

---

# 情報 I No. 22

## データの活用①

---

年	2	組		番		名前	
---	---	---	--	---	--	----	--

## 第4章 情報通信ネットワークとデータの活用 3節 データの活用

### 1. データの収集と整理 2. 数値データの分析 (教 P192-P195)

☞ 問題解決にデータ活用が有効であることを理解しよう。

**【実習】** 様々な都道府県別統計が紹介されている「とどラン」にあるデータを使って、データ同士の関係の深さ（相関関係）を調べよう。

**【実習1】** QRコードを読み込み「とどラン」(<https://todo-ran.com/>)へ移動する。

※ 「とどラン」= さまざまなデータ（国土・インフラ、社会・政治、文化・暮らし・健康、娯楽・スポーツ、店舗分布）について、都道府県別のデータを公開



◇ 手順1 色々なデータの中から自分が興味を持ったデータを3つ書き出そう

◇ 手順2 相関関係（一方が増えれば一方が増える または 一方が増えれば一方が減る関係）がありそうなデータを、仮説（このデータが増えればこのデータは増えるはず）を立てる

仮説	(例) 「高校数が多い都道府県」ほど「甲子園の勝利数」が多い
使うデータ	(例) 「都道府県別高校数」と「都道府県別通算甲子園勝利数」
理由	(例) 甲子園出場までの予選で戦う回数が多いはずだから強いチームと思うから

**【実習2】** 「とどラン」のデータをワークシートに入力し、相関係数を算出する。(Excel を使用)

◇ 手順1

- ① 使うデータの一覧表（地図をスクロールさせた下にある）を表示
- ② 表の「北」ボタンを押し、データを北からの順番に並べる
- ③ 表の「北海道の順位～沖縄の偏差値まで選択してコピーする
- ④ 自分のExcel のワークシートのシートにコピーする。
- ⑤ 単位が入っているデータは「置換」を使って数字のみにする



順位	都道府県	高校数	人口(万人)	偏差値
1	北海道	110	2,200	41.85
2	青森県	62	6,010	46.81
3	岩手県	52	4,420	46.92
4	宮城県	105	4,500	45.77
5	秋田県	62	6,200	48.93
6	山形県	37	3,400	43.73
7	福島県	45	2,400	41.95
8	茨城県	82	2,800	42.61
9	栃木県	100	5,170	46.92
10	群馬県	100	5,400	47.46
11	埼玉県	114	1,500	40.17
12	千葉県	141	2,240	41.46
13	東京都	300	2,200	41.37
14	神奈川県	200	2,200	41.50
15	新潟県	31	1,410	39.90
16	富山県	37	3,570	43.94

◇手順2

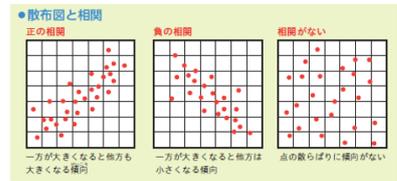
①自動的に計算された平均値、中央値、標準偏差、相関係数を記録する。

	データ①	データ②
①平均値（データ全体の平均）		
②中央値（真ん中の県のデータ）		
③標準偏差（データの散らばり・幅）		
④相関係数（2つのデータの関係ありなし）		

【実習3】仮説（2つのデータの相関関係）が正しかったか、検証しよう。

◇手順1 相関係数をもとに、次の表で相関関係があるかどうか判定しよう。

相関係数	判定
-1~-0.7	強い負の相関がある
-0.7~-0.5	負の相関がある
-0.5~-0.3	弱い負の相関がある
-0.3~0.3	相関はない
0.3~0.5	弱い正の相関がある
0.5~0.7	正の相関がある
0.7~1	強い正の相関がある



相関関係の判定	
考察	

※考察で書くこと ・自分の仮説が正しかったか、違ったのか。  
 ・なぜそういう結果となるのか、自分なりの分析を書く

【知識の整理】

1 データの活用

①データの種類

( )	数値や記号など表形式に整理できるデータ
(非構造化データ)	音声や動画、画像など表形式で表せないデータ

②情報社会の進展で膨大なデータが流通・蓄積されるようになった・・・

・( ) = SNS の書き込みやセンサで集められるデータなど日々生成され記録され増加するデータ

③ ( ) = データをコンピュータや統計学など専門的な知識を使って分析する  
 (例) POS システム = 過去のデータに天候データを加えて分析し、天候によってどの商品が売れるかを予測し、商品発注に活かす

## 2 データ分析の流れ

### ①データの収集

- ・(アンケート調査) = 多くの人に同じ質問を行い、多くの回答を集める方法  
集める方法によって質問紙調査、Web 調査、街頭調査などに分類
- ・(オープンデータ) = 国の行政機関や自治体、研究・教育機関、企業などが広く公開するデータ



### ②データの整理

- 1) データの確認：欠けてる値 (欠損値)、他のものと比べて大きく外れた値 (外れ値) の確認
- 2) 集計する：項目ごとに計算する (単純集計)、2 以上のカテゴリ間で集計する (クロス集計)

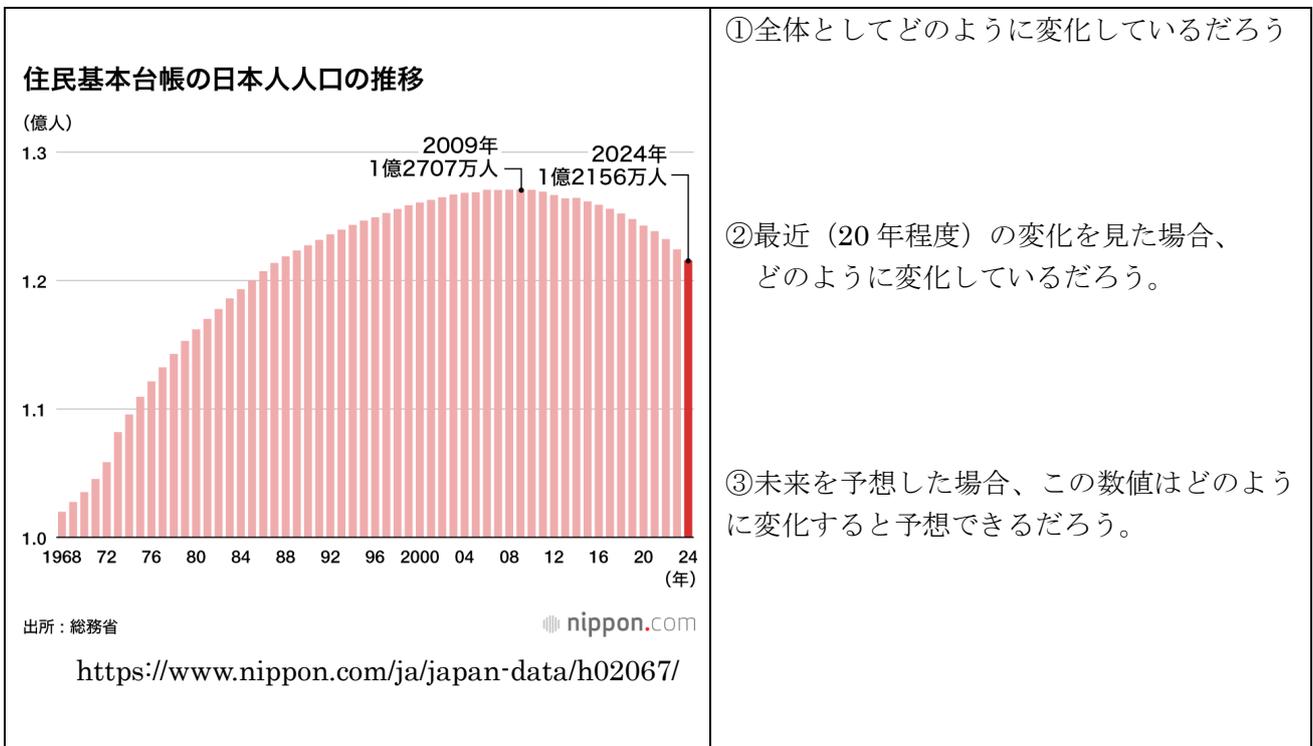


### ③データの分析

( ) = 数値でも計算できない	(名義尺度)	分類のため数値を割り当てた (例) 性別、血液型
	(順序尺度)	数値の順序に意味がある (例) 5 段階評価、おみくじ
( ) = 数値で計算ができる	(間隔尺度)	0 は相対的な意味しかない (例) 温度、西暦
	(比例尺度)	0 は絶対的な意味がある (例) 身長、体重、年齢

## 【確認課題】調べよう・考えよう！

(1) 次のグラフのデータから分析してみよう。



【振り返り】 No.22 の実習・学習で学んだこと、気づいたこと、考えたことを 3 行以上書きましょう。