

令和5年度

「データサイエンス教育」の開発

【令和5年4月～令和6年3月の取組】



文部科学省スーパーサイエンスハイスクール指定校

兵庫県立姫路西高等学校

目 次

第1部 「データサイエンス教育」の概要

1	カリキュラム開発	2
2	統計的探究プロセス「PPDAC サイクル」	3
3	学校設定科目の授業内容と構成例	4
4	評価の方法	6

第2部 資料編

1	授業教材例	7
2	具体的指導例	19
3	PPDACサイクルに対応したチェックリスト	23
4	研究発表会用ルーブリック	24
5	データサイエンスに関する実技テスト(一例)	25

第3部 生徒研究成果事例

・	Hyogo Reconstruction Project	26
・	独自指標作成による地方創生の方法論と兵庫県活性化の提案	27
・	ため池が周辺地域の暑熱環境に及ぼす効果	35
・	電離層総電子数(TEC)データを用いたオープンソース モデリングアルゴリズムの地震予知への有効性に関する考察	37
・	ジビエの利用量を増やすために	46
・	同一系外惑星系内に存在する惑星の種類とその傾向について	52
・	令和5年度 探究活動 受賞一覧	58

第1部 「データサイエンス教育」の概要

「データサイエンス教育」(以下、DS 教育)は、本校の SSH 事業において、国際理学科・普通科の全生徒が3学年にわたり、学校設定科目として受講する、新時代を見据えた独自の教育活動である。学習指導要領の改訂に伴い、各教科を横断的に学ぶことにより、効果を高める教育の実現を目指しカリキュラム開発を行っている。

SSH I 期目では、教科「数学」におけるデータの分析、統計的な推測、教科「情報」におけるデータの活用、情報とデータサイエンス(以下、DS)の分野を組み合わせた内容を学習するとともに、「総合的な探究の時間」における探究のプロセスを実践していくことで、DS を基盤とした探究活動を行う。

1. カリキュラム開発

本校における3年間の DS 教育は、1年次を「探究準備期間」、2年次を「探究実践期間」、3年次を「探究展開期間」と位置づけ、表1のように単位数を設定した。なお、本校では理数に関する学科である「国際理学科」と、「普通科」が設置されている。

表1 DS 教育に関わる科目名と単位数

国際理学科	科目名	単位数	普通科	科目名	単位数
準備(1年)	データサイエンス研究	4	準備(1年)	データサイエンス探究	2
実践(2年)	データリサーチ研究	3	実践(2年)	データリサーチ探究	2
展開(3年)	グローバル研究	2	展開(3年)	グローバル探究	1

各学科において下記のように科目を代替し、3年間の教育課程を編成した。なお、普通科1・2年は新教育課程である。

(1) 国際理学科1年【探究準備期間】

DS 教育の基礎・基本を学ぶ学校設定科目「データサイエンス研究(4単位)」を実施し、それによって「情報Ⅰ(1単位)」「理数探究(1単位)」「総合的な探究の時間(2単位)」の代替とする。

(2) 国際理学科2年【探究実践期間】

DS を基盤とした研究実践を行う学校設定科目「データリサーチ研究(3単位)」を実施し、それによって「情報Ⅰ(1単位)」「理数探究(1単位)」「総合的な探究の時間(1単位)」の代替とする。

(3) 国際理学科3年【探究展開期間】

DS の学びを振り返るための学校設定科目「グローバル研究(2単位)」を実施し、それによって「コミュニケーション英語Ⅲ(1単位)」「課題研究(1単位)」の代替とする。

(4) 普通科1年【探究準備期間】

DS 教育の基礎・基本を学ぶ学校設定科目「データサイエンス探究(2単位)」を実施し、それによって「情報Ⅰ(1単位)」「総合的な探究の時間(1単位)」の代替とする。

(5) 普通科2年【探究実践期間】

DS を基盤とした研究実践を行う学校設定科目「データリサーチ探究(2単位)」を実施し、それによって「情報Ⅰ(1単位)」「総合的な探究の時間(1単位)」の代替とする。

(6) 普通科3年【探究展開期間】

DS の学びを振り返るための学校設定科目「グローバル探究(1単位)」を実施し、それによって「総合的な探究の時間(1単位)」の代替とする。

2. 統計的探究プロセス PPDAC サイクル

平成 30 年度告示高等学校学習指導要領解説理数編における第 1 節数学 I の「3 (4) データの分析」において、統計的探究プロセスが記載されている。統計的探究プロセスの 5 つの段階からなる「問題 (Problem) - 計画 (Plan) - データ (Data) - 分析 (Analysis) - 結論 (Conclusion)」(以下、PPDAC サイクル)に基づき、本校の DS 教育の根幹を組み立てた。

問題 (Problem)

生徒が自分の興味関心のある研究テーマを決め、課題発見に向けて抽象的な内容から具体的な内容へと細分化・具体化していく。

計画 (Plan)

仮説を立てて、必要なデータを考え、仮説を検証するための分析計画を立てる。

データ (Data)

オープンデータや観察・実験などにより実際にデータを収集する。収集したデータを表計算ソフトや hon などのプログラミング等の知識・技能を活かして、データを整理・整形する。

分析 (Analysis)

整理・整形したデータをもとに解析を行う。データ解析に関する知識や技能を活かしてデータを解析し、結論を導く。

結論 (Conclusion)

データ解析によって得られた結果を、先行研究やフィールドワークなどを通して実態と照らし合わせて考察する。

また、具体的な授業を実践していく上では、5 つの段階を細分化し、各教科の内容を割り振りながら授業展開を作成した。PPDAC サイクルを細分化した内容が次の図 1 である。

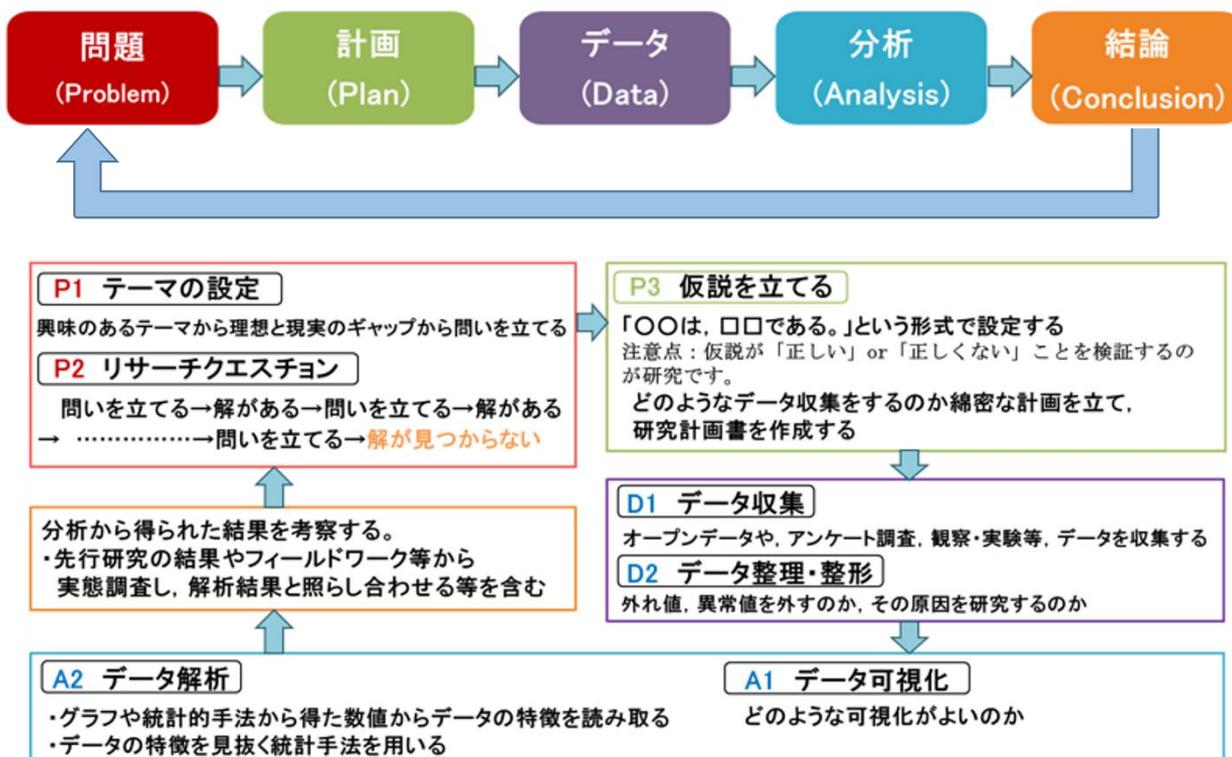


図 1 PPDAC サイクルの詳細

3. 学校設定科目の授業内容と構成例

学校設定科目の授業は、通常講座と集中講座に分けて実施している。通常講座は、時間割内にあるクラス別の授業であり、1・2年では代替している「情報Ⅰ」の内容も扱っている。集中講座は、時間割内ではなく、特定期間を実施している授業であり、探究的な内容を中心に扱っている。研究活動が中心になることから2時間連続で行うことが多く、全クラス同じ時間帯で実施している。

本校には、理数に関する学科である国際理学科があり、普通科よりも各学年1～2単位多く、質・量ともに充実した探究活動を行っているが、以下については、他の学校への普及を図る観点から、普通科における内容を紹介する。

3.1 「データサイエンス探究」の年間計画

(1) 対象生徒 普通科1年生(6クラス, 240名)

(2) 教員担当

通常講座は、各クラス3人で担当する。集中講座は、生徒240名に対して、12名の教員で担当する。1人の教員が4グループ担当し、ファシリテーターとして生徒の探究活動を支援する。

(3) 内容

前期は、企業から提供されたデータをもとに、1班5人グループで、PPDACサイクルに基づく探究活動を進め、夏の研究発表会を経験することで、課題研究に必要な資質・能力を育成する。

後期は、2年からの探究活動の準備段階として、探究プロセスの初期段階(探究テーマの設定、リサーチクエスチョン、先行研究調査、仮説設定)を個人活動として取り組ませる。また、上級生の研究発表会を見学し、質疑応答にも参加することで、研究活動の到達点を意識させる。年間計画を表2に示す。

表2 データサイエンス探究の年間計画

	通常講座	集中講座
4月～9月	<ul style="list-style-type: none"> ・数学Ⅰとの連携「データ分析」(統計量, ヒストグラム, 箱ひげ図) ・DSの初期指導(PPDACサイクルに基づいた問題解決演習の導入) ・情報Ⅰ分野「情報社会の問題解決」(情報の特性, 問題解決の考え方, 法規安全対策, 個人情報の扱い, 知的財産権, 著作権) ・情報Ⅰ分野「問題解決の実践」(データ収集と整理, データ処理, 統計量とデータの尺度) ・データの分布と検定, 時系列分析と回帰分析 ・保健体育科との連携(本校生徒の体力測定データの利用) ・表計算ソフト(データ分析) 	<ul style="list-style-type: none"> ・探究活動のプロセス(PPDACサイクル) Teamsの使い方の技能習得 ・動画を活用したDSの概要理解 ・マンダラートを活用したPPDワークシートの作成, 文献調査 ・文書作成ソフトとプレゼンテーションソフトの活用(ワークシートとスライド作成) ・PBL(問題解決型学習) 企業(株式会社西松屋チェーン)と連携したデータ解析と演習の実施 ・データ分析Ⅰ発表会「西松屋売上向上プラン」 ・データ分析Ⅱ発表会(2年生発表の見学, ルーブリックによる評価, 質疑応答)
10月～3月	<ul style="list-style-type: none"> ・データ分析Ⅰ発表会の振り返り ・情報Ⅰ分野「コミュニケーションと情報デザイン」(メディアの特性, 情報のデザイン化, ピクトグラム, html) ・データ処理の復習と応用 ・情報Ⅰ分野「コンピュータとプログラミング」(アルゴリズムとプログラミング, プログラミングPythonの基本, 配列, 関数, 探索, 整列) ・国語科との連携「データの読み取りと文章表現」 	<ul style="list-style-type: none"> ・PPDACサイクルの確認 ・探究テーマの設定(自己の進路や興味関心と社会や学術分野との関連) ・リサーチクエスチョン(問い立て, 問いの発展, 先行研究や事例による現状把握) ・仮説設定 ・探究計画書の個人制作 ・データ解析Ⅱ発表会(2年生発表の見学, ルーブリックによる評価, 質疑応答) ・SSH成果発表会(2年生発表の見学, 評価, 質疑応答) ・次年度の探究チームの編成

3.2 「データリサーチ探究」の年間計画

- (1) 対象生徒 普通科2年生(6クラス, 240名)
- (2) 教員担当
通常講座は、各クラス3人で担当する。集中講座は、生徒240名に対して、15名の教員で、6クラス同じ時間帯で担当する。1人の教員が4グループ担当し、指導者というより、ファシリテーターとして生徒の探究活動を支援している。
- (3) 内容
4人程度の研究チームを編成し、興味関心に応じて生徒自身が設定した研究テーマについて、データサイエンスを活用した探究活動を行う。研究テーマの内容に応じた専門家の指導を受け、高度で専門的な探究へと深化させる。年2回の校内発表会では、全員が研究成果を発表するほか、希望者は外部発表会やコンテストにも応募する。年間計画を表3に示す。

表3 データリサーチ探究の年間計画

	通常講座	集中講座
4月～9月	<ul style="list-style-type: none"> ・昨年度の振り返り ・PPDACサイクルを基準としたDSの基礎と基本の復習 ・対話的論証を用いた「問題」,「計画」の実践 ・統計手法の理解と習得 (標準化, 重回帰分析, t検定) ・k-means法(Python) ・データ収集と分析 ・数学Bとの連携「統計的な推測」 ・発表会準備 ・研究要綱の作成 ・発表会に向けてのポスター作成 	<ul style="list-style-type: none"> ・マンダラート, PPDワークシート ・対話的論証・三角ロジックについての講義 ・担当教員とチームメンバーの対話を意識した「問題」,「計画」の実践 ・分析方法の復習 経済産業省の動画視聴 ・一次データ(実験等), 二次データ(オープンデータ)の取得 ・「データ収集」,「分析」,「考察」の実践 ・データ分析Ⅱ発表会(中間発表会) 分野別の班に分かれて, チームごとのPPD部分のスライド発表
10月～3月	<ul style="list-style-type: none"> ・統計グラフコンクール応募用ポスター制作 ・情報Ⅰ分野「情報通信ネットワークとデータの活用」 (ネットワーク, データベース) ・研究要綱, 発表会用, 統計グラフコンクール用ポスター制作 ・データ解析Ⅱ発表会の振り返り ・プログラミング(Python) 	<ul style="list-style-type: none"> ・各チームでの探究活動 研究要綱, 発表会用, 統計グラフコンクール用ポスター制作 ・研究要綱の完成 ・データ解析Ⅱ発表会用, 統計グラフコンクール用ポスターの完成 ・希望者による外部コンテストへの出場 ・SSH成果発表会

研究テーマの一例

オフラインでも使える姫路市向け防災アプリの提案 ～複数のプログラミング言語を用いた開発～	日本人にとって使いやすいキーボード配置を見つける
酵母を用いた団粒構造の形成手法について	競技かるたにおける反応速度とその向上について
弁当の食材代用による温室効果ガス削減の提案	緊急地震速報猶予時間の現状とその改善
月面テニスの新たなモデル構築	避難所における生活音漏出の軽減策

3.3 「グローバル探究」の年間計画

- (1) 対象生徒 普通科3年生(6クラス, 240名)
- (2) 教員担当 生徒240名に対して12名の教員で担当する。
- (3) 内容

2年までの探究活動の成果をまとめる。普通科は、全員がポスターを作成し、校外コンクールの一つである全国統計グラフコンクール兵庫県予選に応募する。さらに、全員が英語プレゼンテーション発表を行うことで、国際社会に対応する表現力や発信力を身につける。

グローバル探究発表会後は、研究活動を通して身につけた能力や考え方を整理し、将来、自己の進路にどのように生かせるかを、レポート「私のラーニングプランニング」としてまとめる。後期は探究活動を通して培った英語活用力や情報活用力の振り返りを行い、さらなる力の向上に向けて取り組む。年間計画を表4に示す。

表4 グローバル探究の年間計画

4月～9月	研究活動の振り返り, 研究要綱制作, 英語プレゼンテーション発表に向けた準備 (発表原稿, スライド準備, 発表練習) グローバル探究発表会, データ分析Ⅱ発表会(下級生への指導助言) 「私のラーニングプランニング」作成
10月～3月	研究活動の振り返り, 英語活用力・情報活用力のさらなる向上に向けて

(4) 私のラーニングプランニングからの記述の抜粋

- ・ 物事を多角的に捉えることができるようになった。一年生では主にデータ分析をしたが、目の前にあるデータが示していることは何か、他の研究が示しているデータに穴がないかなどを考えた。
- ・ どうしたら相手に自分たちの研究が魅力的に伝わり面白いと思ってもらえるかを考えた上での発表をすることができ、相手の立場に立って考え伝わりやすいように工夫する力を身につけることができた。
- ・ 発表活動を通して、聴衆に向けて自身の研究をかみ砕き、理解される発表を行う力も向上したと言える。
- ・ 研究班のリーダーになったことで、他者を指導する能力をはぐくむことができた。
- ・ 自分の興味に基づいて学びを深めていくことによって、テストで点数を取るためなどではない、学問の本当の面白さを体感できた。
- ・ それぞれの果たすべき役割は何かを考えその役割をこなして、頼るところは頼って進めていくといったチーム活動の経験をしたことによって協力する大切さを知ることができたため、これからも人との関わり合い、協力といったことを意識しながら物事に取り組もうと思う。
- ・ 何かを学習する際に最も重要なのは意欲である。そして、好奇心はその原動力である。私はその旺盛な好奇心を元に、化学/理系分野にとどまらず幅広い分野を学習したいと考えている。
- ・ 何度も失敗して改善して、と繰り返すのは大変でしたが、過去に調べられた研究はどれも、私たちの何十倍もの失敗を繰り返し、時間をかけられていると思うとそれぞれの研究の見方が変わった気がしました。
- ・ 課題研究を通して、私は客観的に物事を見たり、見聞きしたことに対して本当にそれが正しいのかと批判的に物事を見る力を身につけた。

4. 評価の方法

【知識・技能】

1,2年はDSに係わる実技テストを実施し、3年は提出課題により、知識・技能を評価する。1,2年のテスト内容に係る従来の科目との関係は下表の通りである。

表5 1,2年のテストに含まれる従来の科目内容

従来の科目	1年前半	1年後半	2年前半	2年後半
数学Ⅰ 数学Ⅱ 情報Ⅰ 情報Ⅱ	データの分析①	データの分析② データの活用①	推測統計① 情報とデータサイエンス	データの分析③ 推測統計② データの活用②

【思考力・判断力・表現力】

発表会ごとに教員研修を実施し、ルーブリックにおける基準の相互理解を行い、評価する。

【主体的に学習に取り組む態度】

研究協議(Teamsの討議を含む)による評価を行う。研究グループの生徒同士や、担当教員とのやり取りを踏まえて、主体的に研究に取り組む態度を評価する。

第2部 資料編

1. 授業教材例

(1) 「DS 教育」において価値創造を目指すことを意識させるためのスライド

膨大な情報(ビッグデータ)から、必要なデータを選択し、データが意味を持つように表現(可視化)し、解析結果を融合することで、新たな価値を生み出す過程の獲得、結論の創造を目指す。なお、自然科学分野では、実験・観察によって得られたデータを扱う。

Hyogo Prefecture Himeji Nishi SHS

↑

データサイエンスで目指すこと

現在は「情報化社会」と言われるが、データサイエンスの世界では、**創造性**
「扱えるデータが容易に得られる社会」と捉える。

データサイエンスは、ビッグデータから**有益な情報をあぶり出す**ことである。
多変量のビッグデータからあぶり出す解析を**多変量解析**という。

データ解析で重要なことは？

ヒストグラム・箱ひげ図
散布図をかきまくる
↓
相関をみる
↓
原因となる相関を見抜く

格言1
データサイエンスは、「宝探し」
データとデータを繋ぎ合わせ、筋道をつくる！
異分野のデータや知恵を組み合わせる！

格言2
相関関係と因果関係は異なる
相関関係から因果関係を見抜く！

(2) 「DS」に必要な学問概念をイメージしたスライド

DS は、文系・理系という概念を超えた分野であることを示す。

Hyogo Prefecture Himeji Nishi SHS

↑

データサイエンスとは

データの集積のみでは価値は生まれない
客観的な存在としてのビッグデータを対象
→ **新しい知見の抽出、価値の発見・創造**

統計学と情報学のドッキング
+ 人文知の役割

理系的
収集・加工・処理
データ
エンジニアリング
(コンピューター科学)

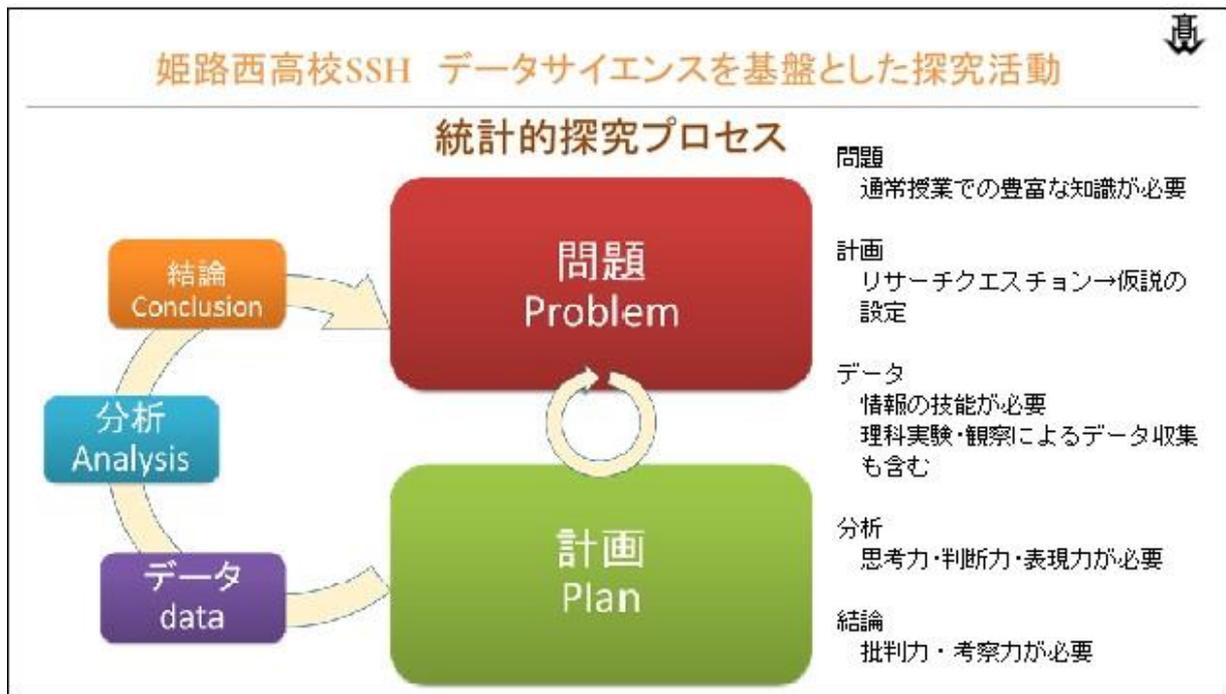
理系的
データ
分析・解析
データ
アナリシス
(統計学)

文系的
諸分野の
知識と理解
人文・社会知
価値創造

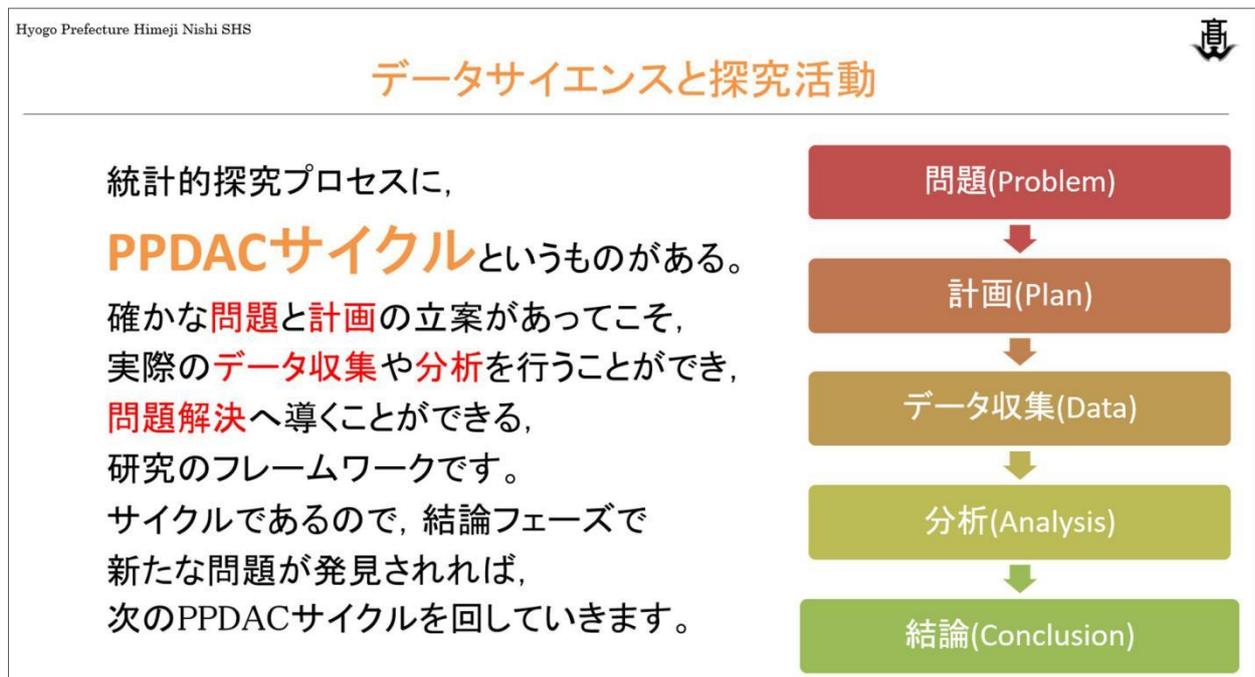
データサイエンティスト

滋賀大学位田隆一学長による講演会資料より抜粋 (2020年7月)

- (3) PPDAC サイクルと各教科の関連を示したスライド
学校の教育活動における各教科の役割を整理している。



- (4) PPDAC サイクルを説明したスライド
統計的探究プロセス，PPDAC サイクルという研究のフレームワークを説明する。



(5) 【問題】に関する授業スライド1

各生徒の興味関心のある内容でテーマを設定させていく。その際に、テーマに対して「理想」を述べさせ、「現実」を調査させると、ギャップが出てくる。そのギャップから問題→課題を見つけるきっかけが生まれることを説明する。

Hyogo Prefecture Himeji Nishi SHS

問題(P1 テーマの設定) 真

研究テーマ→問いを立てる→課題を発見する

- ① 研究テーマを決めよう。【抽象的】
- ② ①に対して、どんな理想をもっているか考えよう。
- ③ ①の「現実」について考えよう。(調査しよう。)**【具体的】**
- ④ ②と③のギャップから問題を見つけよう。【抽象的】
- ⑤ ④の問題から自分が解決したい課題を見つけよう。**【具体的】**

興味・関心があることからテーマを決めよう

(6) 【問題】に関する授業スライド2

研究テーマが抽象的な場合、スライドのように言葉の意味や定義が明確でない場合が多い。そのため、スライドのように言葉の定義を意識させる。

問題(P1 テーマの設定) 真

研究テーマ→問いを立てる→課題を発見する

参考：課題研究メソッド p.31

1. 言葉の意味や定義を問う「問い」
 - 「ニート」って何？
 - きちんと調べる(辞書や国・地方自治体のホームページをみる)
- ・意味や定義が不明確な場合は、自分自身で仮に定義する
 - 研究発表の際の表現を確認しよう。
- ・似ている言葉の違いを比較する。
 - 幼稚園と保育園の違いは何？

資料引用元: 啓林館 岡本尚也(2021)「課題研究メソッド よりよい探究活動のために 2nd Edition」

(7) 【問題】に関する授業スライド3

指導の参考事例として、スライドのように抽象的な内容を具体的な内容へ換言していくと、研究内容が深まる。そうしたマジックワードを意識化させることで、生徒と教員が共有した概念を持つことができる。

Hyogo Prefecture Himeji Nishi SHS

問題(P1 テーマの設定)

研究テーマ→問いを立てる→課題を発見する

参考：課題研究メソッド p.32

マジックワード(抽象的な言葉)を使用していませんか？

安全な社会

言い換えると、
・犯罪のない社会
・戦争のない社会
・災害による被害が少ない社会

災害による被害が少ない社会

言い換えると、
・台風による被害
・大雨による被害
・竜巻による被害

大雨による被害が少ない社会

言い換えると、
・洪水や冠水による
・雨漏りによる浸水被害
・土砂崩れによる被害

大雨による土砂崩れによる被害が少ない社会

マジックワードは、言葉の言い換えを行い、具体化しよう。

資料引用元：啓林館 岡本尚也(2021)「課題研究メソッドよりよい探究活動のために 2nd Edition」

(8) 【問題】に関する授業スライド4

仮説の立て方として、スライドのように「解が見つからない問い」が仮説候補となる。スライドのように問いを深めていくことで、仮説へとつながることを説明する。

Hyogo Prefecture Himeji Nishi SHS

問題(P2 リサーチクエスト)

研究テーマ→問いを立てる→課題を発見する

参考：課題研究メソッド p.44

○ 研究テーマに対して「問い」を立てよう

問いを立てる→解がある→問いを立てる→解がある
→ ……………→問いを立てる→解が見つからない

1. 言葉の意味や定義を問う「問い」 → ○○の意味は？ ○○の定義は？
2. 原因(なぜ)を問う「問い」 → なぜ○○は生じているのか？
3. 信憑性を問う「問い」 → ○○は本当に生じているのか？
4. 比較を問う「問い」 → 他の国・地域ではどのように変化している？
5. 先行研究・先行事例を問う「問い」 → ○○に対して、どのような研究が行われた？
6. 影響を問う「問い」 → ○○によって、どのようなことが起こるのか？
7. 方法や関連性を問う「問い」 → ○○と△△にはどのような関連があるのか？

資料引用元：啓林館 岡本尚也(2021)「課題研究メソッドよりよい探究活動のために 2nd Edition」

- (9) 【計画】に関する授業スライド
 先行研究を調査し、仮説を立てることを示す。

Hyogo Prefecture Himeji Nishi SHS

計画(P3 仮説を立てる) 真

仮説を立てる

○ 仮説

仮説とは、リサーチクエスチョンに対する**予想される仮の答え**である。
 「〇〇は□□である」という形で書こう！

先行研究は必ず調査しよう！
 先行研究に基づいて研究していることを発表で示すこと

1. 先行研究・事例から根拠をもって仮説を示す
2. 複数の仮説を立てる
3. 仮説を検証する方法を考える

参考：課題研究メソッド p.66～70

資料引用元：啓林館 岡本尚也(2021)「課題研究メソッドよりよい探究活動のために 2nd Edition」

- (10) 【データ】に関する授業スライド1 データ収集の方法を示す。

Hyogo Prefecture Himeji Nishi SHS

データ(D1収集) 真

検証のためのデータ収集方法

○ データ収集のポイント

- ・インターネットにあるデータをそのまま使う → ×
- ・インターネットにあるデータの出典をみる → ○
 ⇒ その出典元からデータを取り、自分で**データ整理整形**しよう！
- ・**先行研究**を参考に、データ収集元を探す → ◎
- ・自分でデータを採取する → ◎

- (11) 【データ】に関する授業スライド2
 オープンデータを収集するための集積サイトを活用させる。

Hyogo Prefecture Himeji Nishi SHS

データ(D1収集)

オープンデータのデータ収集方法



JDSSP高等学校データサイエンス教育研究会

高校生のための”DS教育コンソーシアム”
 高校生にデータサイエンス教育を実践する教員のための研究会です。

ホーム 本研究会とは 活動状況 **データサイエンス教材** DS関連サイト お問い合わせ

<https://ds-education.com/>

オープンに利用できる
統計データのサイトが
集まっているページ

DS教材の記事一覧

- 「データ活用」に係わる授業モデル・教材 (2021年2月16日)
- データサイエンスの評価について (2020年10月23日)
- データ収集一覧** (2020年10月23日)

データ収集一覧

2020年10月23日

- DATA GO.JP (日本政府全体のデータ一覧サイト)
- e-Stat (政府統計の総合窓口)
- 統計GIS (地図で見る統計)
- RESAS (リーサス) 地域経済分析システム (データがすぐに可視化できるサイト)
- SSDSE (教育用標準データセット) 統計データ分析
- miripo (マイクロデータ利用ポータルサイト)
- 統計ダッシュボード (総務省統計局)
- なるほど統計学園
- 国土交通省 国土地理院
- 国、都道府県、政令指定都市「統計年鑑(統計書、県勢要覧、統計年報)まとめサイト
- 各府省及び独立行政法人等のページ (国の省庁ごとのデータリンクサイト)
- 国立社会保障・人口問題研究所「将来推計人口データベース」

- (12) 【データ】に関する授業スライド3
 容易にデータが可視化される RESAS(リーサス)というサイトの活用方法を体得させる。

Hyogo Prefecture Himeji Nishi SHS

データ(D1収集)

オープンデータのデータ収集方法

RESAS地域経済分析システムの利用



高校生のためのデータサイエンス入門

滋賀大学

第1週第4回
地域経済分析システム
(RESAS)の利用

DS教材

データ収集一覧

2020年10月23日

- DATA GO.JP (日本政府全体のデータ一覧サイト)
- e-Stat (政府統計の総合窓口)
- 統計GIS (地図で見る統計)
- RESAS (リーサス) 地域経済分析システム** (データがすぐに可視化できるサイト)
- SSDSE (教育用標準データセット) 統計データ分析
- miripo (マイクロデータ利用ポータルサイト)
- 統計ダッシュボード (総務省統計局)
- なるほど統計学園
- 国土交通省 国土地理院
- 国、都道府県、政令指定都市「統計年鑑(統計書、県勢要覧、統計年報)まとめサイト
- 各府省及び独立行政法人等のページ (国の省庁ごとのデータリンクサイト)
- 国立社会保障・人口問題研究所「将来推計人口データベース」

簡単にデータが可視化できます。
この動画をみると使い方がわかります。

本校は滋賀大学と連携協定を締結しているため、動画を閲覧できます。

資料引用元: 滋賀大学 高校生のためのデータサイエンス入門

(13) 【データ】に関する授業スライド4

エクセルファイル・csv ファイルでのオープンデータの取得が可能になる、政府統計の総合窓口(e-stat)というサイトの活用方法を体得させる。

Hyogo Prefecture Himeji Nishi SHS

データ(D1収集)

オープンデータのデータ収集方法

政府統計の総合窓口(e-stat)の利用

滋賀大学

第1週第5回
政府統計の総合窓口
(e-Stat)の利用

高校生のためのデータサイエンス入門

データ収集一覧

- DATA GO.JP (日本政府全体のデータ一覧サイト)
- e-Stat (政府統計の総合窓口)
- 統計GIS (地図で見る統計)
- RESAS (リーサス) 地域経済分析システム (データがすぐに可視化できるサイト)
- SSDSE (教育用標準データセット) 統計データ分析
- miripio (マイクロデータ利用ポータルサイト)
- 統計ダッシュボード (総務省統計局)
- なるほど統計学園
- 国土交通省 国土地理院
- 国、都道府県、政令指定都市「統計年鑑(統計書、県勢要覧、統計年報)」まとめサイト
- 各府省及び独立行政法人等のページ (国の省庁ごとのデータリンクサイト)
- 国立社会保障・人口問題研究所「将来推計人口データベース」

オープンデータがまとまっています。
この動画を見ると使い方がわかります。

本校は滋賀大学と連携協定を締結しているため、動画を閲覧できます。

資料引用元: 滋賀大学 高校生のためのデータサイエンス入門

(14) 【データ】に関する授業スライド5

地図で見る統計(統計 GIS)というサイトによって、地図上に統計データを可視化するソフトの活用方法を体得させる。

Hyogo Prefecture Himeji Nishi SHS

データ(D1収集)

オープンデータのデータ収集方法

地図で見る統計(統計GIS) 【jSTAT MAPの利用】

滋賀大学

地図で見る統計(統計GIS)

各種統計データを地図上に表示し、視覚的に統計を把握できる地理情報システム(GIS)を提供しています。

「お知らせ」

- 2021年4月21日 地図で見る統計(jSTAT MAP)のログイン画面を変更いたしました。
- 2021年3月19日 2018年漁業センサス 都道府県及び市町村の提供を開始いたしました。

> 地図で見る統計(jSTAT MAP)

地図で見る統計(jSTAT MAP)は、誰でも使える地理情報システムです。統計地図を作成する他に、利用者のニーズに沿った地域分析が可能となるようなさまざまな機能を提供しています。防災、施設整備、市場分析等、各種の詳細な計画立案に資する基本的な分析が簡単にできます。

※システムの動作が著しく遅い場合は、システムが混み合っている可能性があります。時間を置いて再度アクセスをお願いいたします。

また、地図で見る統計(jSTAT MAP)起動時にエラーとなる場合は、ブラウザの閲覧履歴の削除を行い再度お試しください。

地図に
様々な統計データを
反映できる

(15) 【データ】に関する授業スライド6

データの特徴を読み取るために、エクセルを活用してデータを整理整形する。

Hyogo Prefecture Himeji Nishi SHS

データ(D2整理整形)

整理整形に活用できるエクセルの関数

統計量とは ... 代表値(最小値・第1四分位数・第2四分位数(中央値)・第3四分位数・最大値)
最頻値, 平均値, 分散, 標準偏差など

求めたい値	エクセルの関数
合計	=SUM(○ : ○)
データ数	=COUNT(○ : ○)
平均	=AVERAGE(○ : ○)
分散	=VARP(○ : ○)
標準偏差	=STDEV(○ : ○)
最小値	=QUARTILE(○:○, 0)
第1四分位数	=QUARTILE(○:○, 1)
第2四分位数(中央値)	=QUARTILE(○:○, 2)
第3四分位数	=QUARTILE(○:○, 3)
最大値	=QUARTILE(○:○, 4)

データ	10
	20
	30
	40
	50
	60
	70
	80
合計	360
データ数	8
平均	45
分散	525
標準偏差	22.91
最小値	10
第1四分位数	27.5
第2四分位数(中央値)	45
第3四分位数	62.5
最大値	80

(16) 【データ】に関する授業スライド7

外れ値や異常値に関するデータの取り扱い方を理解させる。

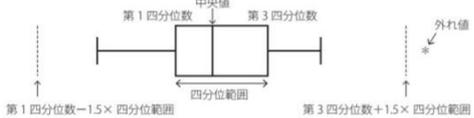
Hyogo Prefecture Himeji Nishi SHS

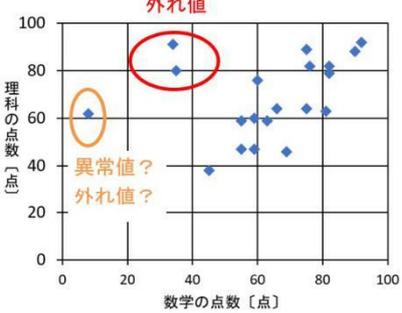
データ(D2整理整形)

データの外れ値・異常値

外れ値 ... 量的データにおいて、他の値から極端にかけ離れたデータのこと
異常値 ... 測定ミスや入力ミスなどによる値

外れ値の判断
 まず、明確な外れ値の定義はない。ただ、1つの目安として、下記を示しておく。
 ・通常四分位範囲の1.5倍以上離れた値(「数学」学習指導要領より)
 ・平均値より±2σ離れた値





「外れ値」のポイント

- ① 外れ値を除いて、研究を進めるのも1つの流れです。
 しかし、研究の流れとしては、一度外れ値を含めた散布図を示し、外れ値として判断した基準を示す。そして、外れ値を除いた散布図を示すことで、きちんと外れ値を判断したことを伝えることが重要です。
- ② 外れ値の原因を研究するのも、1つの研究です。

(17) 【分析】に関する授業スライド1

基本的なデータの可視化について、伝えたい内容に適した方法を理解させる。

Hyogo Prefecture Himeji Nishi SHS

分析(A1 データ可視化)

収集したデータの可視化

棒グラフ

横軸に対する縦軸の値の大きさの変化や違いを明確に示すことができる

円グラフ

全体に対する各項目の割合を明確に示すことができる

項目	割合
野球部	23.0%
サッカー部	21.2%
テニス部	19.3%
ソフトテニス部	12.1%
バドミントン部	8.2%
水泳部	8.1%
バレー部	4.2%
バスケットボール部	3.1%
その他	1.8%

折れ線グラフ

比較対象が多く、時系列など横軸に対する変化をみる際に適している

帯グラフ

割合を示し、複数の項目やその変化を比較できる

学年	10冊以上	6~9冊	3~5冊	1~2冊	ほとんど読まない	不明
高1	15.1	23.4	24.6	22.2	11.2	3.5
高2	12.7	23.1	57.4	6.6	0	0
高3	11.1	27.4	12.1	41.4	0	0

資料引用元: 啓林館 岡本尚也(2021)「課題研究メソッドよりよい探究活動のために 2nd Edition」

(18) 【分析】に関する授業スライド2

解析する際に、単体のデータで判断せずに複数のデータを比較する手法を理解させる。

Hyogo Prefecture Himeji Nishi SHS

分析(A2 データ解析)

グラフの特徴的な部分を読み取る

事例

図1のグラフでは、特徴的な部分はあるが、本当に特徴的な部分であるのか根拠が少ない

特徴的ではあるが、1つのグラフだけで判断はしづらい

図1 2014年夏物服の販売数

毎年同じような傾向がみられる複数のデータから判断する

図2 2014年~2018年の夏物服の販売数

データ解析ポイント

ポイント① 図2のように複数年を重ねると、特徴的であることが明確になる。

ポイント② さらに、2014年は、他の年度より、非常に売り上げが多いのはなぜ？と研究が深まっていく

図2では、他に特徴的な部分がある。気づきますか？

(19) 【分析】に関する授業スライド3

本校で主に指導する統計手法である。ただし、研究でこれらの統計手法をすべて活用しなければならぬわけではない。研究内容に即した手法を選択する力の必要性も理解させる。

Hyogo Prefecture Himeji Nishi SHS

分析(A2 データ解析) 

姫路西高校で学ぶ「統計手法」

回帰分析法(単回帰分析・重回帰分析)

- ・ 未知な事柄を予測するための手法
キーワード: 回帰式・回帰係数・決定係数・寄与率・p値・予測モデル

標準化

- ・ データ群を, 平均0, 標準偏差1にする手法
2つ以上を比較するとき利用する。

仮説検定(平均の差の検定)

- ・ データ群とデータ群を比較し, 平均の差が有意であるかないか判断する手法

クラスタリング(k-means法など)

- ・ 統計的にデータをグループ分けする手法

主成分分析法

- ・ 変数間の関係性から共通する因子(潜在変数)を導き出し,
多くのデータを少ない変数に縮約する分析手法

(20) 【分析】に関する授業スライド4

プログラミングを学ぶ際に、プログラミングを手段として活用することの有効性を伝える。

分析(A2 データ解析) 

プログラミング言語 Python を学ぶ

データの整理・整形, 可視化をエクセルで行うことは非常に手間である
⇒ **プログラミング**を活用すれば, **楽ができる!**

Pythonでプログラミングができる!  **楽ができる!**

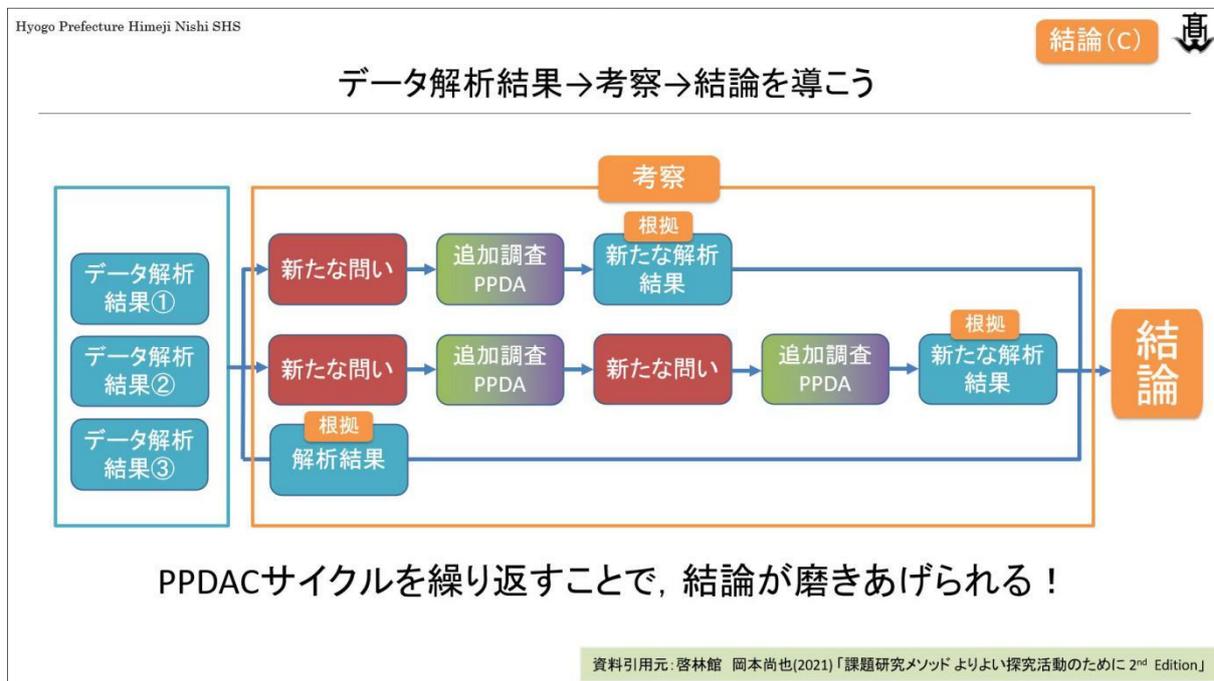
大切なことは,
Pythonを使って何をしたいのか  **Pythonは単なる手段!**

データサイエンスでは、

- ・ データ収集のために, Pythonを使う
- ・ データ研磨のために, Pythonを使う
- ・ Pythonの出力結果から, データ解析する

(21) 【結論】に関する授業スライド1

PPDAC サイクルを何周もまわすことで研究が進化することを理解させる。



(22) 【結論】に関する授業スライド2

新しい価値を生む創造性に関するメカニズムを理解させる。

Hyogo Prefecture Himeji Nishi SHS

結論(C) 眞

【創造力】新たな価値を創造しよう

研究

- ・みんなが当たり前と思っていることをデータで検証する
- ・他の人が思いつかない結論 を導く！
- ・予想外の結論 を導く！

新しい価値を生む創造性には

- ・ある分野の内容を、異なる分野に活用して、
新しいものを創り出すこと ⇒ **リフレーム**
- ・異質なものを結び付けて新たな軸をつくること ⇒ **統合**

解析結果に基づいた解釈していく上で、
リフレームと統合を意識し、新たな価値を創造しましょう！

(23) 【1年生連携企業によるデータ解析 PBL】説明スライド1

1年生前半の学びを実社会のデータの解析に生かす演習。本校と連携しているグローリー(株)の技術を活用した(株)西松屋の顧客データを提携してもらい、そのデータを分析し、(株)西松屋チェーンへ提案する演習である。

子ども服販売の社員になろう！

演習の主旨

- 背景： あなたは、子ども服を販売するお店の店舗で子ども服を販売している社員です。
- ミッション： 社員として働いている店舗の売り上げ高を向上させるよう社長から指令がありました。
- ゴール： **与えられたデータ**に基づいて、**店舗の売上を向上させる施策**を提案してください。

提案方法

与えられたデータに基づき、様々な視点から1つ提案してください。

来店者の地域情報・性別・来客人数・来客頻度などのデータを元に、ターゲットを絞った客層に応じた提案でもよいです。

また、客層に応じたSNSの利用による宣伝方法の提案など、幅広い視点での提案をお願いします。

ただし、どのジャンルにおいても、客観的なデータ分析に基づいたプランにしてください。

(スライド提供：AdInte)

(24) 【1年生連携企業によるデータ解析 PBL】説明スライド2

PPDAC サイクルを活かし、具体的な施策を(株)西松屋チェーンに提案する。

問題解決に向けたデータ解析実践 → スライド発表

データ分析の視点をレクチャー

子ども服販売の社員になろう！

データ分析のプロセス

データ分析のプロセスには、大きく分けて以下の二つがあります。

1. データ分析をして課題を探し施策を検討する（帰納的データ活用）

例えば、「休日に比べ平日の来店数が少なかった。子ども服売上向上のためには、休日来店者のクチコミを広げて平日来店につなげていけばどうだろうか？」など

2. 仮説を立て「仮説の正誤」をデータ分析で明らかにし施策の実施判断をする（演繹的データ活用）

例えば、「子ども服の売上向上にはファミリーの来店を促す必要があるのでは？データ分析の結果、休日に比べ平日の来店数が少なく、「休日来店ファミリーのクチコミを広げて平日来店に繋げたい」など

実際のデータサイエンス・統計の仕事では、1,2.を行き来して提案の説得性と予測性・確度を上げていきます。はじめは、2.から始めるのが発想豊かになりやすいです。

PPDACサイクルの繰り返し

問題(Problem)
... 課題の細分化

計画(Plan)
... 仮説の設定

データ(Data)
... 与えられたデータ +
オープンデータの利用

分析(Analysis)
... 可視化, 解析

結論(Conclusion)
... 解決策の提示

(スライド提供：AdInte)

2. 具体的指導例

《重回帰分析の指導(2時間)》

(1) 重回帰分析法における指導用スライド1

統計手法「重回帰分析法」に関する内容である。単回帰分析と重回帰分析の違いを理解させる。

Hyogo Prefecture Himeji Nishi SHS DR探究・研究

目的変数に対して、説明変数1つだけで十分なの？

質問：ハンドボール投げの結果を予測したい。
ハンドボール投げを予測するための説明変数を見つけよう。

握力 → ハンドボール投げ

「目的変数」に対して、
「1つ」だけの「説明変数」との関係
⇒「**単回帰分析**」という

50m走
握力
上体起こし → ハンドボール投げ

「目的変数」に対して、
「複数」の「説明変数」との関係
⇒「**重回帰分析**」という

(2) 重回帰分析法における指導用スライド2

エクセルの「データの分析」(アドインで設定)を活用して重回帰分析を実行する。その出力結果を読み取り、予測モデルを構築する。回帰係数の有意差の検定を行い、変数減少法を用いて最適なモデルを探索する。

Hyogo Prefecture Himeji Nishi SHS DR探究・研究

MS-Excelを利用して、重回帰式を作成する。

「分析ツール」→「回帰分析」からデータを選択する。

回帰統計	
重相関 R	0.85
重決定 R ²	0.72
補正 R ²	0.72
標準誤差	3.73
観測数	275

「重相関R」→「重相関係数 R」
回帰モデルによる期待値と実測値との相関係数である。

「重決定R²」→「寄与率(決定係数) R²」
0から1の間の値を取り、1に近いほど、当てはまりがよいと判断する

「補正R²」→「自由度調整済み寄与率R²(決定係数)」
複数の回帰式を比較するときに使用する

分散分析表					p値				
	自由度	変動	分散	観測された分散比	有意 F				
回帰	6	9771.33	1628.56	117.28	0.000				
残差	268	3721.40	13.89						
合計	274	13492.73							

変数の係数が0のとき、つまり、この係数がないときの影響はどれほどあるのか判断できる (0.05が基準となることが多い)

	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 99.0%	上限 99.0%
切片	1.44	7.60	0.19	0.850	-13.51	16.40	-18.27	21.15
握力	0.32	0.05	6.78	0.000	0.23	0.41	0.20	0.44
上体起こし	0.21	0.06	3.54	0.000	0.09	0.32	0.06	0.36
反復横とび	0.09	0.03	2.72	0.007	0.03	0.16	0.00	0.18
持久走	-0.97	0.48	-2.00	0.047	-1.92	-0.01	-2.22	0.29
50M走	-1.59	0.54	-2.97	0.003	-2.65	-0.54	-2.99	-0.20
立ち幅跳び	0.04	0.01	2.94	0.004	0.01	0.07	0.01	0.08

係数 正か負であるか符号を確認すること

(3) 演習

授業の目的：過去のデータから、予測モデルを構築する。「重回帰分析」を用いて、まだ測定していないハンドボール投げの予測値を求めよう。

(手順①) 保健体育で測定している過去の生徒の体力測定の結果データをもとにして考える。

部活動	男女	握力	上体起こし	反復横とび	持久走	50M走	立ち幅跳び	ハンドボール投げ
テニス	2	28	26	60	287	8.5	193	17
サッカー	1	32	34	62	343	7.7	223	25
バドミントン	1	34	28	55	378	7.8	230	18
陸上競技	2	28	27	49	262	7.9	191	9
卓球	1	41	27	63	383	7.7	226	21
囲碁将棋	1	47	28	59	391	7.4	224	19
サッカー	1	42	39	64	353	6.7	270	29
E・S・S	2	26	22	42	327	10.4	157	8
音楽 新聞	1	41	36	63	333	7.4	240	28
囲碁将棋	1	34	20	50	412	7.6	233	26
バドミントン	1	25	34	58	332	7.8	180	13

(手順②) 「データ」→「データの分析」を選択する。

「データの分析」が表示されない場合は、オプションのアドインで設定する。



(手順③) 回帰分析を選択する。

(手順④) 「入力Y範囲」に目的変数、「入力X範囲」に説明変数の行データ(複数行が可能)を選択する。ここで、エラーが出る場合はデータが欠損しているなど、データを確認する。



表5 活用するデータ事例

握力	上体起こし	反復横とび	持久走	50M走	立ち幅跳び	ハンドボール投げ
28	26	60	287	8.5	193	17
32	34	62	343	7.7	223	25
34	28	55	378	7.8	230	18
28	27	49	262	7.9	191	9
41	27	63	383	7.7	226	21
47	28	59	391	7.4	224	19
42	39	64	353	6.7	270	29
26	22	42	327	10.4	157	8

説明変数

目的変数

(手順⑤) 分析ツールで実行した数値を読み取り、予測モデルを作る。

表6 重回帰分析の出力結果

	係数	標準誤差	t	P-値
切片	-1.534	9.807	-0.156	0.876
握力	0.345	0.058	5.925	0.000
上体起こし	0.247	0.079	3.128	0.002
反復横とび	0.146	0.048	3.047	0.003
持久走	-1.032	0.520	-1.987	0.049
50M走	-1.594	0.704	-2.264	0.025
立ち幅跳び	0.034	0.019	1.763	0.080

表6から読み取った予測モデル例

ハンドボール投げ(m)の予測値

$$= 0.345(\text{m/kg}) \times \text{握力}(\text{kg}) + 0.247(\text{m/回}) \times \text{上体起こし}(\text{回}) + 0.146(\text{m/回}) \times \text{反復横跳び}(\text{回}) - 1.032(\text{m/秒}) \times \text{持久走}(\text{秒}) - 1.594(\text{m/秒}) \times \text{50m走}(\text{秒}) + 0.034(\text{m/cm}) \times \text{立ち幅跳び}(\text{cm}) - 1.534$$

握力・上体起こし・反復横跳び・持久走・50M走・立ち幅跳びの数値を代入し、ハンドボール投げの予測値を求める。

(手順⑥) 予測モデルの寄与率を読み取る。

表計算ソフトの一例として Excel で示したが、Excel の出力結果は必要に応じて、レポート等では書き直す必要がある。

「重相関 R」→「重相関係数 R」

「重決定 R²」→「寄与率(決定係数)R²」

「補正 R²」→「自由度調整済み R²」

この出力から、モデルの適合度を示す寄与率が 67.6%であることがわかる。

回帰統計	
重相関 R	0.822
重決定 R ²	0.676
補正 R ²	0.665
標準誤差	3.987
観測数	180

(手順⑦) 最適な予測モデルを構築するために数値を読み取る。

複数の説明変数の中には、目的変数の予測に役立たないものが含まれている可能性がある。そこで、モデル式の構築にあたっては、説明変数の取舍選択が重要な課題となる。これを変数選択もしくはモデル選択(モデリング)という。その際、寄与率ができるだけ大きくなることが望ましいが、説明変数を増やせば増やすほど単純に大きくなる。しかし、欠点(過剰適合、過学習)も生じる。ここでは、変数減少法を説明する。

回帰係数の有意差の検定により、回帰係数の有意確率(P-値)の列から、1%もしくは5%以下であるか判断する。表6における青色枠の例では、「握力」「上体起こし」「反復横跳び」は1%有意、「持久走」「50M走」は5%有意であるが、「立ち幅跳び」に有意差はないと判断する。

(手順⑧) 最適な予測モデルを構築する。

有意差が出ない変数は「ハンドボール投げ」の値の変化に影響を与えていない可能性が高いと判断される。

「持久走」を含む結果 「持久走」を含まない結果 「持久走」を含まない回帰係数・p値

回帰統計		回帰統計		係数	標準誤差	t	P-値	
重相関 R	0.822	重相関 R	0.819	切片	9.121	7.769	1.174	0.242
重決定 R2	0.676	重決定 R2	0.671	握力	0.371	0.057	6.530	0.000
補正 R2	0.665	補正 R2	0.661	上体起こし	0.255	0.079	3.217	0.002
標準誤差	3.987	標準誤差	4.011	反復横とび	0.166	0.047	3.555	0.000
観測数	180	観測数	180	持久走	-1.061	0.522	-2.030	0.044
				50M走	-2.314	0.577	-4.009	0.000

【手順⑤に比べ、変数が少ない予測モデル】

ハンドボール投げの予測値(m)

$$= 0.371(\text{m/kg}) \times \text{握力}(\text{kg}) + 0.255(\text{m/回}) \times \text{上体起こし}(\text{回}) + 0.166(\text{m/回}) \times \text{反復横跳び}(\text{回}) - 1.061(\text{m/秒}) \times \text{持久走}(\text{秒}) - 2.314(\text{m/秒}) \times \text{50M走}(\text{秒}) + 9.121$$

この予測モデルに、説明変数(握力・上体起こし・反復横跳び・持久走・50m走)の値を代入した値が、ハンドボール投げの予測値である。そして、予測値と実測値(実際にハンドボール投げをして測定した値)を比較する。

参考資料：文部科学省高等学校情報科「情報Ⅱ」教員研修用教材(本編)

第3章情報とデータサイエンス

3. PPDAC サイクルに対応したチェックリスト

PPDAC サイクルに基づく探究活動を進めるにあたり、PPDAC の各段階において、教員と生徒が共通認識をもって、そのつど活動内容をフィードバックしながら、研究を深めていけるように、チェックリストを作成した。

<input type="checkbox"/> 1 研究テーマやリサーチクエスチョンに用いられている言葉の意味や定義を理解しているか？ <input type="checkbox"/> 2 マジックワードが出てきた際、具体的な文言に言い換えているか？ <input type="checkbox"/> 3 研究テーマで取り上げた課題について、その現状が分かるデータを収集しているか？ <input type="checkbox"/> 4 研究テーマに関する先行研究・事例を理解しているか？ (複数の先行研究を見つけているか？先行研究は信頼できるものか？) <input type="checkbox"/> 5 研究テーマに関する原理や原則、数式を理解しているか？(主に自然科学実践) <input type="checkbox"/> 6 研究テーマと学術分野(文学、経済学、工学など)との関連を理解しているか？ <input type="checkbox"/> 7 (常識的に考えてみれば)答えがすぐに見つかるリサーチクエスチョンになっていないか？	問題 (Problem)
<input type="checkbox"/> 8 研究テーマを細分化し、取り組むことが可能なりサーチクエスチョンになっているか？ <input type="checkbox"/> 9 研究を進めるために何が必要か？また、それらは用意することが可能か？ <input type="checkbox"/> 10 先行研究などを参考にして、仮説の根拠が示しているか？ ※ 「主観」が入っていないか？ <input type="checkbox"/> 11 データの収集が可能な仮説になっているか？ <input type="checkbox"/> 12 グループ内でTeams等を利用して研究の進捗を確認し、定期的に担当教員と情報共有しているか？	計画 (Plan)
<input type="checkbox"/> 13 (統計技法を使用する前に) 代表値・度数分布表・ヒストグラム・箱ひげ図・散布図を用いて、そのデータの特徴を把握しているか？ <input type="checkbox"/> 14 結果を分かりやすい形(グラフ・表・図・文章など)にまとめているか？ <input type="checkbox"/> 15 グラフのタイトル、縦軸・横軸の名前がついているか？ <input type="checkbox"/> 16 軸の単位が書かれているか？	データ (Data)
<input type="checkbox"/> 17 結果に対して立てた問いの答えを見出し、考察ができているか？ ※ 「相関係数」を示した後に、その結果を考察しているか？ <input type="checkbox"/> 18 どこまでが客観的なデータを基にした事実で、どこまでが意見であるのかがはっきりしているか？	分析 (Analysis)
<input type="checkbox"/> 19 結果・考察から論理的に結論(仮説の答え)を導いているか？ ※ 先行研究等を調べずに、結果に対して主観で結論を述べていないか？ <input type="checkbox"/> 20 結果・考察から今後の展望について考えているか？	結論 (Conclusion)

(課題研究に関する用語)	
<ul style="list-style-type: none"> ▪ マジックワード 	聞こえはよいが、抽象的で何を意味するか分からない言葉
<ul style="list-style-type: none"> ▪ リサーチクエスチョン 	課題研究全体で明らかにしたい問い(生徒の問いの多くは先行研究等で明らかになっていることが多い)

4. 研究発表会用ルーブリック

評価点	1	2	3	4	5
観点		標準レベルを達成できたといえない	標準的なレベルを概ね達成できた	標準レベルを越えて達成できた	
課題 発見力	データを集めているが、仮説を立てるために必要なデータを集めることができず、主観的な読み取りをしている。1つ1つのデータにつながりがなく適切に整理がなされていない。データと課題、仮説が対応していない。	必要なデータを集めデータから課題発見し、仮説を立てている。しかし、主観的な読み取りをしている。1つ1つのデータにつながりがなく適切に整理がなされていない。データと課題、仮説が対応していない。	必要なデータを集めデータから課題発見し、仮説を立てている。また、数値を求め、グラフの特徴をきちんと示す等データを客観的に読み取っている。整理したデータをもとに課題を発見し、仮説を設定している。	信頼できることから、必要なデータを(必要に応じて自分たちでも)集めており、仮説に合う適切なデータを抽出している。データを客観的に読み取っており整理したデータをもとに、独自性のある課題を発見し、仮説を設定している。	信頼できることから、必要なデータを(必要に応じて自分たちでも)集めるだけでなく自分たちで数式を作り、データを作成し、仮説に合う適切なデータを抽出するなどの工夫が見られる。データを客観的に読み取り、高度な手法を用いて分析し分かりやすく整理し独自性のある課題を発見し、仮説を設定している。
課題 解決力	インターネットや先行研究等、他者によるデータ解析された解決策を加工した内容である、もしくは、解決策が提示されていない。	データ解析を実践しているが、利用しているデータが課題解決につながりにくく不十分、もしくは、食い違っている。	課題に適したデータによって、データ解析されており、客観的な判断による課題の解決策が提示されている。	複数のデータ解析されており、複数の結果から客観的な判断による課題の解決策が提示されている。	現実に実現可能な解決策を提示し、その解決策に対して自らが解決に向けて実行している。
表現力	データの可視化が不十分である。	データを可視化しているが、適した可視化でなく、グラフ等の特徴が伝わりにくい。	データに適した可視化ができており、グラフの特徴的な部分を示し、データ分析の内容が伝わった。	複数のデータの可視化が論理的に順序立てた構成になっている。	課題発見・解決の全体の構成が聞き手に伝わりやすい可視化ができています。
プレゼンテーション 力	必要以上に資料等を見ながら発表し、聞き手に向かって発表できていない。	聞き手に向かって発表できているが、声の大きさ・速さなどが不十分である。	聞き手に伝わるように発表できている。	ジェスチャー、抑揚間の取り方など、聞き手に伝わりやすい発表ができている。	聞き手の反応を見て臨機応変な発表ができている。

5. データサイエンスに関する実技テスト(一例)

「情報とデータサイエンス」に関わるテスト内容

問題に該当するデータを与え、エクセルを活用して問題を解く形式である。SSDSE(教育用標準データセット)を用いて、エクセルを用いた実技テストを実施している。

兵庫県内の市町村において、2018年度のデータに基づいて、次の問いに答えよ。

- (1) 小学校児童数を、小学校教員数から予測する。
このとき、小学校児童数を(A)変数、小学校教員数を(B)変数という。
次の①～⑥における適した用語を選びなさい。
① 説明 ② 因子 ③ 目的 ④ 決定 ⑤ 潜在 ⑥ 相関
- (2) 小学校児童数を(A)変数とし、小学校教員数を(B)変数とし、回帰分析をしたとき、このモデルの寄与率を次の適切な選択肢から選びなさい。
① 0.119 ② 0.995 ③ 17.119 ④ 678.23
- (3) 小学校教員数が98人の市町村では、小学校児童数は何人であると予想できますか。
次の適切な選択肢から選びなさい。
① 約1000人 ② 約1100人 ③ 約1200人 ④ 約1300人
- (4) ある市町村Aと比べて小学校教員が100人多い市町村Bの小学校児童数は、市町村Aに比べて何人多いと推測できますか。次の適切な選択肢から選びなさい。
① 約680人 ② 約800人 ③ 約1030人 ④ 約1710人
- (5) 次のデータ群はそれぞれ正規分布に従うことを仮定し、平均値の2標本の両側検定、分散は2標本で共通で未知とし、母平均の差を仮説検定で調べた。帰無仮説を「母平均に有意な差はない」、対立仮説を「母平均に有意な差がある」としたとき、有意水準5%とし、帰無仮説を棄却できるデータの組を選びなさい。
① 15歳未満人口の「男」と「女」 ② 75歳以上人口の「男」と「女」
③ 出生数と死亡数 ④ 婚姻件数と離婚件数
- (6) 兵庫県内の市町村における姫路市の婚姻件数において、標準化した値を小数第2位で求めなさい。
- (7) 兵庫県内の市町村における姫路市の婚姻件数と離婚件数は、全体分布の位置から判断し、どちらの方が多いと判断できますか。
① 婚姻件数 ② 離婚件数

独自指標作成による地方創生の方法論と兵庫県活性化の提案

1. 研究のテーマと目的

1.1 問題意識の背景

兵庫県は、摂津・播磨・但馬・丹波・淡路という歴史も風土も異なる5国からできている。阪神地域を中心とする大都市圏に始まり、瀬戸内海に浮かぶ淡路島や城崎温泉で有名な多自然地域、コウノトリや全国屈指の金物生産、重工業企業の参入などで栄える地方都市圏、といったように様々な特色を持つ地域が混在している。それ故に「日本の縮図」と称されるが、抱える問題までも日本を反映して複雑化している。

1.2 現状の把握

県がまとめる課題・現状に関する6つの資料^[1]を効率よく把握するために、テキストマイニングを用いて頻出語・共起語を調べた。(図1) その結果、東西南北を貫くように整備された「基幹道路」の活用法、周辺都市との「交流人口対策」、県内での「繋がり」が今後解決すべき課題ではないかと考えたが、より複雑な課題があると感じた。

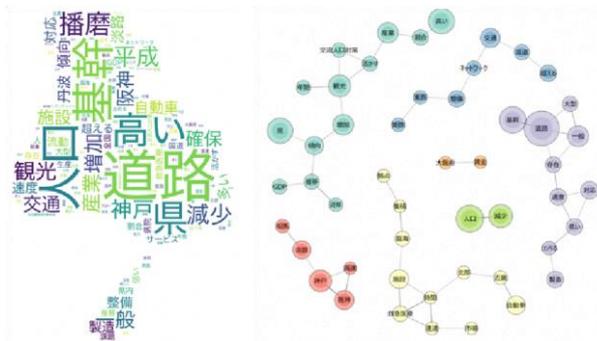


図1：頻出語(左) 共起語(右)

1.3 研究の目的と先行研究

本研究では、「兵庫県の抱える複雑な課題を読み解き、活性化させるにはどうすればよいか。」とリサーチクエスチョンを立てた。従来、地方創生は子育てや経済といった単一側面での課題をきっかけに政策が進められている。ひょうごビジョン2050^[2]によると、100のデータから兵庫の課題が例示されている。しかし、個々の課題を挙げて解決するだけでは意味がなく、最も解決すべき課題は何か、課題どうしの関係性は考えなくて良いかなど、単一側面からでは限界があると私達は考えた。そこで、「統合的アプローチでの方法論を編み出すことで兵庫県の課題発見・解決につなげられる」と仮説を立てた。

ここで参考にした先行研究・取り組みは、ふじのくに少子化突破戦略の羅針盤、AIを活用した長野県の未来に関するシミュレーション、Fragile States Index^[3]である。多側面からデータを収集して独自指標を作成・スコアを算出することで、全国の市区町村間を客観的に比較できるのではないかと考えた。

2. 研究の方法と手順

2.1 研究の概要

本研究は、証拠に基づく政策立案：EBPM (Evidence Based Policy Making) に則り、4ステップで研究を進める。(図2)

- ① データの収集・整形・スケーリング
- ② 評価指標作成
- ③ 地域課題の発見
- ④ 施策提案

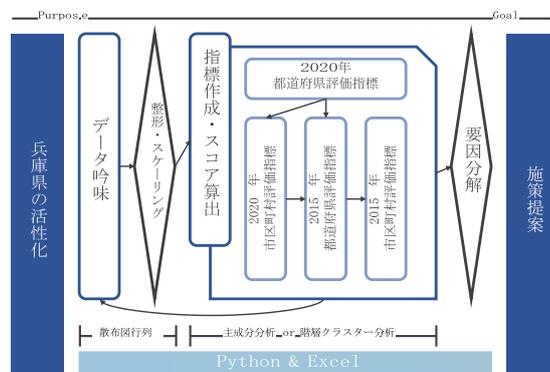


図2：EBPM図

何年も継続して運用・比較が可能な方法論を確立するためには、国勢調査等で収集されたオープンデータが好ましいと考えた。まず、複数観点から各地域を評価できる指標（都道府県評価指標と市区町村評価指標）を作成した。更に、観点ごとに0～10点のスコアを算出する。そのスコアが算出された要因を調べたり、他地域とスコアを比較したりすることで兵庫県の特徴を読み解く。

2.2 方法と手順

4つのステップを詳説する。

① データの収集・整形・スケーリング

SSDSE や e-Stat を利用して市区町村データを収集する。本研究では、都道府県単位で比較する指標と市区町村単位で比較する指標の2種類を作成する。始め、都道府県データを基準にデータ収集したとき、市区町村データが存在しないことや、市区町村データが存在しても欠損値を多く含むデータであることが頻発した。そこで、市区町村データを基準に収集することでこの問題を回避する。しかし、数地域は欲しい年度のデータが欠損値の場合がある。その際には延長推計を用いて推測した。過去5年分のデータから「年度」をXの値、「データ」をYの値に設定して傾きと切片を算出し、欠損値を推計した。(図3)

また、データの整形に関しては、四分位数による外れ値の置換、散布図行列を利用した多重共線性の有無の確認を行った。

② 評価指標作成

収集したデータを元に「都道府県評価指標」と「市区町村評価指標」を作成する。まずは、評価指標作成の手順(図4)に基づいて、都道府県評価指標を作成する。評価指標作成の全体像は(図5)の通りである。観点分けは、統計手法による機械的なグループ分けから、観点到に属するデータや地理的要因等を考慮した意味づけをすることで行う。都道府県評価指標を作成した後、市区町村評価指標、過去年度の指標の作成にあたる。

③ 地域課題の発見 ④ 施策提案

算出したスコアから、都道府県や市区町村をクラスタリングすることで経済圏や都市圏の把握を行う。また、スコアの要因分解をすることで、同じスコアであっても成因が異なる場合にそれぞれの地域の特異性を導けたり、そこから課題点を明らかにできたりする。これらの研究結果から兵庫県の特徴を多面的に読み解き、的確な施策を提案する。

3. データセットの加工

本研究で使用した全データを(表1)に、指標作成に使用したデータセットを(表2)に示す。なお、灰色マーカーの項目は最終的に使用しなかったデータである。

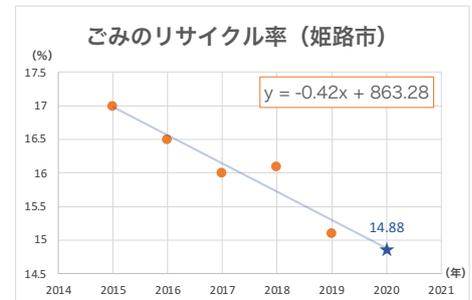


図3：延長推計の例

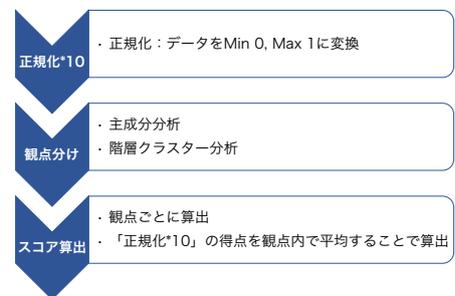


図4：評価指標作成の手順

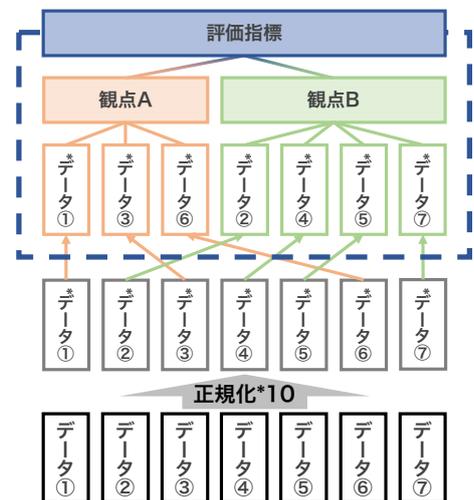


図5：評価指標作成の全体像

表 1：使用した全データ

変数名	単位	出典*1
総人口	人	2023A/2018A
15-64 歳人口	人	2023A/2018A
65歳以上人口	人	2023A/2018A
15歳未満人口	人	2023A/2018A
非労働力人口	人	2023A/e-Stat
転入者数（日本人移住者）	人	2023A/2018A
転出者数（日本人移住者）	人	2023A/2018A
婚姻件数	件	2023A/e-Stat
離婚件数	件	2023A/e-Stat
歳出決算総額（市町村財政）	円	2023A/2018A
民生費（市町村財政）	円	2023A/2018A
教育費（市町村財政）	円	2023A/2018A
経常収支比率（市町村財政）	%	e-Stat/2018A
実質公債費比率（市町村財政）	%	e-Stat/2018A
第3次産業就業者割合	%	2023A/e-Stat
農家数	戸	e-Stat
製造業出荷額等	百万円	e-Stat
製造業就業者数	人	e-Stat
商業年間商品販売額（従業者一人当たり）	万円	e-Stat
課税対象所得	円	e-Stat
10万人あたり病院数	院	e-Stat
有効求人倍率	倍	e-Stat
一人一日当たりの排出量	g	e-Stat
財政力指数	%	e-Stat
昼夜間人口比率	%	e-Stat
完全失業率	%	e-Stat
ごみのリサイクル率	%	e-Stat
合計特殊出生率	%	https://www.mhlw.go.jp/toukei/list/158-60.html
市町村別農業産出額	千万円	https://www.maff.go.jp/2/hokei/kyuhou/syusei_sansyuhou/

表 2：データセット

データ名	計算式	延長推計*2
課税対象所得（人口一人当たり）	[課税対象所得]/[総人口]	
第3次産業就業者割合	[第3次就業者割合]/[就業者割合]*100	
農業産出額（一戸当たり）	[市町村別農業産出額]/[農家数]	[2020]1999-2019
製造品出荷額等（従業者一人当たり）	[製造業出荷額等]/[製造業就業者数]	[2020][2015]2014-2020
商業年間商品販売額（従業者一人当たり）		[2020][2015]2003-2015
財政力指数		[2020]2015-2019
経常収支比率（市町村財政）		[2020]2015-2019
実質公債費比率（市町村財政）		[2020]2015-2019
婚姻率	[婚姻件数]/[総人口]*1000	
離婚率	[離婚件数]/[総人口]*1000	
昼夜間人口比率		
10万人あたり病院数		
有効求人倍率		
完全失業率		[2020]1980-2015
非労働力人口割合	[非労働力人口]/([15~64歳人口]+[65歳以上人口])*100	
合計特殊出生率		[2020]2006-2015
老年化指数	[65歳以上人口]/[15歳未満人口]*100	
転入超過率	[転入者数（日本人移住者）]-[転出者数（日本人移住者）]/[総人口]	
一人一日当たりの排出量		
ごみのリサイクル率		[2020]2015-2019
民生費割合	[民生費（市町村財政）]/[歳出決算総額（市町村財政）]*100	
教育費割合	[教育費（市町村財政）]/[歳出決算総額（市町村財政）]*100	

(※1) 略称を用いて表記した。

2023A：SSDSE-A（市区町村）2023年版

2018A：SSDSE-A（市区町村）2018年版

e-Stat：市区町村のすがた（社会・人口統計体系）

(※2) [推計した年度] 使用したデータの年度

データで見る県勢⁴の経済指標・生活指標を参考に、指標作成に必要なと考えられるデータを選定した。

データの選定ポイントは、①同一年度であること(2020年度, 2015年度) ②優劣(高低)がはっきりしていること ③欠損値のある地域が5つ以内であること の3点で、データ収集地域数は、特別区+市町村の1741地域である。(東京23区は区別するが、大阪市等の区別は区別しない。)

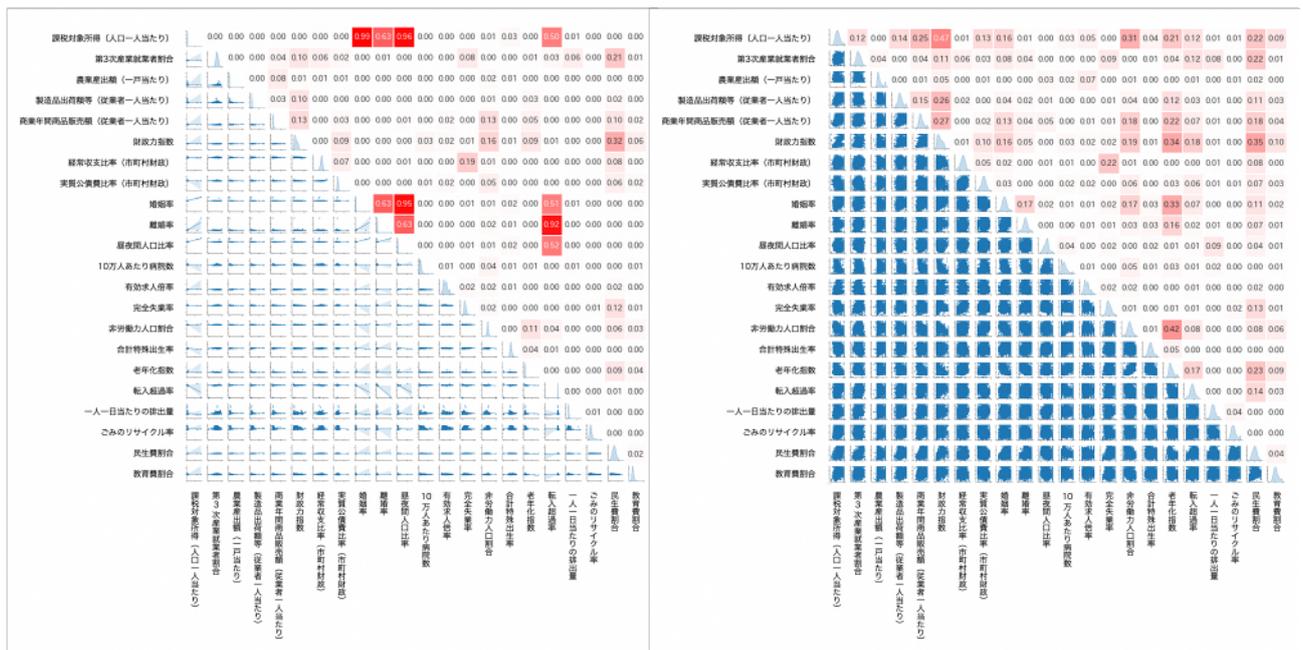


図 6：散布図行列（散布図・ヒストグラム・R2 値）外れ値置換前（左）、外れ値置換後（右）

指標の特性上、同じ性質をもったデータが複数あるのは好ましくない。そこで、重回帰分析で主に利用される多重共線性（VIF）の確認をすることで、データの性質を把握する。VIF \leq 3（R² \leq 0.68）を基準とする。

このために、データセットの散布図行列を作成した。（図 6 左）2015 年度と 2020 年度の全 3482 データの散布図と近似直線、ヒストグラム、R² 値をまとめたものである。これより、「課税対象所得（人口一人当たり）」と「婚姻率」の R² 値が 0.99 であるなど、数データは VIF \leq 3 を満たさないことが分かった。しかし、近似直線に着目すると、外れ値があることに気づいた。これはヒストグラムの裾が軽く、尖度が高いことからわかる。そこで、四分位数を用いて外れ値を除外し置換することにした。（図 7）

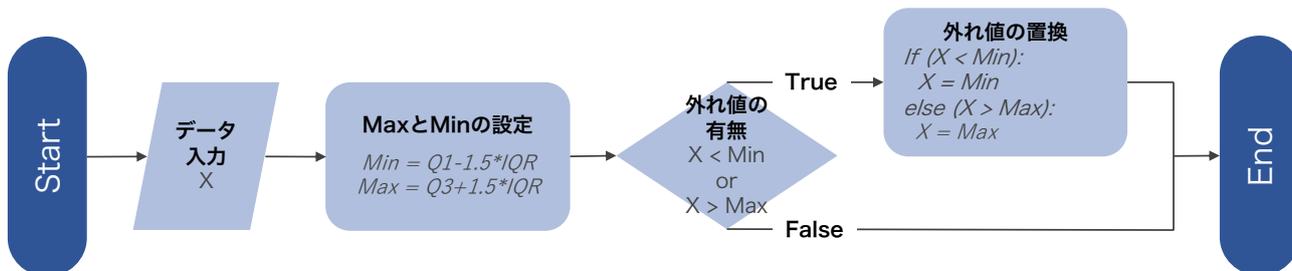


図 7：外れ値置換のフローチャート（ Q1：第 1 四分位数, Q3：第 3 四分位数, IQR：四分位範囲[Q3-Q1] ）

そして、外れ値を置換したデータセットを元に再度散布図行列を作成した。（図 6 右）その結果、このデータセットのデータはほとんど正規分布し、VIF \leq 3 を満たすことが確認できた。従って、外れ値を除外した下ではどのデータも異なる性質を持っていると判断した。

4. データ分析の結果

4.1 評価指標の作成（観点分け）

データセットは 22 項目のデータからなっているが、これをそのまま指標として使い、22 項目を地域ごとに比較するのは研究の効率が悪く、項目間の関わりを見落とす可能性がある。ゆえに、関連性を持つ数データをいくつかの観点に分けることで解決する。

データの観点分けには「主成分分析」と「階層クラスター分析」を検討し、より良い方を採用する。

① 主成分分析

Python を用いて主成分分析を行い、求めた各主成分を観点として利用できるのではないかと考えた。そこで、「2020 年度+2015 年度データセット（外れ値なし・標準化）」のデータセットを作成し、主成分分析を行った。

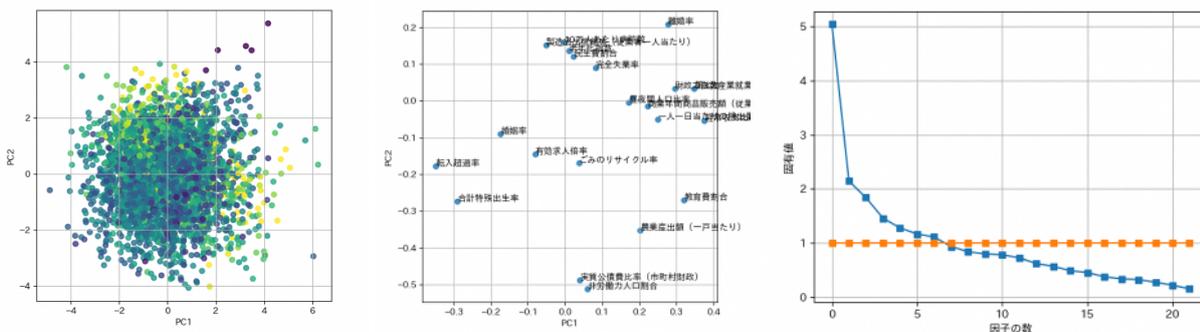


図 8：主成分分析（第 1 主成分と第 2 主成分のプロット図（左）・寄与率（中央）、スクリープロット（右））

外れ値は先程の3. データセットの加工 より、除いた方が良くと判断した。また、データセットを正規化した際、固有値が1より遥かに小さく、元のデータとの関連性が低いと考えられたため、標準化を用いた方が良くと判断した。そして、データ数に幅を持たせるために2015年度と2020年度の両方のデータを用いて主成分を求めることにした。スクリープロットより、第7主成分まで使用して7つに観点分けすることにする。

(図8右)

表3：第7主成分までの主成分負荷量

主成分	課税対象所得 (人口一人当たり)			第3次産業就業 者割合		農業産出額 (戸当たり)		製造品出荷額等 (従業者一人当たり)		商業年間商品販 売額 (従業者一人当たり)		財政力指数		経常収支比率 (市町村財政)		実質公債費比率 (市町村財政)		離婚率		離婚率		昼夜間人口比率																																																							
	PCA1	PCA2	PCA3	PCA4	PCA5	PCA6	PCA7	PCA1	PCA2	PCA3	PCA4	PCA5	PCA6	PCA7	PCA1	PCA2	PCA3	PCA4	PCA5	PCA6	PCA7	PCA1	PCA2	PCA3	PCA4	PCA5	PCA6	PCA7																																																	
PCA1	0.77926	0.45405	-0.11583	0.49748	0.66241	0.83988	0.08974	-0.39049	0.62092	0.38531	-0.00560	0.04954	-0.51577	0.22122	-0.02268	0.04859	-0.07814	-0.71354	-0.13258	0.30450	-0.00713	0.23374	0.14932	-0.03808	-0.44825	-0.16365	-0.27092	0.06724	-0.09075	-0.35450	-0.21815	-0.40137	-0.60399	-0.34316	0.01433	-0.09301	-0.21886	-0.24863	-0.15423	0.01325	0.13373	0.21103	0.46399	-0.44355	-0.05725	0.40787	-0.51678	-0.35699	-0.18609	-0.01813	-0.25166	-0.16295	0.29098	0.11128	0.18574	0.06969	0.37439	0.37902	-0.32713	-0.00124	-0.09796	-0.24392	-0.34842	0.08863	0.08795	-0.07798	-0.03527	0.01666	-0.22510	0.23196	0.05122	0.03654	0.13703	0.16257	0.06334	-0.07914	0.12093
主成分	10万人あたり 病院数		有効求人倍率		完全失業率		非労働力人口割合		会計特殊出生率		老年化指数		転入超過率		一人一日当たり の排出量		ごみのリサイクル率		民生費割合		教育費割合																																																								
	PCA1	PCA2	PCA3	PCA4	PCA5	PCA6	PCA7	PCA1	PCA2	PCA3	PCA4	PCA5	PCA6	PCA7	PCA1	PCA2	PCA3	PCA4	PCA5	PCA6	PCA7	PCA1	PCA2	PCA3	PCA4	PCA5	PCA6	PCA7																																																	
PCA1	-0.18105	0.18322	0.13604	-0.64907	0.02256	-0.78064	0.56033	0.08447	0.05212	0.71822	0.40765	-0.21185	0.13167	-0.74933	-0.40063	0.19813	-0.25881	-0.07479	-0.24653	0.17595	-0.39588	0.00977	-0.46101	0.35615	-0.04251	0.02027	-0.35395	0.09714	0.17482	-0.49404	0.19269	0.09238	0.07670	-0.13130	-0.30435	0.16049	-0.04073	0.54977	-0.31835	0.15589	-0.38479	-0.07388	0.08937	-0.00099	0.04901	0.31745	-0.18869	0.15189	0.11085	0.04235	0.12680	0.38530	-0.26790	-0.06406	-0.21096	0.17444	-0.41354	-0.02623	0.09045	-0.25267	0.08956	0.15968	0.00208	0.47958	0.06803	-0.13824	0.38244	0.26897	-0.10704	0.16055	0.29817	0.02778	0.10954	-0.27617	0.60030	0.11936	-0.35455

そして、第7主成分までの主成分負荷量を求めた。

(表3)しかし、第4主成分以降は意味づけに不十分な負荷量が多いことが判明した。相関係数の絶対値のほとんどが0.3以下の弱い相関で、第3主成分までに比べて各主成分の特異性を説明しにくい。その理由としては、項目間の相

関関係が非常に小さいため、主成分がうまくまとめられていないからではないかと、そもそも扱う主成分が多すぎるからではないかと考えられる。従って、主成分分析での観点分けは断念した。

② 階層クラスタ分析

Pythonを用いて階層クラスタ分析を行い、クラスタリングによって観点分けができるのではないかと考えた。主成分分析同様、「2020年度+2015年度(外れ値なし+標準化)」のデータセットを作成し、クラスタリングを行ったところ、縦軸に示すクラスタ間ユークリッド距離が非常に大きくなってしまった。(図9)そこで、標準化の代わりに正規化でスケールリングすることで解決した。(図10)

このとき、優劣を判定する指標である場合、データの優劣の向きを揃えなくてはならない。例えば、「完全失業率」は低いほど良い。このような項目のデータの優劣の向きを変更する。データセットは正規化されているので、優劣を変えたいデータをXとすると、 $[1-X]$ を計算することで優劣の向きを変更できる。データセットのうち、「経常収支比率」、「実質公債費比率」、「離婚率」、「完全失業率」、「非労働力人口」、「老年化指数」、「転入超過率」、「一人一日当たりの排出量」の優劣を変更した。その後、再度クラスタリングした。(図11)

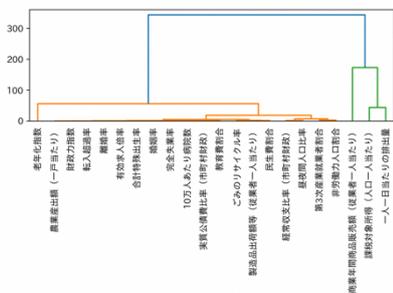


図9：「2020年度+2015年度(外れ値なし+標準化)」

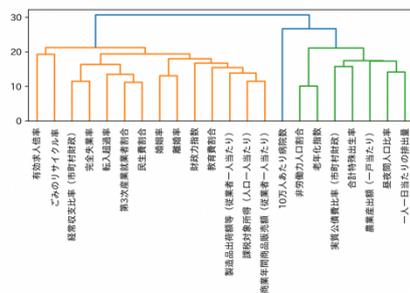


図10：「2020年度+2015年度(外れ値なし+正規化)」

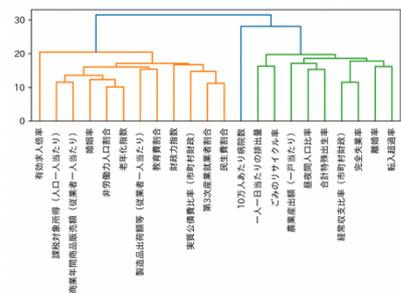


図11：「2020年度+2015年度(外れ値なし+正規化+優劣統一)」

ここで、このデータセットは2つの分野に大別できるのではないかと考えた。また、「有効求人倍率」「財政力指数」「10万人あたり病床数」は孤立しているのではないかと考えられる。そこで、この3つの項目を削除すると共に、分野別に分けてクラスタリングをした。(図12)

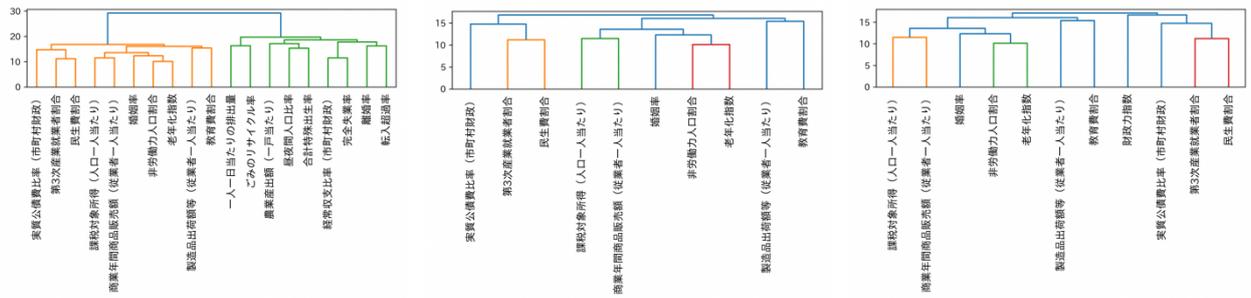


図 12：デンドログラム全体(左)、分野 1 「実質公債費比率（市町村財政）～「教育費割合」（中央）、分野 2 「課税対象所得（人口一人当たり）～「民生費割合」（右）

従って、階層クラスター分析の結果より、19 項目を 2 分野・6 観点に分ける。（表 4）観点の意味づけは、ヒートマップを作成して（図 13）正規化得点の高低や地理的分布から行った。

表 4：評価指標の分野と観点

分野	観点	項目
α	持続可能性	一人一日当たりの排出量
		ごみのリサイクル率
		農業産出額（一戸当たり）
β	第1次産業力	昼夜間人口比率
		合計特殊出生率
		経常収支比率（市町村財政）
γ	地方安定性	完全失業率
		離婚率
		転入超過率
δ	都市集権性	実質公債費比率（市町村財政）
		第3次産業就業者割合
		民生費割合
ε	都市力	課税対象所得（人口一人当たり）
		商業年間商品販売額（従業者一人当たり）
		結婚率
ζ	第2次産業力	非労働力人口割合
		老年化指数
		製造品出荷額等（従業者一人当たり）

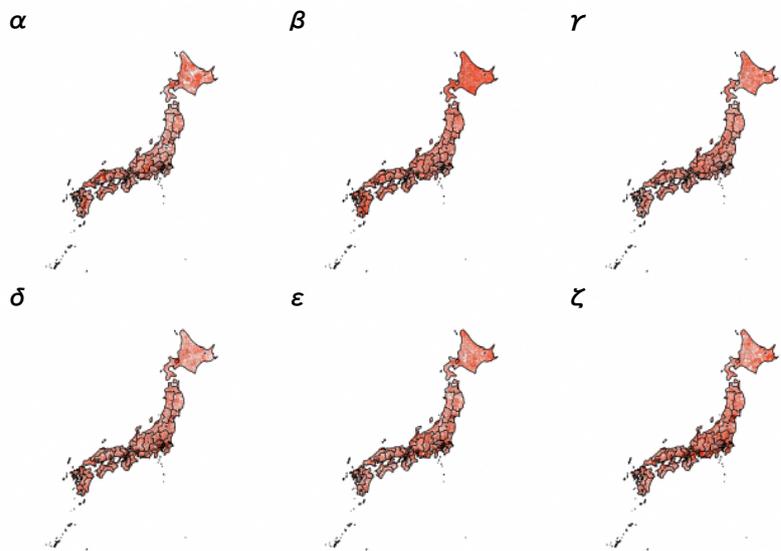


図 13：ヒートマップ（正規化得点が高いほど濃い赤）

4.2 評価指標の作成（スコア算出）

各観点内の項目の正規化得点を単純平均した後、10 倍することでスコアを求める。まずは、2020 年度の都道府県評価指標を作成する。その後、市区町村評価指標も作成し、それぞれ 4 つの地域にクラスタリングし、要因分解や地理的要因から各地域の役割（以後、「地域役割」とする）を把握する。（図 14・15）そして、2015 年度版も作成する。（図 16）

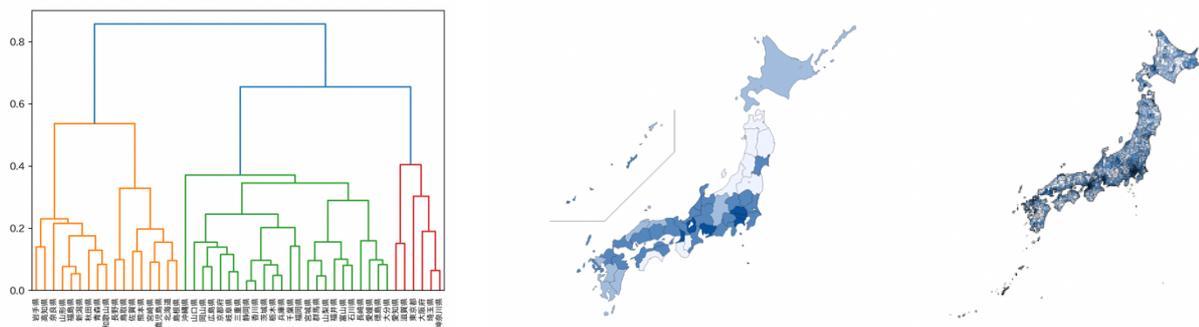


図 14：デンドログラム（左）、都道府県 4 地域分け（中央）、市区町村 4 地域分け（右）

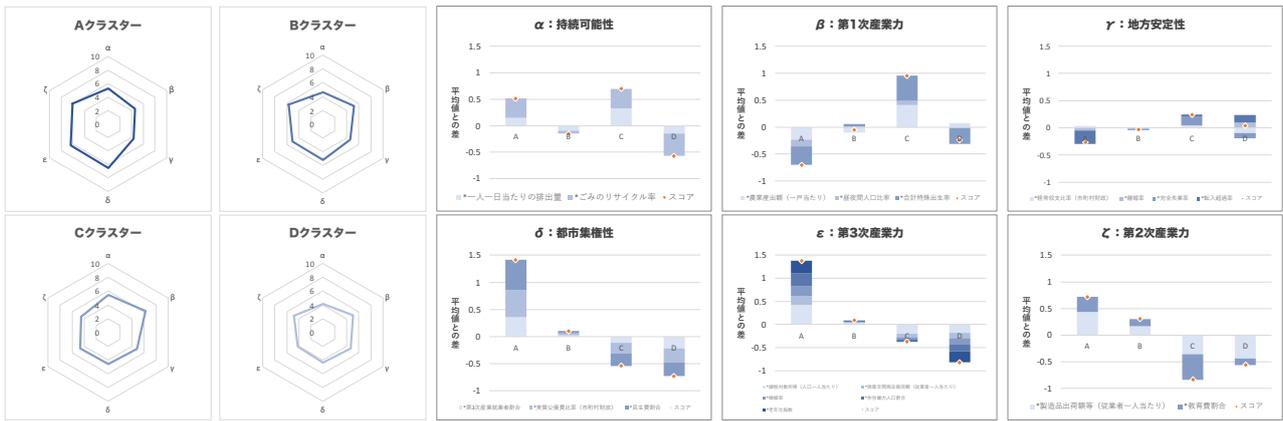


図 15：2020 年度都道府県評価指標のスコアレーダーチャート（左）、要因分解（右）

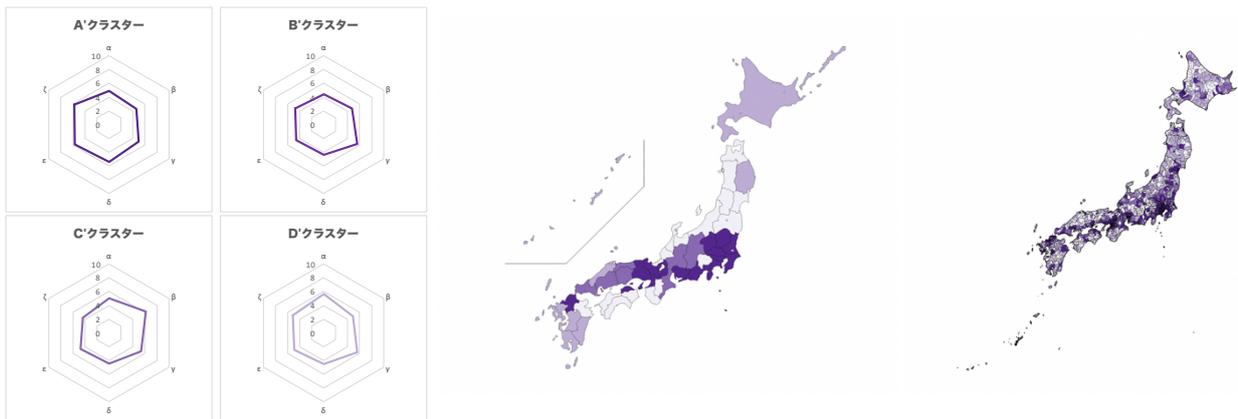


図 16：2015 年度都道府県評価指標のスコアレーダーチャート（左）、都道府県 4 地域分け（中央）、市区町村 4 地域分け（右）

分析結果から 2020 年度と 2015 年度の日本の地域役割を考察した。(表 5) 5 年の内に、都市型の特徴を持つ地域が増加した一方、兵庫県のように都市力が弱まってしまった地域もある。しかし、これは三大都市圏への権力集中が強まったとの見方もできる。事実、都心型の地域は 15 都府県から 6 都府県に減少している。また、地方型の地域の変化が小さいのも問題で、地方活性化はうまく進んでいないと考察できる。

表 5：2020 年度指標による都市圏（左）、2015 年度指標による都市圏（右）

クラスター名	特徴	クラスター名	特徴
都市型 A：都心型	三大都市圏を中心に分布 「地方安定性」が特に低く、「都市集権性」「第3次産業力」が非常に高い	都市型 A'：都心型	三大都市圏を中心に分布 「地方安定性」が特に低く、「都市集権性」「第3次産業力」「第2次産業力」が高い
都市型 B：準都心・工業特化型	Aクラスターからの分離 太平洋ベルト沿い地域で、「第2次産業力」が高め 「第1次産業力」が高い	都市型 B'：地方・都市化型	2020年度指標でBクラスターの地域が多く分布 「都市集権性」が低く、「地方安定性」が高い
地方型 C：地方・農林畜水産特化型	北海道や鹿児島県周辺の農・畜産業、鳥取県や島根県などの水産・農業が盛んな地域	地方型 C'：地方・農林畜水産特化型	2020年度指標のCクラスターに類似 北海道や鹿児島県周辺の農・畜産業、鳥取県や島根県などの水産・農業が盛んな地域
地方型 D：地方型	要因分析より、「地方集権性」のプラス要因もマイナス要因も高い いわゆる少子高齢化が進む過疎地域で、今後の改善が重要	地方型 D'：地方型	要因分析より、「地方集権性」以外はすべてマイナス要因 いわゆる少子高齢化が進む過疎地域で、今後の改善が重要

5. 結果の解釈

5.1 兵庫県の課題把握

現在の兵庫県は、全体として「準都心・工業特化型」地域に当てはまるが、内訳を見てみると、阪神地域や瀬戸内海沿岸部は「都市型」地域、内陸部は「地方型」地域、南北は「地方・農林畜水産特化型」地域である。(図 17) 従って、日本の全ての都市圏を持っていることになる。また、周辺地域に着目してみると、他府県と経済圏を作っていることも分かる。ここで、北部は「京丹後経済圏」、瀬戸内海沿岸の都市型地域は「京阪神経済圏」と定義する。ここで、京阪神経済圏をよりミクロに解析する。クラスタリングの結果、さらに3つの地域に分離し、濃い緑の地域は「地方力」が高めの経済圏であることが分かった。(図 18)

2015 年度指標の結果と比較すると、姫路市周辺の瀬戸内地域や内陸部の変化が大きいことが分かる。(図 19) 現在、内陸部の鉄道路線は赤字が続いており^[5]、地方型都市圏が 2020 年度には増えていることから他府県と隣接する地域よりも内陸部は地方活性化がうまく行っていないのではないかと考察できる。

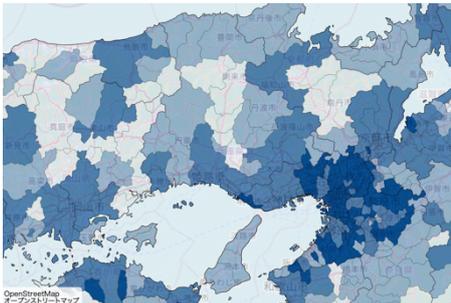


図 17：2020 年度指標

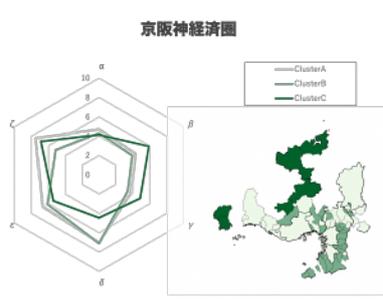


図 18：京阪神経済圏の内訳（色が対応）

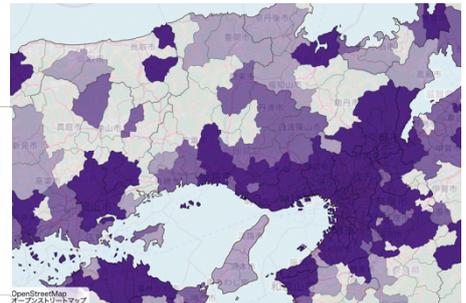


図 19：2015 年度指標

5.2 兵庫県活性化の施策提案

これらの研究結果を踏まえて、私達は兵庫県の「6 次産業都市」化を推進する施策を提案する。兵庫県の課題として内陸部の地方活性化が挙げられた。ここで、県全体で 6 次産業を推進することで各地域の特色をうまく活かし、ピンポイントで各々の課題解決につながるのではないかと考えた。

兵庫県の 6 次産業に関する現状について調べてみると、総合的支援についての指針が示されており^[6]、補助金の交付もあるようだが、現在（2023 年 9 月）公表されている補助金交付事業はたったの 3 件である。そこで、県の 6 次産業の売上について調べてみると、6 次産業関連産業一事業体あたりの売上は全国 43 位、特に農産加工が全校平均より大きく下回っていることが分かった。(図 20) 県は、6 次産業に関して受動的でなく能動的な支援が必要であると考えられる。

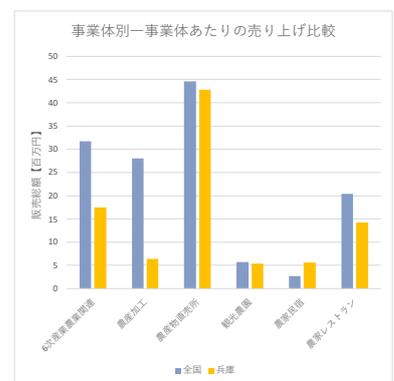


図 20

まず、今まで交付金として用意していた予算を事業費に転換して県内陸部で農産加工の拠点を作成する。更に、農産加工品の販売所を兼ねた地産地消を目的とするレストランやカフェを数カ所に展開する。これは県内のみならず、近隣他府県を含む経済圏単位での出店も視野に入れる。農産加工拠点としては、地方型都市である「朝来市」を推薦する。その理由は、北近畿豊岡自動車の分岐点であることや、周辺地域から岩津ねぎやいちご、丹波の黒豆など様々な特産品を集めやすいことなどである。朝来市を拠点に、協賛していただいた農家の農産物を加工し、基幹道路の弁を活かして各地の販売所・飲食店に輸送する。更に、SNS を通した広告戦略で県内外を問わずにお客さんを集めようという目論見である。農産加工拠点の規模により費用は異なるが、県で問題になっている空き家・空き地を活用することで費用を抑える。設備費用こそ、3 千万円ほどを想定するが、実は既存の補助金^[7]でそのほとんどが賄えることが分かった。つまり、それほどの財源は確保できるはずである。内陸部を活性化させるために、整備された基幹道路を通して特産物を持つ周辺地方部と人が集められる周辺都市部で支える。この施策を通して県全体を更に活性化できるはずだ。

今後の展望は、指標の妥当性評価を地方自治体にしていただき、県政で KPI 化していただくことである。

6. 参考文献・出典

[1] 「神戸・阪神地域における現状・課題と政策目標」 <https://web.pref.hyogo.lg.jp/ks08/documents/2-3-1koube.pdf>
「播磨地域における現状・課題と政策目標」 <https://web.pref.hyogo.lg.jp/ks08/documents/2-3-2harima.pdf>
「但馬地域における現状・課題と政策目標」 <https://web.pref.hyogo.lg.jp/ks08/documents/100issiki-2.pdf>
「兵庫観光の現状と課題について」 <https://web.pref.hyogo.lg.jp/sr16/senryaku/documents/documents/4-siryou2.pdf>
「兵庫経済の現状と課題」 https://web.pref.hyogo.lg.jp/sr07/documents/siryou4_300806.pdf
「本県の医療の現況」 <https://web.pref.hyogo.lg.jp/kf15/documents/kousou2.pdf>
[2] 「未来を考える 100 のデータ」 <https://hyogo-vision.com/wp-content/themes/avant-child/pdf/knowledge100.pdf>

[3] 「ふじのくにに少子化突破戦略の羅針盤」 <https://www.stat.go.jp/dstart/case/06.html>
「AI を活用した長野県の未来に関するシミュレーション」 <https://www.pref.nagano.lg.jp/kikaku/kensei/ai/ai.html>
「FRAGILE STATES INDEX」 <https://fragilestatesindex.org/>
[4] (公益財団法人) 矢野恒太記念会編集：「データで見る県勢 2021」、pp. 14-16、矢野恒太記念会編集(2020)
[5] <https://www.kobe-np.co.jp/news/sougou/202204/0015211113.shtml>
[6] <https://web.pref.hyogo.lg.jp/nk02/hyogo6support.html>
[7] <https://www.smart-hojokin.jp/subsidy/prefectures/28>
「e-Stat」 <https://www.e-stat.go.jp/> 「KH Coder」 <https://kncoder.net/> 「MANDARAJIS」 <https://ktgis.net/mdrjs/>

ため池が周辺地域の暑熱環境に及ぼす効果

1. 研究背景

晴れた夏の日中, 陸より水域の方が低温だということに注目した. そこでため池は周辺地域の高温化抑制に効果があるのではないかと考え研究した.

2. 研究内容

姫路市の四ツ池を研究対象とした. 領域気象モデルWRFを用いて四ツ池がある場合と, 四ツ池を宅地化した場合の気象の違いを比較する. ため池の影響を大きく受けるには風向きと関係があるのか対照実験も施行する.

3. 仮説

- ① 大気に移動する熱量
池を宅地化すると地面から大気に移動する熱量は大きくなる.
- ② 周辺地域の気温変化
池を宅地化すると, 周辺地域の気温は上昇する.
- ③ 風向きと気温の関係
風向きによって池周辺の気温差が大きい場所は変わる.

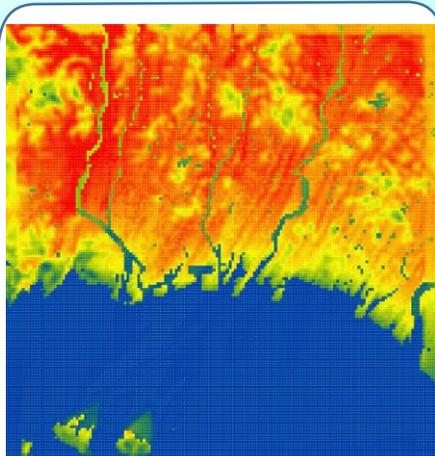
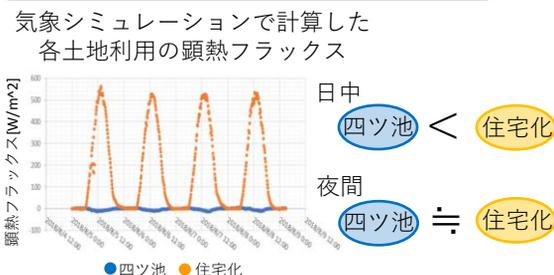


図1. 研究対象地域の
気温メッシュマップ
日時(2022年7月20日14:30)
水平格子点間隔200m

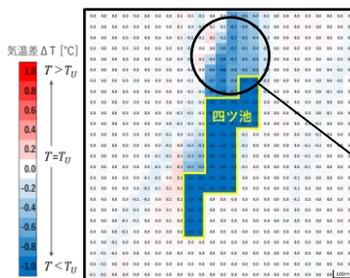
4. 結果

①大気に移動する熱量の変化



結果 1 四ツ池によって大気に移動する熱量は抑えられている

②四ツ池の有無による気温変化



日時(2022年7月20日14:30)
水平格子点間隔50m

四ツ池の有無による気温差[°C]
 $\Delta T = T - T_U$
 T =(四ツ池がある場合の気温)
 T_U =(四ツ池を宅地化した場合の気温)
四ツ池の南側より北側の方が
気温差が大きい

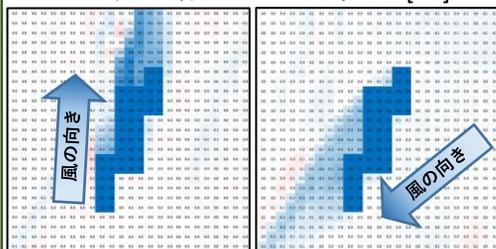
10分間の 積分 → 1日あたりの
気温差 気温差

四ツ池の有無による1日当たりの気温差

$$\sum \Delta T = \int_{t=7:00}^{t=19:00} (T_U - T) dt \quad [^{\circ}\text{C}\cdot\text{h}]$$

③風向きと気温差の関係

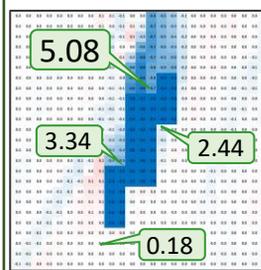
四ツ池の有無による気温差[°C]



日時(2022年7月20日14:30) 日時(2018年8月7日10:10)

	左図	右図
風向き	南→北	北東→南西
色の場所	北	南西
風量	6.19m/s	5.78m/s
気温差	大	小

結果 3 四ツ池の直上と風下側の地域で高温化抑制の効果が波及している



日時(2022年7月20日7:00-19:00)

東西南北 全ての地点で1日
当たりの気温差が正の数

↓
1日を通して四ツ池は周辺
地域の気温を抑えている

結果 2 四ツ池に高温化抑制効果はあるが場所によって差がある

ため池が周辺地域の暑熱環境に及ぼす効果

5. 結論

ため池があることで顕熱フラックスを抑え、池の直上だけでなく風下側の地域で高温化抑制の効果が波及する。現在、田畑の利用減少に伴い多くのため池が埋められている。しかし、周辺地域の気温を抑制する点において自然にあるため池そのものの存在自体がSDGsの目標達成に貢献している。よって私たちはため池を守るべきであると警鐘を鳴らす。

6. 今後の展望

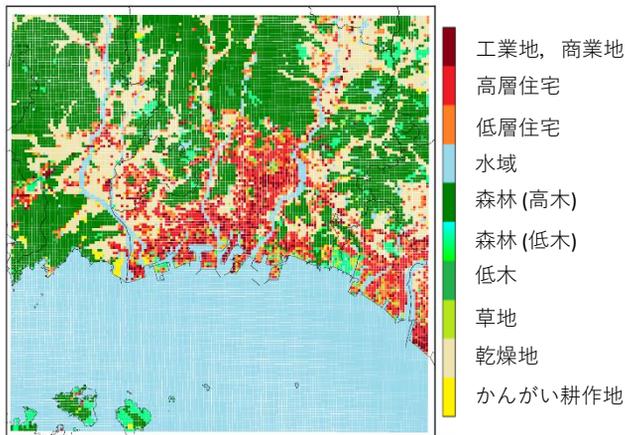


図2. 姫路市南部の土地利用

①池の表面積による高温化抑制効果の違い

疑問点

- ・四ツ池は比較的小さい池
- ・兵庫県には大きい池もある



池の大きさによる違いはどの程度？

対象条件

- ・兵庫県にある池
- ・四ツ池より大きい池
- ・近くにアメダスがある池



比較方法

長池の場所はそのままで以下の高温化抑制効果を比較

- ① 長池の大きさをそのままにする
- ② シミュレーションで長池を四ツ池と同じ大きさにする

②池の場所による高温化抑制効果の違い

(①の予想：池の表面積大 ⇒ 高温化抑制効果大)

疑問点

- ・四ツ池の周辺は住宅地が多い
- ・長池の周辺は住宅地もあるが、田んぼが広がっている



池の周辺地域の環境による違いはどの程度？

比較方法

長池の大きさはそのまま以下的高温化抑制効果を比較

- ① 長池の場所をそのままにする
- ② シミュレーションで長池を住宅地が多い地域(図2.の赤い部分)に移動させる

③ため池周辺地域全体に高温化抑制効果を波及させる手段の提示

(②の予想：田んぼ ⇒ 高温化抑制効果大)

課題点

- ・ため池の風下側でのみ高温化抑制効果がある
- ・住宅地を減らすことは困難



ため池周辺地域を高温化抑制効果の大きい環境にしつつ、風を様々な方向へより遠くまで広げる手段があれば、より広い範囲で高温化抑制効果を波及させる事ができるはず！

提示

- ・ため池周辺に小規模ソーラーパネルと風車を設置することで、再生可能エネルギーによる高温化抑制効果を期待
- ・ため池付近を効率よく低木緑地化することで、環境にやさしい高温化抑制効果を期待

6. 参考文献

- [1]数値予報モデルの引用文献 Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X. Y., Wang, W., Powers, J. G., 2008: A Description of the Advanced Research WRF Version 3. Tech. Note NCAR/TN-475+STR.
- [2]環境省「ヒートアイランド現象とは」 https://www.env.go.jp/air/life/heat_island/guideline/chpt1.pdf
- [3]気象庁「過去の気象データ検索」 https://www.data.jma.go.jp/stats/etm/view/daily_h1.php?prec_no=63&block_no=00&year=2020&month=1&day
- [4]環境省「ヒートアイランドの原因について」 https://www.env.go.jp/air/life/heat_island/manual_01/02_chpt1-2.pdf
- [5]農林水産省「ため池管理保全法に基づく都道府県別の対応状況について」 https://www.maff.go.jp/j/nousin/bousai/bousai_saigai/b_tameike/tameike_taiou.html

兵庫県立姫路西高等学校

謝辞：本研究を遂行するにあたりご指導いただきました
兵庫県立大学 環境人間学部 奥 勇一郎先生に感謝いたします

電離層総電子数 (TEC) データを用いた

オープンソースモデリングアルゴリズムの地震予知への有効性に関する考察

要旨: 地震は壊滅的な影響を発生地帯に及ぼすにも拘わらず、それを防ぐための地震予測の研究はいまだ実用可能な段階には到達していない。近年の多くの研究は、TEC (電離層総電子数) の突発的な変動と地震発生とに関連性があることを指摘している。しかし、TEC には多くの要素が影響しているため、今現在では地震が発生する前に起こる固有の TEC 変動を特定できている研究はない。この問題を解決するため、私たちはオープンソースモデリングアルゴリズムである Facebook Prophet が予測した平常値と地震が起きたときの実測値の比較を行った。この結果、Facebook Prophet は TEC の変動を正しく分析し予測できていることと、すべての地震において発生前に TEC の上昇が見られるわけではないことを私たちは確認した。

キーワード: 地震予測, TEC (電離層総電子数), 電離層, Facebook Prophet

1. 序

1.1. 研究背景

日本は広く認知されている通り地震大国である。2010 年から 2020 年にかけての 10 年間で少なくとも 30 回以上の大規模地震¹が発生しており², それに伴う被害は極めて甚大である。ゆえに、地震をいかに発生前に予測するかについて研究が行われてきた。私たちは、昨年度に地震予知の研究を行っていた本高校 OB の柏木創太先輩を引き継ぐ形で、研究をスタートさせた。

1.2. 先行研究の概略

近年においては、地震と電離層の状態の短期的及び長期的な異常の関係³に焦点を当てている研究が多い。これらの関係は 1964 年 3 月 28 日 3 時 36 分 (UTC) に発生しマグニチュード 9.2 を記録したアラスカ地震 (または聖金曜日地震) の観測を通して、1965 年に Leonard と Barnes によって初めて報告された⁴。これに続いた多くの研究により、電離層擾乱 (Ionospheric disturbances, TEC (後述) の変動) が地震の発生に影響を与えているという結果が示されている¹¹。地球内部の地殻変動によって発生する力が地震発生の原因であるが、この力によって圧縮された岩石中では自由電子が高度に活性化される。この電子群はその後上空に伝播し、これが電離層大気中の電気伝導性の上昇を引き起こすことにより電離層擾乱は発生する¹¹。

具体的な研究事例を挙げると、例えば、柏木 (2021) は東北地方太平洋沖地震が発生するわずか 15~30 分前に電離層擾乱が発生していたことを示した⁵。他にも、珍しいことではあるが本震発生の約 2 か月前に擾乱が確認された事例もある。

¹ M7.0 以上, または M6.0 以上で最大震度 6 弱以上の地震とここでは定義した。

² Wikipedia contributors, 『地震の年表 (日本)#2010 年代』 from Japanese Wikipedia. 2022 年 11 月 23 日 (水) 11:27(UTC) 版の情報に基づく

³ Hammerstrom, Joseph A.; Cornely, Pierre-Richard J. (2016-10-01). *Total Electron Content (TEC) Variations and Correlation with Seismic Activity over Japan*. doi:10.22186/jyi.31.4.13-16.

⁴ Leonard, R. S., & Barnes, R. A. (1965). *Observation of ionospheric disturbances following the Alaska earthquake*. Journal of Geophysical Research, 70(5), 1250–1253. doi:10.1029/JZ070i005p01250.

⁵ 創太, 柏木. (2021) "大気に聞く、地震予測 — GNSS-TEC 法による地震直前予知の可能性 —"

1.3. 用語説明

1.3.1. TEC

電離層の状態を調べるための一つの指標は、特定範囲の電離層に存在する電子の総数である、電離層総電子数 (Total Electron Content, 以下 TEC⁶) である^{7, 8}。TEC はスカラー量であり、その単位は [TECu]⁹ (TEC unit) で、1TECu = 10¹⁶ electron/m² と定義される⁷。TEC には 2 種類あり、それは傾斜電離層総電子数 (Slant Total Electron Content, 以下 STEC) と垂直電離層総電子数 (Vertical Total Electron Content, 以下 VTEC) である。STEC は地球上にある GNSS 観測点と地球上空の GNSS 衛星間斜め方向の電子の数の総数、VTEC は衛星から衛星直下の地表までの間にある 1m² 単位の垂直空間に含まれる TEC の総数と定義される。STEC は、2 つの異なる周波数 f_1 と f_2 を用いた GNSS 観測値の差より得られ、搬送波位相測定値を $P_{4,sm}$ 、真空中の光速を c 、差動符号バイアスの衛星と観測点をそれぞれ DCB_i , DCB_j とすることにより以下の数式のように表される^{10, 17}。

$$STEC = \frac{f_1^2 f_2^2}{40.28(f_1^2 - f_2^2)} (P_{4,sm} + cDCB_j - cDCB_i)$$

STEC はまた、修正単層モデル (MSLM) を用いることによって⁷, R を地球半径 (6371km), H を電離層の最上高度, Z を衛星の地球からの仰角として、以下の等式により VTEC へと変換される¹¹。

$$VTEC = STEC \times \cos \left\{ \arcsin \left(\frac{R \sin Z}{R + H} \right) \right\}$$

以降、特に断りのない限り TEC と言えば VTEC を指す。

$VTEC = STEC \times \cos \left(\sin^{-1} \left(\frac{R \sin Z}{R + H} \right) \right)$ 多くの論文が、正または負の TEC 変動が地震発生の数時間から数週間前に予兆として起こるか、あるいは地震の後に見られることを報告している⁷。例えば、地震の 27 日前に正と負の電離層の変位が起こった例¹², 1998 年~2014 年の地震 ($M_w^{13} \geq 6.0$) 10 日間の統計分析において、正と負の変位が 5 日前に一番大きかった例¹⁴ などがある。しかし、地域別の電離層は様々な要素、例えば太陽放射線、磁場活動、天候事象、人為的効果、そして大気重力波に影響される^{15, 16, 17}。以上のように、様々な要素に影響されるという点で TEC の分析は難しいものとなっている。

⁶ 発音はテック⁸とされる。

⁷ Klobuchar, J. Ionospheric Time-Delay Algorithm for Single-Frequency GPS Users. *IEEE Trans. Aerosp. Electron. Syst.* **23**(3), 325-331.

⁸ 日置 幸介, 菅原 守, 大関 優, 岡崎 郁也 「GPS-TEC 法による地球物理」, 日本測地学会, 2010 年, doi:10.11366/sokuchi.56.125.

⁹ unit の意の u を大文字にした TECU という表記も複数の文献にて見られる。

¹⁰ Jin, R.; Jin, S.; Feng, G. M. DCB: [Matlab code for estimating GNSS satellite and receiver differential code biases](#). *GPS Solut.* 2012, 16, 541-548. doi:10.1007/s10291-012-0279-3

¹¹ Dong, Y.; Gao, C.; Long, F.; Yan, Y. [Suspected Seismo-Ionospheric Anomalies before Three Major Earthquakes Detected by GIMs and GPS TEC of Permanent Stations](#). *Remote Sens.* 2022, 14, 20. doi:10.3390/rs14010020

¹² Guo, J.; Li, W.; Yu, et al. [Impending ionospheric anomaly preceding the Iquique Mw8.2 earthquake in Chile on 2014 April 1](#). *Geophys. J. Int.* 2015, 203, pp. 1461-1470. doi:10.1093/gji/ggv376

¹³ M_w はモーメント・マグニチュードの意であり、マグニチュードと類似した地震の規模の尺度を表す値。

¹⁴ Shah, M.; Jin, S. [Statistical characteristics of seismo-ionospheric GPS TEC disturbances prior to global \$M_w \geq 5.0\$ earthquakes \(1998-2014\)](#). *J. Geodyn.* 2015, 92, 42-49. doi:10.1016/j.jog.2015.10.002

¹⁵ Pulinets, S.A.; Ouzounov, D. et al. [Atmospheric and ionospheric anomalies around the time of the Colima M7.8 earthquake of 21 January 2003](#). *Ann. Geophys.* 2006, 24, 835-849. doi:10.5194/angeo-24-835-2006

¹⁶ Saroso, S.; Liu, J.Y. et al. [Ionospheric GPS TEC Anomalies and \$M \geq 5.9\$ Earthquakes in Indonesia during 1993-2002](#). *Terr. Atmos. Ocean. Sci.* 2008, 19, 481. doi:10.3319/TAO.2008.19.5.481(T)

¹⁷ Yang, S.-S.; Potirakis, S.M.; Sasmal, S.; Hayakawa, M. [Natural Time Analysis of Global Navigation Satellite System Surface Deformation: The Case of the 2016 Kumamoto Earthquakes](#). *Entropy* 2020, 22, 674. doi:10.3390/e22060674

1.3.2. Facebook Prophet

私たちは地震前の TEC を地震予測に活用するために、通常時の TEC と地震前の異常な TEC とを識別できないかと考え、Facebook Prophet (以下 Prophet) という、機械言語 R と Python で提供されているオープンソース・予測モデルの使用を新たに発案した。Prophet に特定の期間のデータを入力すれば、指定した将来の期間のデータの予測が行われる¹⁸。Prophet は、非線的な傾向を持つ時系列データを一般化加法に基づき予測し、年・週・日の季節性 (周期性)、そして祝日効果¹⁹等にも対応することができる。また Prophet は、対象のデータが年・日の周期によって変動することや外れ値等を自動的に考慮し、例外処理をして補正された値を出す機能¹⁴ が自動的に備わっており²⁰、多くの社会科学系の論文²¹において使用されている。

1.4. 研究目的・意義

1.1~1.3 で述べたことより、私たちの研究の目的は以下の二つとなる。

1. Prophet が TEC の傾向を掴み、平常値を予測できるのか確かめること。Prophet が TEC の平常値を予測できるという前提が成り立つならば、先述した通り先行研究より TEC の値は地震発生前後に増減することが示されていることから、Prophet が出力した予測値 (平常値) と実測値にずれ (異常値) が見られる時が地震発生の兆候になっているという仮説を立てることができる。
2. Prophet が出力した予測値と実測値を比較することにより、地震による TEC 変動を、そうでないものと識別すること。

また、私たちの研究により、地震発生を予測しようと試みる研究があることを紹介し人々の関心を高めるとともに、AI アルゴリズムの地震予測への応用が増えることを期待する。

2. 研究方法

Prophet に代入するファイルは日時と TEC の値が記載されたものであるから、これを作成することを目標として進めた。データ整形 4 手順、分析 2 手順である。

2.1. データ整形の手順

1. アメリカの National Geophysical Data Center (NGDC) のサイト²²より 15 分ごとに 1 ファイルとなっている TEC のデータをダウンロードする。私たちが利用可能だと判断したデータベースは NGDC と CODE (Centre for Orbit Determination in Europe)²³の 2 つであるが、NGDC をデータ量とその表示法がよりわかりやすいものであったことを理由として選択した。
2. ファイルが tar.gz 形式で圧縮されているため解凍する。
3. 内包されている txt ファイルより震央の座標の TEC 値を pandas の iat 関数で抽出し、日時とともに csv ファイルに書き出す。
4. 地震発生約 3 か月前までの TEC 値を Python の Prophet 関数に渡し、予測値のグラフを出力させ

¹⁸ “Prophet” 2023 年 4 月 15 日最終閲覧。

¹⁹ 突発的なデータ値の上昇。例えば、遊園地の来客数が祝日に急増することなど。

²⁰ Taylor, Sean J; Letham, Benjamin (2017-09-27). *Forecasting at scale*. doi:10.7287/peerj.preprints.3190v2.

²¹ 降水量の予測に使用されている例: Hossain, Md Monowar; Anwar, A. H. M. Faisal; Garg, Nikhil; Prakash, Mahesh; Bari, Mohammed (2022-06-17). *Monthly Rainfall Prediction at Catchment Level with the Facebook Prophet Model Using Observed and CMIP5 Decadal Data*. pp. 111. doi:10.3390/hydrology9060111.

²² Center, National Geophysical Data. “USDOC/NOAA/NESDIS/NCEI-Ionospheric Data US Total Electron Content (Vertical and Slant)” www.ngdc.noaa.gov. 2023 年 4 月 15 日最終閲覧。

²³ “CODE - Analysis Center”. Astronomical Institute (2017 年 8 月 24 日). 2023 年 4 月 15 日最終閲覧。

る。地震直前までのデータを入力していないのは、先行研究で最大地震 2 か月前に異常が見つかったことによる。

以上の手順で分析するファイルの数は 100 万を超えており膨大であるため、Python で自動化した。

2.2. データ分析の手順

1. 予測値と実測値のグラフを片方を透過させることにより重ね合わせ、地震発生時前後で異常な値が発生しているかどうかを確認する。
2. グラフと同時に出力される”Components”という年、週、日別の季節性分析グラフも合わせて分析し、Prophet が TEC の傾向を掴めているのかも見る。

なお、Prophet からの出力はデフォルトの一日刻みではなく、より詳細な一時間刻みで行っている。また、先行事例²⁴に基づき、TECu は最大値 200、最小値を 0 に設定している。

3. 結果・考察

以下の 4 地震に関して、それぞれ実測値およびそれに Prophet 出力値を重ねたグラフを作成した。なお、データは 2016 年 1 月 1 日からしか存在しないため、選出できる地震はこれ以降に限られる。以降、地震は発生地域名で呼称する。

3.1. 各地震のデータ

地震名	震源	発生時刻 (UTC)	Mw	地震種類	選出理由
2018 年アラスカ湾地震 ²⁵	56.046°N 149.073°W	2018-01-23	Mw 7.9	横ずれ	地震が 2016 年 2 月から 2018 年まで全くない
2019 年リッジレスト地震 ²⁶	35.766°N 117.605°W	A: 2019-07-04 B: 2019-07-05 C: 2019-07-06 D: 2020-06-04 (地震 D はデータの都合上今回分析していない)	A: Mw 6.4 B: Mw 5.4 C: Mw 7.1 D: Mw 5.5	横ずれ	2000 年代以来地震がみられておらず、また人口密集地に近いことより、地震の各情報が正確である可能性が高い
2020 年アイダホ中央地震 ²⁷	44.460°N 115.136°W	2020-03-31	Mw 6.5	横ずれ	比較的地震が少ない地域
2018 年アンカレッジ地震 ²⁸	61.346°N 149.955°W	2018-11-30	Mw 7.1	縦ずれ	観測値がある中のデータで次に Mw が大きい

²⁴ Kashcheyev, Anton. (2019). [Re: What was the highest value of total electron content \(TEC\) ever recorded?](#).

²⁵ [M 7.9 - 261 km SE of Chiniak, Alaska](#), earthquake.usgs.gov. 2023 年 4 月 16 日最終閲覧。

²⁶ M 6.4 - Ridgecrest Earthquake Sequence 以下 [A](#), [B](#), [C](#), [D](#), earthquake.usgs.gov. 2023 年 4 月 16 日最終閲覧。

²⁷ [M 6.5 - Stanley, Idaho](#), earthquake.usgs.gov. 2023 年 4 月 16 日最終閲覧。

²⁸ [M 7.1 - 1 km SE of Point MacKenzie, Alaska](#), earthquake.usgs.gov. 2023 年 4 月 16 日最終閲覧。

3.2. Components の分析

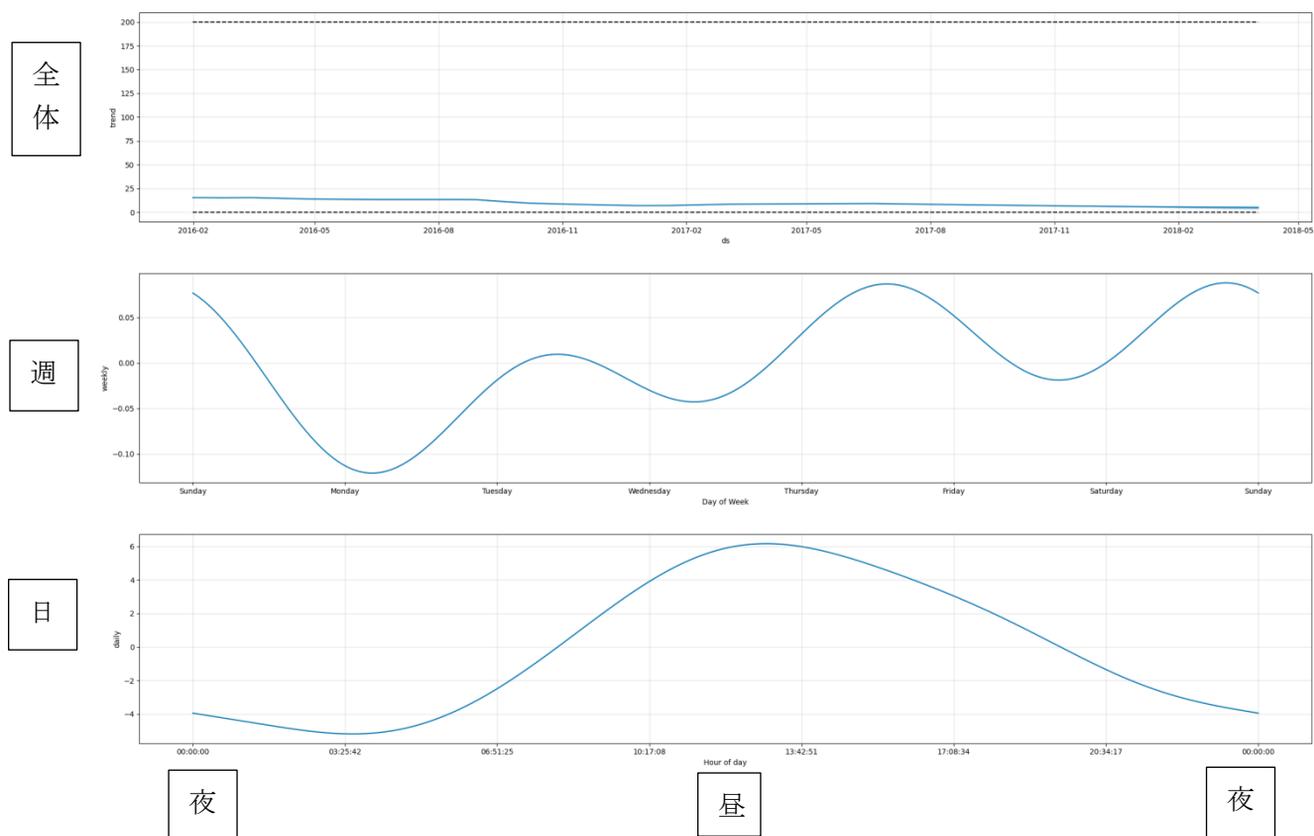


Fig1. アラスカの Components

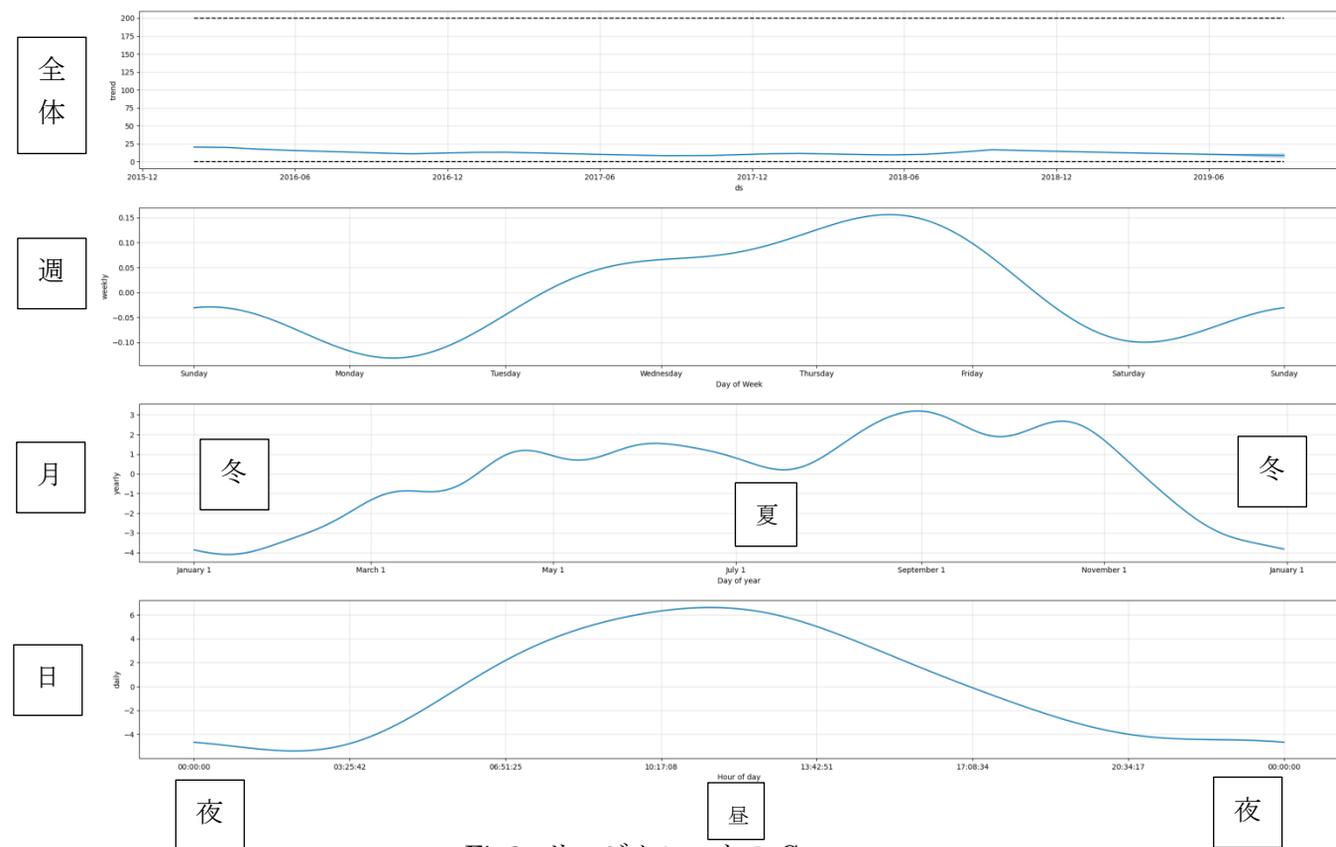


Fig2. リッジレストの Components

注: Fig.1 には月別の分析がなく, Fig.2 にはあるのは, データの入力期間が異なるためである。

まず, Components (Fig1, Fig2)の分析を 2 グラフ総合して行う。

1. 月別の季節性²⁹ (yearly): Fig2 のみ。日照時間が多い夏の方が冬より TEC の値が大きいというおおよその傾向をつかめている³⁰。
 2. 週別の季節性: 日単位のデータと TEC との関連性はないため, ここから得られる情報はない。
 3. 日別の季節性: 日中に大きい値を示し, 夜は逆に小さいという TEC の傾向をつかめている。
- 上記の分析により Prophet は外部からの干渉に影響されずに, TEC 値の予測を正確に行うことができていると考えられる。



Fig3. カリフォルニアの各月の積算日照時間を示したグラフ³¹

また, TEC と光との関係に着目し, 日照時間グラフとの比較を試みた。Fig3 と Fig2 の 月 の Components に類似性がないか考えたが, 関数を方程式で記述するなどして二つの関数の類似性を調べる技術がないため, 現状ではこれを正確に検証することはできていない。

3.3. Prophet 出力グラフの分析

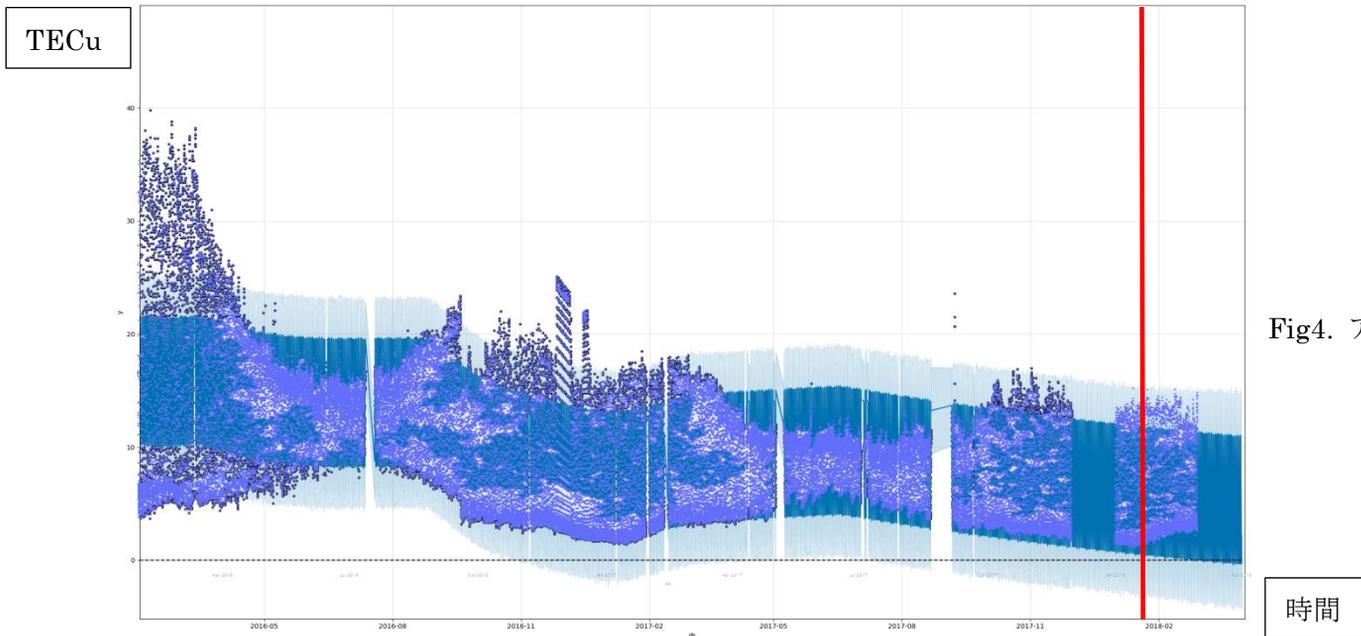


Fig4. アラスカ

²⁹ ここにおける「季節性」というのは季節における各々の性質を述べる言葉ではなく, 単なる周期的な変動を指している。

³⁰ TEC は光を受けてより電離するという事実を踏まえている。

³¹ “Average monthly hours of sunshine in Los Angeles (California), the United States of America”. World Weather & Climate Information.

2023 年 4 月 16 日時点のオリジナルより [アーカイブ](#)。2023 年 4 月 16 日最終閲覧。

TECu

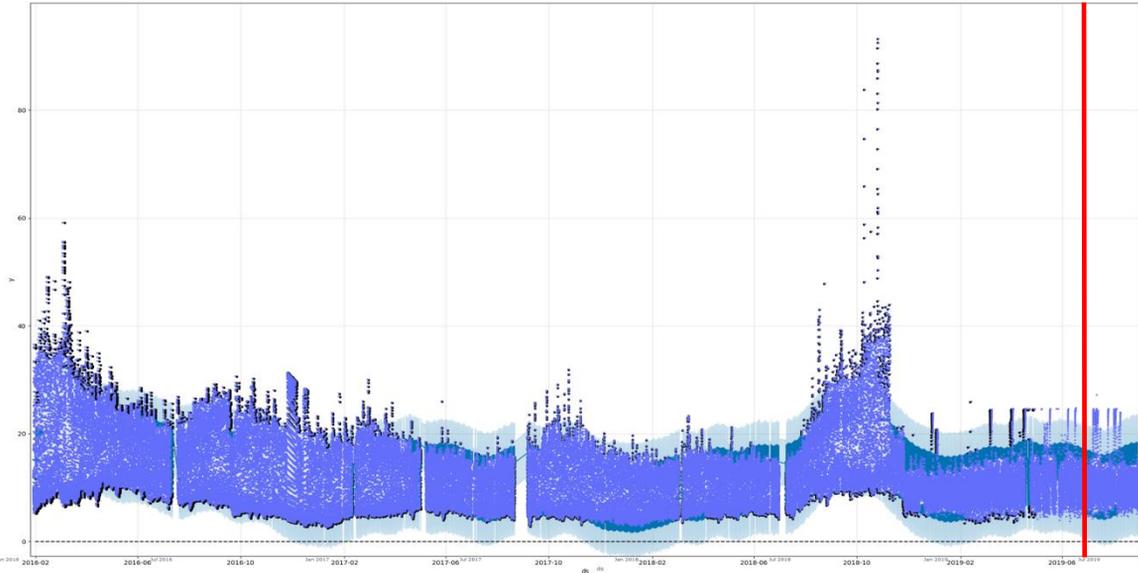


Fig5.
リッジレスト

時間

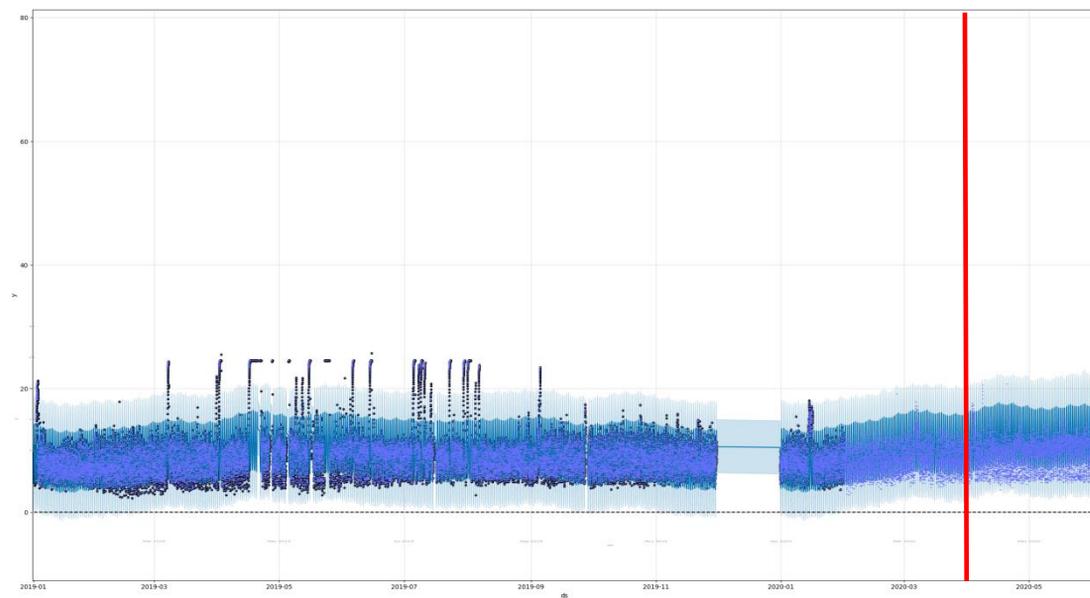


Fig6. アイダホ

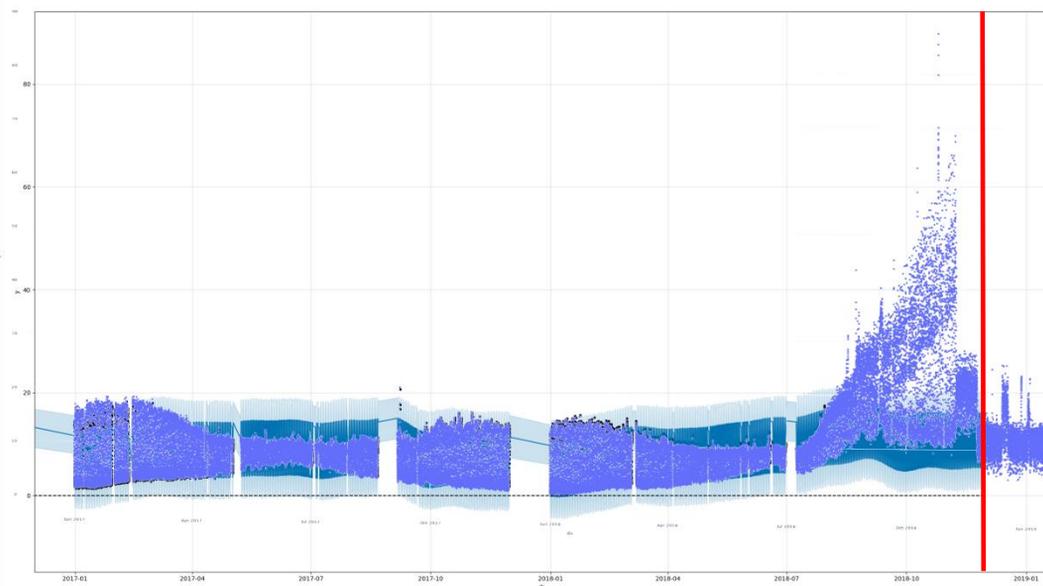


Fig7. アンカレッジ

次に、Prophet が出力した予測値に関して分析を行う。グラフ中の x 軸は時刻、y 軸の単位は TECu である。Prophet グラフが値を黒、実測値グラフが値を紫で表しており、色は異なるが、この二つの点は同じ値を示している。その他、赤線が地震発生日時、濃い青が Prophet の予測値、薄い青の範囲が Prophet の予測範囲（予測最大値～予測最小値）である。先述した通り、Prophet に入力しているデータは地震発生約 3 か月前までであり、それ以降の値は入力していない。もし地震発生前後の値が Prophet の予測範囲を超過していた場合、これを異常と判定する。

3.3.1. 分析の前に

以下の 3 点に注意して分析を行う。

1. すべてのグラフにはデータの欠損がみられるが、グラフには大きな局地的変動は見つからず、これにもある程度は対応できているとみられる。
2. 一部のグラフがもう一つのグラフの範囲を超過しているものがあるが、これは組み合わせた時期が異なることに起因するものであり、超過しているのは実測値のみであるから考察に支障はない。
3. グラフの縦軸の幅がそれぞれで若干異なっていることに留意する。

3.3.2 分析

以下、それぞれの地震について個別に分析を行う。

1. アラスカ: 観測初期~2016-05 や、2016-11~2017-02 の間などにおいて TEC が不明な理由で異常値を示しているが、地震発生時の 2018-01-23 周辺においては Prophet の予測誤差を超えた異常は見られない。これは先行研究と矛盾する結果である。
2. リッジレスト: 2018-07~2018-11 付近など、一部突発的な異常はあり、これは 2018 年のカリフォルニア山火事に影響されたことが可能性として考えられるが、2019-07-04~2019-07-06 の 3 地震直前には予想範囲内にとどまっており、異常値は見られない。一方、先行研究には記載がないことであるが、全体的、特に 2019-03~地震発生にかけて不自然なパルス状（突起状）の異常値が見られている。
3. アイダホ: 2020-03-31 の地震発生直前には全く異常値は見られない。しかし、リッジレストで見られたものと同じ傾向のパルス状異常が 2019-03~2019-10 にかけて見られる。
4. アンカレッジ: 地震発生の 2018-11-30 である 2018-10~2018-11 にかけて急峻な上昇が見られる。
なお、パルスについては私たちは当初は異常だと考えていなかったが、京都大学の梅野教授によると、これは地震の直前に増減する系統の TEC の異常とは別の系統の異常である可能性があるという。

4. 結論

本研究の結論は以下である。

1. Facebook Prophet は、TEC の正常値を予測することができる。
2. 地震発生前後に必ずしも TEC の上昇が見られるわけではない可能性がある。

4.1. 結論の分析

先行研究ではすでに証明されていることとは反する結論に至った原因を考察する。以下、考えられる理由を挙げる。

1. Facebook Prophet が出力する予測上限値と予測下限値との値の幅が大きすぎており、たとえ真実は異常であったとしても、異常値が予測値の間に収まることにより異常と認識できない状況である可能性があること。

2. グラフに散発的に見えるような大規模なデータの欠損があること。これについて私たちは発行機関に問い合わせを行ったが、データを復旧させることは不可能に近いという回答しか得られなかった。
3. 観測している衛星が震央より遠かったことが原因でデータが不正確になっている可能性があること。
4. 1つの衛星のみからのデータを利用したこと。ただしこれに関しては、私たちが探しているのはTECの異常値にすぎないと考えれば、複数検討する必要はなく、1つの衛星だけでも十分だとも考えられる。

5. 今後の展望

以上の結論を踏まえた今後の展望を述べる。

1. 「結論の分析」2~4に挙げた通り、データに不正確性があることが考えられるため、より正確なデータを得るため、私たちはTECデータを自分たちで構築しなければならないと考えた。このため、国土地理院が提供しているGNSS観測データ³²を利用し、GNSSデータをTECデータに変換することを現在試みている。このために、北海道大学の日置教授によるFortranプログラムを応用することを模索している。
2. リッジレスト・アイダホ地震の分析で挙げたパルス状の異常値について、より多くの地震の発生時のTECを分析することにより類似した現象がないか検証する。
3. Componentsの項で挙げた通り、日照時間とTECの関係についてもどこまで相関が見られるか統計的に調査を行う。

6. 謝辞

本論文の執筆にあたっては、京都大学情報学研究科の梅野健教授および梅野研究室の皆様にも専門的な助言を頂いた。また、柏木創太先輩は私たちの研究の原動力となったとともに、多くの指導を頂いた。なにより、課題研究の進行方法や研究の途中で発生した困難への対処法等について、西川昌利、西田宙起両教諭より、貴重なご指導とご助言を賜った。教示頂いたすべての方にこの場を借りて感謝申し上げ、謝辞とする。

³² “[電子基準点 | 国土地理院](https://www.gsi.go.jp)” . www.gsi.go.jp. 2023年4月15日最終閲覧。

ジビエの利用量を増やすために

(How to maximize the amount of gibier used.)

Abstract: 令和3年度のイノシシ・シカの日本全国の捕獲頭数は合わせて約125万頭に上るが、そのほとんどが廃棄されている。その現状を知り、我々は、SDGsの観点から廃棄量を減らし、ジビエとしての利用を進めるべきだと考えた。この課題を解決するため、「ジビエ処理施設が増えればジビエ利用率も上がる」という仮説を立てて都道府県別に研究を行った。その結果、ジビエ処理施設数とジビエ利用率の間にはやや強い正の相関がみられ、ジビエ利用量を増やすためにはまずジビエ処理施設を増やし、イノシシ・シカを処理できる能力を上げていく必要があるということが分かった。

Keywords: #ジビエ利用 #イノシシ・シカの廃棄 #SDGs #ジビエ処理施設

1. 研究背景

令和3年度の野生鳥獣による全国の農作物被害額は約155億円である。その対策として、イノシシ・シカの捕獲が進められ、同年度の捕獲数は合わせて約125万頭に上る。その甲斐あってか、野生鳥獣による全国の農作物被害額は年々減少傾向にある。しかし、その反面で、捕獲されたイノシシ・シカのほとんどが廃棄されている。そこで我々は、SDGsの観点から廃棄量を減らす必要があると考え、ジビエとしての利用を進めるべきだと考えた。まず、現状を知るために、各都道府県において捕獲されたイノシシ・シカに対するジビエ利用率と捕獲頭数との関係を調べたところ、相関係数は0.22127であり、弱い正の相関関係が見られたことから、現在ではジビエとしてあまり利用されていないということが分かった。ゆえに、ジビエ利用率を上げることができれば、廃棄数も減らすことができると考え研究を進めた。

2. 研究目的・意義

イノシシ・シカの廃棄量を減らし、ジビエ利用量を増やすことを目的としている。また、その目的は、SDGsの取り組みとして『2. 飢餓をゼロに』『12. つくる責任・つかう責任』『15. 陸の豊かさも守ろう』の3点から、有効である。

3. 研究方法

主に e-Stat を用いて、農林水産省 調査統計室が提供しているデータをもとに研究を進めた。

また、より正確な相関係数を求めるため、外れ値を「あるデータの第一四分位数～第三四分位数の範囲に収まっていない値」と定義し、箱ひげ図の性質を利用して外れ値を除外した。

ただし、相関関係の基準については、東京大学情報基盤センターの『相関係数の解釈』を基に定め、XとYの間の相関係数を r とし、以下の表 1 のように定義している。

$ r = 0.0$	X と Y の間に相関がない
$0.0 < r \leq 0.2$	X と Y の間にほとんど相関がない
$0.2 < r \leq 0.4$	X と Y の間に弱い相関がある
$0.4 < r \leq 0.7$	X と Y の間にやや強い相関がある
$0.7 < r \leq 1.0$	X と Y の間に強い相関がある

表 1：相関係数と相関関係との関係

4. 結果・考察

まず、ジビエ利用量と関係があるものは何かを調べるため、目的変数をジビエ利用量とし、「都道府県ごとの狩猟免許交付件数」、「処理施設数」、「ジビエの処理・販売によって得られた金額」の 3 項目について重回帰分析を行った。結果は表 2 の通りであり、「施設数」と「ジビエの処理・販売によって得られた金額」で P 値が 0.05 以下になった。

また、「野生鳥獣の解体やジビエ肉等の販売などによって処理施設が得た収益」と「ジビエ利用量」との相関係数は 0.63 であり、やや強い正の相関があることが分かった。散布図は図 1 の通りである。

	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	-45.7194	10.30773	-4.43544	6.27E-05	-66.5069	-24.9319	-66.5069	-24.9319
免許交付	14.76367	9.89633	1.491833	0.143044	-5.19418	34.72152	-5.19418	34.72152
処理施設	3.90687	0.516647	7.561976	1.99E-09	2.864953	4.948788	2.864953	4.948788
得た金額	0.058229	0.017561	3.315748	0.001864	0.022813	0.093644	0.022813	0.093644

表 2：「ジビエ利用量」を目的変数とした重回帰分析の結果

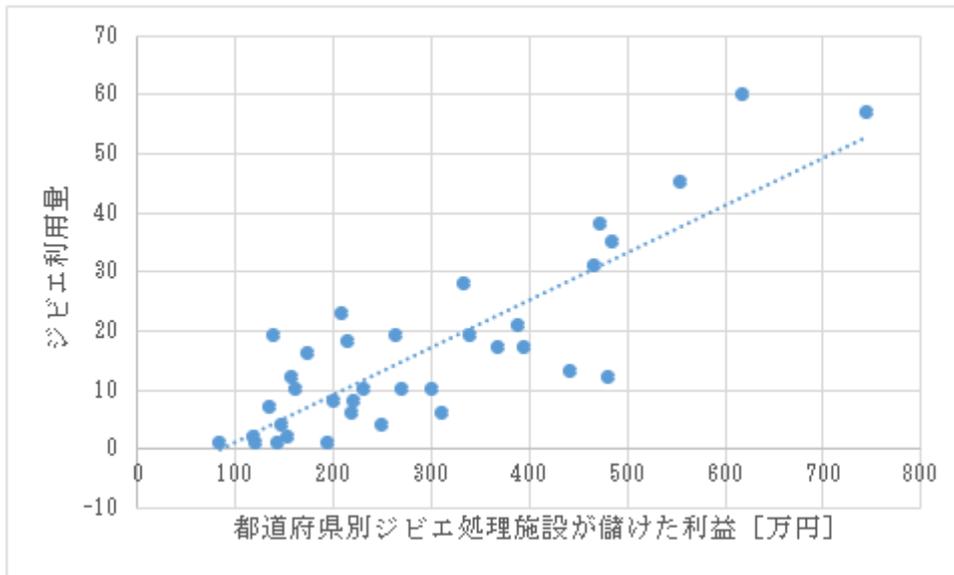


図1：「処理施設が得た収益」と「ジビエ利用量」の関係

次にジビエ利用量とジビエ処理施設数の関係を調べた。

ここで、都道府県によって捕獲頭数に差があり、単にジビエ利用量と施設数を比較することは不可能であると考えたため、独自に「ジビエ利用率^{*1}」を作り、利用した。

〔1〕 ジビエ利用率とジビエ処理施設数の相関係数を求めたところ、0.5399 という値が得られ、やや強い相関関係があるということが認められた。

〔2〕 現在のジビエの処理能力を知るためにジビエ利用率^{*1}が最も高い北海道について調べた。北海道ではジビエ処理施設が合計で85軒あり、その中でもシカの解体頭数規模別施設数^{*2}は、以下の表2の通りになっている。

1施設当たりのシカの処理頭数	該当する施設数
50頭以下	29軒
51頭以上100頭以下	11軒
101頭以上300頭以下	17軒
301頭以上500頭以下	13軒
501頭以上1000頭以下	7軒
1001頭以上1500頭以下	2軒
1501頭以上	6軒

表3：シカの処理頭数別ジビエ処理施設数

表3よりシカの処理頭数を求めたところ、最大で約33,000頭を処理していることがわかるが、令和2年度には126,692頭捕獲されていることから、現状ではジビエ処理施設は不足していると考えられる。

〔1〕〔2〕より、ジビエ利用量を増やすためには、まずジビエ処理施設を増やしてジビエを処理できる能力を十分に上げることが必要だと考えられる。

さらに、どんな施設を立てることで一番ジビエ利用量を増やすことができるのかを調べるため、さまざまなデータとの相関を調べた。

〔3〕各都道府県における面積別施設数の割合*³とジビエ利用量との間の相関関係を調べると、ジビエ処理施設の面積が 50 m²以下、51～99 m²、100 m²以上の 3 つの項目とジビエ利用量の相関関係はそれぞれ 0.2296、0.1965、0.1943 であり、どれも弱い正の相関関係もしくはほとんど相関がみられなかった。

〔4〕年稼働日数別の施設数の割合とジビエ利用量との相関関係を調べたところ、50 日以下、51～100 日、101～200 日、201 日以上の 4 つの項目において相関係数は-0.13、0.084、0.094、0.17 であり、どれもほとんど相関がなかった。全国の年稼働日数別施設数の割合は図 1 のとおりであり、200 日以下が 68%以上をしめている。これは、地域や動物によって異なるが、狩猟期間がおよそ 11 月 15 日～2 月 15 日の 3 カ月間に決められていることに関係があると考えられる。そこで、ジビエ利用量と通年処理の施設数、狩猟期間のみの施設数の二項目でジビエ利用量を調べたが、相関係数はそれぞれ 0.19、-0.083 であり、どちらも相関関係はなかった。

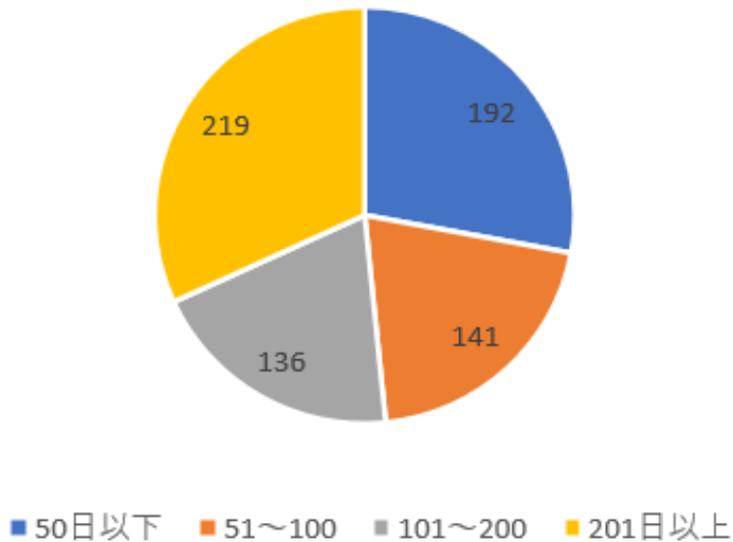


図 2：年間働日数別施設数 [軒]

〔5〕従業員数とジビエ利用量の相関を調べると、1 施設当たりの従業員数が 2 人以下、3～5 人、6～10 人の 3 つの項目において、相関係数はそれぞれ-0.24、0.27、0.027 であった。2 人以下の施設では弱い負の相関が、3～5 人の施設では弱い正の相関関係がみられたが、6～10 人の施設では相関関係はほとんど見られなかったことから、ジビエ利用量を増やす上で従業員を増やすことはあまり有効な手段ではないと考えられる。また、11 人以上の施設があるのは 2 県しかなかったため正確なデータは得られないと判断し、相関関係は調べなかった。散布図は次の図 3～5 の通りである。

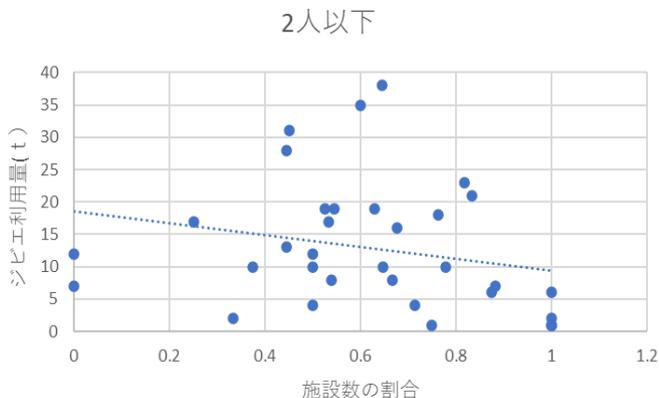


図 3：「従業員が 2 人以下の施設数の割合」と「ジビエ利用量」の関係

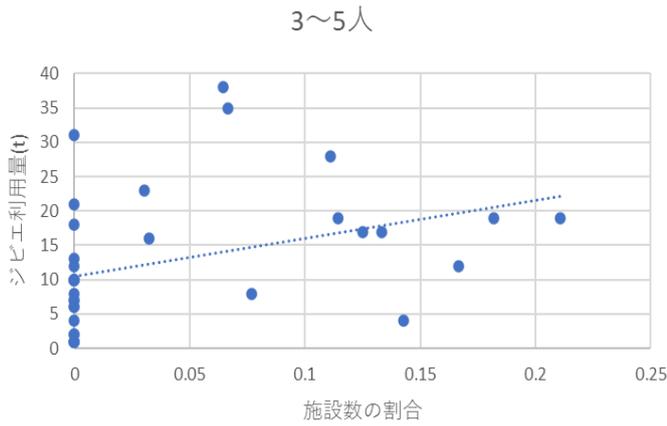


図4：「従業員が3人以上5人以下の施設数の割合」と「ジビエ利用量」の関係

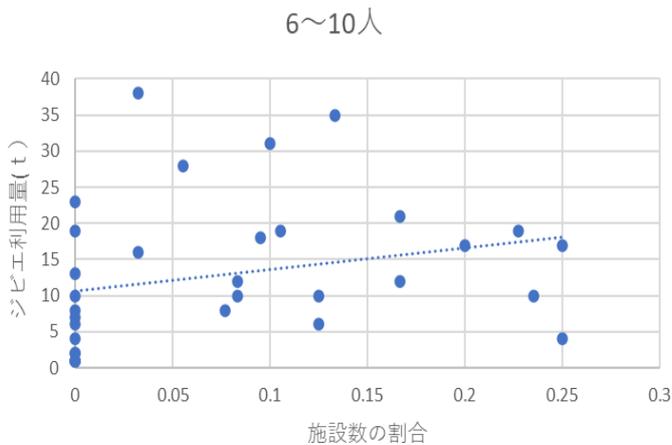


図5：「従業員が6人以上10人以下の施設数の割合」と「ジビエ利用量」の関係

〔6〕 全国でジビエ処理施設が得た収益のうち、88%以上を食肉が占め、食肉以外ではペットフードが90%以上を占めている現在、食用ジビエ肉としての利用は盛んであることから、今後はペットフードとしてジビエを利用することで、ジビエの利用量と消費量を増やすことができるのではないかと考えられる。

〔3〕～〔6〕から、ジビエを利用したペットフードの生産への取り組みが必要であることが分かった。

*1 ジビエ利用率＝イノシシ・シカのジビエ利用量 [t] ÷ イノシシ・シカ捕獲頭数 [頭] ×100

*2 イノシシは北海道には生息していないため、今回はシカの処理施設数のみを調べている。

*3 面積別施設数の割合＝面積ごとの施設数÷処理施設数の合計

*1 の式は今回、我々が独自に作り出した式であり、「×100」をしたのは値を見やすくするためである。そのため、ジビエ利用率自体は正確な値ではないが、都道府県で比較する際の割合は正しいと考えられる。また、ジビエ利用量と捕獲頭数で単位が異なるのは、ジビエ利用量の頭数のデータを得られなかったためである。しかしながら、この式は都道府県単位でほかのデータとの相関関係を求めるために作成したものであり、都道府県別で比較した際の割合は正しいと考えられる。

5. 結論及び今後の展望

この研究からは、ジビエ処理施設を増やし、ペットフードのジビエ利用を増やすことで、ジビエの利用量は増え、イノシシ・シカの廃棄量も減らすことができるということが分かった。

今回、海外のジビエ事情に関するデータ等が少なく、日本国内に限定して研究を行ったが、データは不十分であったと感じている。そこで、今後の展望として、フランスではジビエ料理が高級料理となっており、人気も高いと考えられることから、フランスがどのようにしてジビエを利用しているのかを調べたいと考えている。

また、イノシシ・シカは、日本ではその独特な獣臭さによって食べることを敬遠されがちだが、ジビエ肉は栄養満点であるので、今後は可能であればイノシシ・シカをおいしく調理する方法についても研究したいと考えている。

参考文献・引用文献

農林水産省（2020）「野生鳥獣による被害額の推移（鳥獣種類別）」

https://www.maff.go.jp/j/seisan/tyozyu/higai/hogai_zyoukyou/attach/pdf/index-20.pdf

農林水産省（2020）二ホンジカ・イノシシ捕獲頭数

<https://www.env.go.jp/nature/choju/docs/docs4/sokuhou.pdf>

農林水産省（2020）「野生鳥獣のジビエを巡る最近の状況」

https://www.maff.go.jp/j/seisan/tyozyu/higai/hogai_zyoukyou/index.html

農林水産省（2020）「野生鳥獣利用実態調査」

<https://www.e-stat.go.jp/stat->

[search/files?page=1&layout=datalist&toukei=00500248&tstat=000001112115&cycle=8&tclass1=000001113551&tclass2=000001160526&cycle_facet=tclass1%3Atclass2&tclass3val=0](https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00500248&tstat=000001112115&cycle=8&tclass1=000001113551&tclass2=000001160526&cycle_facet=tclass1%3Atclass2&tclass3val=0)

同一系外惑星系内に存在する惑星の種類とその傾向について

1, 研究背景

1992年に初めての系外惑星が発見されて以降、観測技術などの発展を受けて、多くの系外惑星が発見され、現在では5000個以上の多種多様な系外惑星が発見されている。近年、系外惑星分野の研究も非常に活発である。先行研究より、惑星軌道は主星の質量と関係している点(星野, 2022)やスーパーアースを含む惑星系は惑星保有数が多い点(平野, 2017)、ホットジュピターを含む惑星系は惑星保有数が少ない点(Jonathan, 2021)などの研究がなされている。先行研究の調査を行う中でガス惑星についてはどうなのか、ガス惑星を含む惑星系に存在する惑星数とその種類について、軌道長半径や質量などの惑星の特徴と何らかの関係があるのだろうかという点に注目して研究を行った。また、先の3つの研究にならい、惑星系単位での分類に着目して研究に着手した。

2, 研究目的・意義

本研究では、判明している情報から惑星が存在しているのであるならばどのような惑星が存在しているかについての判断を促し、新たな系外惑星発見の手掛かりになることを目指す。惑星の種類には大きく分けて岩石惑星、ガス惑星、ホットジュピターの3種類がある。研究を始める前に惑星の種類とその惑星系に含まれる惑星の数について事前調査を行った。その結果が次の表1である。

表1 惑星の種類と惑星系に含まれる惑星の数やその割合について

	岩石惑星	ガス惑星	ホットジュピター
合計(個)	3315	783	798
1個のみ含まれる	51%	67%	81%
2個以上の惑星が含まれる	48%	33%	19%

ここで先行研究(Jonathan, 2021)より分かっているホットジュピターを含む惑星系にはそのホットジュピターしか存在しないことが多いのに次いでガス惑星はその惑星単体で存在することが多い。私たちはこの原因としてガス惑星が岩石惑星などを吸収し合体することによる惑星数が減少したと考察した。これらのことから一つ目の仮説「岩石惑星とガス惑星が同一系内に同時に存在する系は少ない」を立てた。また吸収される岩石惑星はガス

惑星の重力の影響を受けたうえで合体すると考察したので、二つ目の仮説「岩石惑星とガス惑星が同一系内に存在するときその惑星系に含まれる岩石惑星の質量は小さくなっている」を立てた。

3, 研究方法

2022年10月13日時点でNASA Exoplanet Archiveに掲載されている5,187個の系外惑星のデータを使用した。岩石惑星、ガス惑星、ホットジュピターのおおよその性質は次のようである。岩石惑星は主に岩石でできており、比較的主星に近いところを公転しており星自体の大きさは小さい。密度は大きい大きさが小さいためガスを主成分とした惑星に比べて質量は小さい。ガス惑星は主に水素やヘリウムを主成分としており半径は大きい。主星から遠いところを公転しており、密度は小さいが体積が大きいため岩石惑星に比べて質量は大きい。ホットジュピターは密度、体積、質量、組成についてはガス惑星と同様だが主星に近い部分で公転している。事前準備としてK-Means法を用いてその系外惑星の質量、軌道長半径について2次元クラスタリングを行い、系外惑星のデータを岩石惑星、ガス惑星、ホットジュピターの3種類に分類した。その結果が図1である。

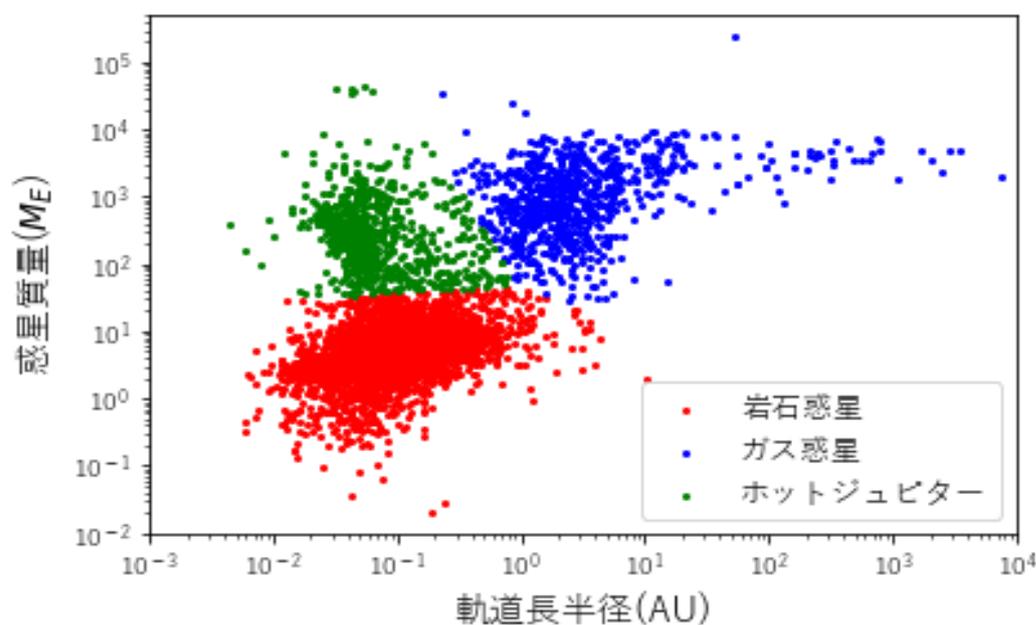


図1：系外惑星の軌道長半径-惑星質量についての散布図

クラスタリング結果と岩石惑星、ガス惑星、ホットジュピターの性質を考慮すると、赤が岩石惑星、緑がホットジュピター、青がガス惑星であると考えられる。以下本研究内で作成する散布図では岩石惑星を赤、ガス惑星を青、ホットジュピターを緑の点で表現する散布図では一つ目の仮説「岩石惑星とガス惑星が同一系内に同時に存在する系は少ない」については取得したデータから現在二つ以上の惑星を持つ惑星系を抽出したうえでその惑星系について岩石惑星とガス惑星が同時に含まれているものの個数と割合を調査した。また、二つ目

の仮説である「岩石惑星とガス惑星が同一系内に存在するときその惑星系に含まれる岩石惑星の質量は小さくなっている」については岩石惑星とガス惑星が同一系内に存在する惑星系に含まれる岩石惑星のみを抽出したうえでそれらについて惑星系に含まれる惑星数毎に惑星質量の箱ひげ図を描画した。また岩石惑星が単体で存在しているものについても同様に箱ひげ図を描画しこれらを比較することによって全体の質量の傾向を調べた。

4 結果・考察

仮説 1 について、調査結果を表に著したところ下の表 2 のようになった。

表 2: ガス惑星を含む惑星系についてその数と岩石惑星と共存する系の個数とその割合

惑星数	2	3	4+
惑星系数 (個)	517	157	72
共存する系の数 (個)	28	3	11
割合	5.4%	1.9%	15.3%

表 2 より正しく岩石惑星とガス惑星が共存するような惑星系の割合は少ないことが分かった。次に仮説 2 について、ガス惑星を含む惑星系に含まれる惑星のみを抽出して軌道長半径と惑星質量について散布図を描画したところ、図 2 のようになった。これに対して、すべての惑星について同様に描画した散布図は図 3 のようになった。この二つの図を比較したところ惑星質量が $0.1 \sim 1.0 M_E$ の岩石惑星が図 2 に存在しないことが分かる。また岩石惑星の質量について箱ひげ図を描画したところ図 4 のようになった。この図からガス惑星が含まれる惑星系に含まれる岩石惑星の質量には明らかに下限が存在しており、質量の平均も大きくなっていることからガス惑星と同じ惑星系に含まれるとき、岩石惑星の質量は大きくなる傾向がある。

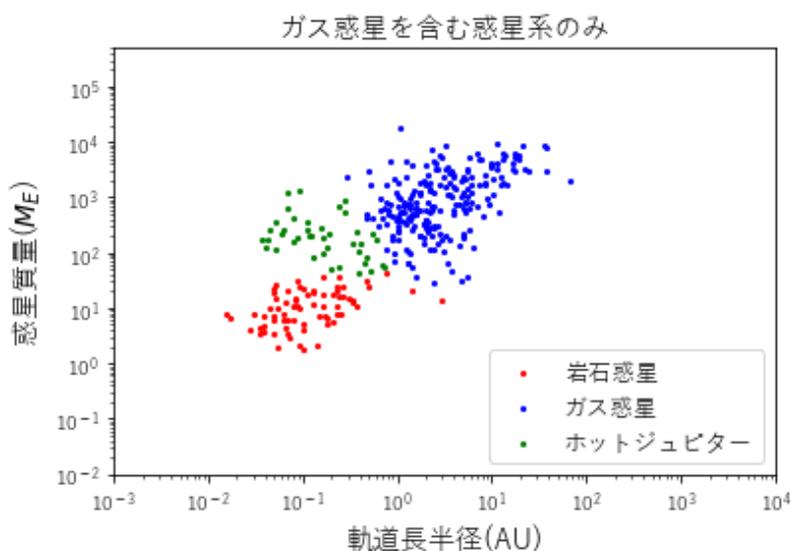


図 2: ガス惑星が含まれる惑星系に存在する惑星について軌道長半径-惑星質量についての散布図

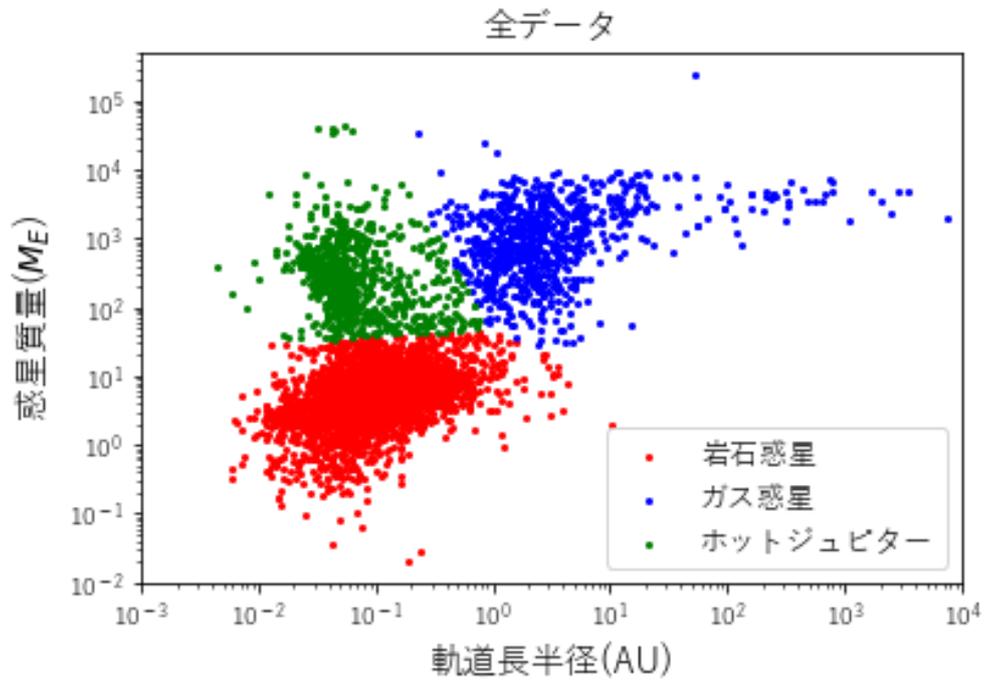


図 3:すべての惑星について軌道長半径-惑星質量についての散布図

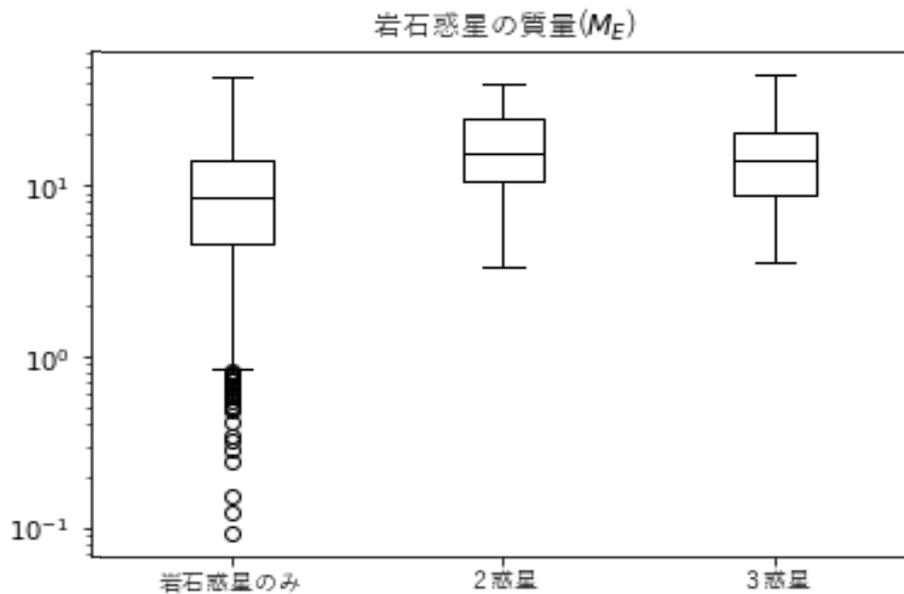


図 4:岩石惑星のみ存在する惑星系の惑星とガス惑星を含み惑星系に含まれる岩石惑星について惑星質量の箱ひげ図

十分な質量をもたない岩石惑星はガス惑星の重力的な影響によって吸収されてしまうことで淘汰されてしまうからではないかと考察した。グラフより完成後の質量が $10^{0.3}$ 地球質量がガス惑星と岩石惑星が共存することができるかどうかのボーダーであるといえる。

5. 結論

岩石惑星とガス惑星が同一系内に存在する割合はかなり少ないことが分かった。また、これらが同一系内に含まれるとき、岩石惑星の質量は大きくなる傾向があることが分かった。これはガス惑星が同じ惑星系に含まれているとき、質量の小さすぎる岩石惑星はガス惑星に吸収されることによって淘汰されたのではないかと考察した。しかし、太陽系には木星や土星などガス惑星が含まれている。にもかかわらず質量 $1.0M_E$ の地球が存在している。これについて私たちは2つの仮説を立てた。一つ目は地球を含む太陽系の軌道が特徴的なものでありその結果ガス惑星と岩石惑星が同時に存在しているというもの。二つ目は観測技術の限界によって遠くのガス惑星を含む惑星系において、質量の小さな岩石惑星を発見できていないだけであるということである。どちらの仮説が正しいのかについては判断することができず、現在のままではこれがこの研究の限界であると考え。今後の展望については、現在スーパーアースなどの異なった種類の惑星についての考察ができていないこと、そしてクラスタリングにより分類した境界付近において正確性がやや欠如しているように感じられる。これを改善するために、岩石惑星の内でも比較的質量の大きなスーパーアースを新たな分類として加えることや、現在の惑星質量、軌道長半径に加えて密度を用いることによる3次元クラスタリングによりさらに分類を絞り込むことで研究の精度を高めることができると考えている。下の図5は惑星質量、軌道長半径に加えて密度を加えた3次元クラスタリングの結果である。また、下の図6は密度を加えた3次元クラスタリングを行うときにクラスタ数を1つ増やしてスーパーアースの分類を増やした散布図である。

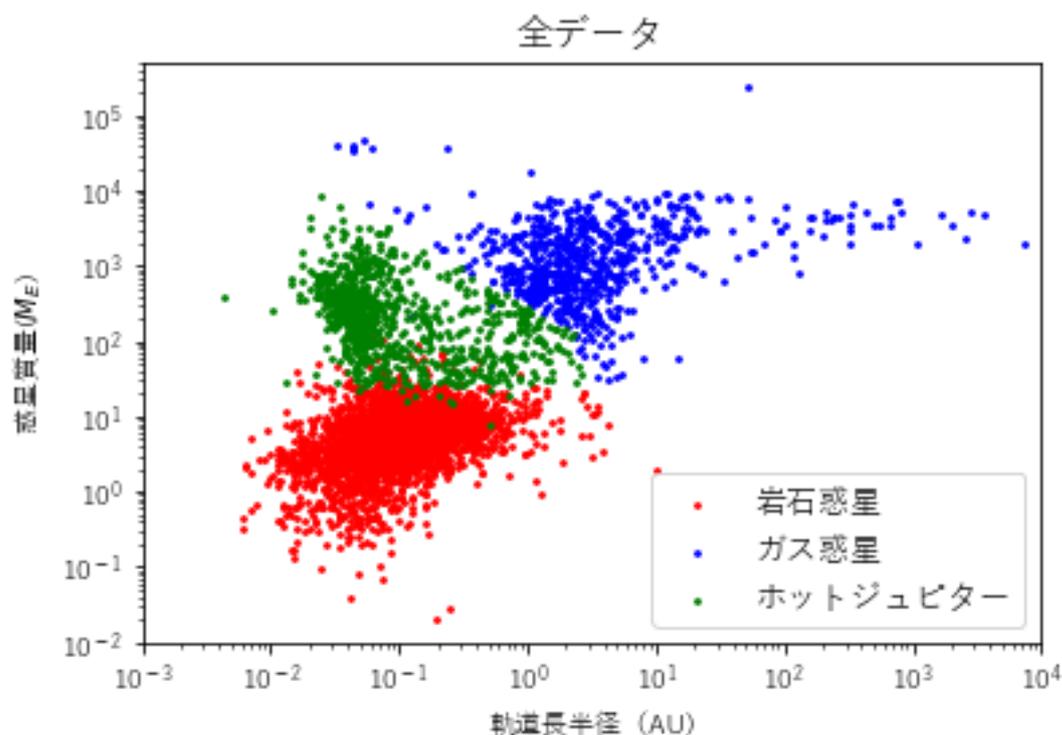


図 5:3次元クラスタリングによるクラスタリング結果

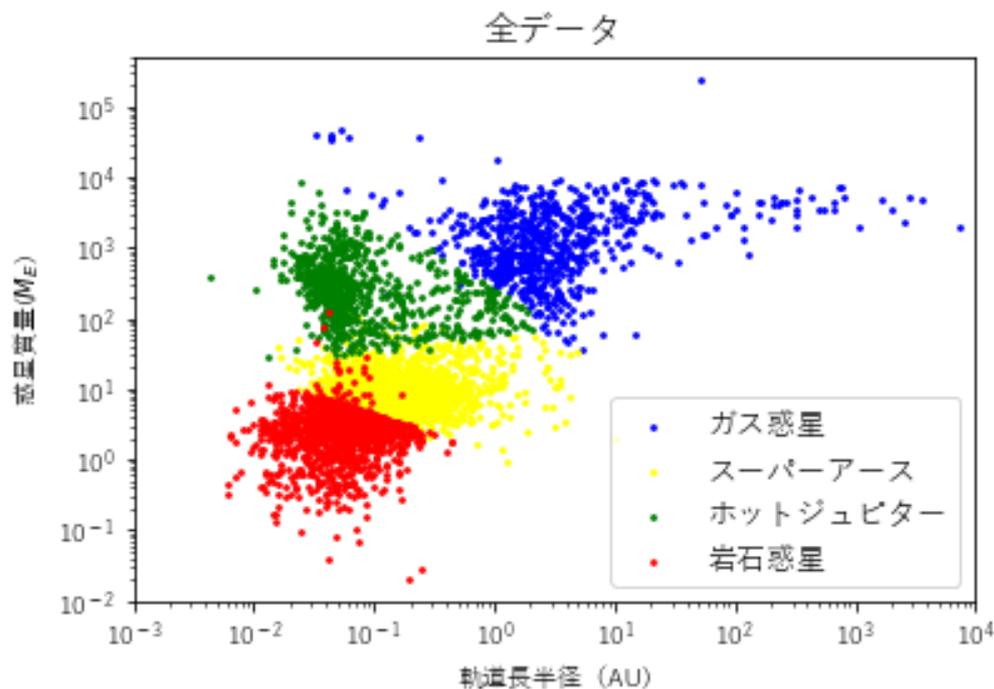


図 6 : 図 3 の分類にスーパーアースを付け足した散布図

この結果ではこれまでと分類が変化した惑星が複数存在しておりさらに私たちの研究を確かなものにできると考えている。この新たな分類を用いてさらに研究を進めていきたい。

6. 参考文献

NExSci (2022). NASA Exoplanet Archive.

<https://exoplanetarchive.ipac.caltech.edu/>. 2022/10/13

Schulze-Makuch, D. (2011) “A Two-Tiered Approach to Assess the Habitability of Exoplanets. *Astrobiology*” 11(10): 1041-1052

成田憲保 (2020). 『地球は特別な惑星か?』. 講談社.

平野照幸 (2017). 『太陽系外惑星探査：見えてきた多様性とその起源』. 『日本物理学会誌』, 72(2), 105-111

成田憲保 (2013). 『観測による巨大惑星の軌道進化理論の検証』. 『日本惑星科学会誌』, 29(4), 242-251

井田茂・渡部潤一・佐々木晶 (2021) 『太陽系と惑星 [第2版] シリーズ現代の天文学』

星野遥・小久保栄一郎 (2022). “Orbital structure of planetary systems formed by giant impacts: stellar mass dependence”

<https://academic.oup.com/mnras/advance-article-abstract/doi/10.1093/mnras/stac3756/6960582>

Jonathan J. Fortney, Rebekah I. Dawson, Thaddeus

D. Komacek. (2021). “Hot Jupiters: Origins, Structure, Atmospheres”

<https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020JE006629>

Jonathan J. Fortney, and Christophe Sotin. (2013). “Structure of exoplanets”

<https://www.pnas.org/doi/10.1073/pnas.1304206111>

<http://www.exoplanetkyoto.org/about/>

令和5年度 探究活動 受賞一覧

(3年生)

- ISLP (International Statistical Literacy Project) 国際統計ポスターコンペティション
2022-2023 日本予選会 高校生部門 (主催: ISLP 国際統計リテラシープロジェクト)
優勝 「Hyogo Reconstruction Project」
→ 国際大会にポスター出品
- 統計データ分析コンペティション 2023
(主催: 総務省統計局等、後援: 文部科学省、国立研究開発法人科学技術振興機構等)
統計活用奨励賞 「独自指標作成による地方創生の方法論と兵庫県活性化の提案」
- 第71回統計グラフ全国コンクール
(主催: 公益財団法人 統計情報研究開発センター、後援: 文部科学省等)
入賞 「3球目攻撃の極意」
- 令和5年度兵庫県統計グラフコンクール パソコン統計グラフの部
兼 第71回統計グラフ全国コンクール兵庫県予選 (主催: 兵庫県)
特選 「3球目攻撃の極意」
入賞 「学級人数から見る中学校不登校率の考察」
佳作 「パッケージの視覚的影響と商品売り上げの関係性」
- 第18回 「科学の芽」賞 (主催: 筑波大学、後援: 文部科学省等)
奨励賞 「電離層総電子数 (TEC) データを用いた オープンソースモデリングアルゴリズム
の地震予知への有効性に関する考察」
- 第14回 坊ちゃん科学賞 研究論文コンテスト 高校部門
(主催: 東京理科大学、後援: 国立研究開発法人科学技術振興機構等)
入賞 「ジビエの利用量を増やすために」
佳作 「姫路観光活性化」
「はちみつを用いた風邪薬の作成」
奨励賞 「色彩とアスペクト比から見る企業のロゴデザインの有効性の考察」
「同一系外惑星系内に存在する惑星の種類とその傾向について」
- 第九回高校生国際シンポジウム スライド発表 社会科学・社会課題分野
(主催: 一般社団法人 Glocal Academy、後援: 文部科学省等)
最優秀賞 「打倒東京一極集中! 京阪神大都市圏活性化プロジェクト」

(2年生)

- 第7回和歌山県データ利活用コンペティション 高校生部門 (主催：和歌山県)
NTT 西日本賞 「アフターコロナに向けた兵庫県観光振興のための施策の提案」
- 高校生ポスター研究発表会 (主催：日本動物学会近畿支部)
最優秀賞 「ミミズ団子の核心に迫る」
- 【人工知能学会・日本統計学会公式】全国中高生 AI・DS 探究コンペティション 2023
(主催：JDSSP 高等学校データサイエンス教育研究会、協賛：人工知能学会、日本統計学会)
日本統計学会特別賞 「関西弁を全国へ」
- 第九回高校生国際シンポジウム スライド発表 人文科学分野 スライド発表
(主催：一般社団法人 Glocal Academy、後援：文部科学省等)
優良賞 「単語の分散表現と主成分分析を用いた日本語形容詞の意味変化の傾向の解明」
- データサイエンスコンテスト (主催：本校)
優勝

(1年生)

- 統計データ分析コンペティション 2023
(主催：総務省統計局等、後援：文部科学省、国立研究開発法人科学技術振興機構等)
審査員奨励賞 「海水温からの降水量予測を目指して」
「兵庫県の宝『いかなご』を守る」

令和6年3月 発行

兵庫県立姫路西高等学校

〒670-0877 兵庫県姫路市北八代2-1-33

TEL : 079-281-6621

FAX : 079-281-6623

<http://www.hyogo-c.ed.jp/~himenisi-hs/>

