

A1-1

「とどらん」でデータサイエンスの授業

アサンプション国際高等学校 岡本 弘之

生徒が興味を持つデータサイエンスの授業として、いろいろな都道府県のデータが紹介されているサイト「とどらん」を使った授業を2年間やってみた。数学科ではなく情報科らしく、仮説→データ分析→考察と問題解決の流れで行った授業と生徒の実際の事例について紹介したい。

1. はじめに

2022年度から始まった「情報Ⅰ」では、情報デザイン、プログラミング、データの活用の3つの内容が新たに加わった。このうちデータの活用について、学習指導要領では「データを表現、蓄積するための表し方と、データを収集、整理、分析する方法について理解し技能を身に付けること。」「データの収集、整理、分析及び結果の表現の方法を適切に選択し、実行し、評価し改善すること。」とあり、データを収集、整理、分析する方法を理解し、実行評価し改善することを求めている。

本実践ではこの内容を意識しつつ、「生徒が興味を持ちそうなデータを使う」、「情報科らしく『仮説』→『分析』→『考察』という問題解決の手順をたどらせたい。」という目標をたて、都道府県別の様々なデータを掲載する Web サイト「とどらん」を使ったデータの活用の授業を考えた。

2. 授業の実践

授業は高校1年生の「情報Ⅰ」の授業3時間を用いて実践した。授業の流れは下表の通りである。

表1 授業の流れ

1時間	【仮説を立てる】 ・「とどらん」にあるデータから相関関係のありそうなデータを選び仮説を立てる
2時間	【分析する】 ・データを収集・整理し、分析を表計算ソフトを使って行う
3時間	【考察する】 ・仮説が正しいか（相関があるか）について検証し、結果について考察する

2.1 仮説を立てる（1時間目）

最初に Web サイト「とどらん」(<https://todoran.com>)を紹介し、どのようなデータがあるかを確認する時間を約10分とった。本サイトを教材として選んだのは、人口・面積と言ったデータだけでなく、「Jリーガーの出身者数」「コンビニの店舗数」など生徒にとって身近な都道府県別のデータが多く掲載されているからである。

次に相関について説明し、本サイトの多くのデータの中から、相関のありそうな2つのデータを選び、その理由も含め仮説を立てる作業を行った。



図1 提示スライド①

考えた仮説は理由とともにワークシートに記録し、次の時間の分析へとつなげた。

2.2 分析する（2時間目）

Excel で作成したワークシートを配布（図2）し、仮説を立てた二つのデータを、「とどらん」のサイトからコピーしてワークシートに貼り付け、加工するように指示した。本サイトのデータは単位まで入っていて文字列のデータもあるため、「置換」を使ってデータを加工する必要がある。

ワークシートにはあらかじめ2つのデータの平均値・中央値・標準偏差と2つのデータの相関係数を計算する関数と散布図を入れており、データをコピーすれば自動でこれらの数値を算出するようにしている。



図3 提示スライド③

2.3 考察する（3 時間目）

前時の Excel シートの結果から、相関係数を読み取り、自分が選んだ 2 つのデータについて相関があるかどうかを判定し、その結果について考察する授業を行った。考察には「自分の仮説が正しかったのか」、「なぜそのような結果となったのか」について分析して書くように説明した。



図 3 提示スライド③

途中のクラスからはまとめとして、他クラスの仮説と相関があったかどうかの結果をスライドにまとめ、紹介した。当該クラスのデータも次の授業の冒頭に紹介し共有した。

3. 授業の結果

生徒がどのような仮説を立て、どのような結果となったか、事例を紹介する。

表 2 生徒の仮説と結果（抜粋）

・全国学力テストが高いほど校外学習率が高い → 弱い正の相関
・平均気温が高いほどアイスクリームショップが多い → 相関なし
・テニスコートが多いほどテニスプレーヤーが多い → 弱い負の相関
・ゲームプレイ率が高いほど宿題実行率が低い → 相関なし
・高校テニス部が多いほど大人でテニスする人が多い → 強い正の相関
・ゲーム利用率が低いほど学力テスト正答率は低い → 弱い正の相関
・塾の多い県ほど東大合格者が多い → 強い正の相関
・家族との会話が多いほど、自己肯定感が強い → 弱い正の相関
・人口密度が高いほど自家用車の保有台数は少ない → 正の相関
・本を読む県ほど学力が高い → 相関なし
・労働時間が長いほど、休養時間は短い → 弱い負の相関
・晴が多いほど病気になる人は少ない → 相関なし
・人口密度が高いほどコロナ感染者が多い → 強い正の相関
・犯罪件数が多いほど弁護士が多い → 相関なし
・テレビの視聴時間が長いほど学力テスト結果は低い → 相関なし

表 2 のデータからわかるように生徒の仮説は様々であり、自分が興味のある分野について仮説を立てて検証したことがわかる。これは「とどラン」にあるデータが多く分野にわたるため、生徒は自分の興味のある分野のデータを選択し、自分なりの仮説を立てることができるからである。

これら結果をスライドにまとめ教室で共有し

た際にも、「このデータとこのデータは関係がある（ない）のか」と、仮説・結果とも興味を持って聞いていた。

4. まとめ

4.1 授業の効果

本実践で活用した Web「とどラン」には生徒にとって身近で興味を引くデータが多くあり、そのデータから自分なりの仮説を立てることで、生徒は興味を持って授業に取り組むことができた。

自分で考えた仮説なので正しいかどうか気になり、主体的にデータ分析・考察に取り組むことができ、その中で体験的にデータ分析の流れを学ぶことができた。2 つのデータに関係があるかどうか考え検証することは、予想外の結果もあり生徒は楽しく学ぶことができた。また授業を行ったのが 3 学期だったため、数学 I でのデータ分析の授業とも重なり、教科横断的に学ぶこともできた。

4.2 情報科らしいデータサイエンスの授業

2022 年度から実施されている情報 I では、問題解決の視点が柱となっている。本実践でも仮説を立てる、分析する、考察・検証するという問題解決の流れを意識して授業を組み立てた。それが情報科らしいデータサイエンスの授業と考えるからである。（下図 4）

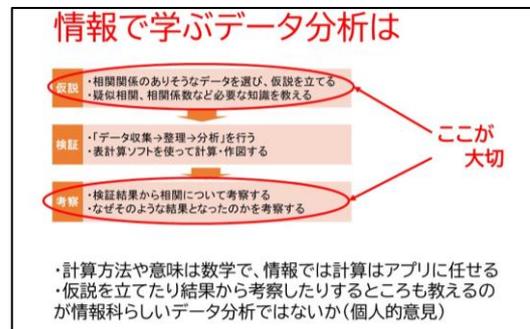


図 4 情報科らしいデータサイエンスとは

データ活用や相関については前述のように数学科でも扱う。理屈や計算方法は数学科で扱い、情報科では計算はアプリに任せてしまい、最初の仮説を立てるところ、計算結果から考察するところ = 「考えるところ」を重視した授業とすることは大切ではないだろうか。

参考文献

- (1) 黒上晴夫, 堀田達也, 村井純編:「情報 I」教科書, 日本文教出版, pp.000-000 (2023)
- (2) 久保哲郎:「とどラン」, <https://todo-ran.com> (参照 2023-05-31)
- (3) 岡本弘之「情報科の授業アイデア」, <https://okamon.jp> (参照 2023-05-31)