



連載



## 情報の授業をしよう！

本コーナー「情報の授業をしよう！」は、小学校や中学校で情報活用能力を育む内容を授業で教えている先生や、高校で情報科を教えている先生が、「自分はこの内容はこういう風に教えている」というノウハウを紹介するものです。情報のさまざまな内容について、他人にどうやって分かって

もらうか、という工夫やアイデアは、読者の皆様にもきっと役立つことと思います。そして「自分も教え方の工夫を紹介したい」と思われた場合は、こちらにご連絡ください。

(E-mail : editj@ipsj.or.jp)

# 探究的データサイエンス



## —「GIS × 統計分析」で仮説検証サイクルを回す授業実践—

西出新也 | 関西学院千里国際中等部・高等部

### 自由な校風と ICT 教育環境

私が勤務する関西学院千里国際中等部・高等部は、大阪府箕面市に位置し、同じキャンパス内に関西学院大阪インターナショナルスクールを併設したユニークな学校である。1987年に文部省（当時）が中等教育課程における新国際学校設置の提言を行い、それを受けた形で、1989年に第1号として都立国際学校が、次いで、1991年に本校の前身である学校法人千里国際学園が設立された。開校当初から帰国生の受け入れを積極的に行い、現在も一般生徒と海外生活をバックグラウンドに持つ生徒がほぼ半数ずつ在籍している。関西学院大阪インターナショナルスクールと一部の授業では「シェアードプログラム」として、両校の生徒がともに学び、行事やクラブでは、一緒に活動できる環境が用意され、まさに“Two Schools Together”を体現している。

本校には校則はなく、「5 リスペクト（自分、他人、学習、リーダーシップ、環境の5つを大切にする）」という独自の指針に照らし合わせて、生徒自らが考

えて行動することを重んじる校風がある。ICT活用においても、生徒に規制をかけるのではなく、活用することを前提とした上で、「どのようにすればテクノロジーの善き使い手になれるか」について、教員・生徒と一緒に考え、議論する機会を増やす取り組みを行っている。

ICT教育環境では、生徒が自分自身のPCやタブレットを持参するBYOD (Bring Your Own Device) を導入しており、ほとんどの授業においてICTを活用した授業が展開されている。また、本校ではGoogle Workspace for Education が導入されているため、授業コンテンツの管理はGoogle Classroomで行われ、コロナ禍でもスムーズな授業運営が可能であった。

本校では「情報I」の扱いについて、情報科の10種類の開講科目から履修選択をする形式を採用している。プログラミングに重点を置いた科目や、関西学院大学との高大連携授業であるAI活用に関する科目など、情報Iの分野を足がかりに、生徒たちの学びへの意欲を刺激するバラエティに富んだ内容を展開している。今回は、その中の1つである「データサイエンス」



の授業実践に焦点を当てて紹介したい。

## 情報Ⅰ「データの活用」の位置づけ

情報Ⅰで学ぶ4つの項目のうち、データの活用を取り扱う「情報通信ネットワークとデータ活用」の学習目標(平成30年告示高等学校学習指導要領)では、「データを蓄積、管理、提供する方法、データを収集、整理、分析する方法、情報セキュリティを確保する方法を身に付けるようにし、目的に応じて情報通信ネットワークや情報システムにより提供されるサービスを安全かつ効率的に活用する力やデータを問題の発見・解決に活用する力を身に付ける」ことが示されている。データを分析した結果の「意味」について考察することが特徴であり、まさに「活用」に主眼を置いた内容である。一方、情報Ⅱの「データサイエンス」における学習目標では、「不確実な事象を予測するなどの問題発見・解決を行うために、データの収集、整理、整形、モデル化、可視化、分析、評価、実行、効果検証など各過程における方法を理解し、必要な技能を身に付け、データに基づいて科学的に考えることにより問題解決に取り組む力を養う」とされている。情報Ⅰでのデータの「収集・整理・分析」から、情報Ⅱでは、「収集・整理・整形・モデル化・可視化・分析・評価・実行・効果検証」と専門的に学ぶ内容が大きく増えている。もちろん、数学Bにおける「統計的な推測」との関連も深いため、教科横断的な学習を展開することで、より効果的に進める工夫も考えられる。しかし、本校では学期ごとの帰国生を受け入れるため「学期完結制」のカリキュラムとなっており、数学の分野と連携することは容易ではない状況にある。こうした状況を踏まえた上で、本校が情報Ⅱの開講について具体的に検討を始めるにあたり、あらためて「データの活用」と「データサイエンス」の学習内容を整理した(表-1)。それにより、従来の「データの活用」に、情報Ⅱで扱う統計・分析の一部と探究的な視点を加えることで、「データサイエンス」へ自然な形で歩みを進めることができ、シームレスな接続が可能になるのではないかと考えた。

## シームレスな接続と探究的視点

これまでのデータを活用する授業では、こちらでサンプルとなるデータを用意し、それらについて「散布図作成・相関係数・単回帰分析」などを扱っていた。特に、相関分析から考察するプロセスにおいては、提供するデータの種類により、生徒たちの興味関心に大きな影響が出ることが、これまでの授業実践から明らかであった。また、考察の幅も限定されやすく、問題発見や解決の視点につながる展開ができていなかった。

探究的な視点を取り入れるため、データを提供することから一歩踏み出し、生徒自らが課題を発見し、データを取得・分析することを目指した。たとえば、生活圏にある身近な問題や地理的な課題などは、生徒たちの当事者意識が高まるテーマである。「地理総合」でも取り扱う、GIS(地理情報システム)には、デジタル地図上にさまざまな地理情報を重ねて可視化する機能が備わっており、これらは生徒による探究的な課題発見においてまさにうってつけの機能である。さらに、その課題に関連するデータを取得することで、探究的なデータサイエンス学習へつなげることが可能になると見込んだ。

## ツールの選定(GIS・統計分析)

### jSTAT MAP

GISのソフトウェアを選定する上では、本校の

■表-1 情報Ⅰと情報Ⅱ「データの活用」比較

	情報Ⅰ	情報Ⅱ
統計*	統計指標(分散・標準偏差・相関係数など)、散布図、検定の考え方など	統計的な推測(標本調査・母集団の特徴や傾向)、仮説検定の方法など
分析*	クロス集計、仮説検定、単回帰分析、これらを通じたデータの可視化、現象のモデル化と予想	重回帰分析、分類、クラスタリング、これらを通じた可視化、現象のモデル化と予測及びモデルの評価、機械学習
データ*	量的データ、質的データ、外れ値、欠損値	多様かつ大量データを扱い、バイアスなどデータの信頼性にかかわることに配慮する
尺度	尺度水準(名義・順序・間隔・比例)の違いを扱う	
データベース	情報を収集・蓄積・提供する方法として全員が学ぶ	データの整形など、データを扱うプログラミングにも触れる

\*「情報Ⅰ」の一部内容は中学数学「Dデータの活用」、高校「数学Ⅰ」の「データの分析」と関連している。

BYOD環境が1つのハードルとなる。生徒は多様な機種やスペックの端末を持ち込むため、統一したアプリ・ソフトウェアのインストール指示ができない。そのため、ブラウザアプリで実現可能なことが条件となった。そこで、WebGISである「jSTAT MAP」(総務省統計局および独立行政法人統計センター提供)を採用した(図-1)。このGISは、政府統計のデータを地理的な情報と組み合わせることで視覚化することができ、さまざまな統計データを地理的な文脈で解析するために設計されている。主な機能は次のとおりである。

### 1. 統計データの地理的表示

人口統計、経済活動、社会インフラなどの各種政府統計データを地図上に表示する。これにより、データが具体的な地理的位置と結びつけられ、分析や比較が容易となる。

### 2. 多様なデータレイヤーの重ね合わせ

ユーザーは、異なる統計データを重ね合わせて表示することが可能となる。これにより、地域間の比較や、異なるデータセット間での関連性について探索することができる。

### 3. カスタマイズ可能なデータの可視化

データの表示方法(色分け・サイズ・形状など)

をカスタマイズすることができ、より直感的で理解しやすいビジュアルの作成が可能となる。また、特定の条件に基づいたデータのフィルタリングも行うことができる。

### 4. データのエクスポート

分析結果をCSVファイルとしてエクスポートする機能があり、これにより分析ツールへのスムーズなデータ移行を行うことができる。

この他、リッチレポート<sup>☆1</sup>の出力機能などもあり、統計データと地理情報の統合から新たな洞察を引き出すことが可能である。

## XLMiner Analysis ToolPak

GISから抽出したデータを分析する手法については、主に情報IIで扱う「重回帰分析」を採用した。これは複数データの関連性を明らかにする手法であり、各説明変数が目的変数にどの程度の影響を与えているかを捉えることができる。本校では、Google スプレッドシートを表計算・分析ツールとして活用しているため、重回帰分析を実施するにあたり、多様な統計分析機能を提供する「XLMiner Analysis ToolPak」

☆1 指定したエリア内の統計量を出力するエリア分析レポート

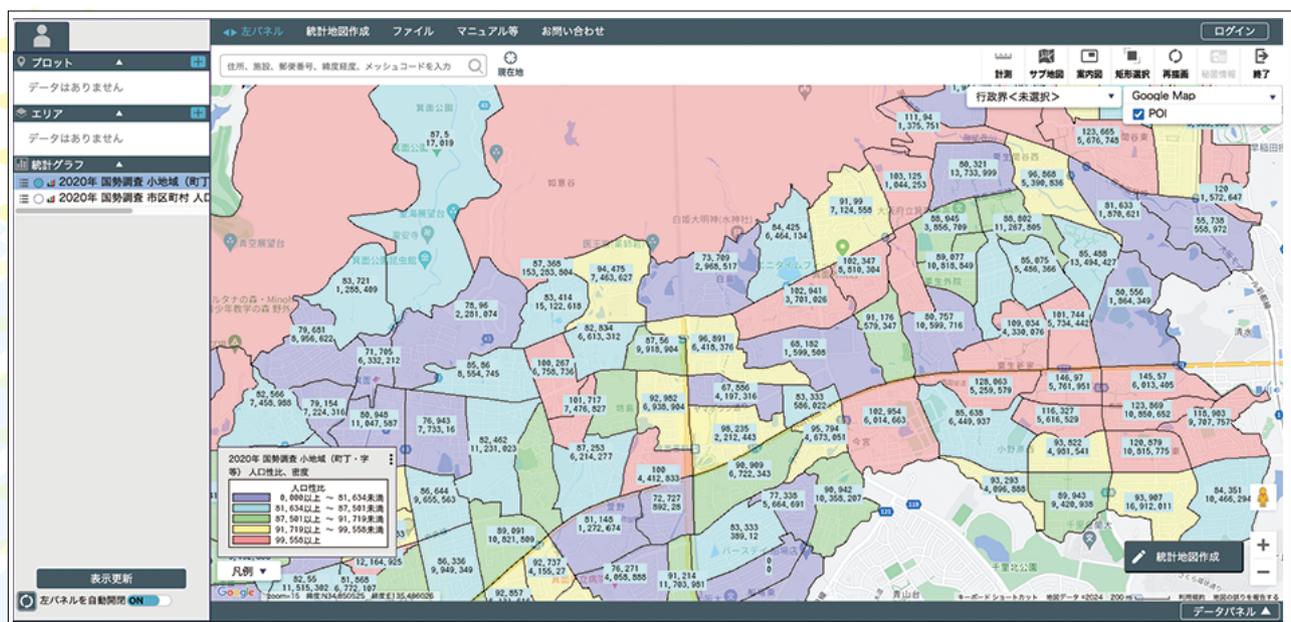


図-1 jSTAT MAP



のアドオンを使用させることにした(図-2)。このツールの主な機能は次のとおりである。

### 1. 記述統計

データセットの基本的な記述統計(平均・中央値・標準偏差・分散など)を提供する。データの概要を把握することができる。

### 2. 回帰分析

単回帰や多重回帰分析を行い、変数間の関係性を解析する。予測モデルを構築することができる。

### 3. t検定とANOVA

2つのサンプル間の平均値の差の検定(t検定)、複数のグループ間での平均値の比較(ANOVA)を行う。データ間の統計的差異を明確にすることができる。

### 4. 相関分析

2つの変数間の相関係数を算出する。各変数の関連性の強さと方向を評価することができる。

### 5. ヒストグラム

データの分布や外れ値に対応したグラフィカルツールの提供がされている。視覚的に分析することができる。

このXLMiner Analysis ToolPakの導入により、

Google スプレッドシートでも、複数データの統計分析を直感的に行うことが可能となる。

## 授業実践

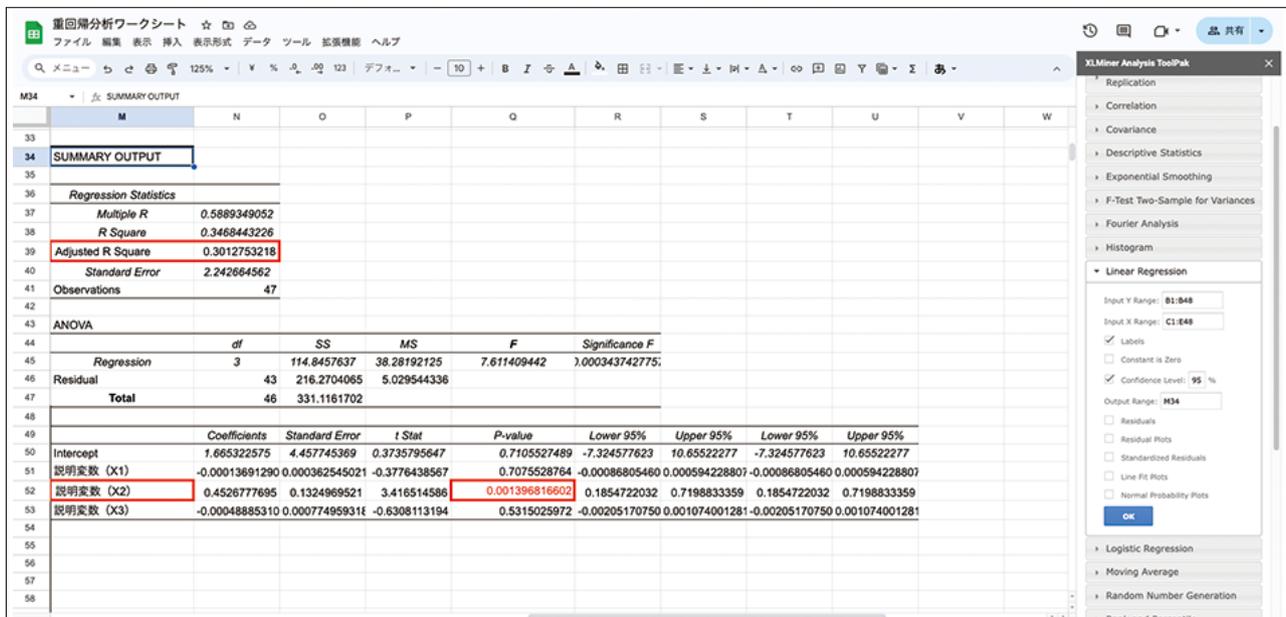
### 導入「データの活用」

本校の学期完結制カリキュラムでは、1学期に35時間で1単位の修得となる。それに基づき、探究的データサイエンス全体の学習活動を表-2のように設定した。

このうち、表の項目1に挙げた学習活動では、情報Iの「データの活用」を意識して導入を進めた。

■表-2 全体の学習活動(35時間)

項目	学習活動
1	<ul style="list-style-type: none"> <li>・イントロダクション</li> <li>・データの活用</li> <li>・形態素解析(2h)</li> <li>・社会における情報システム</li> <li>・基本統計・相関分析(3h)</li> <li>・単回帰分析(3h)</li> </ul>
2	<ul style="list-style-type: none"> <li>・jSTAT MAP 演習(3h)</li> <li>・重回帰分析(4h)</li> </ul>
3	<ul style="list-style-type: none"> <li>・データサイエンス探究・前半(5h)</li> <li>・中間発表・フィードバック</li> <li>・データサイエンス探究・後半(4h)</li> </ul>
4	<ul style="list-style-type: none"> <li>・研究ポスター・プレゼン動画制作(3h)</li> <li>・知識理解到達度テスト</li> <li>・完成作品の共有</li> <li>・相互評価</li> <li>・リフレクション</li> </ul>



■図-2 XLMiner Analysis ToolPak

生徒は、形態素解析の学習活動において、Google Colaboratory に MeCab をインストールして、ワードクラウドの生成を使った歌詞の年代別比較を体験した。また相関分析では、都道府県別のデータを活用して、散布図の作成と相関係数の算出を行い、その結果から考察をスライドにまとめて、グループ内で発表を行った (図-3)。

## GIS と重回帰分析の理解

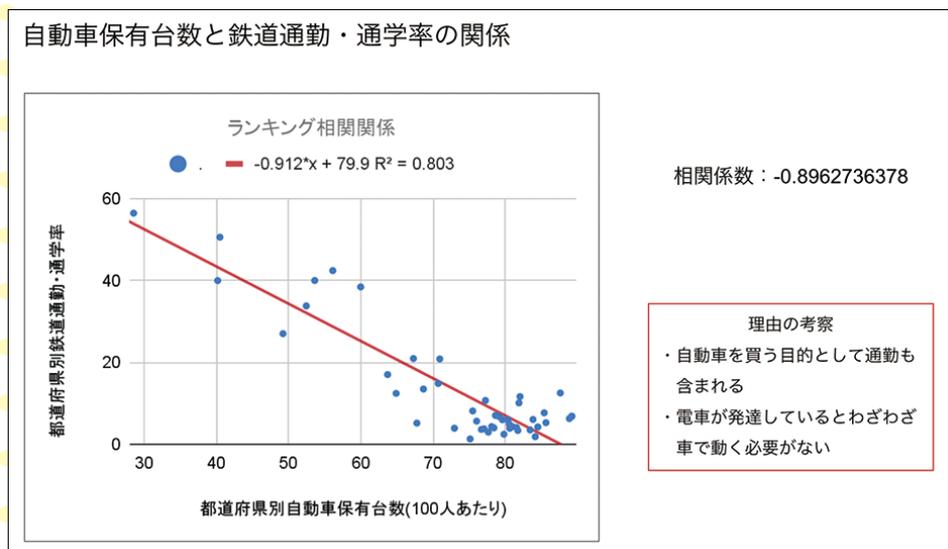
次に、「探究的データサイエンス」の学習へ一歩踏み込むための準備として、GIS と重回帰分析の理解を目指し、学習を展開した。jSTAT MAP 演習

(3 時間) では、生徒がこの地理情報システムの使い方と特性を実践的に理解することに主眼を置いた。ユーザー登録、統計データのプロット読み込み、リッチレポートの作成といった一連のプロセスを体験的に学習することで、jSTAT MAP を用いた地理的データの視覚化と簡単な分析レポート作成のスキル習得につながった。このことにより、地理情報に基づくデータ解釈の重要性について、生徒の理解はできたものと思われる。

重回帰分析 (4 時間) のセッションでは、複数の説明変数を用いた回帰式の分析結果の解釈方法に焦点を当てた。補正済み決定係数 (補正 R<sup>2</sup>) や p 値

といった重要な指標の意味と解釈については、講義スライドと動画資料を用いて、生徒が繰り返し見直すことができる環境を準備した (図-4)。その後、ワークシートを使って実際のデータを分析しながら、重回帰分析の概念と実践的な運用方法の学習を行った。これにより、多変量データの複雑な関係性を探索・解釈するための基礎的な知識を身につけるとともに、それらを適切に活用するための理解を深めることができた。

jSTAT MAP を用いたデータの可視化と、重回帰分析を用いた多変量データの関係性の探索は、本実践のデータサイエンスの探究的アプローチにおいて重要なポイントになると捉えている。これらの演習と学習を通して、



■ 図-3 相関分析による考察スライドの例

係数の有意性をみるには「P-値」を確認します。一般的にP-値が0.05未満であれば、その説明変数は目的変数に対して「**関係性がある**」という判断をします。0.05以上の場合は「関係がない」と捉えることができます。

概要								
回帰統計								
重回帰 R	0.98750518							
重決定 R <sup>2</sup>	0.97516649							
補正 R <sup>2</sup>	0.95529968							
標準誤差	168.715095							
観測数	10							
分散分析表								
	自由度	変動	分散	測された分散	有意 F			
回帰	4	5588799.68	1397199.92	49.085212	0.00033411			
残差	5	142323.916	28464.7832					
合計	9	5731123.6						
	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
切片	7201.43848	651.867691	11.0473929	0.00010583	5525.75923	8877.11772	5525.75923	8877.11772
席数(席)	16.9404013	11.840429	1.43072529	0.21191423	-13.496391	47.3771931	-13.496391	47.3771931
徒歩時間(分)	-85.748922	30.3393891	-2.8263233	0.0368336	-163.7388	-7.7590393	-163.7388	-7.7590393
モーニング有無	14.835115	433.89296	0.03419073	0.97404801	-1100.5222	1130.19248	-1100.5222	1130.19248
限定商品数(個)	168.845795	43.4929829	3.88213877	0.01161678	57.0435234	280.648067	57.0435234	280.648067

**p < 0.05**  
「帰無仮説が棄却された」  
「対立仮説が支持された」

■ 図-4 重回帰分析の講義スライド

生徒はGISの役割と重回帰分析の基本的な概念を理解し、情報IIの内容を見据えた学びに入る準備を整えることができた。

## データサイエンス探究の実践

データサイエンス探究の流れは、大きく3つのステップとした。まず、Step1では、日常生活やニュースから疑問や興味を引き出すことから始めた。生徒同士のディスカッションの中から、「空き家が最近増えているって聞いたけど?」や「ゴミの排出量は増えているの?」といった疑問を拾い出し、それらを探究の出発点とした。

Step2では、生まれた疑問に関連するデータを「SSDSE(教育用標準データセット)」や「e-Stat(政府統計の総合窓口)」といったデータポータルから見つけ出し、jSTAT MAPにインポートすることで地図上に可視化した。この過程で、数字だけでは見えなかった地理的な傾向や特徴から、たとえば、「この地域において〇〇が多くあるのは△△の影響ではないか?」といった仮説を見出すように導いた。

最後にStep3では、重回帰分析を用いて、「目的変数」に影響を与えている「説明変数」について、

補正済み決定係数(補正R2)やp値といった指標を読み取った。そして、その分析結果から示唆されることを生徒自身で考察し、結論を導き出していく。ただし、この時点では因果関係までは断言ができないという注意点については、中間報告の議論の中で理解を深めていくこととした。考察には、jSTAT MAPのリッチレポート分析を取り入れる工夫など、見る者に納得感を与える結果を導き出すことも評価の観点として挙げた(表-3)。

## 探究ポスター・プレゼン動画の共有

中間報告では、アイデアをスライドにまとめて、5人程度のグループ内で行った。その時点で、お互いに気づいた点や疑問点を挙げて議論を行った。また、教員からのフィードバックを得て、軌道修正をしながら完成につなげていった。最終的には探究的な分析をポスターにまとめ(図-5)、それらについてのプレゼンテーションを各自、動画に収めた(図-6)。

GISに統計データを通して可視化することで、数字だけでは気づくことのできなかった地理的な要因や、それらに関連する要素が浮かび上がってくる感覚を掴んだ生徒も多かったようだ。授業が後半に進

表-3 評価の観点

評価の観点	
データ収集と前処理	<ul style="list-style-type: none"> <li>データの収集方法と選択</li> <li>データの整形</li> <li>空間データの整理と正確性</li> </ul>
データの可視化	<ul style="list-style-type: none"> <li>地図上での分かりやすいデータの可視化</li> <li>適切な色別・ラベル・凡例の表示</li> <li>地理的パターンや相関の視覚的表示</li> </ul>
重回帰分析	<ul style="list-style-type: none"> <li>モデルの選択と正当化</li> <li>目的変数と説明変数の選択理由</li> <li>統計的有意性の評価</li> </ul>
結果の解釈	<ul style="list-style-type: none"> <li>結果指標の適切な読み取り</li> <li>地理情報と統計結果の結びつき</li> <li>課題の目的に対する適切な分析</li> </ul>
結論と発表	<ul style="list-style-type: none"> <li>明示的な洞察を含む</li> <li>グラフや図表など視覚的な訴求力</li> <li>考察と結論から提言につながる</li> </ul>
イノベーションと改善	<ul style="list-style-type: none"> <li>新しいアイデアとアプローチ</li> <li>課題解決への改善策とフィードバック</li> </ul>

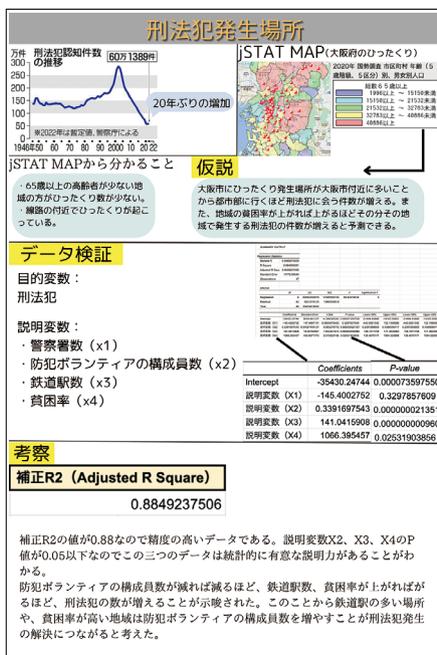


図-5 探究ポスターの完成例



むにつれ、自然と生徒同士で、分析結果についての話をすることや、意見を求める姿がよく見られた。

本実践の終了後に、この授業を通してどのような学びを得たか、ということについて、生徒ヘリフレクションを求めた(図-7)。「データ採取の大変さ」や、「データ整形の大切さ」について学びを深めることができた、という意見も複数見られた。

## 「情報II」接続への展望

今回は、情報I「データの活用」から、情報II「データサイエンス」へのシームレスな接続を意識した、試行的な教育実践を紹介した。この取り組みでは、データの「整形・可視化・分析・効果検証」まで範囲を広げて扱ったが、情報IIを開講する上では、さらに分析のアプローチ(分類・クラスタリングなど)や、回帰分析の内生性への対処など、盛

り込むべき内容は多く残されている。しかし、この実践を通して、GISを活用したデータの可視化を手法として取り入れることで、探究的な切り口で学びを広げることできる、という一定の手応えを感じた。

高等学校DX加速化推進事業(DXハイスクール)も開始され、今後さらに「情報II」を開設する高等学校は増加すると考えられる。今回の授業実践の紹介が、その開設を見据えた「シームレスな接続」を検討する上の一助になれば幸いである。

(2024年5月1日受付)



西出新也(正会員)  
snishide@kwansei.ac.jp

関西学院千里国際中等部・高等部(技術科・情報科教諭)、神戸大学大学院国際文化学研究所情報コミュニケーションコース(博士後期課程在学中)。

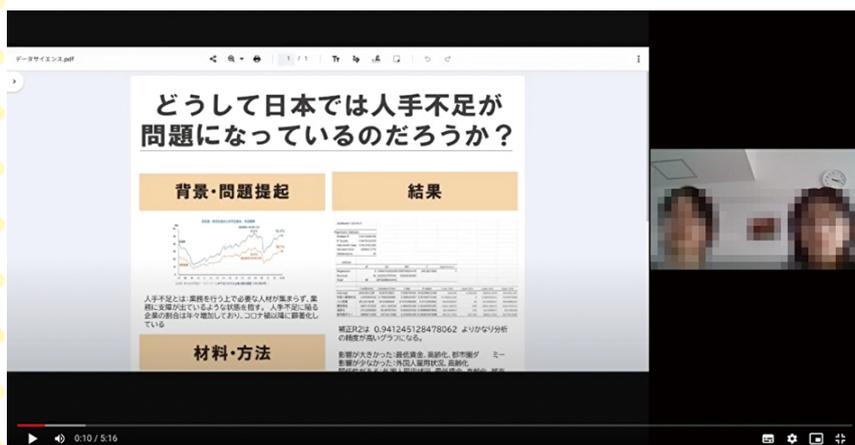


図-6 プレゼンテーション動画の例

### 2. 生徒リフレクション (GIS × 重回帰分析の探究課題を通してどのような学びを得たか)

→ GIS × 重回帰分析の研究課題を通して、まず基本的な重回帰分析の使い方や、p値や修正R2などどこを見れば良いのかということや、実際に興味を持ったことを研究してみて、調べてもなかなか出てこないなどといったデータの採取の大変さやデータ検証を行って見た時に発生した不具合などのトラブルから解決した時に判明した研究課題の結果までを知ることができ、今まで何も知らなかったことを新たな知識として取り入れることができました。また、自分たちの研究課題の結果だけでなく、みんなの発表を聞いて新しい疑問が生まれたり、その疑問が解決したりなど周りの人からも興味深い研究テーマについて知ることができ、この授業を通して様々なことを学ぶことができました。

図-7 生徒リフレクションの例