都道府県中)



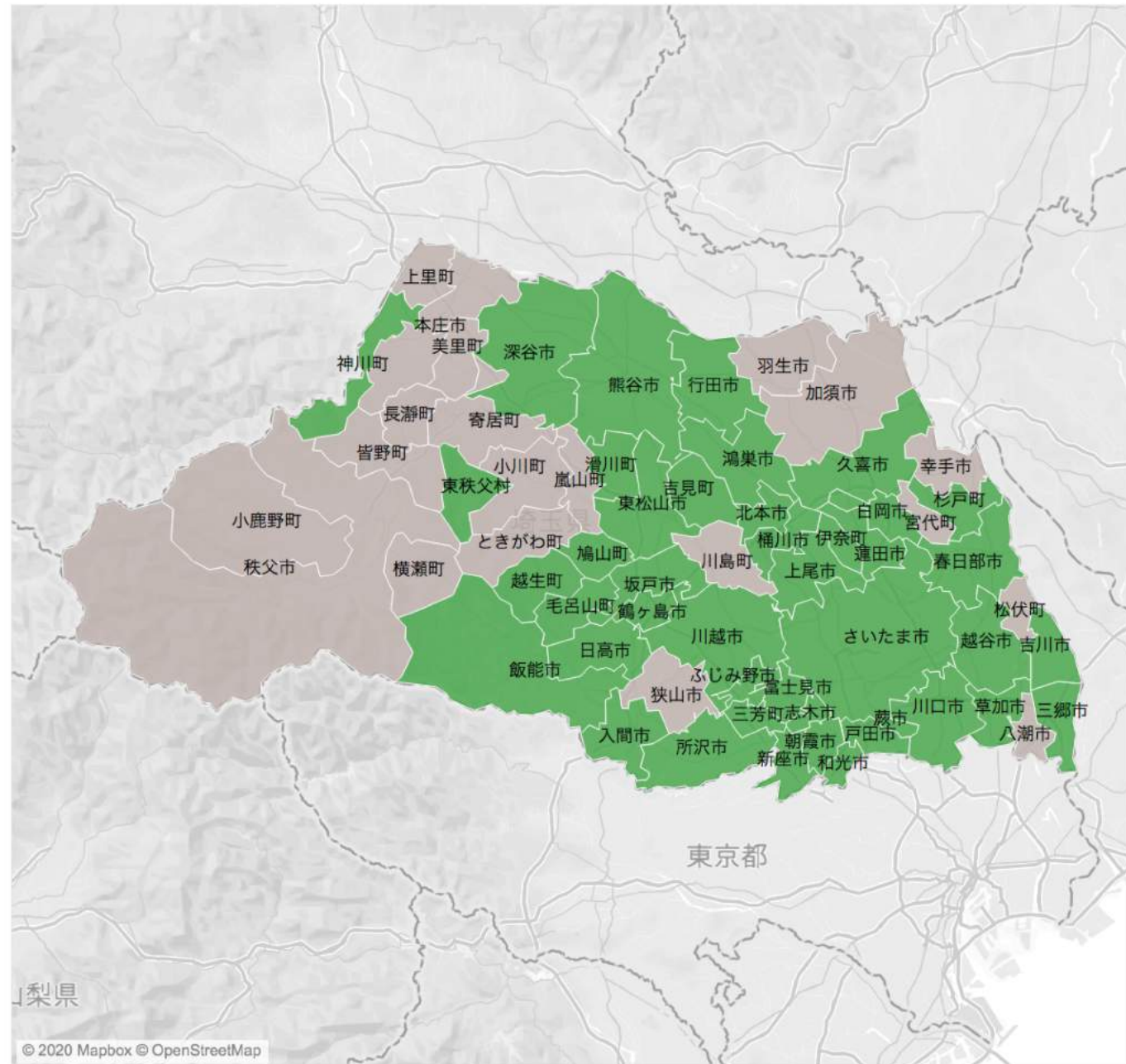
<https://public.tableau.com/profile/sayoko.shimoyama.linkdata#!/vizhome/5946/sheet0>

# 埼玉県内基礎自治体の オープンデータ推進状況 (2019年12月16日時点)

# 68%

(43/63自治体)

[https://public.tableau.com/shared/MBGTPWHPZ?:display\\_count=y&:origin=viz\\_share\\_link](https://public.tableau.com/shared/MBGTPWHPZ?:display_count=y&:origin=viz_share_link)



# 3. 政策立案のための データ分析の流れ



データ



情報

違いを聞かれたら説明できますか？

# 「データ」とは？

## ISO及びJIS規格による定義

"A reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing."

情報の表現であって、  
伝達、解釈または処理に適するように形式化され、  
再度情報として解釈できるもの

ISO/IEC 2382-1:1993 <https://www.iso.org/obp/ui/#iso:std:iso-iec:2382:-1:ed-3:en>

日本工業規格「X0001 情報処理用語-基本用語」<http://kikakurui.com/x0/X0001-1994-01.html>

# 「情報」とは？

## ISO及びJIS規格による定義

"Knowledge concerning objects, such as facts, events, things, processes, or ideas, including concepts, that within a certain context has a particular meaning.

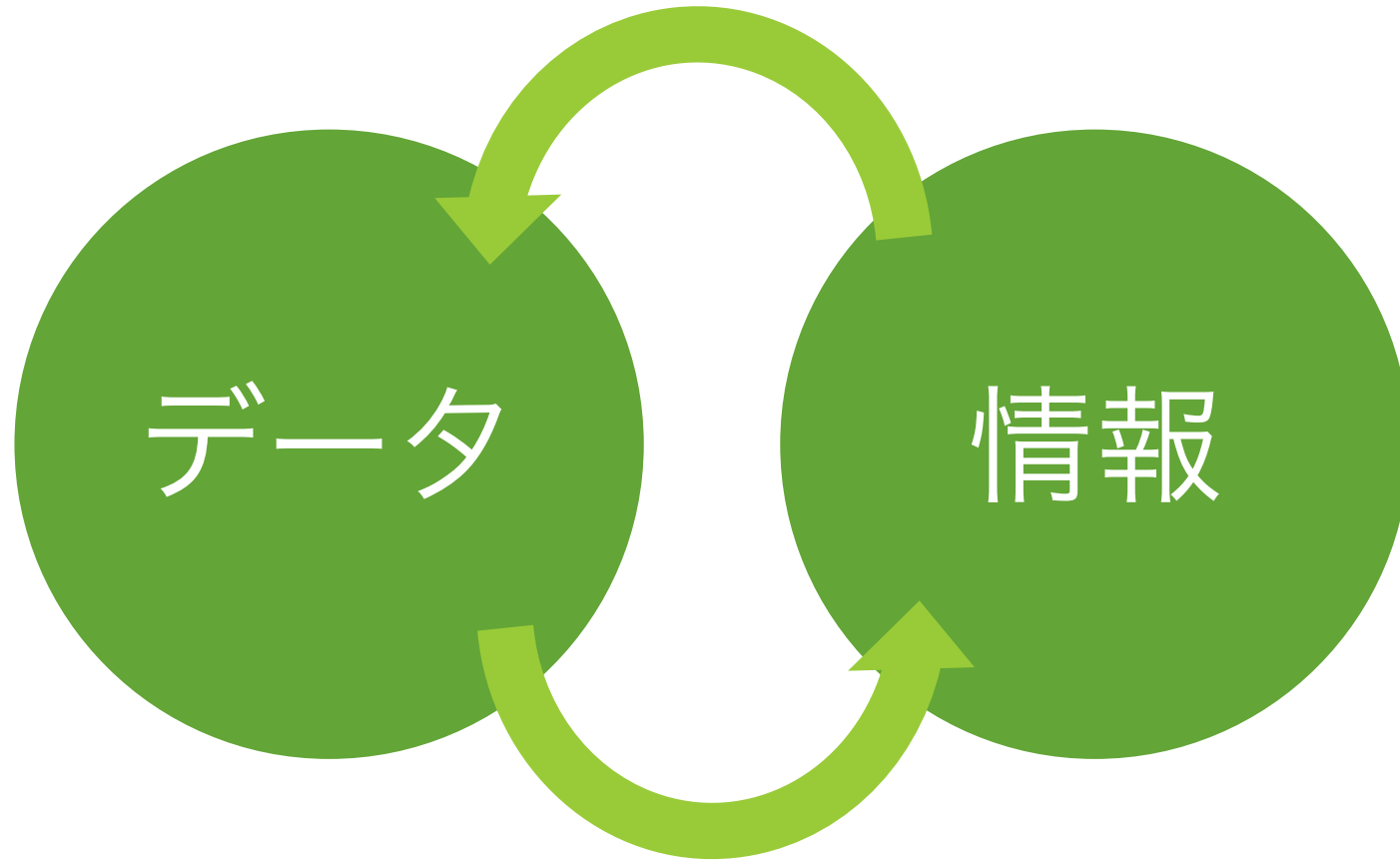
事実，事象，事物，過程，着想などの  
対象物に関して知り得たことであって，  
概念を含み，一定の文脈中で特定の意味をもつもの。

ISO/IEC 2382-1:1993 <https://www.iso.org/obp/ui/#iso:std:iso-iec:2382:-1:ed-3:en>

日本工業規格「X0001 情報処理用語-基本用語」 <http://kikakurui.com/x0/X0001-1994-01.html>



データは、情報を形式化したもの



再度**情報**として解釈できるもの

40

数字

横浜市役所で  
2018年11月23日に開催された  
データ分析ワークショップの  
参加登録者数は  
**40**人です。

## 情報

ワークショップ 開催日	ワークショップ 開催場所	参加登録者 数
2018-11-23	横浜市役所	40

データ

# 情報⇔データ 双方向に変換可能

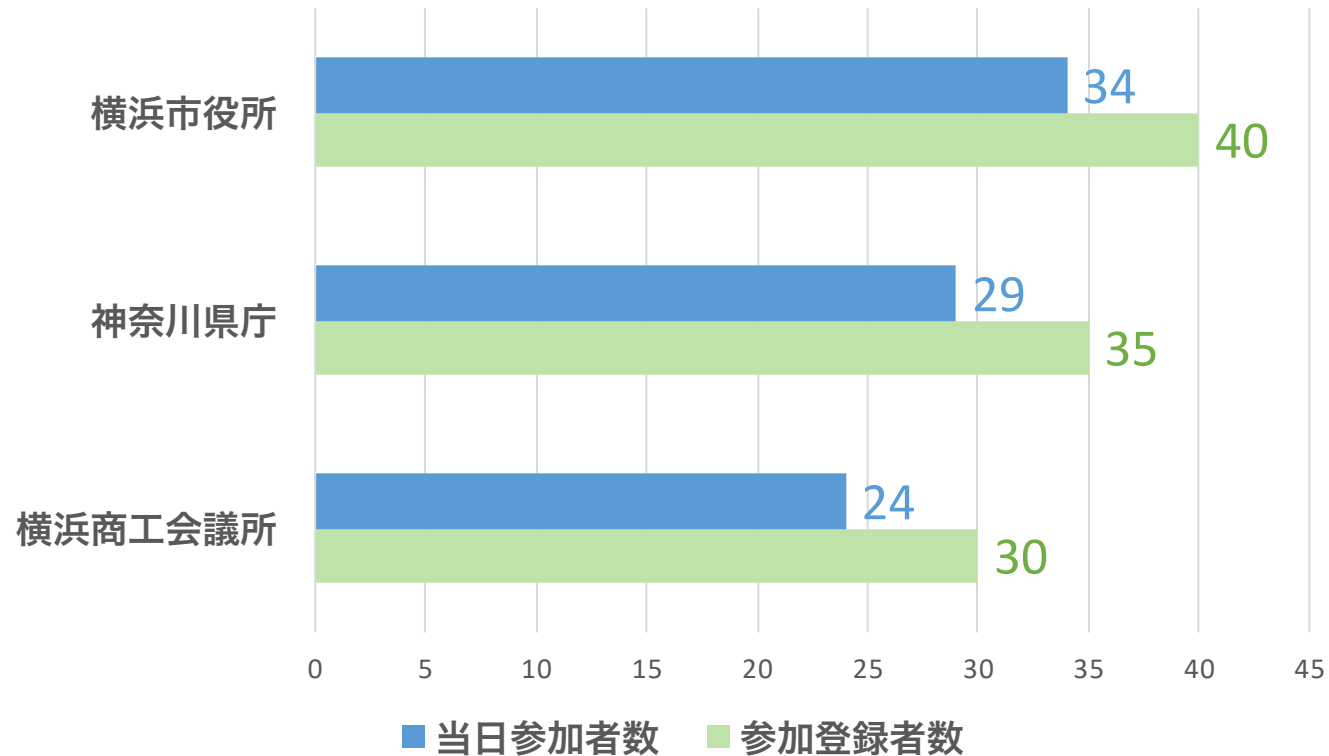
横浜市役所で  
2018年11月23日に開催された  
データ分析ワークショップの  
参加登録者数は  
**40**人です。



ワークショップ 開催日	ワークショップ 開催場所	参加登録 者数
2018-11-23	横浜市役所	<b>40</b>

# データ化することで 他のデータと組み合わせられる

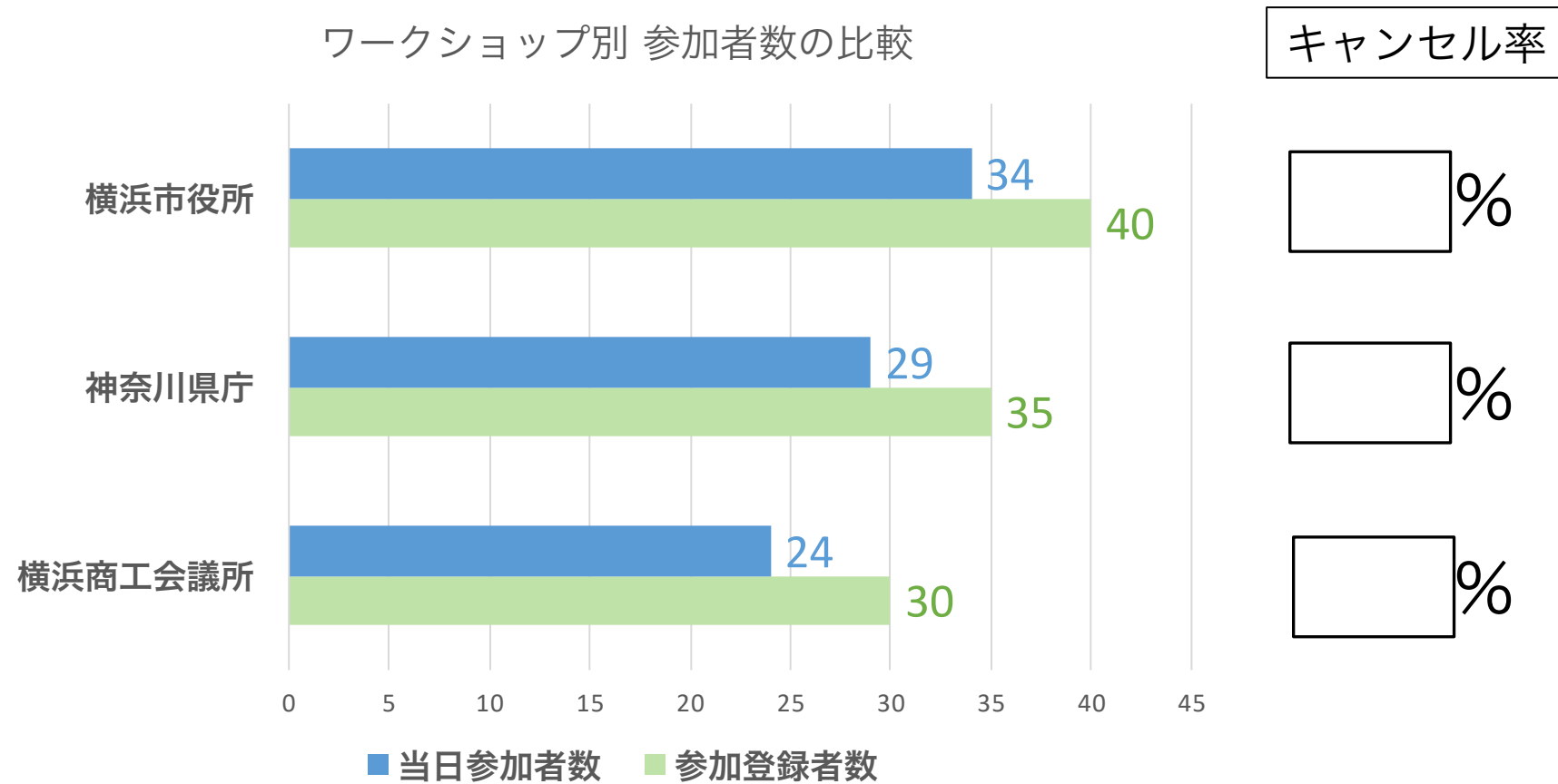
ワークショップ別 参加者数の比較



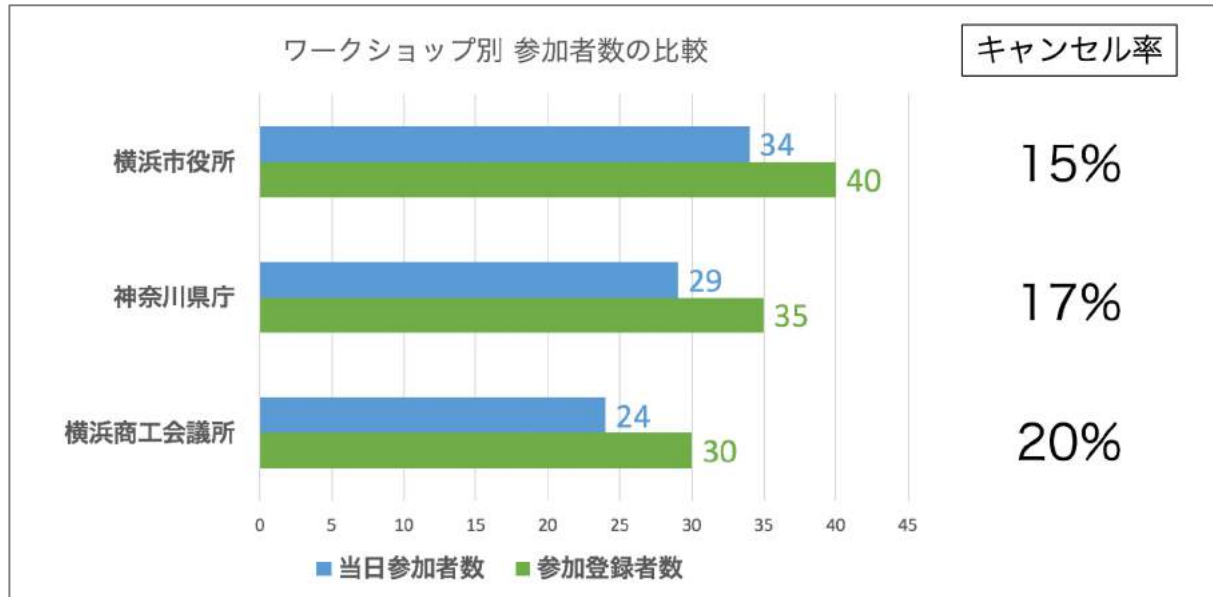
新しい「情報」が  
取り出せる

Q：各ワークショップの  
「キャンセル率」は  
どのくらい？

# データ分析 = データから情報を引き出すこと



# データ分析は 目的を定めておかないと意味がない



誰にどんな情報を提供するのか、  
誰のどんな課題を解決したいのか、  
データ分析をする前に設定しておくことが重要

## イベント運営者



次から  
定員+15~20%  
くらい多めに申込を  
受け付けても  
大丈夫そう!

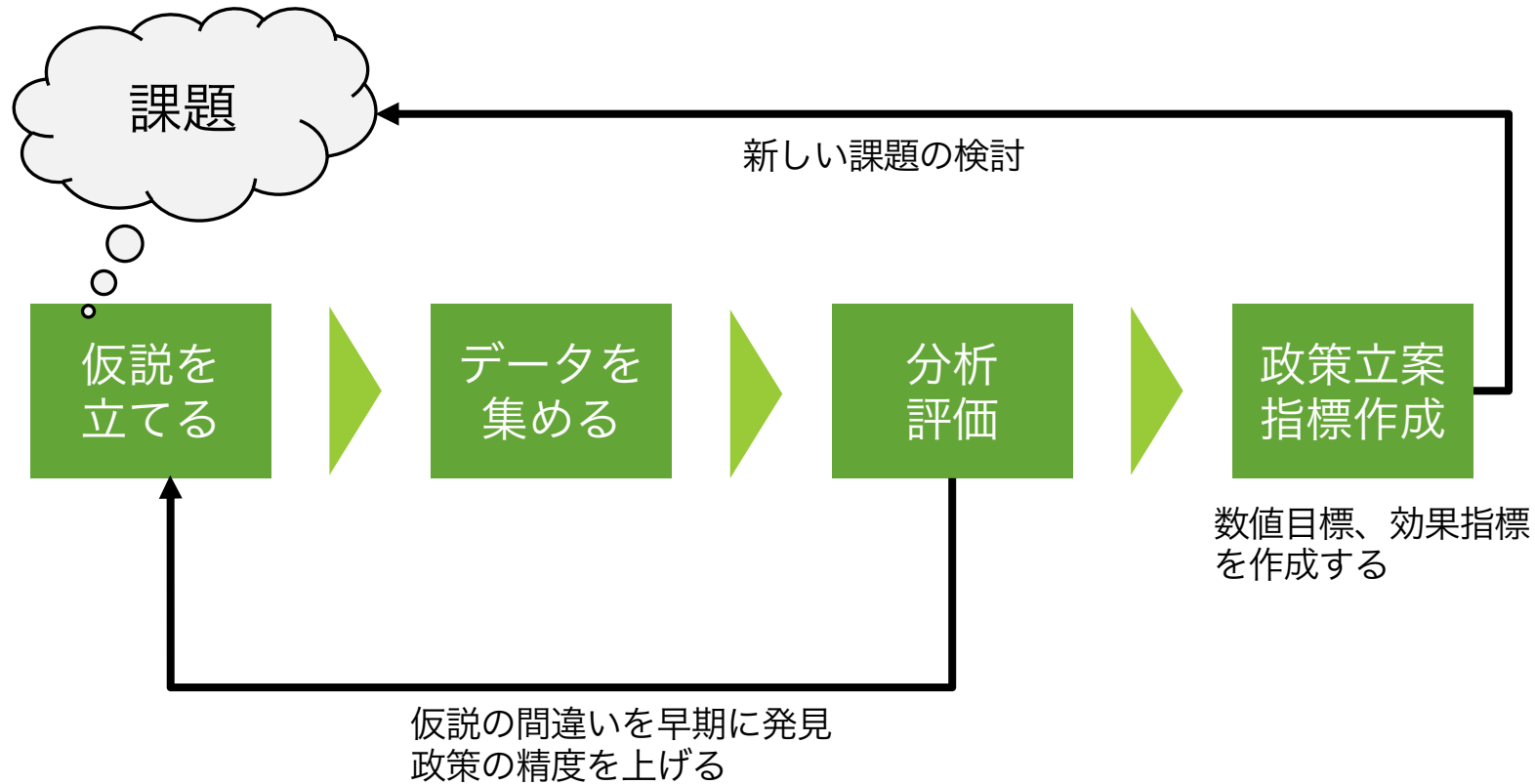
## その他の人



何の役に立つ  
情報なんだろう?

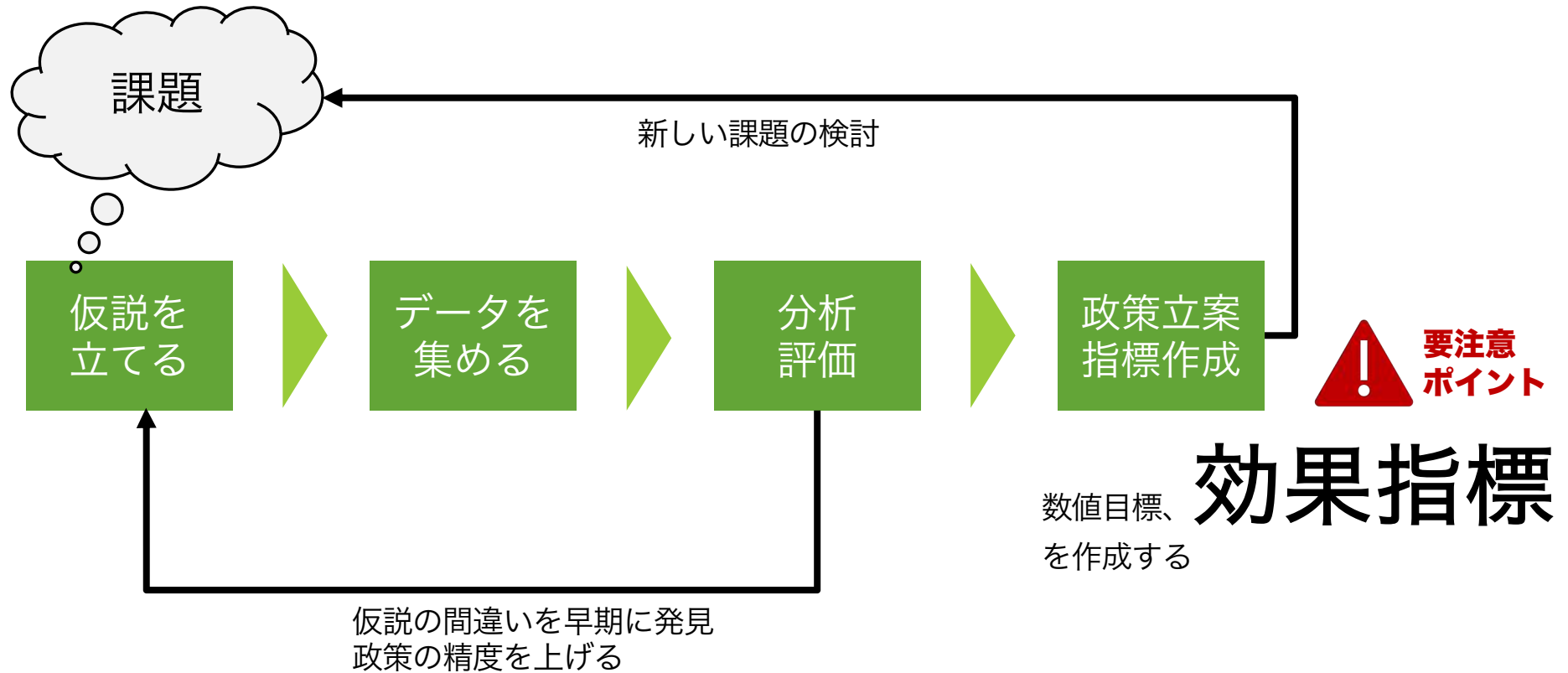


# データ分析による政策反映の流れ



出典元：総務省 地方公共団体におけるデータ活用ガイドブック Ver. 2.0  
別添資料4. データアカデミー（データ分析編）の教材・資料等, CC BY 2.1 JP

# データ分析による政策反映の流れ



出典元：総務省 地方公共団体におけるデータ活用ガイドブック Ver. 2.0  
別添資料4. データアカデミー（データ分析編）の教材・資料等, CC BY 2.1 JP

# 効果指標：定性的効果をどう測る？

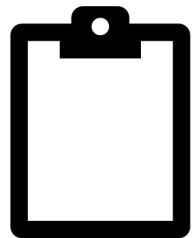
住みやすさ

まちの  
にぎわい

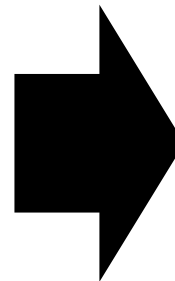
老後の安心

子育ての  
しやすさ

男女平等



アンケートで聞く  
という方法がよく取られる



指標の1つにするのはOK



アンケート結果のみで測ると  
客観性に欠ける場合がある

ポイント：  
定性的効果は  
定量的効果に  
分解して  
指標化する

例：住みよさランキング（東洋経済）

■「住みよさランキング」総合評価トップ50(1~25位)

順位	都市名(都道府県名)	総合評価 偏差値	順位			
			安心度	利便度	快適度	富裕度
1	白山(石川)	54.231	116	308	8	77
2	文京区(東京)	54.099	535	9	53	68
3	野々市(石川)	54.052	121	11	182	488
4	福井(福井)	54.039	165	42	213	67
5	倉吉(鳥取)	53.873	5	60	276	654
6	敦賀(福井)	53.756	104	110	154	130
7	黒部(富山)	53.734	150	626	5	35
8	能美(石川)	53.730	84	632	11	70
9	魚津(富山)	53.687	192	24	339	92
10	駒ヶ根(長野)	53.652	33	92	209	376
11	四万十(高知)	53.625	46	18	22	775
12	郡上(岐阜)	53.462	37	227	114	317
13	かほく(石川)	53.332	127	117	52	290
14	印西(千葉)	53.224	205	161	229	37
15	七尾(石川)	53.153	16	89	582	332
16	小松(石川)	53.144	125	127	306	103
17	北杜(山梨)	53.140	201	105	7	413
18	渋谷区(東京)	53.129	798	1	117	71
19	合志(熊本)	53.096	29	804	15	181
20	金沢(石川)	53.077	300	32	175	184
21	にかほ(秋田)	53.058	25	310	387	223
22	下松(山口)	52.988	369	106	6	204
23	大阪(大阪)	52.931	796	7	2	111
24	新宿区(東京)	52.893	790	3	70	71
25	武蔵野(東京)	52.884	793	4	104	62

TOYOKEIZAI ONLINE

# 定性的効果は定量的効果に分解して指標化する

## 例：住みよさランキング2019（東洋経済）利用データ

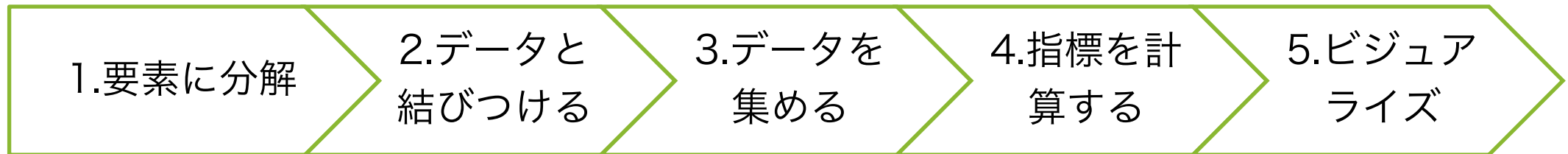
住みよさ	A. 安心度	1 人口当たり病院・一般診療所病床数（2017年10月）：厚生労働省「医療施設調査」
		2 老年人口当たり介護老人福祉・保健施設定員数（2017年10月）：厚生労働省「介護サービス施設・事業所調査」
		3 20～39歳女性人口当たり0～4歳児数（2018年1月）：総務省「住民基本台帳に基づく人口、人口動態及び世帯数」
		4 子ども医療費助成（対象年齢・所得制限の有無）（2019年6月）：東洋経済調べ
		5 人口当たり刑法犯認知件数（2017年）：各都道府県警察調べ
		6 人口当たり交通事故件数（2017年）：交通事故総合分析センター調べ
	B. 利便度	7 人口当たり小売販売額（2015年）：総務省・経済産業省「経済センサス活動調査」
		8 人口当たり大規模小売店店舗面積（2018年）：東洋経済「全国大型小売店総覧」
		9 可住地面積当たり飲食料品小売事業所数（2016年6月）：総務省・経済産業省「経済センサス活動調査」
		10 人口当たり飲食店数（2016年6月）総務省・経済産業省「経済センサス活動調査」
	C. 快適度	11 転出入人口比率（2015～2017年）：総務省「住民基本台帳に基づく人口、人口動態及び世帯数」
		12 人口当たり財政歳出額（2017年度）：総務省「市町村別決算状況調」
		13 水道料金（2019年6月）：東洋経済調べ
		14 汚水処理人口普及率（2018年3月）：国土交通省、農林水産省、環境省調べ
		15 都市計画区域人口当たり都市公園面積（2017年3月）：国土交通省「都市公園整備水準調書」
		16 気候（月平均最高・最低気温、日照時間）（1981～2010年）：気象庁「メッシュ平年値データ」
	D. 富裕度	17 財政力指数（2017年度）：総務省「市町村別決算状況調」
		18 1事業所当たり売上高（2016年6月）：総務省・経済産業省「経済センサス活動調査」
		19 納税義務者1人当たり所得（2017年）：総務省「市町村税課税状況等の調」
		20 1住宅当たり延べ床面積（2013年10月）：総務省「住宅・土地統計調査」
		21 持ち家世帯比率（2015年10月）：総務省「国勢調査」
		22 住宅地平均地価（2018年7月）：国土交通省「都道府県地価調査」

# 指標は定期的に見直して改善することも重要

## 例：住みよさランキング2018（東洋経済）利用データ

住みよさ	安心度	1	病院・一般診療所病床数（人口当たり）：2016年10月1日現在 厚生労働省「医療施設調査」
		2	介護老人福祉施設・介護老人保健施設定員数（65歳以上人口当たり）：2016年10月1日現在 厚生労働省「介護サービス施設・事業所調査」
		3	出生数（15～49歳女性人口当たり）：2016年 総務省「住民基本台帳に基づく人口、人口動態」
		4	年少人口（0～14歳人口）増減率：2017年1月1日÷2014年1月1日 総務省「住民基本台帳に基づく人口、人口動態」
	利便度	5	小売業年間商品販売額（人口当たり）：2016年6月 総務省・経済産業省「経済センサス活動調査（商業統計）」
		6	大型小売店店舗面積（人口当たり）：2017年5月 東洋経済「全国大型小売店総覧」
		7	飲食料品小売事業所数（可住地面積当たり）：2016年6月1日 総務省・経済産業省「経済センサス活動調査（商業統計）」
	快適度	8	汚水処理人口普及率：2017年3月末 国土交通省・農林水産省・環境省「汚水処理人口普及状況」、各都道府県資料
		9	都市公園面積（人口当たり）：2016年3月末 国土交通省調べ
		10	転入・転出人口比率：2014～2016年 総務省「住民基本台帳に基づく人口、人口動態」
		11	新設住宅着工戸数（世帯当たり）：2014～2016年度 国土交通省「建築着工統計調査」
	富裕度	12	財政力指数：2016年度 総務省「市町村別決算状況調」
		13	地方税収入額（人口当たり）：2016年度 総務省「市町村別決算状況調」
		14	課税対象所得額（納税義務者1人当たり）：2016年 総務省「市町村税課税状況等の調」
	住居水準 充実度	15	住宅延べ床面積（1住宅当たり）：2013年10月1日 総務省「住宅・土地統計調査」
		16	持ち家世帯比率：2015年10月1日 総務省「国勢調査」

# 指標作成のステップ



構成要素として考えられるものを列挙してツリー構造になるように整理する

それぞれの構成要素を測るためにどのようなデータが必要か考える

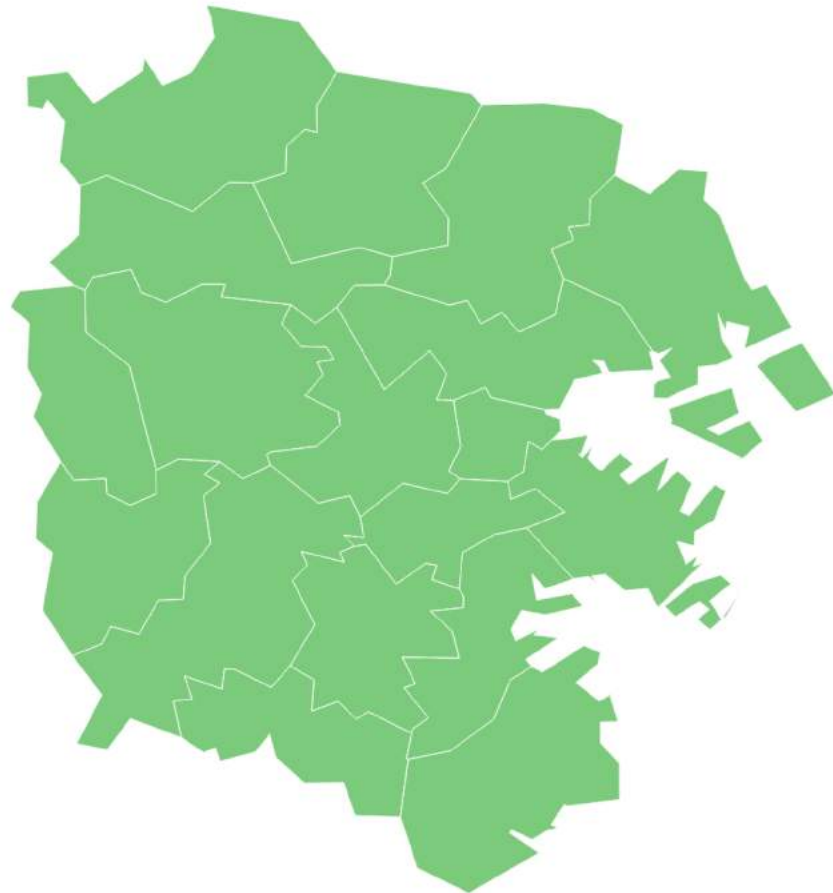
2. で検討したデータが実際にあるか探して収集する

収集したデータを用いて、比較対象別に指標を計算する

算出した指標値を比較しやすくするために可視化する

# 指標設計ワークショップにおける テーマの設定例

---

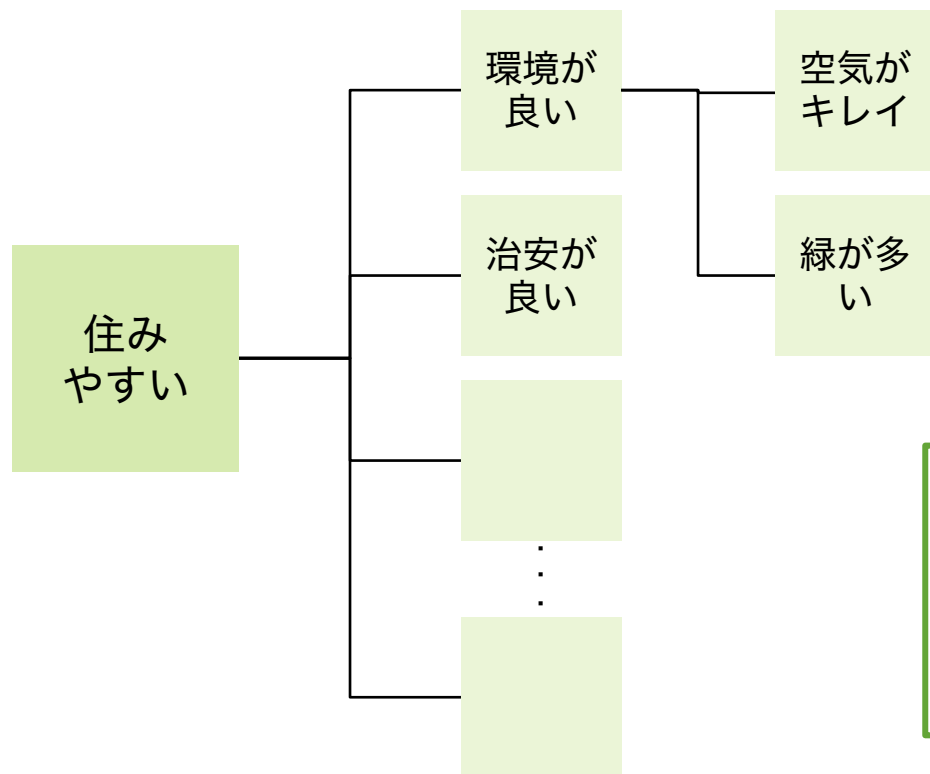


横浜市18区の  
「住みやすさ」  
を比較する  
指標を作ろう



# Step 1. 要素に分解

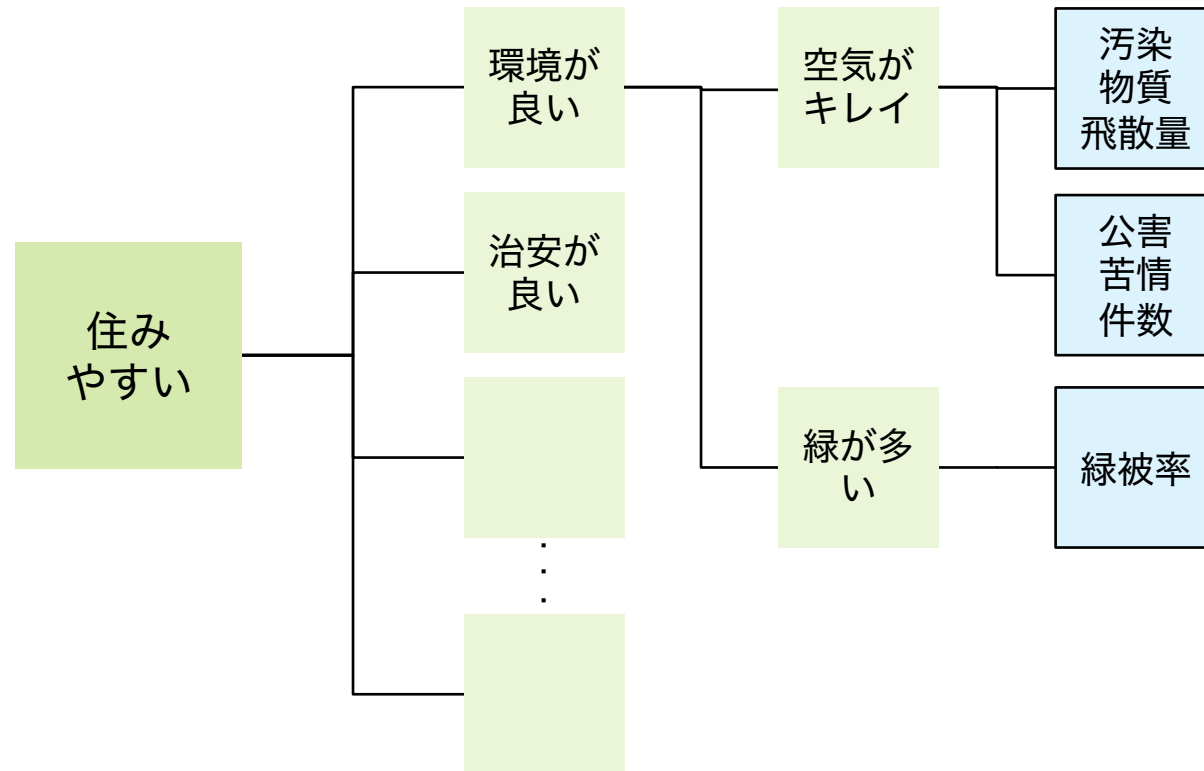
「住みやすい」という状況はどのような要素で成り立っているか、付箋紙を使って挙げてみましょう



- 「これが正解」というものは無いので、思いつくものをどんどん挙げてみましょう
- ここでは一旦データがあるかどうかは気にしなくてOKです

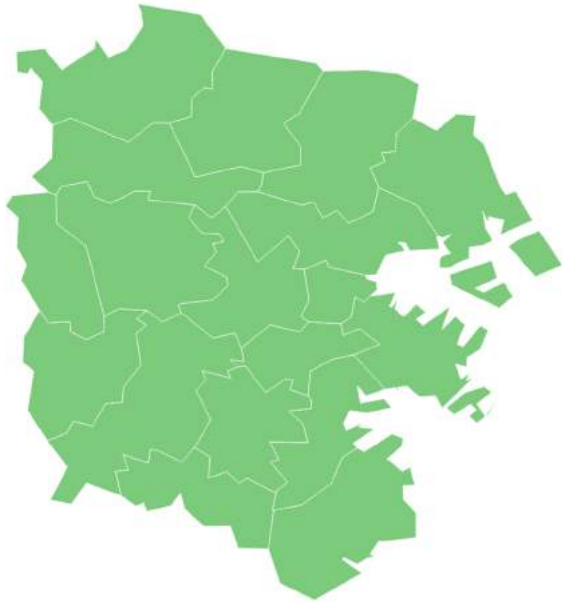
# Step2. データと結びつける

Step1で挙げた各要素は、どんなデータで測ることができそうか  
挙げてみましょう



# Step3. データを集める

---



- Step2で挙げたデータが実際にあるか探してみましょう
- 見つけたデータが18区別のデータになっているか、確認しましょう
- データを探す中で、「このデータも使えそう！」というものがあったら追加してもOK

# Step4. 指標を計算する

1. 指標に使うデータの種類ごとに各区の **偏差値** を計算する
2. 各区の **総合偏差値** を計算する

(計算のイメージ)

	偏差値_緑被率	偏差値_汚染物質	偏差値_XXX	総合偏差値
中区	41.23	55.93	...	52.55
西区	34.24	37.71	...	34.89
南区	58.13	35.03	...	53.76
...	...	...	...	...

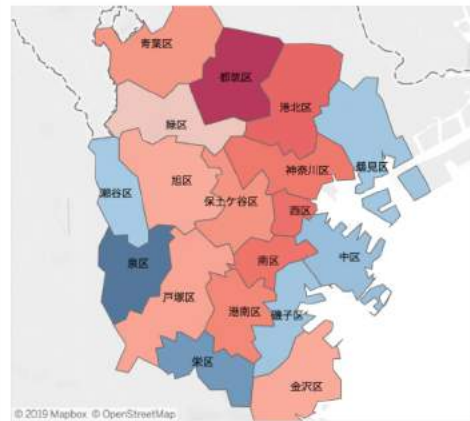
# Step5. ビジューアライズ

算出した指標値を比較しやすくするために可視化する

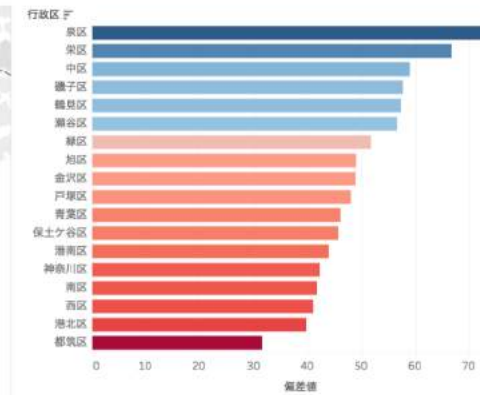
※横浜情報科学専門学校の学生チームによる「横浜市18区の住みやすさ」指標の可視化例

<https://public.tableau.com/profile/.48315447#!/vizhome/7877/sheet7>

住みやすい区(地図)



住みやすい区の偏差値



出典元  
横浜市 統計ポータル 横浜市統計書  
第2章 人口  
3 行政区別世帯数及び人口  
16 人口の行政区別状況(4)労働力状態、男女別15歳以上人口  
<https://www.city.yokohama.lg.jp/city-info/yokohamashi/tokei-chosa/portal/tokeisho/02.html>  
第15章 衛生及び環境  
1 医療施設(1)医療施設数  
5 公園緑地(2)行政区別都市公園数及び面積  
<https://www.city.yokohama.lg.jp/city-info/yokohamashi/tokei-chosa/portal/tokeisho/15.html>  
第17章 司法及び治安  
6 交通事故発生状況(1)行政区別  
<https://www.city.yokohama.lg.jp/city-info/yokohamashi/tokei-chosa/portal/tokeisho/17.html>

指標を選んだ理由  
病院と一般診療所の数→病院が近くにあると急な病気や体調不良にも迅速に対応でき、数が多いとインフルエンザが流行するよう時期でも待ち時間を相対的に減らすことができるから  
公園面積の割合→緑のある区の指標になり、遊樂場所にもなり得るため高ければ高いほど住みやすい区と考えたから  
1件あたりの犯罪件数→治安の悪い所には安心して住めないから  
労働者人口に対する失業者の割合→これが高い区は仕事をするには不都合な区であり、住みにくい原因の一つになると考えたから  
人口密度→通勤地帯に住むのは負担だと考えたから(便利な区に人が集中するという見方もできるから、プラス評価でも良かったかもしれない)

気づいたこと  
住みやすい区トップ3は南区、中央区、ワースト3は西区、港北区、都筑区だった。  
偏差値45以上55以下の区が6区しかなく、住みやすい区と住みにくい区が明確に別れた。

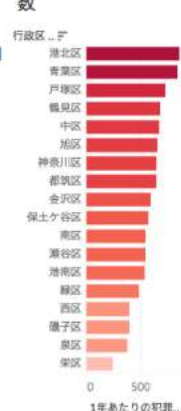
病院と一般診療所の合計



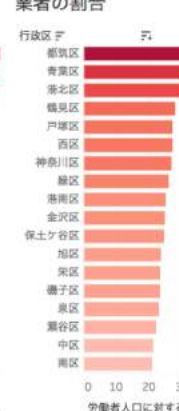
面積に対する公園面積の割合



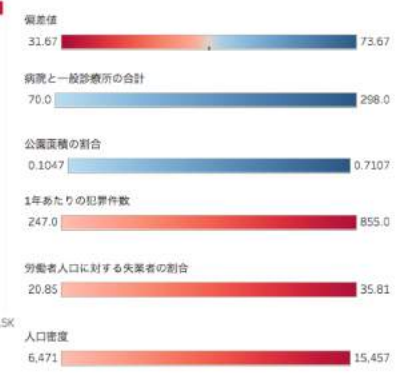
1年あたりの犯罪件数



労働者人口に対する失業者の割合



人口密度



## 指標設計体験ワーク

次のうち一番興味のあるテーマを選んでください

---

住みやすさ

まちの  
にぎわい

老後の安心

子育ての  
しやすさ

男女平等

# 指標設計体験ワーク

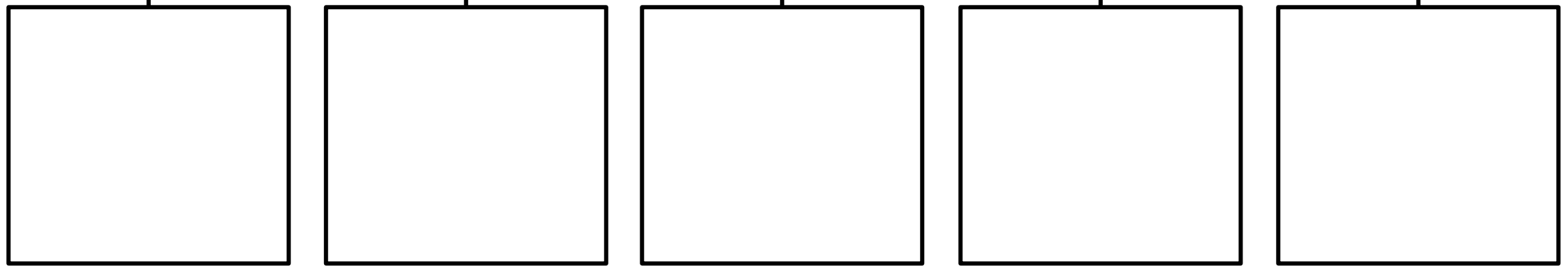
テーマ

※今回はStep2まで

Step1：要素に分解



Step2：データと結びつける

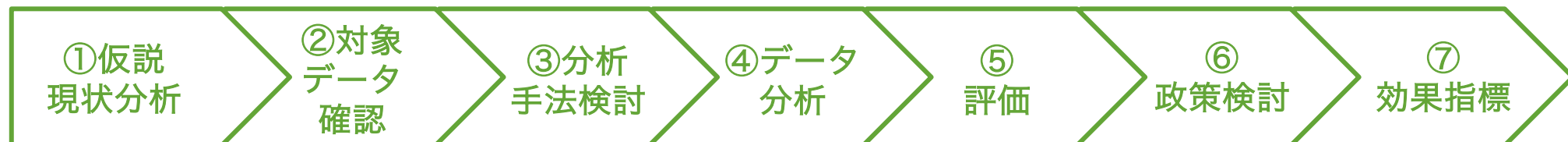


データ分析の流れ全体を実際に体験したい方にオススメ

# データアカデミー



- ITで地域課題の解決を進める団体「Code for Japan」が実施する、行政職員向けデータ利活用研修プログラム
- 2017年度の総務省事業においてカリキュラム開発を実施
- これまでに60以上の自治体が参加
- 従来の座学による研修とは異なり、職員が普段の業務で使用しているデータを用いて、政策の立案や検証をおこなうことが最大の特徴
- データ分析による政策反映の7stepの流れを、半日×4日間の研修で習得する



詳細はこちら→データアカデミー研究会 <https://da.code4japan.org/>



# 4. データ活用における よくある勘違い

# データ活用におけるよくある勘違い

---

1. 定性データは役に立たない
2. 相関があれば必ず因果関係がある
3. エビデンスを求められた時は、データを示せばOK！

# データ活用におけるよくある勘違い

---

1. 定性データは役に立たない
2. 相関があれば必ず因果関係がある
3. エビデンスを求められた時は、データを示せばOK！

# データは性質で大きく2つに分けられる

## 定量データ

- 数値で大小を表すことができるデータ
- 客観性が高い  
(誰が見ても大体同じように解釈できる)
- マクロな視点での分析に用いる  
(鳥の目)

## 定性データ

- 数値で表現できない文章などのデータ
- 主観性が高い  
(解釈はその人の持つ尺度に委ねられる)
- ミクロな視点での分析に用いる  
(虫の目)

# 定量データの例： アンケートの選択項目

某市における行政職員向けデータ活用研修のアンケートより

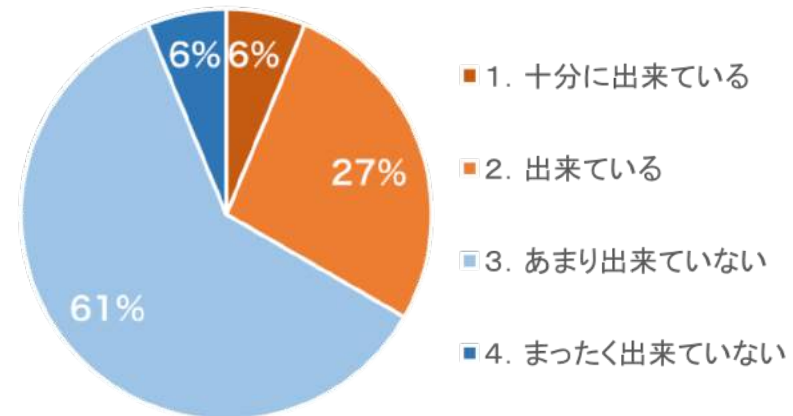
あなたは、日々の業務の中で、データを活用出来ていると思いますか？

1. 十分に出来ている
2. 出来ている
3. あまり出来ていない
4. まったく出来ていない



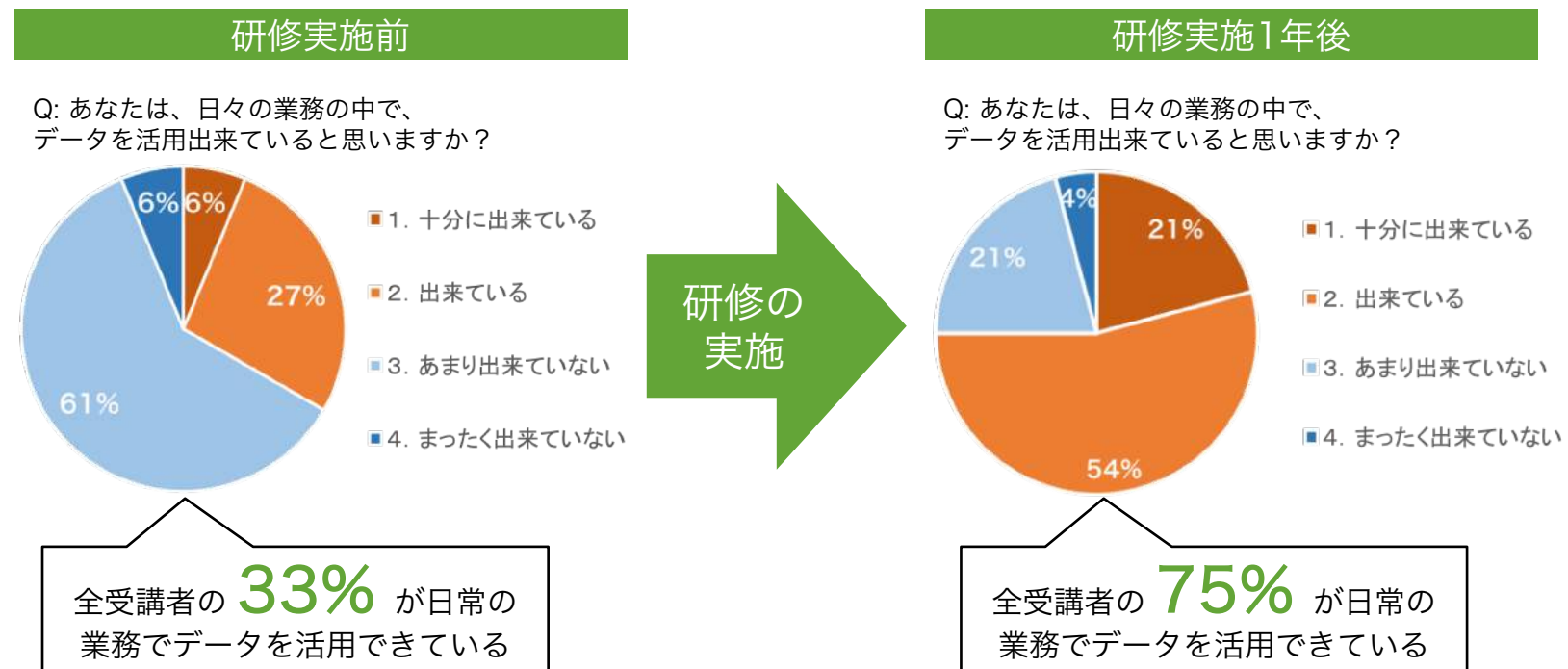
集計して数値化できる = **定量データ**

1. 十分に出来ている	3
2. 出来ている	13
3. あまり出来ていない	29
4. まったく出来ていない	3



# 定量データの使いどころは「俯瞰」

- 数値で比較することで、状態を俯瞰的に把握できる
- 統一された基準で比較できる



# 定性データの例： アンケートの自由記述項目

某市における行政職員向けデータ活用研修のアンケートより

データ活用について、課題等があれば教えてください。（自由記述）



数値化できない  
= **定性データ**

データ収集、活用、提供のノウハウ不足
現在の業務内容上、あまり多くのデータを扱わないため、活用方法に悩んでいる
オープンデータ概念が全く浸透していない、仕事もデータもタテ割り
PC等、ハード・ソフト環境が不足している（処理能力等）
普段の業務から活用した後のイメージがわからない
データを見て何かを感じとる、把握することに苦手意識があり活用まで至らない、センス？慣れ？
未知の有効活用されていない庁内データが存在すると思われる
...

# 定性データの使いどころは「深掘り」

- 定量データよりも情報量は多いため、個々の状態をより深掘りして把握できる
- 次のアクションを検討する上での示唆が得られる

Q: データ活用について、課題等があれば教えてください。

データ収集、活用、提供のノウハウ不足
現在の業務内容上、あまり多くのデータを扱わないため、活用方法に悩んでいる
オープンデータ概念が全く浸透していない、仕事もデータもタテ割り
PC等、ハード・ソフト環境が不足している（処理能力等）
普段の業務から活用した後のイメージがわからない
データを見て何かを感じとる、把握することに苦手意識があり活用まで至らない、センス？慣れ？
未知の有効活用されていない社内データが存在すると思われる
...



- データを自由に使える環境が整っていない
- 自分の業務と結び付けて考えることが困難
- データを使うことに対する漠然とした苦手感覚がある



個別のインタビュー等でより状況を深く把握



# データ活用における よくある勘違い

---

1. ~~定性データは役に立たない~~  
→ 定量データと定性データは目的に応じて使い分ける
2. 相関があれば必ず因果関係がある
3. エビデンスを求められた時は、  
データを示せばOK!

# データ活用におけるよくある勘違い

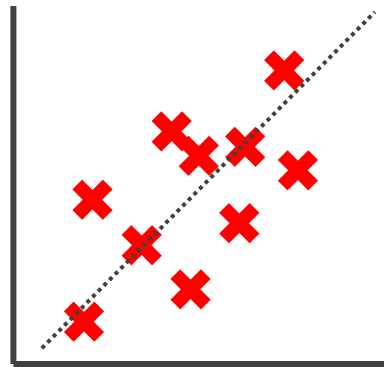
---

1. ~~定性データは役に立たない~~  
→ 定量データと定性データは目的に応じて使い分ける
2. 相関があれば必ず因果関係がある
3. エビデンスを求められた時は、  
データを示せばOK!

# 相関とは

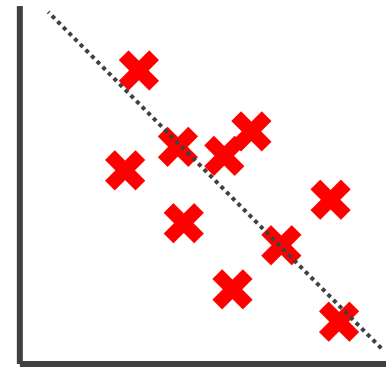
二つの事象が密接にかかわり合い、一方が変化すれば他方も変化するような関係。

正の相関



一方の値が大きくなると、  
もう一方の値も大きくなる

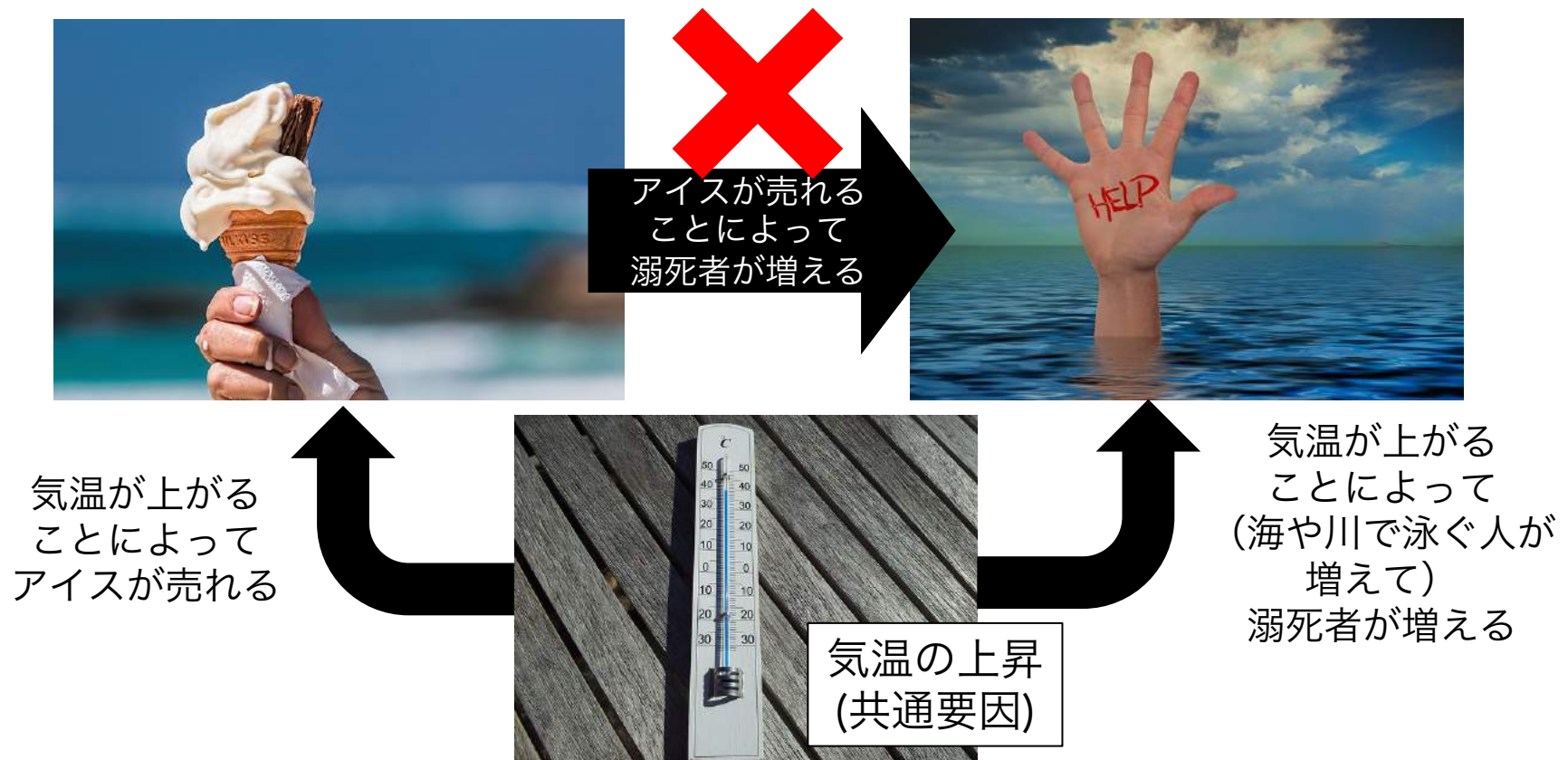
負の相関



一方の値が大きくなると、  
もう一方の値は小さくなる

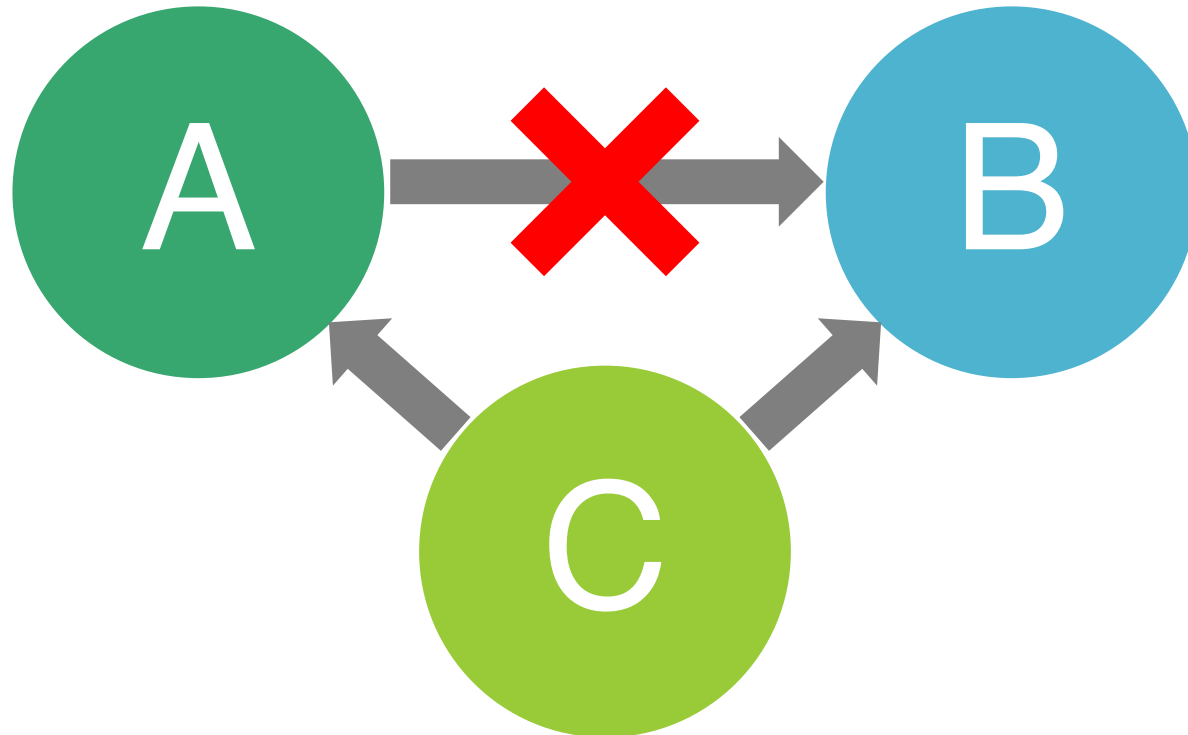
# 相関≠因果関係

例) アイスクリームの売り上げと溺死者の数に相関があった



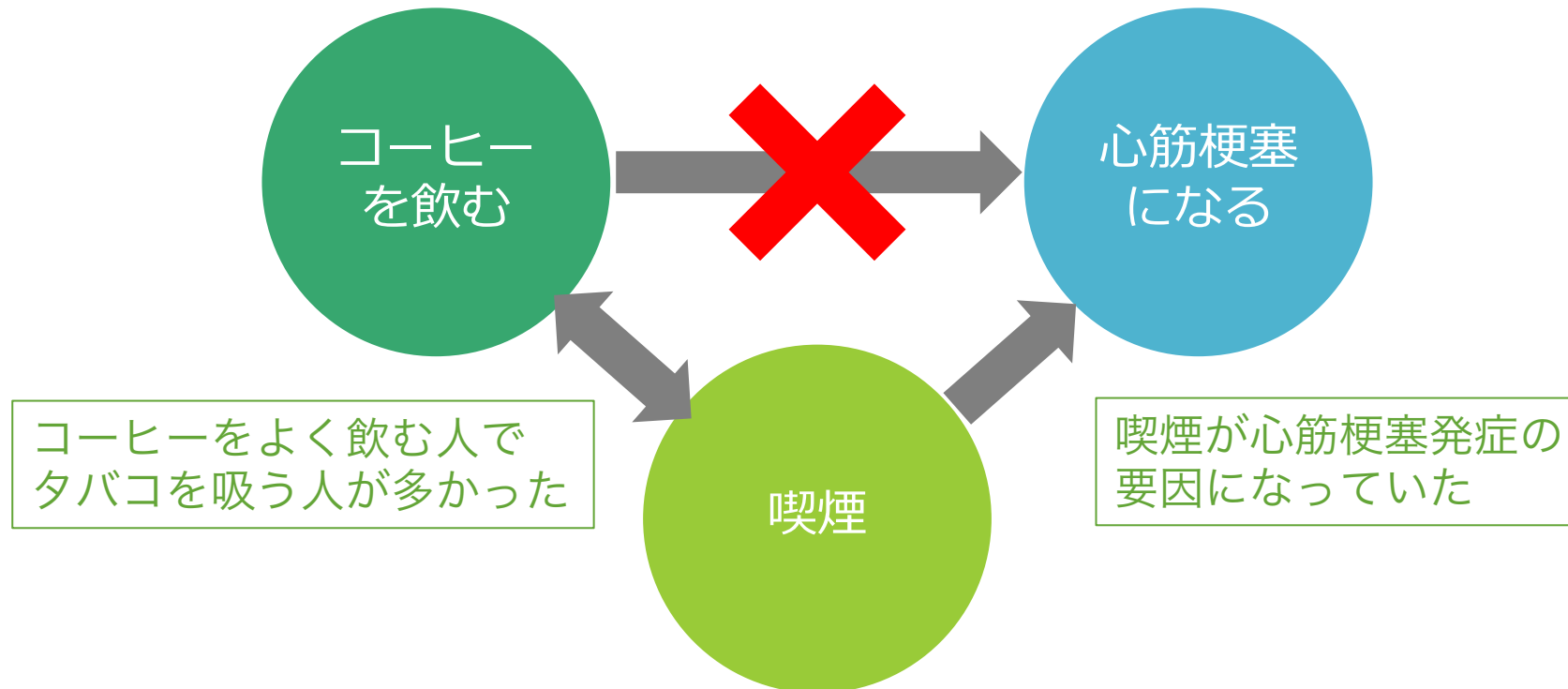
# 相関があっても因果関係がないケース①： 共通の要因がある（交絡）

- CがAとBに影響を与えている状態
- AとBの間に直接的な関係性はない



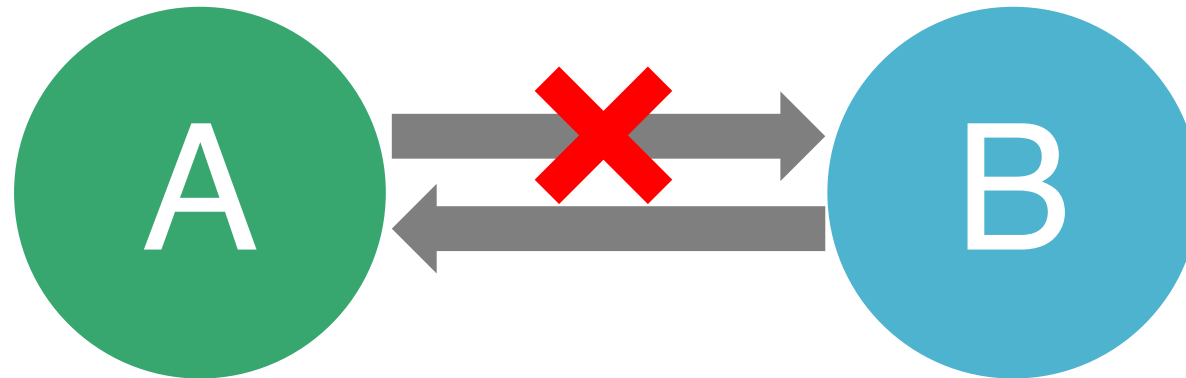
# 相関があっても因果関係がないケース①： 共通の要因がある（交絡）

- 例) 「コーヒーを飲む量」と「心筋梗塞の発症率」に正の相関があった



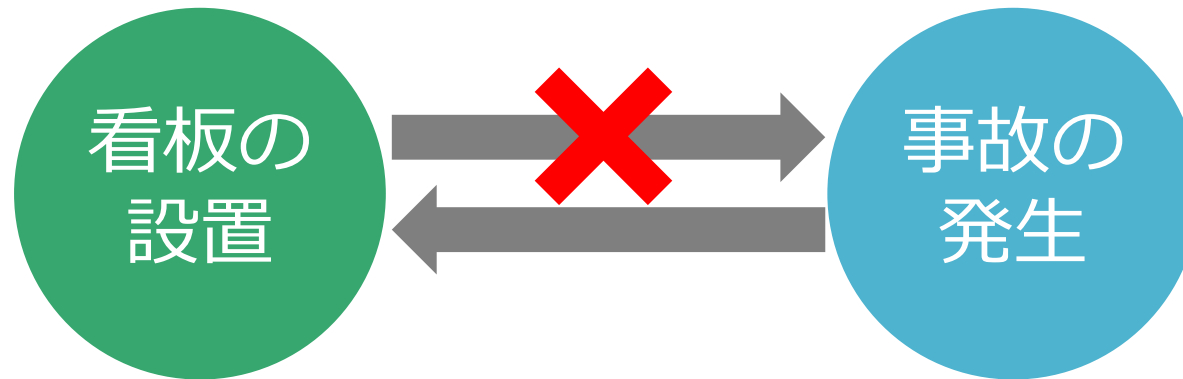
## 相関があっても因果関係がないケース②： 因果関係の方向が逆

- AによってBが引き起こされたのではなく、BによってAが引き起こされた
- 相関分析では時間的な前後関係が見えないため、気づきにくい



# 相関があっても因果関係がないケース②： 因果関係の方向が逆

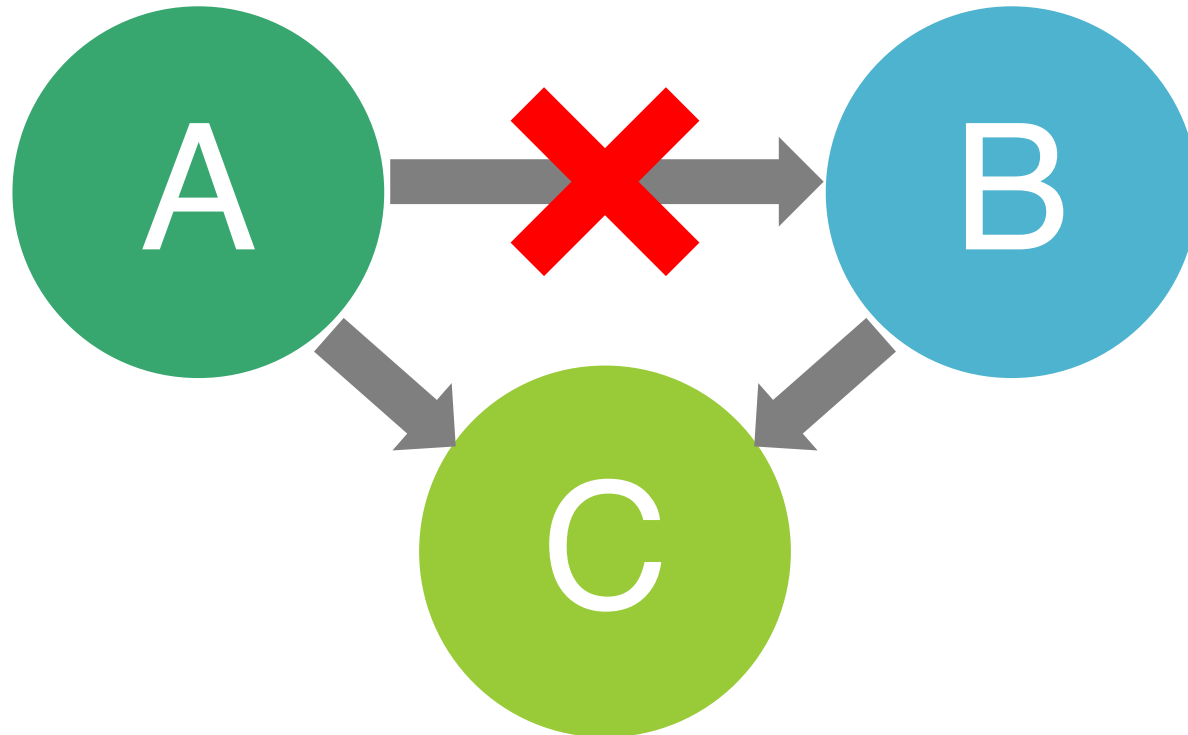
- 例) 「事故の発生件数」と「注意喚起の看板の設置数」に正の相関があった





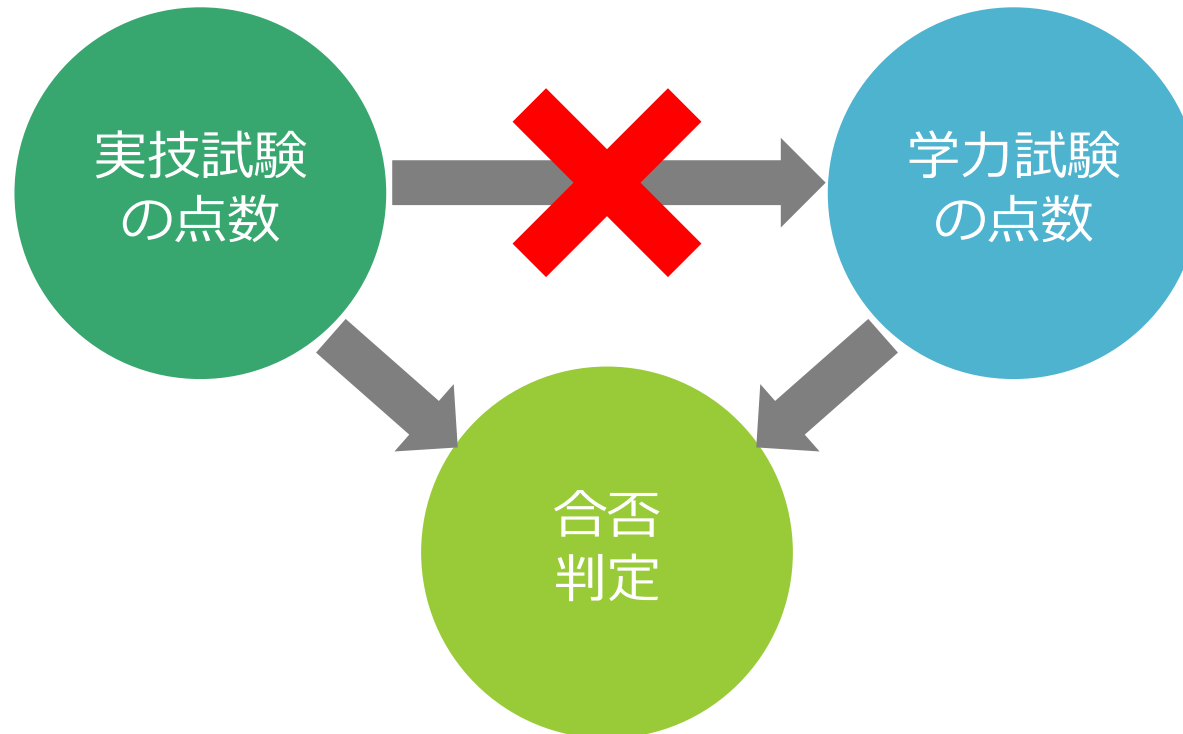
# 相関があっても因果関係がないケース③： 選択バイアス

- AとBの分析対象データが因果の合流点Cにおいて、選抜／層別／調整されてしまっている影響で、因果関係がないのに相関が生じてしまうケース



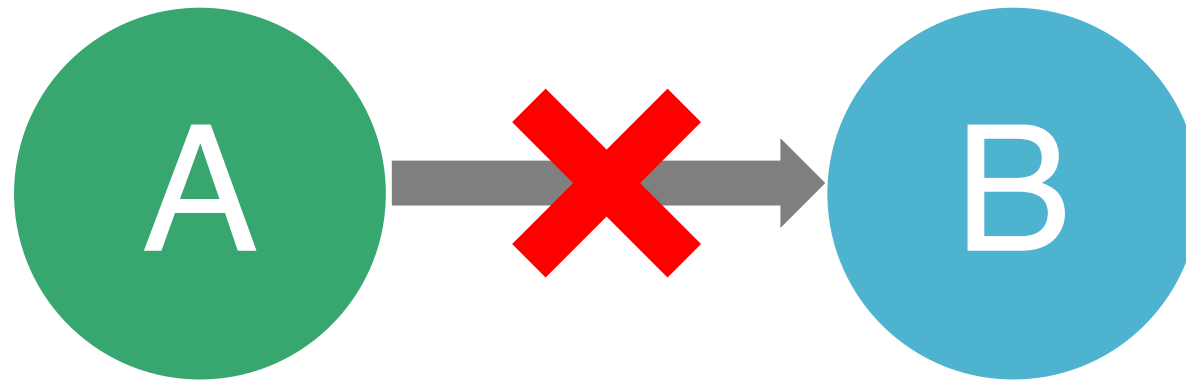
# 相関があっても因果関係がないケース③： 選択バイアス

- 例) 実技試験の点数と学力試験の点数の合計点で合否が決まる試験において、合格者の「実技試験の点数」と「学力試験の点数」に負の相関があった



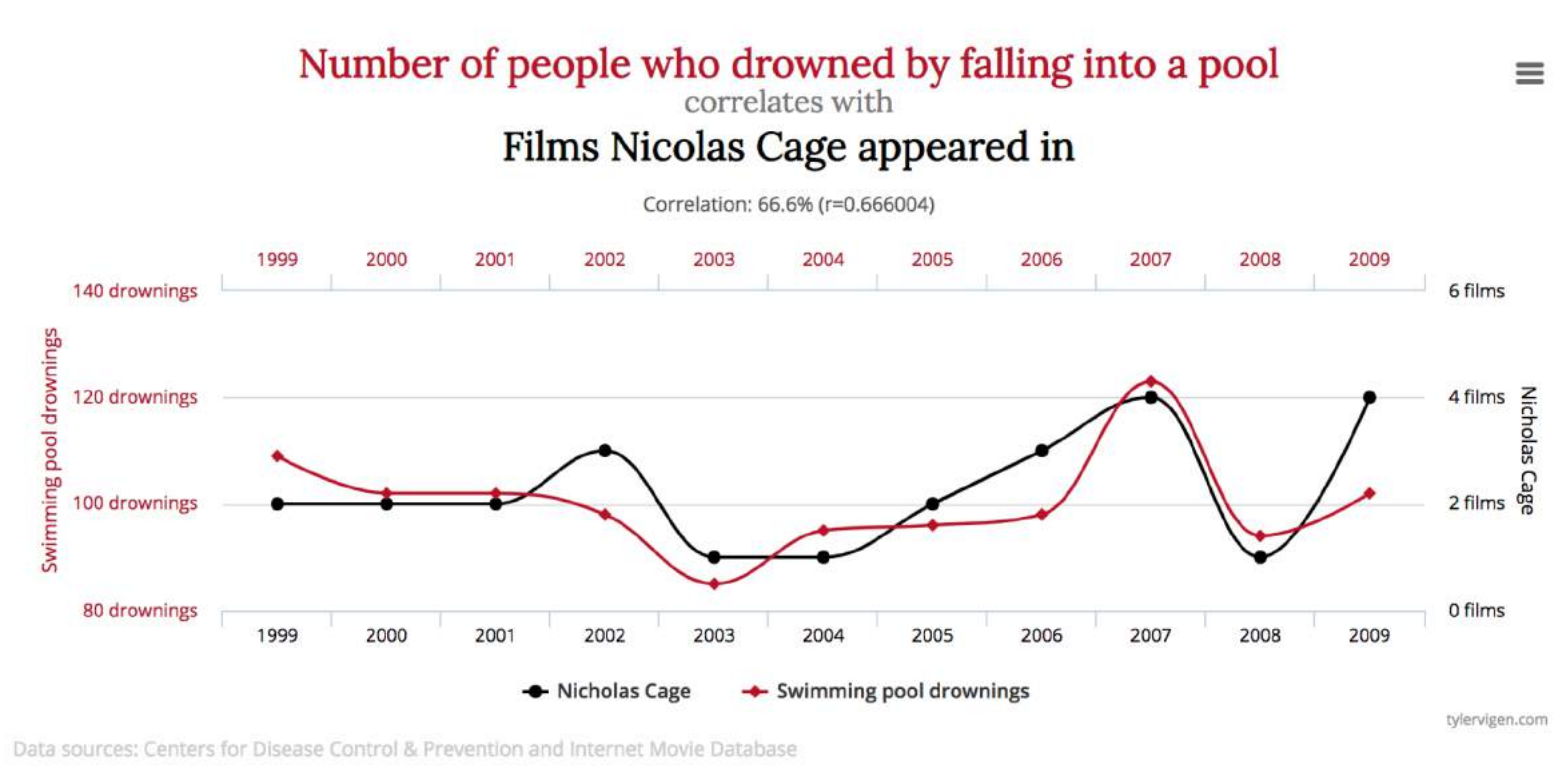
# 相関があっても因果関係がないケース④： 全くの偶然

- 全く何の関係もないAとBであっても、偶然相関が生じてしまうケース



# 相関があっても因果関係がないケース④： 全くの偶然

- 例) 「プールにおける溺死者数」と「ニコラス・ケイジの映画出演数」に正の相関があった



出典：<http://www.tylervigen.com/spurious-correlations>

# データ活用におけるよくある勘違い

---

- ~~1. 定性データは役に立たない~~  
→ 定量データと定性データは目的に応じて使い分ける
- ~~2. 相関があれば必ず因果関係がある~~  
→ 相関があっても因果関係がない場合がある
3. エビデンスを求められた時は、  
データを示せばOK!

# データ活用におけるよくある勘違い

---

- ~~1. 定性データは役に立たない~~  
→ 定量データと定性データは目的に応じて使い分ける
- ~~2. 相関があれば必ず因果関係がある~~  
→ 相関があっても因果関係がない場合がある
3. エビデンスを求められた時は、  
データを示せばOK!

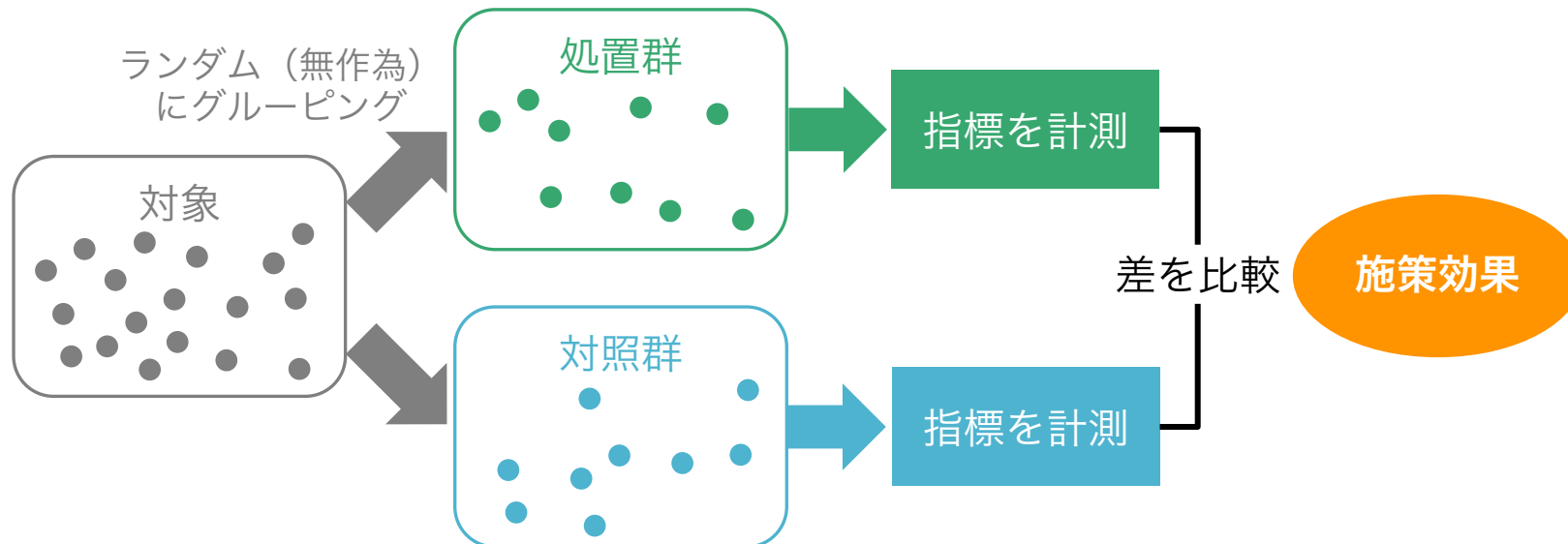
# Before/Afterデータだけで 因果関係を推定するのは困難

- 実験室内なら、影響を及ぼす可能性のある要素を排除するように条件をコントロールできる
- しかし、実社会では特定の要素を排除することは困難
- Before/Afterのデータだけでは、施策によって課題が改善されたかは分からない（他の要因の可能性を排除できない）



# ランダム比較化試験 (Randomized Controlled Trial: RCT)

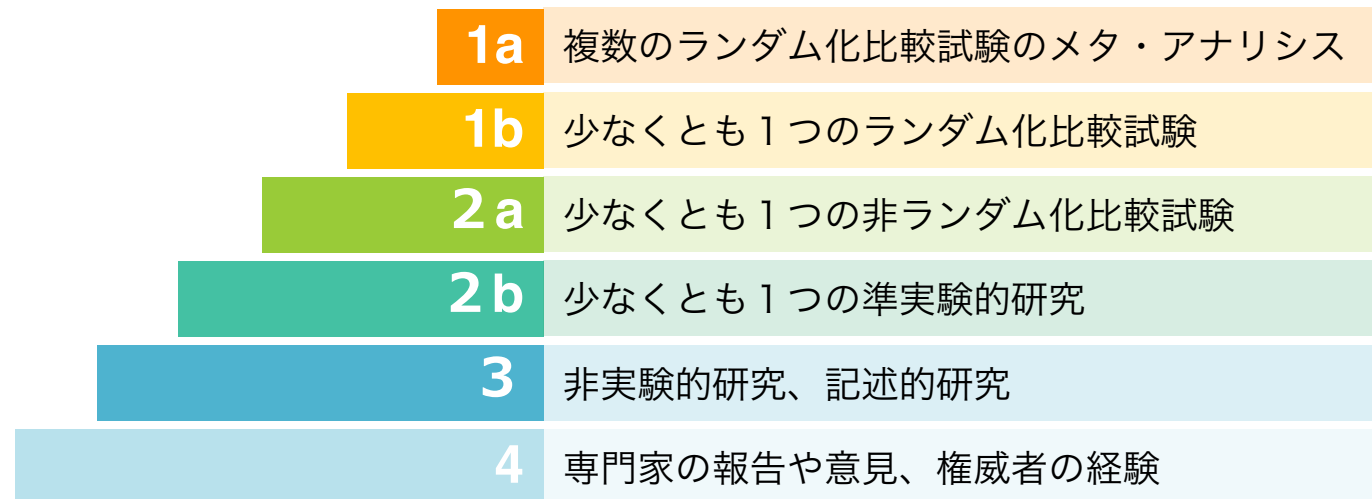
- 対象をランダムにグループ分けして指標を計測し比較
  1. 施策を実施するグループ (処置群)
  2. 施策を実施しないグループ (対照群)
- ランダム・グループピングによって他の要因による影響を排除





# エビデンスの強さの階層

- 「エビデンスに基づく医療 (Evidence-Based Medicine)」に端を発したエビデンスの階層 (レベル) が社会科学にも応用されている
- レベルが上がるほどより正確に因果関係を推定できると見なされる

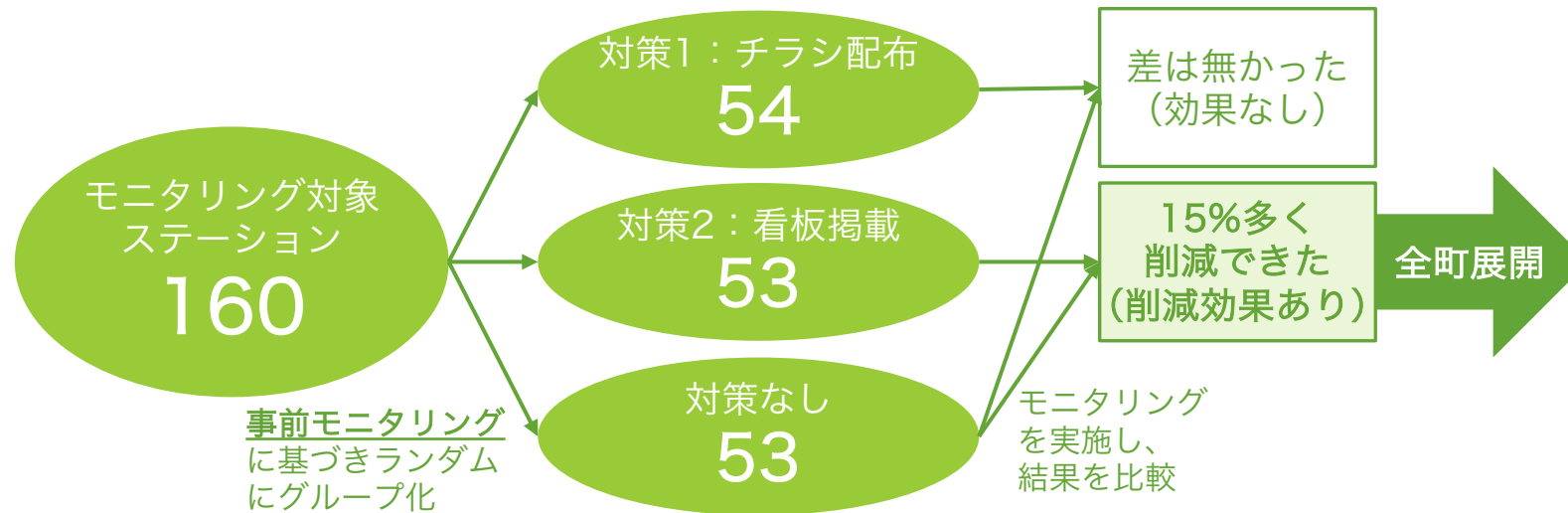


出典：正木朋也、津谷喜一郎（2006）「エビデンスに基づく医療（EMBの系譜と方向性：保健医療評価に果たすコクラン共同計画の役割と未来）」『日本評価研究 第6巻第1号』pp.3-20

# 事例：神奈川県葉山町の 放置ごみ削減施策の効果検証

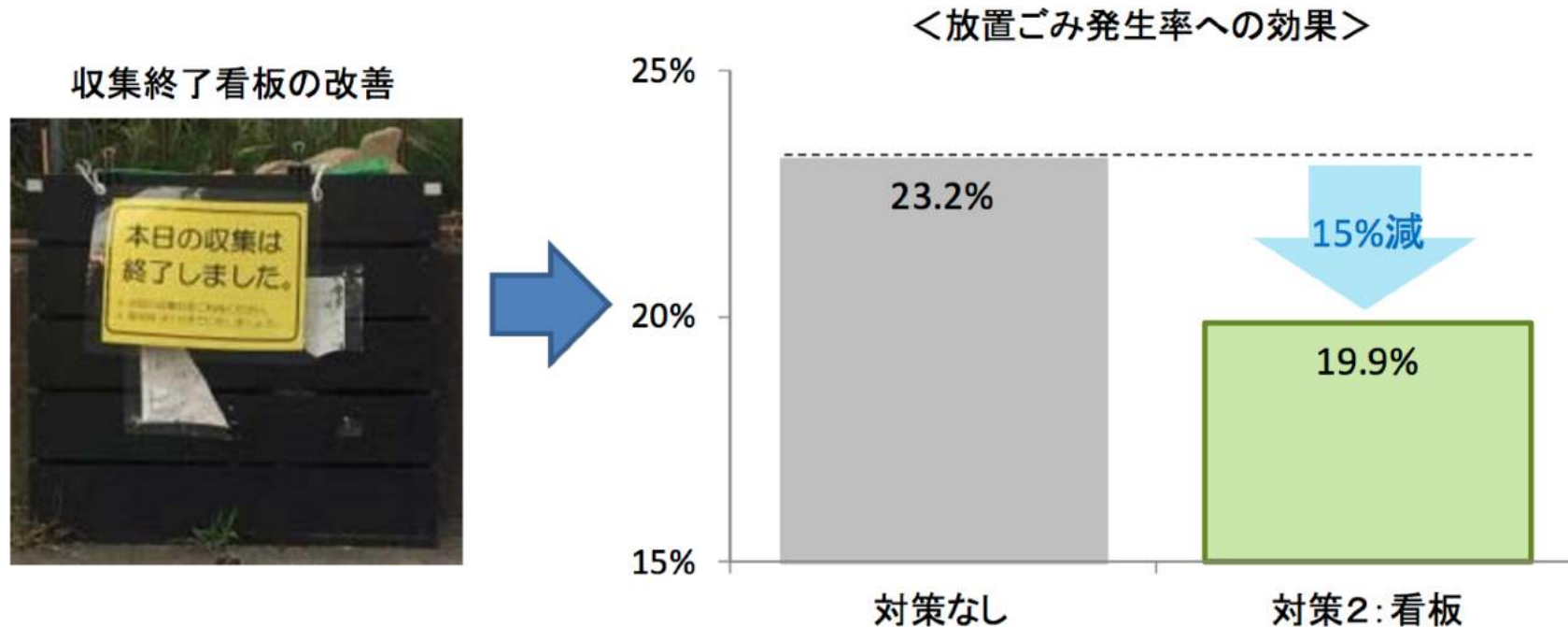
- 葉山町は2008年、日本で3番目に「ゼロ・ウェイスト政策」（ごみの徹底的な資源化・減量化）を導入
- 資源ごみの収集拠点である資源ステーションの管理は、ゼロ・ウェイストの達成に向けた重要施策
- 資源ステーションでの放置ごみが減らないことに町・住民が頭を悩ませていた

## → 施策実験を実施



出典：政策の効果をどう測定するか？：海外における「エビデンスに基づく政策」の最新動向  
<https://www.rieti.go.jp/ip/events/bbl/16102501.html>

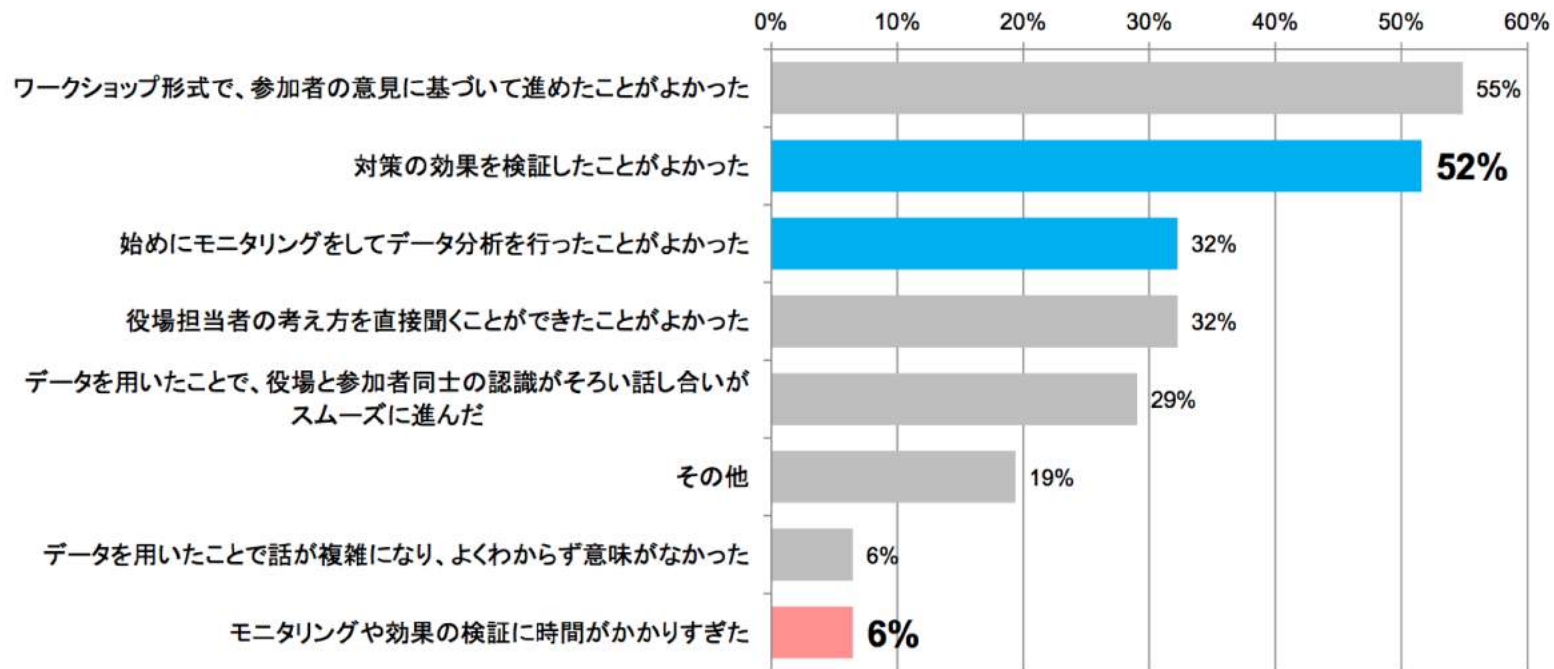
# 事例：神奈川県葉山町の 放置ごみ削減施策の効果検証



出典：「エビデンスを政策にどう使うか -イギリスの動向と日本への適用事例からの示唆-」，小林庸平，2016  
<https://www.rieti.go.jp/jp/events/bbl/16102501.pdf>

# 事例：神奈川県葉山町の 放置ごみ削減施策の効果検証

＜参加者アンケート：今回の取組の進め方について、あてはまるものすべて教えてください＞



出典：「エビデンスを政策にどう使うか -イギリスの動向と日本への適用事例からの示唆-」, 小林庸平, 2016  
<https://www.rieti.go.jp/jp/events/bbl/16102501.pdf>

# 事例：神奈川県葉山町の 放置ごみ削減施策の効果検証

- 取組の効果を学術的な方法で調べることで、自己満足ではない形で成果を実感することができる。

- これまではモデル地区で施策を実施して成果があがっても、「あの地区にはあの人がいるから」「地域によって温度差があるから他の地区では無理」という反応があり、他地域への展開に対する住民や議会の理解が得るのに時間がかかった。
- ランダム化比較試験では、こういう属人的な影響が排除された形で、取り組みの効果を測ることができ、成果とその原因が明確になるので、他地域への展開がより短時間で可能になり、時間の無駄を省くことができると期待している。

- 他の施策分野でも、同じような発想が使えるかも？とよく考えるようになった。



環境課 両宮課長補佐

出典：「エビデンスを政策にどう使うか -イギリスの動向と日本への適用事例からの示唆-」, 小林庸平, 2016  
<https://www.rieti.go.jp/jp/events/bbl/16102501.pdf>

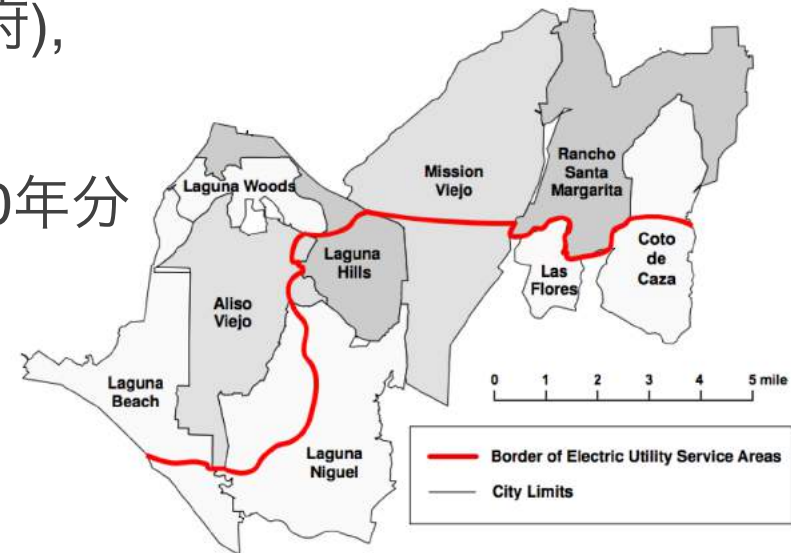
# RCTが実施できない場合、 質の高いエビデンスを得ることは不可能？

---

- 答えは「No」
- ランダム比較化試験 (RCT) を行なったのと同様の状況のデータが入手できれば可能
- 既存の行政データが使える場合がある

# 例) カリフォルニア州における 電力消費量の分析

- 分析の目的：  
「電力の料金を上げると、消費者がどのくらいの量を節電するのか知りたい」
- 分析実施者：シカゴ大学公共政策大学院 伊藤公一郎氏
- 協力機関：California Energy Commission (政府),  
SCE, SDG&E (電力会社)
- 利用データ：世帯ごとの月間電力消費データ10年分
  - カリフォルニア州のある都市では南北で電力会社が異なる
  - 両方の電力会社の電力料金が同様に変化していた時期と、南部だけ電力料金が上がった時期がある

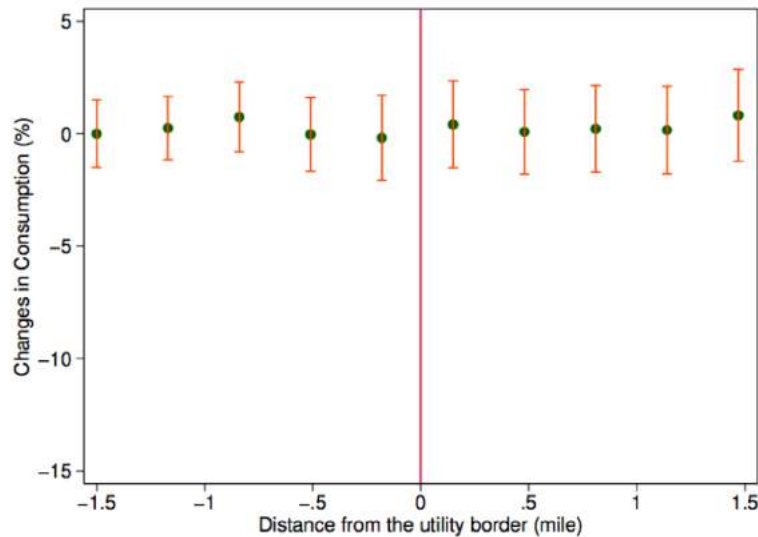


出典：政策の効果をどう測定するか？：海外における「エビデンスに基づく政策」の最新動向  
<https://www.rieti.go.jp/jp/events/bbl/16102501.html>

# 例) カリフォルニア州における 電力消費量の分析

両方の電力会社の電力料金が同様に変化していた時期と比較すると、南部で電力料金が上がった  
時期の電力消費量は12-13%減少していた  
→電力料金が上がると消費者は節電する

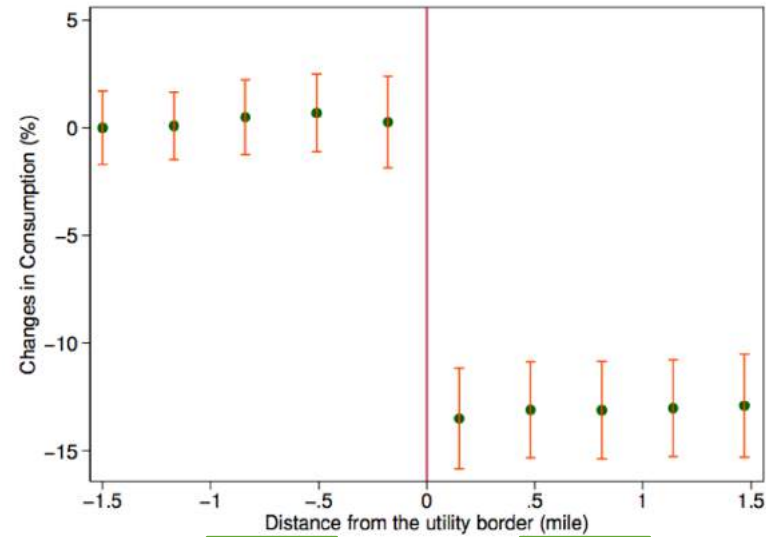
両地域で料金が同様だった時期



北部

南部

南部だけ料金が上がった時期



北部

南部

データ出典 : Koichiro Ito, American Economic Review 2014, 104(2): 537-563  
<http://dx.doi.org/10.1257/aer.104.2.537>



# データ活用におけるよくある勘違い

---

- ~~1. 定性データは役に立たない~~  
→ 定量データと定性データは目的に応じて使い分ける
- ~~2. 相関があれば必ず因果関係がある~~  
→ 相関があっても因果関係がない場合がある
- ~~3. エビデンスを求められた時は、データを示せばOK!~~  
→ エビデンスにも質のレベルが存在するため、ただデータを示すだけでは科学的根拠として弱い場合がある

# 5. お薦めの文献・資料

# データを活用した政策立案プロセスを学びたい方へ データ活用で地域のミライを変える！ 課題解決の7Step

NEW

おすすめ 試し読み



データ活用で地域のミライを変える！  
課題解決の7 Step

紙：¥ 2,310

電子：¥ 2,310

数量  
1

🛒 カートに入れる

♡ お気に入りに追加

クレジットカード払い・コンビニ支払い・請求書払い対応

編著者名 : 一般社団法人 コード・フォー・ジャパン / 編著

判型 : A5・196ページ

商品形態 : 単行本

雑誌コード :

図書コード : 5108579-00-000  
8172327-00-000

ISBNコード : 978-4-324-10764-5

発行年月 : 2019/12

販売価格 : 2,310円 (税込み)

試し読み >

- 50以上の自治体を受講したデータ活用研修  
“データアカデミー”のエッセンスを1冊に集約
- 2019年12月に発売されたばかりの新刊

<https://shop.gyosei.jp/products/detail/10196>

# 人に「伝える」ためのデータ可視化手法を学びたい方へ グラフをつくる前に読む本



グラフをつくる前に読む本 一瞬で伝わる表現はどのように生まれたのか 単行本 (ソフトカバー) - 2017/9/23

松本 健太郎 (著)

★★★★☆ 31件のカスタマーレビュー

> その他 (2) の形式およびエディションを表示する

Kindle版  
¥ 1,642

単行本 (ソフトカバー)  
¥ 1,728 ✓prime

今すぐお読みいただけます: 無料アプリ

¥ 667 より 26 中古品の出品  
¥ 1,728 より 4 新品



[https://www.amazon.co.jp/dp/4774192198/ref=cm\\_sw\\_r\\_tw\\_dp\\_U\\_x\\_JgZnDbPN2PAXN](https://www.amazon.co.jp/dp/4774192198/ref=cm_sw_r_tw_dp_U_x_JgZnDbPN2PAXN)

# 効果的なアンケート調査の手法を学びたい方へ 社会教育調査ハンドブック



- 「社会教育調査」と題されているが、実際はアンケート調査全般に使える考え方やtipsが解説されている
- 特に「第Ⅳ部 調査票の作り方」はアンケートを設計する方必見！
- PDFでダウンロード可能



<https://www.nier.go.jp/jissen/chosa/handbook1-23.htm>

# 個人情報情報の扱いに心配を持っている方へ 個人情報を含むデータ活用検討のためのワークシート

個人情報を含むデータ活用検討のためのワークシート

項目		内容				
1. 目的	個人情報を利用したい部署					
	活用の目的・具体的方法(必要に応じて図などを別添)					
	「統計目的」か「統計目的以外」か 「利用」か「提供」か(以下「活用」と総称)					
項目		個人情報①	個人情報②	個人情報③	個人情報④	個人情報⑤
2. 活用したい個人情報	個人情報保有部署					
	個人情報取扱事務名					
3. 個人情報保有部署との協議	当該事務に記載されている目的 (右の欄に各個人情報取扱事務記載の目的を記入)	①	①	①	①	①
	今回の活用は目的の範囲内か判断	②	②	②	②	②
	その理由	③	③	③	③	③
	活用にあたっての条件	④	④	④	④	④
	その理由	⑤	⑤	⑤	⑤	⑤
4. 個人情報保護所管部署との協議	個人情報保護条例の目的外活用の条項(下記の各号のうち、該当または検討するものに○)					
	個人情報保護条例の目的外利用・提供に関する条項					
	①					
	②					
	③					
	④					
	⑤					
	⑥					
	⑦					
	⑧					
今回の活用が目的外として認められるか判断						
その理由						
活用にあたっての条件						
5. その他関連法令等の確認	個人情報保護条例以外に関連法令等による利用制限はないか。(例: 地方自治法など)					
	所管部署 活用可否の判断					
	その理由					
6. 個人情報活用関係の庁内手続	活用にあたっての条件					
	情報保有部署への申請など、庁内ルールに則って。					
7. 情報システム関係の庁内手続	情報システムの当該データへのアクセス権限設定など。					

- 総務省「地方公共団体におけるデータ利活用ガイドブック」の付録として公開されている
- 個人情報を含むデータ活用を検討する際に、最低限確認すべき事項を整理するためのワークシート
- 千葉市と姫路市での実証の際の検討例を掲載

[http://www.soumu.go.jp/menu\\_seisaku/ictseisaku/ictriyoubigdata.html](http://www.soumu.go.jp/menu_seisaku/ictseisaku/ictriyoubigdata.html)

Link Data Now!!



**LinkData.org**

本資料に関するご質問・ご意見や、  
データ活用研修・ワークショップ開催のご相談などは  
こちらへお寄せください。

一般社団法人リンクデータ  
代表理事 下山 紗代子  
Email: [support@linkdata.org](mailto:support@linkdata.org)