# 情報 I No. 17 データの活用

# 第4章 情報通信ネットワークとデータの活用 第3節 データの活用

- 【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】「とどラン」(https://todo-ran.com/)で検索し、 サイトへ移動する。
- ※「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、その他)について、都道府県別のデータを公開している



◇手順1 どのようなデータがあるか3つ書き出してみよう

◇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

| - 0) / (  | . 7.6 / | アと、仮記(この) アル 塩 たんじるこの アーブ は 塩 た し は チェ と 立 こ し |
|-----------|---------|--|
| 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<br>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する。相関係数を算出する。

### ◇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) を表示させる。

②「北」ボタンを押し、データを北からの都道府県順に並べる

③データを見ながら順番に数値を入力する。※手で入力

4)もう一つのデータも同様に入力する。



|   | 順位      | 都道府県 | 勝利数   |               |       |
|---|---------|------|-------|---------------|-------|
|   |         |      | 総数    | 人口10万人<br>あたり | 偏差值   |
|   | 310 640 | 匪 南  | 路順 四順 | 路頂 買順         | 路頃 男頃 |
|   | 43      | 北海道  | 11889 | 2.26勝         | 41.49 |
|   | 28      | 青森県  | 628#  | 5.01.勝        | 46.61 |
|   | 29      | 岩手県  | 5689  | 4.629時        | 45.90 |
|   | 30      | 宮城県  | 1058時 | 4.56膊         | 45.77 |
|   | 18      | 秋田県  | 608#  | 6.25勝         | 48.93 |
| ľ | 38      | 山形県  | 3786  | 3.46勝         | 43.73 |
|   | 41      | 福島県  | 4589  | 2.45勝         | 41.85 |
|   | 40      | 茨城県  | 8289  | 2.86勝         | 42.61 |
|   | 26      | 柩木県  | 10086 | 5.17勝         | 46.92 |
|   | 24      | 都馬県  | 106卿  | 5.46颜         | 47.46 |
| Ī | 46      | 埼玉県  | 11486 | 1.55膊         | 40.17 |
| ľ | 44      | 千葉県  | 14198 | 2.24勝         | 41.46 |
| ľ | 45      | 東京都  | 30986 | 2.20勝         | 41.37 |
|   | 42      | 神奈川鳳 | 20989 | 2.26肺         | 41.50 |
|   | 47      | 新潟県  | 3186  | 1.41膊         | 39.90 |
|   | 36      | 富山県  | 378#  | 3.575#        | 43.94 |

□ 都道府県別甲子園通算勝利数

### ◇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。



- ※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】(教 P192~P193)

①データの活用

- 1) (データサイエンス ) = コンピュータや専門的な知識を使ってデータを分析する取組み
- 2) データ分析の流れ:  $(データの収集) \rightarrow (データの整理) \rightarrow (データの分析)$

# ②データの収集

- 1) (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調査
- 2) (オープンデータ) = 公開され自由に利用できるデータ(例)行政機関、研究・教育機関

# ③データの整理

・値の確認の方法:(欠損値) =欠けている値、(外れ値) =極端に外れた値 を確認

| 【知識の                             | 整理】(教 P194~P197)  |  |  |  |  |  |  |
|----------------------------------|---|--|--|--|--|--|--|
| ④数値デ                             | ータの分析   |  |  |  |  |  |  |
|                                  |   |  |  |  |  |  |  |
| ・( 量的データ )=数値として意味のあるデータ(例)身長、点数 |   |  |  |  |  |  |  |
| • (                              | <mark>質的データ</mark> ) =データをカテゴリで示したもの(例) h   | 生別、名前  |  |  |  |  |  |
| 2)分                              | 析の手法  |  |  |  |  |  |  |
| • (                              | <b>折れ線グラフ</b> ) =時系列で変化するデータを分析するの  | のに適している  |  |  |  |  |  |
| • (                              | 帯グラフ )・( 円グラフ ) =データの割合を見るのに  | こ適している   |  |  |  |  |  |
| • (                              | 散布図 $)$ ・( 相関係数 $)=2$ つのデータの関係性を気   | 分析するのに適している  |  |  |  |  |  |
|                                  |   |  |  |  |  |  |  |
| ⑤テキス                             | トデータの分析   |  |  |  |  |  |  |
| • ( 5                            | テキストマイニング ) =大量のテキストから何らかの特征  | 数を分析する処理   |  |  |  |  |  |
|                                  |   |  |  |  |  |  |  |
| 【実習】自                            | 自分の好きなアーチスト・曲の歌詞を分析しよう  |  |  |  |  |  |  |
|                                  |   |  |  |  |  |  |  |
| ◇手順 1                            | 自分の好きなアーチスト、曲を選ぶ(アートストの場合に  | は複数の曲を選ぶ)  |  |  |  |  |  |
|                                  |   | © User Local АРФХЬФИСЕР АФИ / ТОКПТ БЕГИ В В В В В В В В В В В В В В В В В В В |  |  |  |  |  |
|                                  |   | <b>プキがは、</b> 821 年   |  |  |  |  |  |
|                                  |   | MATCH:cofficiritimのでは、4. (New!) 解析したいテキストを入力する                                 |  |  |  |  |  |
|                                  |   | 存在   (プレン・イス・アセス・ブラ を)   (1.00(2)  |  |  |  |  |  |
|                                  | →歌詞のデータがあるところを探しておく。  | 927年 1 大地市 ECTAID A  |  |  |  |  |  |
|                                  |   |  |  |  |  |  |  |
| ◇手順 2                            | ブラウザの検索で「テキストマイニング」と入力し、  |  |  |  |  |  |  |
|                                  | スクロールさせた先にある「UserLocal (AI テキスト   | v. 19800.  |  |  |  |  |  |
|                                  | マイニングサービス <a href="https://textmining.userlocal.jp">https://textmining.userlocal.jp</a> | <u>/</u> )   |  |  |  |  |  |
|                                  | 」のページへ行く。   | <b>つらい</b> おめでとう   |  |  |  |  |  |
|                                  | 歌詞のデータをコピーまたは入力しに結果を確認する。   | 死ぬ <sub>いま</sub> 行く かさい ありがたい<br>くれる 羊 <sup>4</sup>                            |  |  |  |  |  |
|                                  |   | がい また 走る のまた 教 沈む 得っ ありがとう 大きい pro まま メート 本人 群衆 つきり                            |  |  |  |  |  |
|                                  |   | ロ惜しい 濁流 g セリヌンティウス Les   |  |  |  |  |  |
|                                  |   | * 暴君 身代り 友 佳い 結婚式 たまらばい はまる ままま ままま ままま エー・・・・・・・・・・・・・・・・・・・・・・・・・            |  |  |  |  |  |
|                                  |   | 描い 東る 物東 三日 笑う 高い だい ない おい ない 石い ない 石い ない 石い                                   |  |  |  |  |  |
| ◇手順 3                            | 結果から分かったことを分析する。  |  |  |  |  |  |  |
|                                  |   |  |  |  |  |  |  |
|                                  |   |  |  |  |  |  |  |
|                                  |   |  |  |  |  |  |  |
|                                  |   |  |  |  |  |  |  |
|                                  |   |  |  |  |  |  |  |
| 【振り返り                            | )】No.17 の実習・学習で学んだこと、気づいたこと、考え  | えたことを 3 行以上書きましょう。   |  |  |  |  |  |
|                                  |   |  |  |  |  |  |  |
|                                  |   |  |  |  |  |  |  |
|                                  |   |  |  |  |  |  |  |
|                                  |   |  |  |  |  |  |  |
|                                  |   |  |  |  |  |  |  |