

高等学校 第2学年 情報科 学習指導案

「データサイエンス入門」(1単位)

期 日：2023年12月18日(月)

時 間：第3校時(11:15~12:05)

対 象：第2学年E組生命科学コース21名

学校名：清心女子高等学校

場 所：普通教室棟3階 高2E教室

授業者：坂部 高平

1 単元名

「データの活用」

(参考教科書)「最新情報I」(実教出版)

(参考図書)「絵と図でわかるデータサイエンス」上藤一郎著(技術評論社)

「完全独習統計学入門」小島寛之著(ダイヤモンド社)

「エピソードで学ぶ統計リテラシー」山田剛史/金森保智編著(北大路書房)

「文系でも仕事に使えるデータ分析はじめの一步」本丸諒(かんき出版)

2 単元の目標

- (1) 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。
- (2) 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。
- (3) 問題の発見や解決にデータを活用するために、適切なデータの選択や、分析の仕方、解釈の仕方について、粘り強く取り組み、試行錯誤を通じて改善しようとする。

3 単元の評価規準

知識・技能	思考・判断・表現	主体的に学習に取り組む態度
データを問題の発見・解決に活用するために、データの収集、整理、分析の要点を理解し、データの特徴を表す指標と、その評価について理解している。 「質的データ」と「量的データ」の性質や扱い方の違いを理解した上で、データの内容や形式を踏まえて処理する技能を身に付けている。 データの特徴と目的に応じた基礎的な分析及び可視化の方法を理解するとともにその技能を身に付けている。	必要なデータの収集について、選択、判断し、目的に応じて適切なデータの整理や変換の方法を判断することができる。また、その結果について多面的にアプローチし、可視化を行うことにより、データに含まれる傾向を見いだすことができる。 データの傾向に関して評価するための客観的な指標を選択して利用し、自身の考えを基にした適切な解釈を行うことができる。	情報社会の中でのデータと人との関わりについて考えながら、問題の発見・解決に向けて主体的にデータを活用するために、適切なデータの選択や、分析の仕方、解釈の仕方について試行錯誤し、粘り強く取り組んで改善しようとしている。

4 指導観

(1) 単元観

情報社会のさらに先の社会においては、モノのインターネット(Internet of Things=IoT)によるつながりとデータ収集が進む。現実社会のあらゆるものがセンサーなどでインターネットとつながるので、その情報量は飛躍的に増大し、その処理は人の能力を超えたものになるものと考えられる。そこでその処理のために人工知能(AI)の活用が進む。それだけでなく、AI ツールが身近な道具として休息に広がりを見せている現在、それら社会を支えるしくみについて、その構成員が無知でいてよいはずはない。情報が経済的な発展の道具から、環境面や人の暮らしを含めた社会基盤を支える道具になる世界での活躍が期待される世代に向けて、統計に関する内容を中心に扱う中で、データの利活用に関する科学的な知識や理解を促し、問題の発見や解決に向けて適切な手法でデータを手に入れ、それを効果的に用いる技能と、それをさらに粘り強く改善していく姿勢を養うことを目指す。

(2) 教材観

学校設定科目であり、教科書の内容を扱う科目にはなっていないので、教材はオリジナルで作成したものを使用する。ただし、情報科の科目「情報 I」の内容の一部も扱う必要があるため、「情報 I」の授業と連携する。本単元はデータの活用がテーマであるため、統計に関する内容が中心である。ICT 機器やそれを利用するアプリケーションはあくまでも道具であり、今手にしている道具を効果的に利用していくための基本的な知識・技能と、それを自分自身で必要に応じて発展させることのできる基礎となる論理的な思考を可能とするように理解を深めてもらうことを意識する。今年度から高校生は全員が ICT 機器としての同一機種 of PC である Chromebook を手にしており、校内においていつでもどこでもインターネットに接続された同じ道具を利用できる環境が整った。データサイエンスについて考え、学んでいく上で表計算は避けて通ることのできないものツールであることもあり、実習は Google スプレッドシートを利用したものを中心に授業を展開する。教材のやりとりについては Google Classroom を利用し、オンラインでの小テストはほぼ授業ごとに実施して Classroom で返却し、ファイルによる課題提出を Classroom で回収することで、学習の過程が自動的にオンライン上に蓄積されて散逸しない状態で授業を進めている。

データの処理に取り組むにあたり扱うデータについては、模式的なものを準備して、その処理のしくみの根本的な理解を促し、その次の段階では公的な実際のオープンデータを用いた作業も取り入れ、データサイエンス、特に統計のことを身近に感じ、自分たちにも使うことができるものであり、有用な道具・手法であるという意識を持たせたい。

(3) 生徒観

全員が ICT 機器として Chromebook を手にして、各教科の授業でも利用されていることから、PC の操作に不安がある生徒は皆無である。特に、この授業対象となっている生命科学コースの生徒は、どちらかという理系の科目や進路に興味をもつ生徒が多くを占めていることもあり、PC 利用にアレルギー反応を示すような生徒は見られず、課題にも比較的真面目に取り組んでいる。本単元も終盤となり、Google スプレッドシートを統計の道具として利用することに、生徒も少しは慣れてきたものと思われる。

本科目の授業時間内においては、個人作業においては各自が静かに取り組んでおり、ほぼ毎時間実施するオンラインによる確認テストにも真面目に取り組んでいるようすが窺える。グループワークを実施すると意見交換によって雰囲気や和み、お互いの情報交換、教え合い、助け合いの場面を作りやすいことから、協働による学習意欲の向上と技能習得の促進のため、要所においてグループワークを取り入れて学習効果の向上を図りたい。

5 年間指導計画における位置付け

学年	第2学年 (1単位)「データサイエンス入門」	第2学年 (1単位)「情報I」〔並行履修〕
単元名	①データの活用 (←本時の単元) ②モデル化とシミュレーション ③アルゴリズムとプログラミング	①情報社会 ②メディアとデザイン ③システムとデジタル化 ④ネットワークとセキュリティ
主な内容	①データの種類や分析の手法を知り、統計のための基礎技能の習得 ②モデル化の手法を知り、シミュレーションによる問題解決を考える ④Pythonの基礎的な技能の習得と合わせたアルゴリズムの理解と体得	

6 単元の指導計画と評価計画

指導計画	第一次	データ種類と特徴.....	2時間
	第二次	データの集め方と仮説.....	2時間
	第三次	データとバイアス.....	2時間
	第四次	データの可視化と代表値.....	4時間
	第五次	データの分布と相関関係.....	4時間
	第六次	データからの予測.....	3時間 (第3時 本時)
	第七次	データの倫理.....	1時間

〔第六次 指導計画〕

時	★目標 ○学習内容 ・学習活動	■評価規準 (評価方法)
第1時	★データの相関関係と回帰分析について知り、単回帰分析を行うことができる。	
	○相関関係と回帰分析について知る。 ・政府統計のオープンデータを用いて、Google スプレッドシートで相関係数や決定係数を求め、散布図に回帰直線を描かせてみる。	■知識・技能 スプレッドシート上で相関係数、決定係数を求め、回帰直線を描かせることができる。 ■思考・判断・表現 相関のあるデータ間の因果関係を考えることができる。
第2時	★データの中から相関関係を見つけ出し、因果関係を考えて単回帰分析をすることができる。	
	・政府統計のオープンデータから複数の組合せを選び出し、Google スプレッドシートで散布図を作成して単回帰分析を行う。	■知識・技能 スプレッドシート上で回帰分析をすることができる。 ■思考・判断・表現 数あるデータの中から相関のあるものを適切に選択することができる。 ■主体的に学習に取り組む態度 課題に粘り強く取り組むことができる。

第3時 (本時)	★データ間の因果関係を深く考え、要因の可能性を多面的に考えることができる。	
	○関連のあるデータから因果関係を確認する。 ○地域に細分化したデータを元に地域差を考え、さらなる要因を予想する。 ・政府統計のオープンデータを用いて、Google スプレッドシートで複数のデータを比較し、地域差の情報を加味して、追加要因を予想する。	■知識・技能 スプレッドシート上でスムーズに複数の散布図を作成して比較することができる。 ■思考・判断・表現 複数のデータを比較し、因果関係を考察して複数の要因の可能性を示唆することができる。 ■主体的に学習に取り組む態度 データと知識を照らし合わせ、目的を達するために協働して粘り強く取り組むことができる。

7 指導に当たって

- (1) 本単元第四次の目的は縮約の方法を知ることである。具体的には、グラフ化と各種統計量について、その種類と特徴を知って、適切なものを選んで提示することができることを目標とする。グラフについてはスプレッドシートでいくつか作成をしてみることで、表とグラフの間のつながりを概ね把握させたい。ただし、箱ひげ図については現時点のスプレッドシートでは作成できないが、統計グラフとしては身につけておいてもらいたい。手書きでの作成に取り組んでおく。統計量については、それぞれの意味するものができる限り丁寧に伝えることで、その必要性を把握できるように進めたい。
- (2) 課題解決につながる思考法としての論理的思考・批判的思考・水平的思考の3つを養う授業展開を目指している。それらの思考を促す場面の設定を意識したい。生徒個人で作業や思考をおこなう場面において論理的思考を促すだけでなく、生徒同士での結果の共有と意見交換ができるように進めることで、批判的思考や水平的思考を促したい。

8 本時案（第六次 全3時間中の3時間目）

(1) 本時の目標

- 政府統計のオープンデータを使った散布図の作成を通して、相関の強さとデータの分布との関係性をつかむ。
- 異なる視点から見たデータを比較することから、複数の要因の可能性を考え、多角的な視点をもってデータを見る重要性を知る。

(2) 本時の展開

時間	○学習内容 ・学習活動	・指導上の留意事項	■評価規準（評価方法）
導入 (8分)	<ul style="list-style-type: none"> ○相関と因果関係について振り返り、本時はデータの相関を見ながら追加の要因について考えることを知る。 ・オンラインで回帰分析に関する小テストをして復習する。 ・本時のプリント配布 ・アイスクリーム類消費金額と気温のデータの散布図を提示し、相関について確認する。 	<ul style="list-style-type: none"> ・相関関係について思い出させる。 	<ul style="list-style-type: none"> ■知識・技能 相関や線形回帰について理解している。
展開 (40分)	<ul style="list-style-type: none"> ○政府統計のオープンデータから得た家計調査の食料品の品目別データの地域別消費金額ランキングを比較する。 ・提示された各品目の地域別棒グラフを見て、上位に入っている地域の共通性を考える。 ○他の例と異なる特徴をもつアイスクリーム類の消費金額で金沢市が上位になっていることから、地域別のデータを確認する。 ・配付データを利用し、散布図を作成して、その違いから地域差を考察する。 ・その他欲しい情報があれば、Webで検索する。 ・生徒が予想した要因を情報共有する。 	<ul style="list-style-type: none"> ・消費金額上位の地域は、その品目の特産地であるものが多いことに気づかせる。 ・机間指導する中で、必要に応じて生徒同士での相談を促す。 ・調査地域の地理的な位置関係も確認させる。 ・配付データにヒントとなりうる他品目も含める。 ・積極的に、生徒同士での相談を促し、机間指導する。 ・情報共有は共有ファイルを利用する。 	<ul style="list-style-type: none"> ■知識・技能 ・散布図を作成できている。 ・相関関係が理解できている。 ・地域の地理的な位置関係が把握できている。 ■思考・判断・表現 論理的に考えをまとめ、他者の意見も適切に受け入れながら、情報をまとめることができている。 ■主体的に学習に取り組む態度 個人及びグループで積極的に作業や議論ができている。
まとめ (2分)	<ul style="list-style-type: none"> ○企業団体による調査報告を知る。 ・一般社団法人日本アイスクリーム協会による調査報告書の存在を紹介する。 	<ul style="list-style-type: none"> ・データを利用した実際の報告書の例として報告書の内容を提示する。 	

(3) 板書計画 (今回配付する授業プリント : A4 裏表 2 ページ)

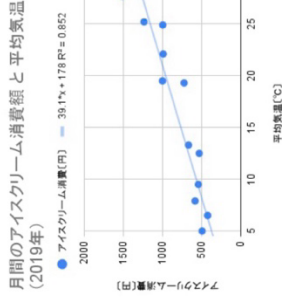
データサイエンス入門 授業プリント No.17

データの因果関係をさらに考える

データには相関関係はあっても、因果関係がない場合があるので、分析するときには注意が必要です。さらに、その要因(原因)が一つとは限らないので、そのことを考えてみましょう。

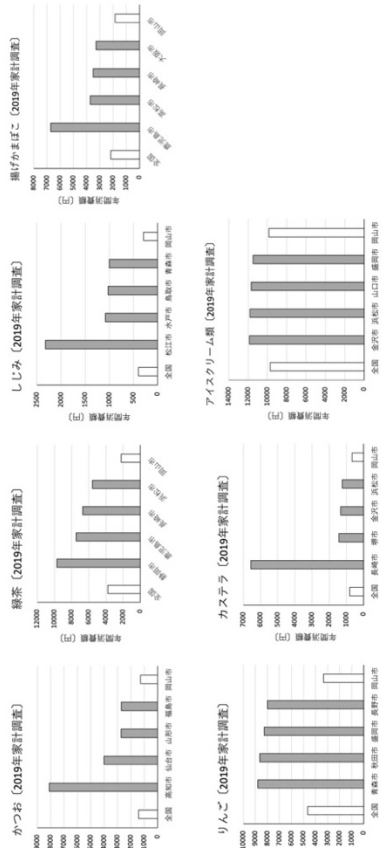
まずは一つの相関関係の確認

以前も出てきましたが、政府統計の2019年の家計調査と気象情報から、アイスクリーム類の世帯当たり月間消費金額と月平均気温の間には(正)の相関がありました。(復帰)直線をひくと右図のようになり、回帰式と(決定)係数は図の中に表示されているようになります。そこには因果関係がありますので、これによって予測が可能になります。



【考えてみよう15】

食品の品目別世帯当たり年間消費額を地域(都市)別に見ると、地域の特徴が見えてくるはずです。そこで、いくつかの品目の消費額上位を取り出してグラフにしたものが下図になります。上位、特に1位になっている地域の共通点は何なのか考えてみましょう。(下図のグラフが小さくて見にくいので、Classroomで配付するスプレッドシート「No17 考えてみよう15」のシート1で確認してください。)



要因は1つではない

アイスクリーム類の消費額だけは地域との関係性がわかりません。しかし、1位の金沢市が他の地域より気温が高いとは考えにくいでしょう。では、偶然かというところ、右図のデータを見る限り、偶然ではなさそうです。この結果には気温の高さに比例する以外の要因が影響していることが予想されます。では、そのような要因にはどんなものがあると考えられるでしょうか。

【練習13】

アイスクリーム類の消費額に対する気温以外の要因を予想してみましょう。Classroomからスプレッドシート「No17練習13シート」にいくつかのデータを追加に入れて配付するので、それらを踏まえて、気温の高さと比例する以外の要因として何が考えられるか、次の手順で検討してください。(データはいずれも政府統計のオープンデータからです。)

手順①: シート1の東京, 金沢, 山形, 岡山それぞれのアイスクリーム類の世帯当たり月間消費額と月平均気温のデータ (2019年1月~2023年9月) から、それぞれ散布図を作成して比較します。(相関係数の値や回帰直線の表示を入れると、なお良い。)

手順②: シート2にある2019年の地域別の世帯当たり年間都市別消費額のデータを、フィルタ機能を使って並べ替えてみて、金沢と他の地域特性を確認して比較します。

【フィルタ機能】1つの列の値から、昇順・降順など条件によって行を並べ替えることができます。スプレッドシートで「データ」メニューから「フィルタ作成」を選択すると、表の一番上の行のセルの中に「≡」マークが現れるので、条件になるセルの「≡」マークをクリックして並べ替える条件を指定する。メニューから「フィルタ削除」を選択すればマークが消える。

手順③: ①②から予想される要因に関係して、金沢に関して知りたい情報があればWeb検索で探出して、予想を補強する材料にする。

手順④: 予想した要因についてClassroomからシートと一緒に配信して共有している「No17練習13報告」を開いて左上のセルから下へ順に入力する。(いくつでも。他の人と内容が同様であってもよい。)

	1位	2位	3位
2022年	金沢市	さいたま市	堺市
2021年	さいたま市	福島市	山形市
2020年	金沢市	宮崎市	千葉市
2019年	金沢市	浜松市	山口市
2018年	浜松市	山形市	名古屋市
2017年	金沢市	福島市	青森市
2016年	富山市	さいたま市	静岡市
2015年	金沢市	岡山市	横濱市
2014年	金沢市	川崎市	富山市