

# 10章 データ分析力

1年 組 番 氏名

## 今回の授業の位置づけ

○探究活動(課題設定→情報収集→分析と整理→まとめ・表現)の過程にあつて、課題に対して立てた仮説の検証に向けて、収集したデータの真偽や妥当性を判断する際に駆使するのが「データ分析力」である。

## ねらい

- データの種類(質的・量的)と4つの尺度を理解し、適切に分類できる。
- 相関関係と因果関係の違いを理解し、「見せかけの相関」を見抜く批判的思考力を養う。
- 公的統計(統計ダッシュボード、e-Stat)の可視化データから、グラフの特徴を読み解く。

## 準備

○クラスを6つのグループに分けておく。各グループに進行役を1名決めておく。

## 1 「統計を使った嘘」の3つのパターン

「本製品は利用者の90%の満足度をいただいております」という宣伝文句を信用する？しない？

利用客から返送されたアンケートハガキによる数値だったならばどうだろうか。アンケートハガキを返送すること自体がそもそもこの製品に好意的であるとも考えられるのではないだろうか。

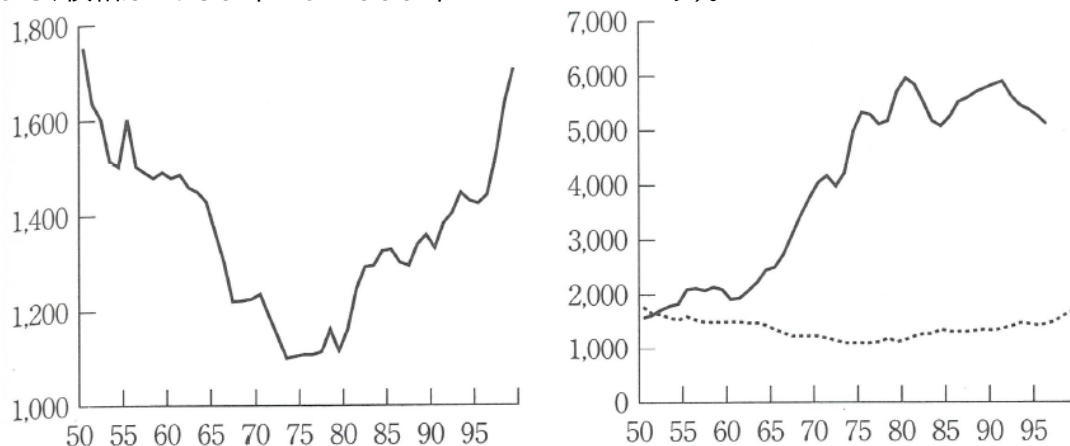
あるいは、店頭でその製品を買おうとした客をねらって満足度を聞いたのかもしれない。その製品に不満ならば、そもそも買おうとしないのではないだろうか。

人数でなくパーセント表示したのも疑わしい。少ない母集団から全体を推測することには無理がある。

### (1) 見せ方による嘘

よくあるのが、グラフの描き方によって、目盛りの取り方によって嘘をつくものです。

下の図は、日米の犯罪率を比較したグラフです。左が日本の犯罪率で、右がアメリカの犯罪率を表しています(どちらも、横軸は1950年から2000年までをとっています)。



(『大学生のためのリサーチリテラシー入門』p179のより引用・一部改変)

このグラフをぱっと見ると、「アメリカ(右)は、1960年代に数値が上昇し、1980年代以降は横ばいで推移している。それに対して、日本(左)は、1980年代まで下がってきて、その後急激に値が上昇し続けているのだな」と思うでしょう。しかし、このグラフには嘘があります。

1つ目の嘘は左のグラフにおいて「日本の V 字型の左側は粗暴犯や凶悪犯が多かったのに対し、反対に右のグラフは窃盗が圧倒的に多い」という数字の性質の違いです。

そして、もう1つは目盛りの違いです。縦軸の目盛りが左と右のグラフで異なっているのです。右の図で点線で描かれているのは、左のグラフを右の目盛りに合わせて書き直したものです。点線を見ると、つまり、同じ目盛りで比較すると、日本は 1950 年代以降、1,000 と 2,000 の間を安定して推移していることがわかります。これが見せ方による嘘の例です。

## (2) データ選択の嘘

もともとの条件が異なるものを比較しようとする、おかしな結果が導かれる場合があります。これを利用して、嘘をつくことが可能です。

たとえば、年間の自動車事故による死亡者数と、飛行機事故による死亡者数を比較したところ、自動車事故による死亡者数の方が多かったとします。この結果から、「飛行機の方が自動車よりも安全な乗り物だ」と判断してよいのでしょうか。それはまずいですよね。

なぜなら、自動車は毎日乗る人も多いでしょうし、利用される機会＝事故に遭う可能性のある機会が、そもそも、飛行機のそれよりもはるかに多いからです。そうした前提条件が異なるものを単純に比較してはいけません。

さらに、こんな例も考えられます。都会にある A 市と田舎にある B 市で、中学生の子どもがいる家庭のうち、何%が子どもを塾へ通わせているかを調べました。すると A 市では68%が、B 市では44%が子どもを塾に通わせているという結果になりました。この結果を見て、「都会のほうが教育熱心な親が多く、子どもを塾へ通わせる家庭が多いのだ」と考えたとしたらどうでしょうか。それは正しいといえるでしょうか。

A 市と B 市ではそもそも塾の数が違うかもしれません。都会の A 市には塾がたくさんあって、田舎の B 市には塾そのものの数が少ないということも考えられます。また、年収が多い人のほうが子どもを塾に通わせる割合が高いということもあるかもしれません。同じ年収で比較したら塾へ通わせる割合は変わらないのに、都会である A 市の方が高収入の人が多いために、A 市と B 市で塾へ通わせる割合に差が生じているという可能性もあります。

## (3) データ収集の嘘

どのようにデータを集めるかによって、まったく違った結果が得られることがあります。

高校生の化粧使用頻度についてアンケート調査を行う場合に、渋谷駅周辺の高校生にアンケートを実施した場合と、地方の農村部でアンケートを実施した場合とでは、まったく違った結果になるでしょう。たとえば、前者だけのデータでは「日本の女子高校生は化粧ばかりしている」という結論が得られるかもしれないし、後者だけのアンケートでは「日本の女子高校生は今なお質素で清楚である」という結論が得られるかもしれません。いずれの場合も、全国の高校生を代表する集団とは見なせないですね。

あるいは、「大学生の授業関心度」を調べるための質問紙調査を、大学の 1 限で実施したとします。その調査結果を一般的な大学生の授業関心度と見なすのは問題がありそうです。というのは、1限の授業にきちんと出てくる学生は、もともと授業への関心が高い可能性があるからです。1限の授業の受講生だけを調べたとしたら、本来の調査対象全体をうまく反映していないことになりかねません。

ここで取り上げた2つの例は、いずれも関心のある調査対象全体の一部からデータを収集して、そのデータを分析することで、調査対象全体についての様相を明らかにしようとするものです。こうした調査方法は、標本調査(サンプル調査)といわれるものです。関心のある調査対象全体を母集団といいます。そして、母集団の一部であり、実際に調査を行った集団を標本(あるいはサンプル)といいます。

ここでの2例は、どちらも、偏った標本(サンプル)を分析しても、その結果は、母集団の様子を正確に表すものにはならないということを伝えてくれるものです。

『大学生のためのリサーチリテラシー入門』(山田剛史・林創著 ミネルヴァ書房 2011)より  
抜粋

## 2 相関関係と因果関係

夏の暑い日に外へ出たところ、団扇(うちわ)や扇子であおいでいる人を多く見かけたでしょう。

その時、「団扇や扇子であおいでいる人たちがいるから、暑くなる」と判断する人はいないだろう。暑くて汗が出るから、団扇や扇子であおいで涼んでいる(気化熱を利用している)のは言うまでもない。

では、クルマの保有率が減少し、地価が高騰した地域があったとしたらどうか？さらに、その地域で商店街が急に衰退してしまったとしたら？

クルマと地価に直接の関係は無いはずだし、一般的には、商店街が衰退すれば地価が下落するだろうと思われがちだが、その地域に利便性の高い鉄道が敷設されて駅ビルができたという事実を聞けば、納得するのではあるまいか。クルマ、地価、商店街という要素の変化に、駅ビルという第四の変数が介在したということである。

こうしたさまざまな関係性についての説明を、谷岡一郎著『データはウソをつく』から以下に引用する。

〈「相関関係(correlation)」と「因果関係(causality)」とは、よく混同される概念ですが両者は同じではありません。まず両者の定義を見てもらいます。

☆相関関係…変数の一方□が変化するとき、もう一方の変数□の変化がランダムでない関係(または同じことだけど、変数□が変化するとき、変数□の変化がランダムでない関係)。(記号で□↔□のように、両方向の矢印を使用する)

☆因果関係…変数の一方□の変化によって、もう一方の変数□の変化を引き起こす関係。つまり、変数□が原因で、変数□が結果。(記号で□→□のように、一方向の矢印を使用する)

相関関係と因果関係の違いをわかりやすく説明すると、因果関係の変数□、□は、時間的に□が□に先行しますが、相関関係ではどちらが先かわかりません。□が□に先行するかもしれないし、その逆かもしれないし、はたまた同時かもしれないのです。

「(統計学上有意な)相関関係が存在する」ことは、因果関係の前提です。つまり有意な相関関係は、因果関係があることの必要条件となります。逆に言えば、相関関係がある、というだけでは、因果関係についてはまだ何も言えません。



- ②グループの最優秀作を話し合って決める。  
 ③各グループの最優秀作を、それぞれ発表する。

|                            |  |
|----------------------------|--|
| 「見せかけの因果関係」について、バズセッションでメモ |  |
|----------------------------|--|

|           |  |
|-----------|--|
| グループの最優秀作 |  |
|-----------|--|

**【課題2】** ※先生はタイムキーパー役をお願いします

ねらい

- 因果関係と相関関係について理解する。

内容

- 実際によく使われる論法に対して、クリティカル・シンキング(批判的思考力)を働かせる。

**ワーク1** クリティカル・シンキング問題

文部科学省は、「全国学力・学習状況調査」という学力テストの結果を用いて、子どもの学力と家庭環境にどのような関係がみられるかを分析しています。その分析によると、「親の年収や学歴が低くても学力が高い児童の特徴は、家庭で読書をしていること」だとされています。

この結果を受けて、多くのメディアは「子どもに読書をさせることが学力向上に重要だ」と報道をしています。

↓

この報道の間違ひの可能性を指摘してください。

『学力の経済学』（中室牧子 2015 p20～23）より改題

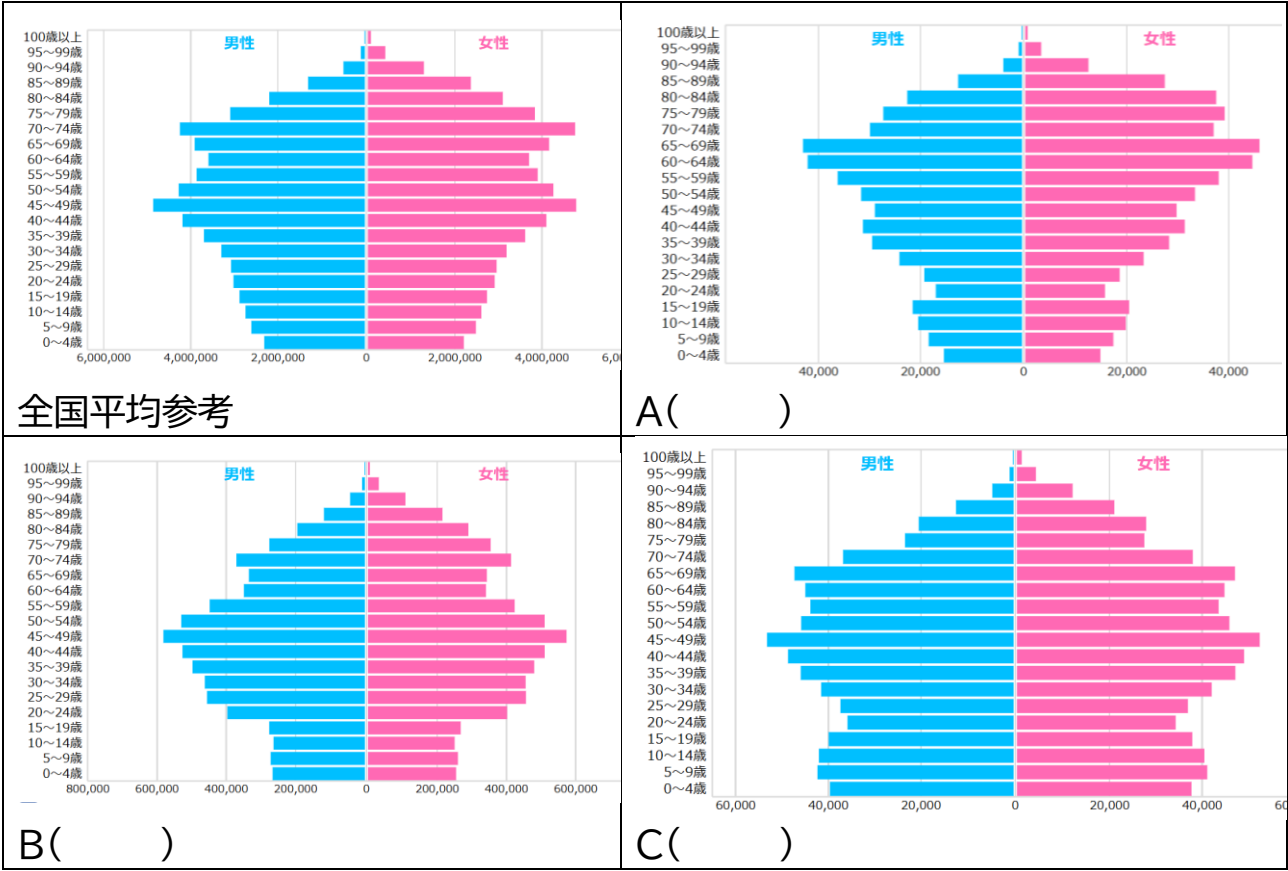
報道の間違ひの可能性について、自分の考えを下欄に記入する。

|  |
|--|
|  |
|--|

解答例を聴き、要点をメモする。

【課題3】 2015 年のある都道府県の人口ピラミッドです。全国の人口ピラミッドと比べると形がそれぞれ全くちがうことがわかります。どこの都道府県のものでしょうか？また形の特徴は何に由来しているのでしょうか？

ア:東京都    イ:沖縄県    ウ:秋田県  
(それぞれ理由もお答えください)



|       |  |
|-------|--|
| A の理由 |  |
| B の理由 |  |
| C の理由 |  |

統計局：統計ダッシュボードを利用し作成しました。( <https://dashboard.e-stat.go.jp/pyramidGraph?screenCode=00570&regionCode=00000&pyramidAreaType=2>)

# 教師用資料 10章 データ分析力【授業デザイン】

## 今回の授業の位置づけ

○探究スパイラル(課題設定→情報収集→分析と整理→まとめ・表現)の過程にあって、課題に対して立てた仮説の検証に向けて、収集したデータの真偽や妥当性を判断する際に駆使するのが「データ分析力」である。

## ねらい

- データの種類(質的・量的)と4つの尺度を理解し、適切に分類できる。
- 相関関係と因果関係の違いを理解し、「見せかけの相関」を見抜く批判的思考力を養う。
- 公的統計(統計ダッシュボード、e-Stat)の可視化データから、グラフの特徴を読み解く。

※クラスを6つのグループに分けておいてください。

| 時間                       | 内容・留意点                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|--------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 【導入】<br>(10分)            | 第10章①データの理解と尺度の理解の視聴。【動画は3分】<br>「1『統計を使った嘘』の3つのパターン(1)(2)(3)」、「2相関関係と因果関係」を読んで理解する。<br>※「このあと、見せかけの因果関係を考えたり、事柄の関係を見抜く問題を実際に解いたりするので、そのつもりで内容理解に努めてください。」と指示してください。<br>※「相関関係」と「因果関係」の違いがわかることに力点を置いてください。「9つの可能性」に関しては参考までに紹介してもらえればと思います。                                                                                                                                                                                                                                                                                                              |
| 【課題1】<br>【ワーク1】<br>(15分) | ①「見せかけの因果関係」をバズセッションで出し合う。(6分)<br>②各グループの最優秀作を決める。(3分)<br>③発表し合う。6分<br>※先生方はタイムキーパーをお願いします。                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 【課題2】<br>(13分)           | ①各自で報道の間違ひの可能性を考える (5分)<br>②グループで考えを共有する。(5分)<br>③解答例を聴き、要点をメモする (3分)<br>※上記③については、先生が解答例を読み上げてください。<br>第一に、読書と学力の間に因果関係があるようにとらえているのが誤りです。「因果関係」と「相関関係」、どちらも2つの出来事の関係を示すときに使われる言葉ですが、決定的に違う点があります。<br>因果関係は「Aという原因によってBという結果が生じた」ことを意味しています。しかし、相関関係は単に、「AとBが同時に起こっている」ことを意味しているにすぎません。相関関係は2つの出来事のうちどちらが「原因」で、どちらが「結果」であるかを明らかにするものではないのです。「相関関係」があるということは、必ずしも「因果関係」があることを意味しません。<br>つまり、読書をしているから子どもの学力が高い(因果関係)のではなく、学力の高い子どもが読書をしているにすぎない(相関関係)可能性があるのです。<br>「読書をする」ことが原因で、「学力が高くなる」という結果がもたらされていることがはっきりしないのに、本を買い与えたり、読み聞かせをしたりしたら、お金や時間の無駄遣いになっ |

|                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                        | <p>てしまうかもしれません。</p> <p>また、この報道にはもうひとつの誤りがあります。それは、「見せかけの相関」の可能性を検討していないことです。つまり、読書にも学力にも影響するような「第三の要因」があるかもしれないのに、そのことを考慮していないのです。</p> <p>「第三の要因」としては、たとえば、「子どもに対する親の関心の高さ」などが考えられます。子どもに対する関心が高い親は、子どもを勉強するように促すでしょうし、同時に子どもに本を買い与えたりもするでしょう。その「関心の高さ」こそが両方の変化を同時に引き起こしているにもかかわらず、あたかも、読書と学力の間に相関関係があるかのように見えてしまう。これが「見せかけの相関」です。</p> <p>間違っ、見せかけの相関を因果関係と解釈してしまうと誤った判断のもとになります。</p>                                                                                                                                                          |
| <p>【課題3】<br/>(12分)</p> | <p>①各自でグラフに対応する都道府県とその理由を考える。(5分)</p> <p>②グループで考えを共有する。(5分)</p> <p>③解答例を聴き、要点をメモする。(2分)</p> <p>2015年のある都道府県の人口ピラミッドです。全国の人口ピラミッドと比べると形がそれぞれ全くちがうことがわかります。どこの都道府県のものでしょうか？ 形の特徴は何に由来しているのでしょうか？</p> <p>ア:東京都 イ:沖縄県 ウ:秋田県 (それぞれ理由もお答えください)</p> <p>正解と解説</p> <p>A-秋田県(高齢者人口が多く、子ども人口が少ない)</p> <p>*子ども(15歳未満)の割合が47都道府県で最も小さく(10.5%)、高齢者(65歳以上)の割合が最も大きい(33.8%)。</p> <p>B-東京都(20歳～50歳までの割合が多い)</p> <p>*生産年齢(15歳以上 65歳未満)人口割合が47都道府県で最も大きい(65.9%)。</p> <p>C-沖縄県(子ども人口が多い)</p> <p>*子ども(15歳未満)の割合が47都道府県で最も大きく(17.4%)、高齢者(65歳以上)の割合が最も小さい(19.6%)。</p> |
| <p>振り返り<br/>(5分)</p>   | <p>Classi アンケートに回答。</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |