

高等学校

探究的アプローチで学ぶ データサイエンス

【情報II×探究】GISと重回帰分析を活用した授業案

本授業案の一部もしくは全部を無許可で転載、複製、複写、転用することを禁止します。

本授業案について



西出新也

関西学院千里国際中等部・高等部(技術科・情報科主任)。業務と並行し、神戸大学大学院 博士後期課程にてHCI(ヒューマンコンピュータインタラクション)分野の研究に従事する。

ごあいさつ

情報Iと情報IIにおけるデータ活用の学習内容のレベル差について、授業現場で課題感を持っておりました。そのギャップを自然に埋め、学習者が自らデータを探し、分析したくなるような題材として、本授業案『探究的アプローチで学ぶデータサイエンス』を作成いたしました。

本授業案では、GISを用いてデータを可視化し、日経電子版でのディスカッションを通じて社会課題の発見と仮説構築を促します。その後、重回帰分析を用いて検証・考察を行い、データサイエンスの基礎と応用のスキルを実践的に身につけることを目指しています。

DXハイスクールでの情報IIの開設を見据え、データサイエンスを探究的に学ぶ本授業案が先生方のお役に立てば幸いです。

本授業案について

本授業案は、高等学校DX加速化推進事業(DXハイスクール)で求められる「情報II等」の設置の取り組みとして、情報IIと総合的な探究を接続する授業実践の内容とポイントをまとめています。

本授業案は全4章から成り、各章は章のまとめと、それぞれの章で行った授業展開で構成されています。本書では本校で実施した全36時間の授業をまとめていますが、各章や授業展開の一部を切り出して使えるように工夫しています。各校の授業計画に応じて必要な部分を参考にしてください。

各章のまとめについて

各章の冒頭には、下図のような章のまとめがあります。

- 1 章の目標を記載しています。学習目標に加え、各章で身につけたいスキルや能力などを記載しています。
- 2 各授業展開の概要、目標、学習活動、評価、留意点をまとめています。予めこの部分を確認して、参考にしたい授業展開を選ぶことができます。
- 3 章における授業展開のポイントを記載しています。各授業展開の中にもポイントをまとめていますが、それぞれの章を実施するにあたって特に重要なポイントを記載しています。

各授業展開について

各章は3つから4つの授業展開で構成されています。授業展開は下図のようにまとめられています。

- 1 各節の目標、評価規準、授業の準備について記載をしています。
- 2 各節の授業の組み立てについて記載しています。導入、展開、まとめそれぞれの概要と時間の目安を記載しています。
- 3 具体的な学習活動についてまとめています。ここに記載されている学習活動のいくつかを取り出してアレンジして実施することや、いくつかの学習活動を省略して授業を行うこともできます。
- 4 各授業展開のポイントや留意点についてまとめています。実施する学習活動に応じて参考にしてください。

1 章 データへの関心

本節の目標は、データや事実に基づいて世界を正しく理解し、データサイエンスの基本的な考え方や手法を体験的に学ぶことである。Gapminderを活用して、教育、貧困、環境など幅広い分野のデータを分析し、主観的な思い込みと事実との乖離を理解する。さらに、Pythonを用いて言語データを処理し、テキストマイニング技術を習得する。これにより、データ分析が社会の意思決定や問題解決にどのように活用されるかを考察する能力を養う。

1-1 Gapminder	1-2 テキストマイニング	1-3 ネガボジ判定
時間: 50分 概要: Pythonを活用し、データや事実に基づいて世界を理解する。教育、貧困、環境、エネルギー、人口など多様な分野のデータを取り上げ、主観的な思い込みやバイアスと事実との乖離を考察する。	時間: 150分(3時間) 概要: Pythonを用いたテキストマイニングを実践し、言語データの処理と分析方法を学ぶ。ビジュアル形式でデータを収集し、形態素解析やワードクラウドによる可視化を行い、テキストデータの特徴を考察する。	時間: 50分 概要: ネガボジ判定(感情分析)を実践し、データサイエンスの基本的な考え方や手法を体験的に学ぶ。具体的には、Pythonを用いて形態素解析を行い、日本語評価感情辞書を活用してテキストデータの感情を数値化する。これにより、データ分析が社会における意思決定や問題解決にどのように活用されているかについて考えを深める。
目標: データや事実に基づいて世界を読み解き、データへの関心を高めることを目標とする。本日の世界の姿を知るために、教育、貧困、環境、エネルギー、人口など幅広い分野を取り上げる。思い込みやバイアスと事実に基づいてデータの間にどれほど乖離があるかを理解する。	目標: 分析する元になるデータを整理し、Pythonを用いて言語データを処理する能力を身につけることを目標とする。適切なライブラリやテキストマイニング技術を活用し、得られた結果を考察する思考力を養う。具体的には、ビジュアル形式で収集したデータを使い、データ収集から前処理、分析、可視化、考察までの一連のプロセスを体験する。	目標: ネガボジ判定(感情分析)の実践を通して、データサイエンスの基本的な考え方や手法を体験的に学ぶことを目標とする。具体的には、Pythonを用いて形態素解析を行い、日本語評価感情辞書を活用してテキストデータの感情を数値化する。これにより、データ分析が社会における意思決定や問題解決にどのように活用されているかについて考えを深める。
学習活動: まず、目の錯覚の例を通じて主観的な認識の正確さを理解する。次に、ハンズ・ロスリング氏の動画を視聴し、世界に対する誤解を知る。Gapminderのクイズに挑戦し、感覚的な認識と事実との乖離を体験する。最後に、日経電子版を活用したディスカッションを通じて、事実に基づく世界観の重要性を学ぶ。	学習活動: まず、テキストマイニングの基本概念と社会での活用例を紹介し、興味を喚起する。次に、自分の好きな本をビジュアル形式で紹介し、その発表内容をGoogleドキュメントで文字起こししてデータを収集する。その後、Pythonと適切なライブラリを用いて形態素解析やワードクラウドによる可視化を行い、結果を考察する。また、日経電子版を用いてテキストマイニングの事例をリサーチする。	学習活動: まず、感情分析の基本概念と社会での活用例を紹介し、興味を喚起する。次に、前回のビジュアル形式で収集したデータを用いて、自分たちが感じた感情を数値化する。その後、Pythonと適切なライブラリを用いて形態素解析やワードクラウドによる可視化を行い、結果を考察する。また、日経電子版を用いて感情分析の事例をリサーチする。
評価: 事実に基づくデータと感覚に基づくデータの違いを理解し、Gapminderの活用方法を習得している。 思考・判断・表現: データの乖離を考察し、自分の意見を論理的に表現できる。 主体的な学習態度: 積極的にクイズやディスカッションに参加し、データサイエンスへの関心と探究心を示している。	知識・技能: Pythonを用いたテキストマイニングの基本技術を理解し、適切なライブラリを活用できる。 思考・判断・表現: テキストデータを分析し、重要な情報を抽出し、結果を的確に考察・表現できる。 主体的な学習態度: 自ら興味を持ち、テキストマイニングを積極的に学び、実社会との関連を探究している。	知識・技能: Pythonを用いて感情分析のプログラムを正しく実行し、形態素解析や日本語評価感情辞書の活用方法を理解している。 思考・判断・表現: 定性的・定量的なデータの結果を比較し、分析結果を的確に考察・表現できる。 主体的な学習態度: グループ活動に積極的に参加し、感情分析の社会的活用について自ら情報を収集・共有する姿勢を示している。
留意点: 事前にハンズ・ロスリング氏の動画(後述)を視聴し、Gapminder財団サイトにアクセスしておくこと。Gapminder財団サイトは英語表記のため、あらかじめ翻訳ツールなどを使用し、日本語表記しておくことなど準備を行う。著作権に配慮し、授業で使用する資料や画像は適切な範囲で利用する。	留意点: 事前にGoogle ColaboratoryとGoogleドキュメントの使用法を確認しておくこと。ビジュアル形式で紹介する本の内容は、著作権に配慮し適切な引用を避ける(青空文庫などを活用する)。また、収集したデータのプライバシーや倫理的な取り扱いに留意する。	留意点: 事前にGoogle ColaboratoryとGoogleドキュメントの使用法を確認しておくこと。感情分析に使用するデータはプライバシーに配慮し、個人情報が含まれないように注意する。また、分析結果の解釈において偏見や誤解が生じないように、倫理的な視点を持って取り組むこと。

授業展開のポイント

- 1 **主観的認識から客観的視点への転換**
学習者が自身の先入観やバイアスに気づき、主観的な認識を客観的な視点へと転換することを促す。目の錯覚や心算テスト、統計データと直感的な回答の比較などを通して、認知がいかに主観的であるかを理解する。例えば、世界の貧困率や人口増加に関するクイズを通じて、自分の思い込みと実際のデータとの乖離を実感する。この気づきを基に、世界を見る際には事実に基づいた情報収集と分析が重要であることを学ぶ。
- 2 **データリテラシーの基礎の習得**
統計データや事実に基づいて世界を正しく理解するためのデータリテラシーを育成する。ハンズ・ロスリング氏の動画やGapminderのクイズを活用し、感覚的な認識と実際のデータとのギャップを実感する。これにより、統計データの読み取り方や信頼性の高い情報源の重要性を理解し、データに基づく意思決定の基礎を築く。また、データの収集・分析・解釈の一連の流れを学び、データサイエンスの基礎的なスキルを身につける。
- 3 **批判的思考と多様な視点の涵養**
データ分析結果を基に、ディスカッションや考察を行い、批判的思考力を高める。他者との意見交換を通じて、多様な視点の社会的意義や応用範囲を理解する。例えば、日経電子版での記事リサーチを行い、信頼性の高い情報源からデータを収集し、情報を批判的に評価する能力を育成する。これにより、自分の思い込みを見直し、偏りのない視点をもつとともに、コミュニケーションスキルも向上させる。
- 4 **データサイエンスの社会的応用の理解**
データ分析が社会の意思決定や問題解決にどのように活用されているかを学ぶ。具体的な活用例を通じて、データサイエンスの社会的意義や応用範囲を理解する。例えば、マーケティングや顧客満足度の向上にどのように役立っているかを日経電子版からリサーチする。これにより、学習者の興味・関心を高め、主体的な学習態度を促進するとともに、データサイエンスの実用性を実感する。

1-2 テキストマイニングの授業展開

本節の目標
分析する元になるデータを整理し、Pythonを用いて言語データを処理することを通じて、適切なライブラリやテキストマイニング技術を活用できる能力を身につける。テキストマイニングから得られた結果を考察する思考力を養うことを目標とする。

授業の組み立て

導入(10分): テキストマイニングへの興味を喚起し、文章データの可視化や社会での活用例を紹介する。
展開1(40分): 自分の好きな本をビジュアル形式で紹介する。
展開2(50分): 紹介した重要な情報から考察することや、不要語の除去からデータ分析の精度を高める体験を通して理解する。
まとめ(15分): 結果を共有し、実社会での活用事例を調べ、データサイエンスの実用性を考察する。

評価規準
知識・技能: Pythonを用いたテキストマイニングの基本技術を理解し、適切なライブラリを活用できる。
 思考・判断・表現: テキストデータを分析し、重要な情報を抽出し、結果を的確に考察・表現できる。
 主体的な学習態度: 自ら興味を持ち、テキストマイニングを積極的に学び、実社会との関連を探究している。

準備
Google Colaboratory: プログラミング言語のPythonを、Webブラウザから直接記述・実行できるサービス
<https://s.nikkei.com/3U96Bdu>
 Googleドキュメント: 文書作成ツール
<https://s.nikkei.com/3YSBwb>

展開3: データクレンジング
取得した文字起こしデータの確認を行い、誤認識された単語の修正、「スー」と「あー」などの不要な話し言葉の除去を行い、テキストデータをクレンジングします。これにより、分析の精度を高めることができます。データクレンジングの過程で、ワードクラウドによる可視化を行います。生成されたワードクラウドを通じて、文章の特徴やキーワードを視覚的に把握します。

単語	修正前	修正後
おたすた	あすた	あすた
かたじうた	かたじうた	かたじうた
そんなことばく	そんなことばく	そんなことばく
それおもしろい	それおもしろい	それおもしろい
それおもしろい	それおもしろい	それおもしろい

展開4: テキストマイニングの実践
クレンジングしたテキストデータを用いて、テキストマイニングを行います。まず、形態素解析と、文章を単語に分割し、その品詞を特定する技術であること説明します。Google Colaboratory上でJanomeライブラリを用いて形態素解析を実践します。解析結果を基に、抽出した単語を抽出し、ワードクラウドによる可視化を行います。生成されたワードクラウドを通じて、文章の特徴やキーワードを視覚的に把握します。

まとめ: 結果の考察と活用議論
ワードクラウド結果から考察を行い、抽出ワードが自分の予想と一致しているかを確認します。また、意外な単語や関連性のあるワードが見つかった場合、その背景や理由を考察します。最後に、この技術がどのような場面で活用できるか、日経電子版から関連記事リサーチして、ディスカッションを行います。例えば、マーケティングでの顧客ニーズの分析、教育分野での学習内容の把握、医療分野での症例記録の分析など、多様な応用事例の共有し、データサイエンスの可能性を探究します。

授業展開のポイント

- 1 **テキストマイニングの基本理解と社会的応用の理解**
テキストマイニングの基本概念を理解し、SNSの投稿分析や顧客レビューの分析、ニュース記事の分類など、具体的な活用例を通じてその社会的意義を学びます。これにより、データサイエンスが実社会でどのように役立っているかを日経電子版からリサーチする。これにより、学習者の興味・関心を高め、主体的な学習態度を促進するとともに、データサイエンスの実用性を理解します。
- 2 **データ収集と前処理の実践的スキルの習得**
ビジュアル形式を通じてデータを収集し、Google Colaboratory上のJanomeライブラリを用いて形態素解析を行い、テキストデータを単語に分割し、品詞を特定する技術を習得します。さらに、ワードクラウドを用いて抽出単語を可視化し、テキストデータの特徴やキーワードを視覚的に把握する学習方法を学びます。
- 3 **分析結果の考察と応用可能性の探究**
生成されたワードクラウドを基に、予想していた抽出ワードと一致しない単語や意外な単語について考察します。テキストマイニングの結果から得られる洞察を深め、この技術が他のどのような場面で活用できるかをディスカッションし、データサイエンスの応用範囲を探究します。
- 4 **主体的な学習態度とコミュニケーション能力の向上**
ビジュアル形式や日経電子版を活用したワードクラウドやディスカッションを通じて、自ら積極的に学習に取り組み、他者との意見交換や協力を進め、コミュニケーション能力やプレゼンテーションスキルを向上させます。

本授業案について

本授業の目標

本授業では、データに基づく地域課題の発見と考察を目指し、データサイエンスを探究的に学び、実践的なスキルの習得を目指します。学習者は、e-Stat やオープンデータを活用し、GIS（地理情報システム）ツールでデータの収集・可視化・モデリングを行い、地域課題に関する仮説を構築します。また、重回帰分析を用いて複数の変数間の関係性を統計的に分析し、データに基づく論理的な考察を深めます。さらに、得られた結論を批判的に評価し、説得力のある方法で表現する力を身につけます。プレゼンテーションを通じて、他者にわかりやすく情報を伝える能力も高めます。探究活動を通じて粘り強く課題に取り組み、多様な意見を尊重しながら議論を進める協働的な姿勢を育み、フィードバックを活用して自身の分析や結論を発展させる力を身につけます。

探究的データサイエンスの評価

◆知識・技能

データサイエンスの基礎である目的の設定、データの収集・可視化・モデリングなど、統計的分析手法を体系的に理解することができる。GIS（地理情報システム）ツールを活用して、e-Stat やオープンデータを可視化し、地域課題の仮説を組み立てることができる。重回帰分析を用いて設定した複数の変数間の関係性を統計的に分析することができる。

◆思考力・判断力・表現力

データ分析の結果にもとづいて論理的に考察することができる。また、社会的背景などのリサーチから、分析を批判的に評価することができる。得られた結論を明確かつ説得力のある方法で表現ことができ、プレゼンテーションや視覚資料を用いて他者に理解しやすく情報を伝えることができる。

◆学びに向かう力・人間性

データ分析に対して粘り強く取り組み、柔軟に対応しながら解決策を見いだすことができる。また、グループでのディスカッションなどを通じて、多様な意見や視点を尊重しつつ、自身の考えを明確かつ論理的に表現することができる。そして、他者のリサーチや考察に対しても適切に議論を進めることができる。フィードバックを前向きに受け入れ、自らの分析や結論を改善・発展させることで、探究活動を通じてより深い学びを追求する姿勢で取り組むことができる。

中単元のテーマと学習活動(36時限配分の場合)

1. データへの関心(9時限)

- 1-1 Gapminder (2時限)
- 1-2 テキストマイニング (3時限)
- 1-3 ネガポジ判定 (4時限)

本章の目標は、データや事実に基づいて世界を正しく理解し、データサイエンスの基本的な考え方や手法を体験的に学ぶことである。Gapminder を活用して、教育、貧困、環境など幅広い分野のデータを分析し、主観的な思い込みと事実との乖離を理解する。さらに、Python を用いて言語データを処理し、テキストマイニング技術を習得する。これにより、データ分析が社会の意思決定や問題解決にどのように活用されるかを考察する能力を養う。

2. データの整理・分析(7時限)

- 2-1 基本統計量 (2時限)
- 2-2 散布図作成 (2時限)
- 2-3 相関係数 (1時限)
- 2-4 単回帰分析 (2時限)

本章の目標は、基本統計量の理解を深め、散布図や相関係数を用いてデータ間の関係性を把握し、さらに単回帰分析を通じて予測モデルの基礎を学ぶことである。これにより、データの特性や傾向を的確に捉え、統計的手法を活用して客観的な判断や意思決定ができる力を養う。データリテラシーを高め、高等学校「情報II」のデータサイエンスの範囲に対応した実践的なスキルを身につけることを目指す。

3. データを使った探究(6時限)

- 3-1 GIS の活用 (2時限)
- 3-2 統計データの活用 (2時限)
- 3-3 地域課題の発見・考察 (2時限)

本章では、GIS（地理情報システム）であるjSTAT MAPとe-Statの社会統計データの基本操作を学び、統計データを地理的に可視化する意義と活用方法を理解することを目指す。生徒はe-Statから都道府県や市区町村のデータを取得し、jSTAT MAP上で地図データと属性データを組み合わせることで、地理的なパターンや傾向を視覚的に把握し、社会課題への応用力を養う。防災や都市計画などの具体例を通じ、GISと統計データの組み合わせが問題解決にどう貢献するかを実感させる。データサイエンスの基礎的な知識と多角的な解釈力を身につけ、データに基づいた考察力や問題解決力を身につける。

4. データの分析・解釈(14時限)

- 4-1 重回帰分析 (3時限)
- 4-2 データの整形 (1時限)
- 4-3 探究課題：GIS×統計分析 (8時限)
- 4-4 検証と分類 (2時限)

本章の目標は、高校生が「情報II」で学ぶデータサイエンスにおける統計分析やデータ可視化の実践的なスキルを探究的な学びを通して身につけることである。重回帰分析や分類などの統計手法を理解し、さらにGISを活用したデータの地理的可視化を通じて、多面的な分析能力を養うことを目指す。これにより、現実社会の課題を客観的かつ論理的に捉え、データを活用した意思決定や課題解決の力を育成する。また、データの取得から分析、解釈、応用に至るプロセスを探究的に学習し、データリテラシーを深化させる。

中単元のテーマと学習活動(30時限の場合)

1. データへの関心(5時限)

- 1-1 Gapminder (1時限)
- 1-2 テキストマイニング (3時限)
- 1-3 ネガポジ判定 (1時限)

本章の目標は、データや事実に基づいて世界を正しく理解し、データサイエンスの基本的な考え方や手法を体験的に学ぶことである。Gapminderを活用して、教育、貧困、環境など幅広い分野のデータを分析し、主観的な思い込みと事実との乖離を理解する。さらに、Pythonを用いた言語データの処理を体験し、基本的なテキストマイニング技術を習得する。これにより、データ分析が社会の意思決定や問題解決にどのように活用されるかを考察する能力を養う。

2. データの整理・分析(6時限)

- 2-1 基本統計量 (1時限)
- 2-2 散布図作成 (2時限)
- 2-3 相関係数 (1時限)
- 2-4 単回帰分析 (2時限)

本章の目標は、基本統計量の理解を深め、散布図や相関係数を用いてデータ間の関係性を把握し、さらに単回帰分析を通じて予測モデルの基礎を学ぶことである。これにより、データの特性や傾向を的確に捉え、統計的手法を活用して客観的な判断や意思決定ができる力を養う。データリテラシーを高め、高等学校「情報II」のデータサイエンスの範囲に対応した実践的なスキルを身につけることを目指す。

3. データを使った探究(5時限)

- 3-1 GISの活用 (1時限)
- 3-2 統計データの活用 (2時限)
- 3-3 地域課題の発見・考察 (2時限)

本章では、GIS(地理情報システム)であるjSTAT MAPとe-Statの社会統計データの基本操作を学び、統計データを地理的に可視化する意義と活用方法を理解することを目指す。生徒はe-Statから都道府県や市区町村のデータを取得し、jSTAT MAP上で地図データと属性データを組み合わせることで、地理的なパターンや傾向を視覚的に把握し、社会課題への応用力を養う。防災や都市計画などの具体例を通じ、GISと統計データの組み合わせが問題解決にどう貢献するかを実感させる。データサイエンスの基礎的な知識と多角的な解釈力を身につけ、データに基づいた考察力や問題解決力を身につける。

4. データの分析・解釈(14時限)

- 4-1 重回帰分析 (3時限)
- 4-2 データの整形 (1時限)
- 4-3 探究課題:GIS×統計分析 (8時限)
- 4-4 検証と分類 (2時限)

本章の目標は、高校生が「情報II」で学ぶデータサイエンスにおける統計分析やデータ可視化の実践的なスキルを探究的な学びを通して身につけることである。重回帰分析や分類などの統計手法を理解し、さらにGISを活用したデータの地理的可視化を通じて、多面的な分析能力を養うことを目指す。これにより、現実社会の課題を客観的かつ論理的に捉え、データを活用した意思決定や課題解決の力を育成する。また、データの取得から分析、解釈、応用に至るプロセスを探究的に学習し、データリテラシーを深化させる。

中単元のテーマと学習活動(10時限の場合)

1. データへの関心(2時限)

- 1-1 Gapminder (1時限)
- 1-2 テキストマイニング (1時限)

本章の目標は、データや事実に基づいて世界を正しく理解し、データサイエンスの基本的な考え方や手法を体験的に学ぶことである。Gapminderを活用して、教育、貧困、環境など幅広い分野のデータを分析し、主観的な思い込みと事実との乖離を理解する。さらに、テキストマイニングに触れ、データ分析が社会の意思決定や問題解決にどのように活用されるかを考察する能力を養う。

2. データの整理・分析(3時限)

- 2-1 基本統計量 (1時限)
- 2-2 散布図作成 (0.5時限)
- 2-3 相関係数 (0.5時限)
- 2-4 単回帰分析 (1時限)

本章の目標は、基本統計量の理解を深め、散布図や相関係数を用いてデータ間の関係性を把握し、さらに単回帰分析を通じて予測モデルの概要を学ぶことである。これにより、データの特性や傾向を的確に捉え、統計的手法を活用して客観的な判断や意思決定の基礎を身につける。データリテラシーを高め、高等学校「情報II」のデータサイエンスの範囲に対応した実践的なスキル習得への接続を目指す。

3. データを使った探究(2時限)

- 3-1 GISの活用 (1時限)
- 3-2 統計データの活用 (1時限)

本章では、GIS(地理情報システム)であるjSTAT MAPとe-Statの社会統計データに触れ、統計データを地理的に可視化する意義と活用方法を理解することを目指す。e-Statの都道府県や市区町村のデータや、jSTAT MAP上で地図データと属性データとの組み合わせを示すことで、地理的なパターンや傾向を視覚的に把握し、社会課題への応用力を養う。防災や都市計画などの具体例によって、GISと統計データの組み合わせが問題解決にどう貢献するかを考察させる。データサイエンスの基礎的な知識と多角的な解釈の必要性を理解し、データに基づいた考察や問題解決の姿勢を養成する。

4. データの分析・解釈(3時限)

- 4-1 重回帰分析 (2時限)
- 4-2 データの整形 (1時限)

本章の目標は、高校生が「情報II」で学ぶデータサイエンスにおける統計分析やデータ可視化の実践的なスキルを探究的な学びを通して理解することである。重回帰分析や分類などの統計手法の実践を通じて、データを活用した意思決定や課題解決に触れる。また、データの取得から分析、解釈に至るプロセスを体験し、応用可能性を確認することでデータリテラシーを深化させる。

目次

本授業案について	1
1章 データへの関心	5
1-1 Gapminder	
1-2 テキストマイニング	
1-3 ネガポジ判定	
2章 データの整理・分析	10
2-1 基本統計量	
2-2 散布図作成	
2-3 相関係数	
2-4 単回帰分析	
3章 データを使った探究	16
3-1 GISの活用	
3-2 統計データの活用	
3-3 地域課題の発見・考察	
4章 データの分析・解釈	21
4-1 重回帰分析	
4-2 データの整形	
4-3 探究課題:GIS×統計分析	
4-4 検証と分類	



1章

データへの関心

- 1-1 Gapminder
- 1-2 テキストマイニング
- 1-3 ネガポジ判定

1 データへの関心

章

本章の目標は、データや事実に基づいて世界を正しく理解し、データサイエンスの基本的な考え方や手法を体験的に学ぶことである。Gapminderを活用して、教育、貧困、環境など幅広い分野のデータを分析し、主観的な思い込みと事実との乖離を理解する。さらに、Pythonを用いて言語データを処理し、テキストマイニング技術を習得する。これにより、データ分析が社会の意思決定や問題解決にどのように活用されるかを考察する能力を養う。

	時間・概要	目標	学習活動	評価	留意点
1-1 Gapminder	時間：50分 概要： Gapminder を活用し、データや事実に基づいて世界を理解する授業である。教育、貧困、環境、エネルギー、人口など多様な分野のデータを上げ、主観的な思い込みやバイアスと事実との乖離を考察する。	データや事実に基づいて世界を読み解き、データへの関心を高めることを目標とする。 本当の世界の姿を知るために、教育、貧困、環境、エネルギー、人口など幅広い分野を取り上げる。思い込みやバイアスと事実に基づくデータの間でどれほど乖離があるかを理解する。	まず、目の錯覚の例を通じて主観的な認識の不確かさを理解する。次に、ハンス・ロスリング氏の動画を視聴し、世界に対する誤解を知る。Gapminder のクイズに挑戦し、感覚的な認識と事実との乖離を体験する。最後に、日経電子版を活用したディスカッションを通じて、事実に基づく世界理解の重要性を学ぶ。	知識・技能： 事実に基づくデータと感覚に基づくデータの違いを理解し、Gapminder の活用方法を習得している。 思考・判断・表現： データの乖離を考察し、自分の意見を論理的に表現できる。 主体的な学習態度： 積極的にクイズやディスカッションに参加し、データサイエンスへの関心と探究心を示している。	事前にハンス・ロスリング氏の動画(後述)を視聴し、Gapminder財団サイトにアクセスしておくこと。Gapminder財団サイトは英語表記のため、あらかじめ翻訳ツールなどを使用して、日本語表記にしておくことなど準備を行う。著作権に配慮し、授業で使用される資料や画像は適切な範囲で利用する。
1-2 テキストマイニング	時間：150分(3時限) 概要： Python を用いたテキストマイニングを実践し、言語データの処理と分析方法を学ぶ授業である。ビブリオバトル形式でデータを収集し、形態素解析やワードクラウドによる可視化を行い、テキストデータの特徴を考察する。	分析する元になるデータを整理し、Python を用いて言語データを処理する能力を身につけることを目標とする。 適切なライブラリやテキストマイニング技術を活用し、得られた結果を考察する思考力を養う。具体的には、ビブリオバトルで収集したデータを使い、データ収集から前処理、分析、可視化、考察までの一連のプロセスを体験する。	まず、テキストマイニングの基本概念と社会での活用例を紹介し、興味を喚起する。次に、自分の好きな本をビブリオバトル形式で紹介し、その発表内容をGoogle ドキュメントで文字起こししてデータを収集する。その後、Python と適切なライブラリを用いて形態素解析とワードクラウドによる可視化を行い、結果を考察する。また、日経電子版を用いてテキストマイニングの事例をリサーチする。	知識・技能： Python を用いたテキストマイニングの基本技術を理解し、適切なライブラリを活用できる。 思考・判断・表現： テキストデータを分析し、重要な情報を抽出し、結果を的確に考察・表現できる。 主体的な学習態度： 自ら興味を持ち、テキストマイニングを積極的に学び、実社会との関連を探究する姿勢を示している。	事前にGoogle Colaboratory とGoogle ドキュメントの使用方法を確認しておくこと。ビブリオバトルで紹介する本の内容は、著作権に配慮し過度な引用を避ける(青空文庫などを活用する)。また、収集したデータのプライバシーや倫理的な取り扱いに留意する。
1-3 ネガポジ判定	時間：50分 概要： ネガポジ判定(感情分析)を実践し、データサイエンスの基本的な考え方や手法を体験的に学ぶ授業である。前回のビブリオバトルで収集した文字起こしデータを用いて、Python で感情分析を行い、データ分析が社会の意思決定や問題解決にどのように活用されるかを考察する。	ネガポジ判定(感情分析)の実践を通して、データサイエンスの基本的な考え方や手法を体験的に学ぶことを目標とする。 具体的には、Python を用いて形態素解析を行い、日本語評価極性辞書を活用してテキストデータの感情を数値化する。これにより、データ分析が社会における意思決定や問題解決にどのように活用されているかについて考えを深める。	まず、感情分析の基本概念と社会での活用例を紹介し、興味を喚起する。次に、前回のビブリオバトルで収集した文字起こしデータを用いて、自分たちが感じた感情を整理する。その後、Python と適切なライブラリを用いて形態素解析とネガポジ判定を実施し、結果を数値化・可視化する。最後に、結果を考察し、日経電子版を活用して、データ分析の社会的意義についてのディスカッションを行う。	知識・技能： Python を用いて感情分析のプログラムを正しく実行し、形態素解析や日本語評価極性辞書の活用方法を理解している。 思考・判断・表現： 定性的・定量的なネガポジ結果を比較し、分析結果を的確に考察・表現できる。 主体的な学習態度： グループ活動に積極的に参加し、感情分析の社会的活用について自ら情報を収集・共有する姿勢を示している。	事前にGoogle Colaboratory とGoogle ドキュメントの使用方法を確認しておくこと。感情分析に使用するデータはプライバシーに配慮し、個人情報が含まれないように注意する。また、分析結果の解釈において偏見や誤解が生じないように、倫理的な視点を持って取り組むこと。

授業展開のポイント

主観的認識から客観的視点への転換

学習者が自身の先入観やバイアスに気づき、主観的な認識を客観的な視点へと転換することを促す。目の錯覚や心理テスト、統計データと直感的な回答の比較などを活用し、認知がいかに主観的であるかを理解する。例えば、世界の貧困率や人口増加に関するクイズを通じて、自分の思い込みと実際のデータとの乖離を実感する。この気づきを基に、世界を見る際には事実に基づいた情報収集と分析が重要であることを学ぶ。

データリテラシーの基礎の習得

統計データや事実に基づいて世界を正確に理解するためのデータリテラシーを育成する。ハンス・ロスリング氏の動画やGapminder のクイズを活用し、感覚的な認識と実際のデータとのギャップを実感する。これにより、統計データの読み解き方や信頼性の高い情報源の重要性を理解し、データに基づく意思決定の基礎を築く。また、データの収集・分析・解釈の一連の流れを学び、データサイエンスの基礎的なスキルを身につける。

批判的思考と多様な視点の涵養

データ分析結果を基に、ディスカッションや考察を行い、批判的思考力を高める。他者との意見交換を通じて、多様な視点を理解し、柔軟な世界観を構築する。例えば、日経電子版での記事リサーチを行い、信頼性の高い情報源からデータを収集し、情報を批判的に評価する能力を育成する。これにより、自分の思い込みを見直し、偏りのない視点を養うとともに、コミュニケーションスキルも向上させる。

データサイエンスの社会的応用の理解

データ分析が社会の意思決定や問題解決にどのように活用されているかを学ぶ。具体的な活用例を調べ、データサイエンスの社会的意義や応用範囲を理解する。例えば、感情分析技術がマーケティングや顧客満足度の向上にどのように役立っているかを日経電子版からリサーチする。これにより、学習者の興味・関心を高め、主体的な学習態度を促進するとともに、データサイエンスの実用性を実感する。

1-1 Gapminderの授業展開

本節の目標

データや事実に基づいて世界を読み解き、データへの関心を高めることを目標とする。本当の世界の姿を知るために、教育、貧困、環境、エネルギー、人口など幅広い分野を取り上げる。思い込みやバイアスと事実に基づくデータの間にどれほど乖離があるかを理解する。

評価規準

知識・技能：事実に基づくデータと感覚に基づくデータの違いを知り、Gapminder財団サイトの活用方法を理解している。

思考・判断・表現：データの乖離を考察し、感覚的な認識と事実の違いを判断できる。ディスカッションを通じて、自分の意見を論理的に表現できる。

主体的に学習に取り組む態度：積極的にクイズやディスカッションに参加し、データサイエンスへの関心と探究心を持って学習に取り組もうとしている。

準備

「TED日本語 - ハンス・ロスリング、オーラ・ロスリング: 世界について無知にならないために」視聴 (0:00 ~ 8:30)

<https://s.nikkei.com/4h8SGhw>

Gapminder財団サイト(英語)

<https://s.nikkei.com/4dOpZ6D>

授業の組み立て

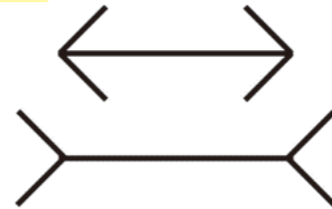
導入 (5分)：データの信憑性と信頼性について関心を持たせる。体験に基づくデータと事実に基づくデータは、どちらが信じるに値する科学的根拠となり得るかを問いかける。

展開 (30分)：Gapminder財団サイトのクイズを通じて、感覚的な回答と事実に基づく回答の乖離を体験しディスカッションを行う。また、「寿命と所得」のビジュアルグラフから読み取れることを議論する。

まとめ (15分)：話し合いの結果を発表し、これから始まるデータサイエンスの学びへの関心を高める。

導入: 私たちは世界をどう見ているのか?

まず、二本の横線のうちどちらが長いかを問いかける目の錯覚の例を見ます。多くの人が実際には同じ長さであるにもかかわらず、片方が長く見えると答えます。これは、私たちの認知がいかに主観的であるかを示しています。同様に、世界に対する認識も先入観や偏見によって歪められている可能性があります。このステップでは、自分たちの見方が必ずしも正確でないことを理解し、客観的な視点の重要性を学びます。



ミュラーリヤー錯視
出典: ナショナルジオグラフィック

展開3: 日経電子版で事実調査

Gapminderのクイズで間違えた問題について、日経電子版で関連記事を調べます。信頼性の高い情報源から最新のデータや専門家の分析を得ることで、自分たちの誤解を修正します。このステップでは、情報リテラシーを高め、正確な情報に基づいて世界を理解する方法を身につけます。



展開1: ハンス・ロスリング「世界について無知にならないために」TED動画視聴

ハンス・ロスリング氏は世界の人口、健康、収入に関する統計データを、長い差し棒とカラフルなグラフを使ってわかりやすく解説します。多くの人が持つ世界のイメージと実際のデータが大きく異なることを示し、私たちの無知や誤解を明らかにします。この動画を通じて、データに基づいた正確な世界理解の重要性を感じ取ります。(0:00 ~ 8:30 まで視聴)



展開2: 「世界は分断されている」という思い込み

Gapminder財団のサイトでクイズに挑戦し、自分たちの世界観を検証します。例えば、地球温暖化や極度の貧困に関する質問で、実際の状況と自分たちの認識がずれていることに気づきます。これにより、「先進国と発展途上国」という二分法的な見方が誤解を生む原因であることを理解します。この活動を通じて、固定観念を捨て、世界をより正確に見ることの大切さを学びます。

You are probably wrong about



まとめ1: ディスカッション

グループで関連する記事を共有し、なぜ多くの人が世界について誤解しているのかを話し合います。メディアの影響、教育の偏り、情報のアクセス格差など、多様な視点から原因を探ります。このディスカッションを通じて、自分たちの認識を深め、より広い視野で世界を見る力を養います。



まとめ2: 10の本能について提示(参考資料: ファクトフルネス)

最後に、書籍「FACTFULNESS(ファクトフルネス)」で紹介されている「10の本能」を学びます。特に「分断本能」に注目し、私たち自身が文系や理系といったカテゴリーで自他を分けていないかを考えます。他にも「ネガティブ本能」や「恐怖本能」など、自分たちの思考パターンを客観的に見つめ直し、偏った認識から脱却する方法を探ります。これにより、事実に基づいたバランスの取れた世界観を築くことを目指します。

分断本能	ネガティブ本能	直線本能	恐怖本能	過大視本能
「世界は分断されている」という思い込み	「世界はどんどん悪くなっている」という思い込み	「世界の人口はひたすら増え続ける」という思い込み	危険でないことを、恐ろしいと考えてしまう思い込み	「目の前の数字がいかに重要だ」という思い込み
パターン化本能	宿命本能	単純化本能	犯人探し本能	焦り本能
「ひとつの例がすべてに当てはまる」という思い込み	「すべてはあらかじめ決まっている」という思い込み	「世界はひとつの切り口で理解できる」という思い込み	「誰かを責めれば物事は解決する」という思い込み	「いますぐ手を打たないと大変なことになる」という思い込み

授業展開のポイント

主観的認識と客観的視点の重要性

目の錯覚を通じて、私たちの認知が主観的であり、先入観や偏見によって歪められる可能性があることを理解します。この気づきを基に、世界を見る際には客観的な視点が重要であり、事実に基づいた情報収集と分析が必要であることを学びます。

データに基づく世界理解の必要性

ハンス・ロスリング氏のTED動画やGapminderのクイズを活用し、感覚的な認識と実際のデータとのギャップを実感します。これにより、統計データや事実に基づいて世界を正確に理解することの重要性を認識し、データリテラシーの基礎を築きます。

情報リテラシーと批判的思考の育成

日経電子版での記事リサーチやディスカッションを通じて、信頼性の高い情報源からデータを収集し、批判的に評価する能力を高めます。情報を分析し、自分の意見を論理的に表現することで、批判的思考力とコミュニケーションスキルを向上させます。

多様な視点の理解と柔軟な世界観の構築

グループで情報を共有したり話し合ったりすることで、さまざまな人の意見や考え方を理解する機会を持ちます。これにより、自分の思い込みを見直し、多角的に世界を見ることの大切さを学びます。また、自分とは異なる背景や考えを持つ人たちと、根拠のあるデータに基づいて一緒に取り組むことの価値を実感します。

認知バイアスと固定観念の克服

「分断本能」などファクトフルネスで紹介されている「10の本能」を理解し、自分たちの思考パターンを客観的に見る機会を持ちます。無意識のバイアスや固定観念が世界理解を妨げていることを理解し、それらを軽減する方法を探ることで、偏りのない視点を養います。

1-3 ネガポジ判定の授業展開

本節の目標

ネガポジ判定（感情分析）の実践を通してデータサイエンスの基本的な考え方や手法を体験的に学び、データ分析が社会における意思決定や問題解決にどのように活用されるかについて考えを深めることを目指す。

評価規準

知識・技能：ビブリオバトルの文字起こしデータを使って、Pythonで感情分析するプログラムを正しく実行することができる。この体験を通して、感情分析がどのようなものか理解できる。

思考・判断・表現：定性的・定量的なネガポジ結果の違いを比較検討し、自分の言葉で考察をまとめて表現できる。

主体的に学習に取り組む態度：グループ活動に積極的に参加し、感情分析の社会的活用について自ら情報を収集・共有することができる。

準備

Google Colaboratory：プログラミング言語のPythonを、Webブラウザから直接記述・実行できるサービス

<https://s.nikkei.com/3U96Bdu>

Googleドキュメント：文章作成ツール

<https://s.nikkei.com/3Y5Bwbj>

授業の組み立て

導入（5分）：感情分析への関心を高めるため、ワードクラウドの結果を提示し、直感的なポジティブ・ネガティブ判断の正確性や、適切な分析データについて問いかける。

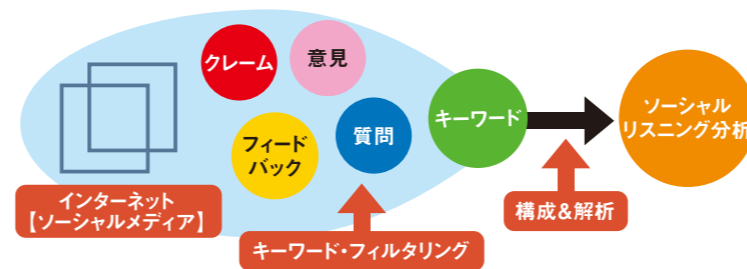
展開（30分）：前回のビブリオバトルの文字起こしデータを用いてネガポジ分析を実施。定性的なネガポジと定量的なネガポジの違いを比較する。

まとめ（15分）：グループで分析結果を共有。日経電子版で感情分析に関する記事を検索し、社会での活用事例をリサーチする。

導入：ネガポジ判定（感情分析）への興味関心

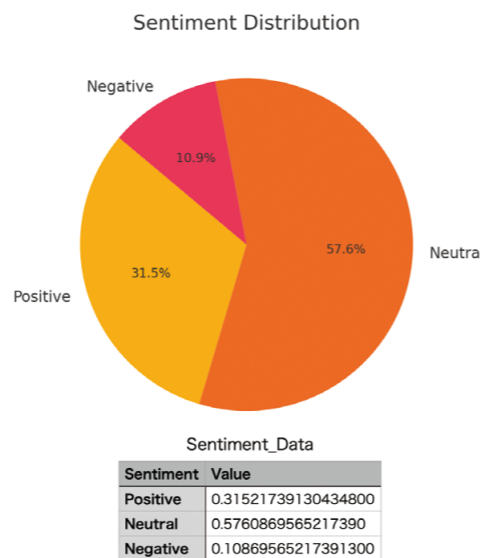
ネガポジ判定（感情分析）とは何かを理解するために、その基本概念を説明します。感情分析は、テキストデータから人々の感情や意見を解析し、ポジティブ、ネガティブ、ニュートラルといったカテゴリに分類する技術です。例えば、SNSの投稿や商品レビューから消費者の満足度を数値化することが可能です。この技術は企業のマーケティング戦略や顧客サービスの改善、世論調査など、幅広い分野で活用されています。その活用例を示す図や表を用いることで、理解を深めます。

ソーシャルな情報をさまざまなメディアから指定のキーワードで抽出します



展開2: Google Colabでネガポジ判定

Google Colaboratory を用いて実際にネガポジ判定を行います。まず、日本語評価極性辞書を読み込み、感情分析のための準備をします。その後、ビブリオバトルの文字起こしデータを読み込み、Pythonでプログラムを実施します。形態素解析を行い、テキスト内の単語を評価極性辞書と照合して感情スコアを計算します。結果をグラフや表で視覚化し、データの傾向を把握します。



展開3: 形態素解析からネガポジの数値化

プログラムを実行すると、形態素解析の後、ネガポジ判定が数値化されます。プログラムがポジティブな評価を持つ単語の数やその合計スコアを算出し、ネガティブな単語についても同様に計算します。また、日本語評価極性辞書にないニュートラルな単語の数も算出します。これらの結果を数値として視覚的に理解します。

```
# ビブリオバトル文字起こしデータを読み込む
fp = open("talk.txt", "rt", encoding="shift_jis")
text = fp.read()

# 形態素解析
from janome.tokenizer import Tokenizer
tok = Tokenizer()

# 数える
res = {"p":0, "n":0, "e":0}
for t in tok.tokenize(text):
    bf = t.base_form # 基本形
    # 辞書にあるか確認
    if bf in np.dic:
        r = np.dic[bf]
        if r in res:
            res[r] += 1

# 結果を表示
print(res)
cnt = res["p"] + res["n"] + res["e"]
print("ポジティブ度", res["p"] / cnt)
print("ニュートラル度", res["e"] / cnt)
print("ネガティブ度", res["n"] / cnt)
```

```
{'p': 29, 'n': 10, 'e': 53}
ポジティブ度 0.31521739130434784
ニュートラル度 0.5760869565217391
ネガティブ度 0.10869565217391304
```

展開1: ビブリオバトルの文字起こしデータから受けた感情

自分のビブリオバトル文字起こしデータを読み、その内容から受けた感情について考えます。ポジティブな要素としては、興味深いテーマや新たな発見が挙げられます。ネガティブな要素としては、難解な内容や批判的な意見が含まれる場合があります。また、特に強い感情を引き起こさないニュートラルな情報も存在します。これらの感情を整理し、自分自身の感じたことを明確にします。



まとめ: 結果から考察

得られた結果を基に考察を行います。事前に予想していたネガポジ判定と実際の結果を比較し、その違いや共通点について分析します。特に感情スコアに大きく影響した関連ワードを特定し、それらがテキストの感情にどのように寄与しているかを考察します。この技術をどのような場面で活用できるかをグループでディスカッションし、日経電子版から具体的な活用例や可能性についてリサーチを行います。



授業展開のポイント

感情分析の基本概念を理解する

感情分析は、テキストデータから人々の感情や意見を解析し、ポジティブ、ネガティブ、ニュートラルなどに分類する技術です。SNSの投稿や商品レビューなど、実社会で広く活用されているその基本的な仕組みや活用例を学ぶことで、データサイエンスの基礎知識を身につけます。

ビブリオバトル文字起こしのデータでネガポジ判定を体験する

前回のビブリオバトルで収集した文字起こしデータを使用し、実際にネガポジ判定を行います。自分たちが発表した内容がどのような感情を引き起こすのかを分析し、定性的な感覚と定量的なデータの違いを体験的に理解します。

Pythonを用いた形態素解析とネガポジの数値化を学ぶ

Google Colaboratory上でPythonを使い、形態素解析を行ってテキストデータを細かく分析します。日本語評価極性辞書を活用し、各単語の感情スコアを算出してネガポジを数値化することで、プログラミングによるデータ処理の手法を学びます。

分析結果の考察と社会での感情分析の活用例を探る

得られた分析結果を基に、自身の予想との比較や関連ワードの影響を考察します。また、日経電子版を利用して感情分析の社会的な活用事例を調べ、データ分析が社会の意思決定や問題解決にどのように役立っているかディスカッションを通して理解を深めます。

2章

データの整理・分析

- 2-1 基本統計量
- 2-2 散布図作成
- 2-3 相関係数
- 2-4 単回帰分析

2章

データの整理・分析

本章の目標は、基本統計量の理解を深め、散布図や相関係数を用いてデータ間の関係性を把握し、さらに単回帰分析を通じて予測モデルの基礎を学ぶことである。これにより、データの特性や傾向を的確に捉え、統計的手法を活用して客観的な判断や意思決定ができる力を養う。データリテラシーを高め、高等学校「情報II」のデータサイエンスの範囲に対応した実践的なスキルを身につけることを目指す。

	時間・概要	目標	学習活動	評価	留意点
2-1 基本統計量	時間：50分 概要： 基本統計量の概要を学び、平均値と中央値の違いを考える授業である。ワークシートを用いた演習を通じて、実際のデータから平均値、中央値、最頻値、分散、標準偏差などの統計量を算出する。最後に、グループで結果を共有し、統計の実社会での応用例を調べる。	データリテラシーを身につけ、平均値、中央値、最頻値、分散、標準偏差などの基本統計量を理解することを目指す。 これらの統計量を用いてデータの分析や解釈を行い、データの特性を正確に捉え、客観的な指標として統計量を適切に扱う力を養う。 また、統計的な思考を通じて、データに基づいた意思決定の重要性を認識する。	導入では、基本統計量の概要を説明し、身近な例を用いて平均値と中央値の違いを考えさせる。展開では、ワークシートを用いて各統計量の意味と算出方法を実際のデータで学ぶ。個別演習により、データの「位置」と「広がり」の概念を理解する。まとめでは、グループで結果を共有し、日経電子版を使って統計の実社会での応用例を調べる。	知識・技能： 基本統計量の概念を理解し、データから平均値、中央値、最頻値、分散、標準偏差を正確に算出できる。 思考・判断・表現： データの特性や目的に応じて適切な統計量を選び、その意味を解釈してデータの傾向や特徴を説明できる。 主体的な学習態度： 統計的な考え方を日常や社会問題に応用し、データに基づいた判断の重要性を認識している。	生徒が興味を持つテーマを選べるよう、様々なデータセットを用意することが重要である。計算だけでなく、結果の解釈やデータの背後にあるストーリーを考えさせる。また、グループワークを通じて多角的な視点を養い、データの質や信頼性について議論する機会を設ける。
2-2 散布図作成	時間：50分 概要： データの分布や相関関係を視覚化する散布図の意義と作成方法を学ぶ授業である。Google スプレッドシートを用いて二つの数値データの関係を散布図で表現し、データの傾向を解釈する。最後に、社会での散布図の活用例をリサーチする。	散布図の作成と相関関係の理解を通して、 データサイエンスの基本的な考え方や手法を学ぶことを目標とする。 具体的には、Google スプレッドシートを用いて二つの数値データの関係を散布図で表現し、グラフの編集やカスタマイズ方法を習得する。さらに、データの傾向を解釈し、データ分析の基本スキルを身につける。	まず、散布図と相関関係の基本概念を説明する。次に、ワークシートの例題を通じて強い相関と弱い相関の違いを学ぶ。その後、各自がデータを収集し、Google スプレッドシートで散布図を作成する。最後に、結果を考察し、グループで共有し、フィードバックを行う。	知識・技能： 散布図の作成方法と相関関係の理解ができる。 思考・判断・表現： 散布図を分析し、データ間の関係性を考察し、自らの仮説を根拠に基づいて説明できる。 主体的な学習態度： 自主的にデータを探索し、積極的にディスカッションに参加する姿勢を示す。	事前にGoogle スプレッドシートの基本操作やSSDSEのデータセットの利用方法を案内しておくこと。生徒がデータ収集や散布図作成に困難を感じないようにサポートする。また、生徒間の理解度の差に配慮し、適切な指導を行う。
2-3 相関係数	時間：100分（2時限） 概要： 相関係数の概念と計算方法を学び、実際のデータを用いて演習を行う授業である。生徒は興味のあるデータを選び、相関係数を算出し、結果を考察・共有する。	相関係数の概念とその意味を理解し、 データの収集、整理、分析の技能を身につけることを目指す。 表計算ソフトを用いて相関係数を算出し、散布図や回帰直線との関連を考察する。また、相関と因果の違いを認識し、データを客観的に分析する視点を養う。	導入では相関係数の概念を紹介し、日常生活の例で興味を喚起する。展開では都道府県データを用いて相関係数の計算方法を学び、ワークシートで演習を行う。その後、生徒は興味のあるデータを選び、相関の仮説を立てて相関係数を算出する。最後に、グループで結果を共有し、相関と因果の違いについてディスカッションを行う。	知識・技能： 相関係数の概念を理解し、正確に計算できる。 思考・判断・表現： 相関係数を用いてデータの関係性を分析し、結果を具体的に説明できる。相関と因果の違いを適切に判断できる。 主体的な学習態度： 積極的にデータを収集・分析し、結果を批判的に考察する姿勢を示す。	生徒がデータ収集や相関係数の計算に困難を感じないように、事前にGoogle スプレッドシートの操作方法や都道府県データの利用方法を案内する。また、相関と因果を混同しないよう指導し、データの解釈における注意点を強調する。
2-4 単回帰分析	時間：100分（2時限） 概要： 単回帰分析の概念と手法を学び、Pythonを用いて為替レートと日経平均株価のデータを分析する授業である。データの関係を数式で表し、予測に活用する力を養う。最後に、他の指標を探し、手法の応用を考える。	データの形式や分析目的に応じた可視化方法を選択し、 相関係数や相関関係・因果関係、交絡因子を理解する手がかりをつかむことを目指す。 Python プログラムでの単回帰分析を通じてデータの関係性を数式で表し、データの予測に活用できる力を養う。	まず、散布図と相関係数を復習し、それが予測につながることを説明する。次に、Pythonを用いて為替レートと日経平均株価のデータを取得し、散布図を作成する。単回帰分析を実施し、回帰式を求め、データの予測を行う。最後に、他の指標を探し、手法の応用をグループでディスカッションする。	知識・技能： 単回帰分析の手法を理解し、Pythonで回帰式を求め、予測値を計算できる。 思考・判断・表現： データの相関関係を考察し、適切な分析方法を選択し、結果を明確に表現できる。 主体的な学習態度： データ分析に積極的に取り組み、他の指標を探究し、学習を振り返る姿勢を示す。	生徒がPythonでの単回帰分析に困難を感じないように、事前にGoogle Colaboratoryの操作方法やデータ取得方法を案内する。また、相関関係と因果関係、交絡因子を混同しないよう指導し、結果の解釈に注意を促す。

授業展開のポイント

データの特性理解とストーリーの読み解き

平均値、中央値、最頻値、分散、標準偏差などの基本統計量を理解し、それらがデータの背後にあるストーリーを読み解くための重要なツールであることを学ぶ。具体的な例を用いて各統計量の違いや適用場面を説明し、データの分布やばらつきを正確に捉える力を養う。これにより、データを客観的に分析し、適切に統計量を扱えるようになる。

ツールを用いたデータ分析と可視化の習得

Google スプレッドシートやPythonなどのツールを活用し、自らデータを収集・整理し、散布図や回帰直線を作成する。興味のあるテーマを選んで分析し、計算結果からデータの特性を説明する練習を行う。Pythonを用いて単回帰分析を実施し、データの関係性を数式で表現する技術を身につける。これにより、データの相関関係や傾向を視覚的・数値的に理解し、データ分析の技術を深める。

相関関係と因果関係の理解と仮説検証

相関係数や単回帰分析を通じて、データ間の関係性を数値化し、その意味を考察する。具体的な例を用いて相関関係の概念を理解し、相関関係と因果関係の違いを明確に認識する。データから仮説を立てて検証するプロセスを経験し、論理的思考力と分析力を養う。さらに、データの信頼性や偏りについても考慮し、データ分析の限界や留意点を理解する。

データに基づく意思決定と社会的応用の認識

実社会のデータを用いた分析を通じて、データに基づいた意思決定の重要性を深く認識する。日経電子版を活用し、データサイエンスが社会やビジネスでどのように活用されているかを学ぶ。さらに、データの質や信頼性、倫理的な側面についても考察し、情報リテラシーを向上させる。これにより、データ分析の結果を基に仮説を立て、問題解決に取り組む力を育てる。

2-1 基本統計量の授業展開

本節の目標

データリテラシーを身につけ、平均や分散、標準偏差などの基本統計量を理解し、様々なデータの分析や解釈に活用できるようになることを目指す。データの特性をつかみ、客観的な指標としての扱いを理解する。

評価規準

知識・技能: 基本統計量の概念を理解し、データから平均値、中央値、最頻値、分散、標準偏差などを正確に算出できる。

思考・判断・表現: データの特性や目的に応じて適切な統計量を選択し、その意味を解釈して、データの傾向や特徴を説明できる。

主体的に学習に取り組む態度: 統計的な考え方を日常生活や社会の問題に応用しようと、データに基づいた判断や意思決定の重要性を認識している。

準備

Google スプレッドシート: Google 社が提供している表計算ツール
<https://s.nikkei.com/3YIAw4C>

XLMiner Analysis ToolPak: Google スプレッドシートの分析ツール
<https://s.nikkei.com/4f3Q3LR>

SSDSE: 教育用標準データセット
<https://s.nikkei.com/4f7aSpV>

※XLMiner Analysis ToolPak の利用が難しい場合は、Excel などの表計算ソフトで代用できる

授業の組み立て

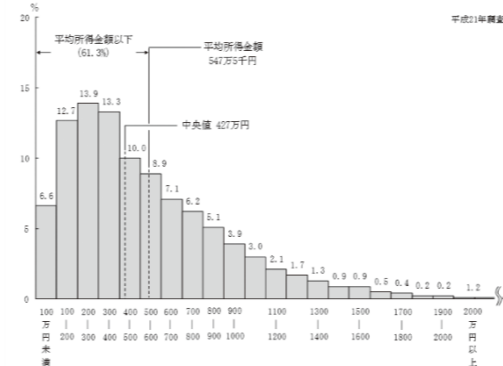
導入 (5分): 基本統計量の概要を説明し、データ全体の傾向把握への活用を紹介する。身近な例で平均値と中央値の違いを考えさせ、関心を喚起する。

展開 (30分): ワークシートを用いて基本統計量の例題を解説する。「位置」と「広がり」の概念を説明し、個別に演習に取り組ませる。実際のデータを用いて、各統計量の意味と算出方法を学ばせる。

まとめ (15分): グループでワークシートの結果を共有し、日経電子版で代表値やヒストグラムなどを検索し、統計の実社会での応用例を調べさせる。

導入: 基本統計量の概要

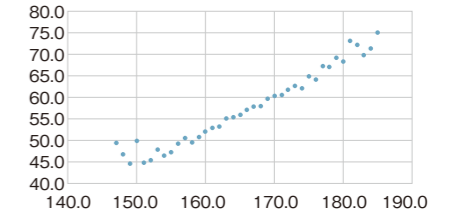
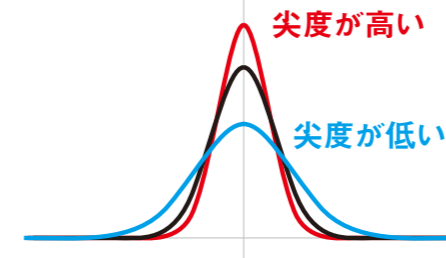
基本統計量は、データ全体の傾向を効率的に把握するための重要な指標です。代表値として平均値、最頻値、中央値を紹介し、それぞれの特徴と用途を説明します。例えば、日本人の年収データを用いて、平均値と中央値の違いを具体的に示します。この比較を通じて、データの分布形状によって適切な代表値が異なることを理解させます。また、外れ値の影響や、データの偏りによる代表値の変化についても触れ、単一の統計量だけでなく、複数の視点からデータを捉えることの重要性を強調します。さらに、基本統計量が実社会でどのように活用されているかの事例も紹介し、学習の動機づけを行います。



出典:厚生労働省 2023(令和5)年 国民生活基礎調査の概況 所得金額階級別みた世帯数の相対度数分布 <https://s.nikkei.com/4h2mY5I>

展開1: 基本統計量ワークシート演習1

標準偏差、尖度、歪度といったより高度な統計量の概念を導入します。これらの統計量がデータの散らばりや分布の形状をどのように表現するかを説明し、データ分析における重要性を理解させます。Google スプレッドシート拡張ツールを用いて、実際のデータセットから各統計量を算出する方法を学びます。例えば、身長や体重のデータを使用し、平均値と標準偏差の関係、尖度や歪度がデータの分布形状をどのように反映するかを視覚的に示します。また、これらの統計量を用いて、正規分布との比較や外れ値の検出方法についても触れます。さらに、産業や学術研究などの分野で、これらの統計量がどのように活用されているかの具体例も紹介し、実践的な理解を促します。



出典:文部科学省 令和4年度 学校保健統計調査 身長と体重の相関表及び身長別体重の平均値 男性15歳 を加工して作成 <https://s.nikkei.com/3Blgs3w>

展開2: 基本統計量ワークシート演習2

SSDSE (教育用標準データセット) を活用し、生徒自身でデータを選択して基本統計量を算出する実習を行います。SSDSE の特徴や利用方法を説明し、生徒が興味を持つデータセットを選択できるようガイドします。例えば、都道府県別の人口データや教育関連データなどから、生徒自身で演習データを設定させます。XLMiner Analysis ToolPak を使用して、選択したデータセットの基本統計量を算出する方法も学びます。算出した統計量の意味を解釈し、データセットの特徴や傾向を説明する練習を行います。これにより、統計的思考力と実践的なデータ分析スキルを養成します。

都道府県	推定人口 (10歳以上)	スポーツの総数	野球(キャッチボールを含む)	ソフトボール	バレーボール	バスケットボール	サッカー(フットサルを含む)	卓球	テニス	バドミントン
全国	112462	66.5	6.3	1.5	3.5	3.6	4.7	4.9	3.4	6.1
北海道	4614	62.2	7.3	1.1	4.1	3.8	4.4	4.8	2.5	6.1
青森県	1090	52.1	5.8	0.6	2.9	4.0	3.5	2.8	2.1	4.4
岩手県	1075	59.1	5.8	1.7	3.6	4.2	4.3	5.0	2.6	5.8
宮城県	2065	64.4	7.0	1.2	4.1	4.7	5.0	4.9	2.7	7.4
秋田県	849	57.1	7.0	0.9	3.8	5.3	3.7	4.8	2.2	5.2

出典:「SSDSE-社会生活(SSDSE-D)」(独立行政法人統計センター) <https://s.nikkei.com/4h7fluH> より抜粋して作成

展開3: 結果の考察

算出したデータから見える傾向や共通項を詳細に検討します。各生徒が選択したデータセットの特徴をワークシートに記述します。例えば、平均値と中央値の差から、データの偏りを推測したり、標準偏差の大きさからデータのばらつきの程度を解釈したりします。また、尖度や歪度を用いて、データ分布の形状を正規分布と比較し、その特徴を考察します。さらに、日経電子版などの信頼できるソースを用いて、自分が分析したデータに関連する最新の記事や研究結果を検索させます。これにより、実社会での統計の活用例を学び、自身の分析結果と比較検討します。この過程で、データ分析の結果が社会にどのような影響を与えるか、また、どのように意思決定に活用されているかを考察させ、データサイエンスについての理解を深めます。

まとめ: グループディスカッション

ワークシートの結果をグループで共有し、互いにフィードバックを行う活動を実施します。各生徒が選択したデータセットと算出した基本統計量、そこから導き出した考察をグループ内で発表します。他のメンバーは、発表内容に対して質問やコメントを行います。この過程で、同じデータセットでも異なる視点からの解釈が可能であることを学び、多角的な分析力を養います。また、各自が選んだデータセットの特徴や傾向を比較し、共通点や相違点を見出す作業を通じて、データの文脈や背景の重要性を理解します。さらに、グループ内で最も興味深い分析結果を選び、クラス全体に発表する機会を設けます。これにより、プレゼンテーションスキルやデータを用いた説得力のある主張の組み立て方も学びます。

授業展開のポイント

データの特性を捉え、客観的な指標としての扱いを理解する

基本統計量を単なる計算結果として捉えるのではなく、データの背後にあるストーリーを読み解くための重要なツールであることを理解させます。平均値、中央値、最頻値といった代表値の違いを、具体的な例を用いて丁寧に説明し、データの分布形状によって適切な代表値が異なることを強調します。また、分散や標準偏差といった散布度を導入することで、データのばらつき具合を定量的に評価できることを示します。

演習を通して統計量の計算と解釈を習得する

Google スプレッドシートなどのツールを活用し、生徒が自ら手を動かしながら統計量を計算する機会を設けます。ワークシートには、様々な種類のデータセットを用意し、生徒が興味を持つテーマを選んで分析できるようにします。計算結果を基に、データの特徴を説明する練習を行い、統計的な思考力を養います。また、グループワークを取り入れ、他の生徒の分析結果と比較することで、多角的な視点からデータを見ることの重要性を認識させます。

実社会のデータに触れる機会を提供する

教育用標準データセット(SSDSE)を用いることで、生徒が実際の社会現象に関わるデータに触れ、統計学の応用性を体感できるようにします。生徒自身が興味を持つテーマを選び、データ分析を行うことで、学習意欲を高めます。また、日経電子版などのメディアで公開されているデータも活用し、統計情報がどのように社会で利用されているかを知る機会を提供します。

データに基づいた意思決定の重要性を認識する

統計学の知識を単に習得するだけでなく、その知識を活かして問題解決に取り組む力を育成します。データ分析の結果を基に、仮説を立てたり、結論を導き出したりする練習を行います。また、データの質や信頼性についても考察し、データに基づいた意思決定の重要性を認識させます。さらに、データサイエンスが社会に与える影響について議論し、倫理的な側面についても触れます。

2-2 散布図作成の授業展開

本節の目標

データの分布や相関関係を視覚化する散布図の意義と作成方法を理解する。Google スプレッドシートを用いて、2つの数値データの関係を散布図で表現し、グラフの編集やカスタマイズ方法を習得する。さらに、データの傾向を解釈し、基本的なデータ分析スキルを身につけることを目指す。

評価規準

知識・技能：散布図の作成方法と相関関係の理解ができる。データセットを適切に扱い、正確なグラフを作成できる。

思考・判断・表現：散布図を分析し、データ間の関係性を考察する。自らの仮説を根拠に基づいて説明できる。

主体的に学習に取り組む態度：自主的にデータセットを探索し、積極的に散布図作成と考察についてのディスカッションに取り組む姿勢を示す。

準備

Google スプレッドシート：表計算ツール
<https://s.nikkei.com/3YIAw4C>

SSDSE：教育用標準データセット
<https://s.nikkei.com/4f7aSpV>

未インストールの場合は

XLMiner Analysis ToolPak：スプレッドシートの分析ツール
<https://s.nikkei.com/4f3Q3LR>

※XLMiner Analysis ToolPak の利用が難しい場合は、Excel などの表計算ソフトで代用できる

授業の組み立て

導入 (5分)：正の相関、負の相関、無相関のグラフを見たことがあるか質問する。それらの例としてどのような事象が当てはまるかを考えさせる。この問いかけを通して、相関関係の概念に触れさせる。

展開 (30分)：Google スプレッドシートを使用し、散布図を作成する演習を行う。共通のデータセットを使用して散布図を作成する演習後に、各自でデータセットを探し、個別に散布図の作成に進む。データの相関関係を観察し、仮説を立てながら演習を進める。

まとめ (15分)：各生徒が使用したデータセットと結果をグループで共有する。最後に、日経電子版で散布図に関する記事を検索し、社会でどのように散布図が活用されているかをリサーチする。

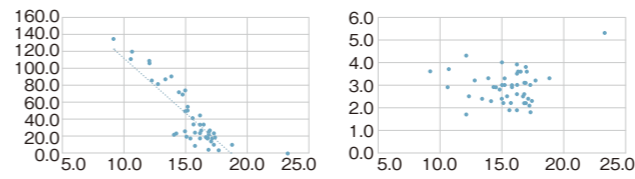
導入：散布図について

データ間の相関関係を視覚的に表現するグラフである散布図について説明します。正の相関では片方の変数が増加するにつれもう一方も増加します(例：身長と体重)、負の相関では片方が増加するにつれもう一方が減少します(例：運動時間と体脂肪率)。無相関の場合は、変数間に関連が見られない(例：身長と読書時間)など、**データの傾向やパターンを簡単に把握することが可能であることについて、基本的な概念から、実際のグラフを見せることで生徒に理解を促します。**



展開1：ワークシート例題から学ぶ

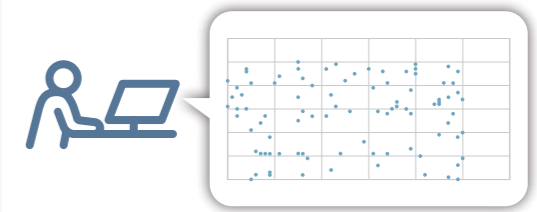
ワークシートの例題から、強い相関と弱い相関の違いを学びます。強い相関では、データ点が直線に近い形で並び、変数間の関係が明確に見えます。**一方、弱い相関ではデータ点がばらばらに散らばっているが、ある程度関係が確認できます。**この過程で、単なるグラフ作成ではなく、グラフから情報を読み取るスキルを養うことを目指します。また、具体的な数値や事例を用いることで、相関関係の理解を深めることができます。



出典：「SSDSE-気候値 (SSDSE-F)」(独立行政法人統計センター)
<https://s.nikkei.com/4eZ4H6Y> より抜粋して作成

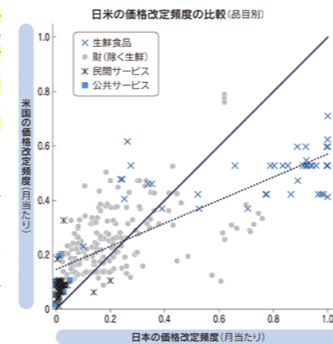
展開2：散布図ワークシート演習

ワークシート例題に取り組んだ後、各生徒は自分でデータを探して散布図を作成します。**教育用標準データセット (SSDSE) を活用し、教育関連のデータを基に散布図を作成する演習を行います。**この作業を通じて、生徒は自分でデータを収集し、分析に取り組む体験を得ます。また、データの扱い方や散布図作成の技術を学ぶと同時に、データに基づいた意思決定の基礎を学びます。



展開3：結果の考察

演習後、作成した散布図を基に、データからどのような傾向や共通点が見られるかを考察します。**具体的には、算出したデータが社会的にどのように活用できるかなど、日経電子版の記事を検索し、散布図がビジネスや社会でどのように使われているかを調べ、実際の応用例を探します。**このリサーチ活動を通じて、データ活用の現実的な側面を理解することができます。



出典：日経電子版「物価の現状と展望 サービス価格の硬直性、顕著」2024年1月12日
<https://s.nikkei.com/4eLwRTq> より抜粋

まとめ：グループディスカッション

グループ内で各自が作成したワークシートを共有し、相互にフィードバックを行います。他の生徒が発見したことや感じたことを参考にし、各自のデータ分析に対する視点を広げます。**また、フィードバックを通じて自分の考えやアプローチを客観的に見直す機会を得ます。**このような相互評価を通じて、生徒たちは自分の学習を深め、より良いデータサイエンスのスキルを身に付けることができます。



授業展開のポイント

相関関係の基本概念的徹底理解

生徒が正の相関、負の相関、無相関の概念を理解できるよう、具体的な事例や視覚的なグラフを活用します。例えば、身長と体重の関係や運動時間と体脂肪率の変化など、日常生活に即した例を用いることで、抽象的な概念を具体化し理解を深めます。また、実際の散布図を示し、各相関の特徴を視覚的に捉えることで、概念の定着を図ることができます。

データセット選択の指導と難易度調整

各自でデータセットを探す際に、信頼性の高いデータを選ぶ方法を指導します。教育用標準データセット (SSDSE) などの信頼できる資料を活用し、生徒が適切なデータを選択できるようサポートします。また、生徒の理解度に応じてデータの難易度を調整し、無理なく学習を進められる環境を整えることが重要です。

データ分析力と仮説検証の育成

散布図の作成に留まらず、データから読み取れる傾向や相関関係を考察する力を養うことを目指します。生徒が自ら仮説を立て、その仮説をデータで検証するプロセスを経験させることで、論理的思考力と分析力を向上させます。また、データの解釈や仮説の検証方法について指導し、実践的なデータサイエンスのスキルを身につけさせることが目的です。相関係数についての具体的な学習は次回に実施することを予告します。

社会的応用の理解と批判的思考の促進

散布図がビジネスや社会でどのように活用されているかを具体的に学ぶことで、学習内容の実社会への関連性を示すことができます。日経電子版の記事を活用し、実際の応用例をリサーチさせることで、データ活用の現実的な側面を理解させます。また、データの解釈における注意点や限界について議論させ、批判的思考を促進することで、情報リテラシーの向上を図ることができます。

2-3 相関係数の授業展開

本節の目標

相関係数の概念と意味を理解し、演習を通じてデータを収集、整理、分析する方法についての技能を身につける。表計算ソフトを使用して相関係数を算出し、散布図や回帰直線との関連を考察する。また、相関と因果の違いを認識し、データの特徴を客観的に分析する視点を養う。

評価規準

知識・技能：相関係数の概念を理解し、散布図から相関の強さを判断できる。相関係数を計算し、その値から関係性を解釈できる。

思考・判断・表現：データの関係性を相関係数を用いて分析し、具体的な数値で説明できる。相関と因果関係の違いを理解し、適切に判断できる。

主体的に学習に取り組む態度：相関分析の意義を理解し、積極的にデータを収集・分析しようとする。結果を批判的に考察し、学びを深めようとする。

準備

Google スプレッドシート：表計算ツール
<https://s.nikkei.com/3YIAw4C>

都道府県データランキング：人口・交通・災害など多彩なデータを対象にしたWeb サイト
<https://s.nikkei.com/3YkylZz>

未インストールの場合は

XLMiner Analysis ToolPak：スプレッドシートの分析ツール
<https://s.nikkei.com/4f3Q3LR>

※XLMiner Analysis ToolPak の利用が難しい場合は、Excel などの表計算ソフトで代用できる

授業の組み立て

導入 (10分)：相関係数の概念を紹介。日常生活での例を示し、興味を喚起。

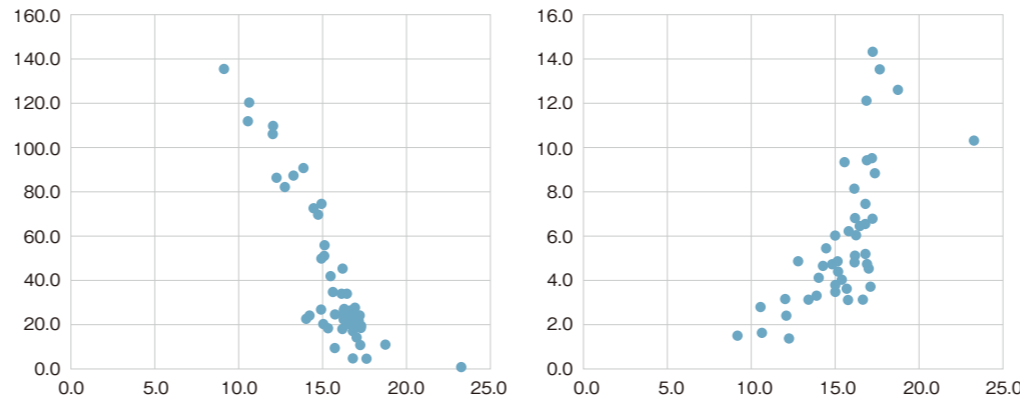
展開1 (40分)：都道府県データランキングを紹介。相関係数の計算方法を説明し、ワークシートで例題を演習。その後、生徒に興味のあるデータを選ばせ、相関についての仮説を立てて相関係数の算出を実施。

展開2 (35分)：グループで分析結果を共有し、考察についてのディスカッションを行う。

まとめ (15分)：各グループの発見と考察を発表。相関と因果関係の違いを強調し、データ分析の重要性をまとめる。

導入1：散布図と相関係数

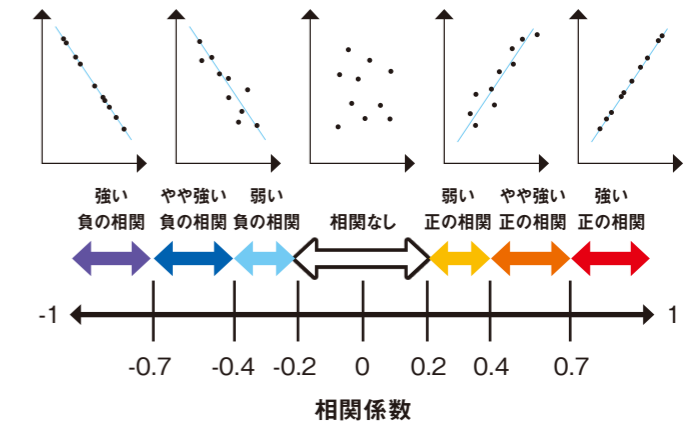
散布図は、2つの変数間の関係性を視覚的に示すグラフです。データ点が右上がりに配置されていれば**正の相関**、右下がりであれば**負の相関**、データがばらついていれば**無相関**を示す内容について、**前回の復習**を。これにより、データ間の関連性を直感的に理解でき、相関係数は、この関係性を数値化し、散布図でのパターンを定量的に評価する手段であることを理解させます。



出典：「SSDSE-気候値(SSDSE-F)」(独立行政法人統計センター) <https://s.nikkei.com/4eZ4H6Y>より抜粋して作成

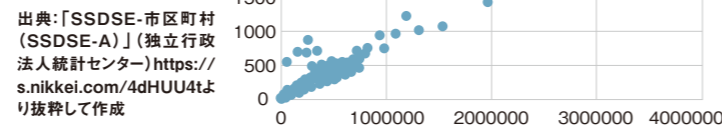
導入2：相関係数の算出

相関係数は、2つの変数間の直線的な関係の強さを示す指標で、値は1から-1の間をとります。計算には表計算ソフトのツールを活用すると効率的です。相関係数が1に近いほど強い正の相関、-1に近いほど強い負の相関、0に近いほど相関がないことを示します。



展開1：都道府県データから演習

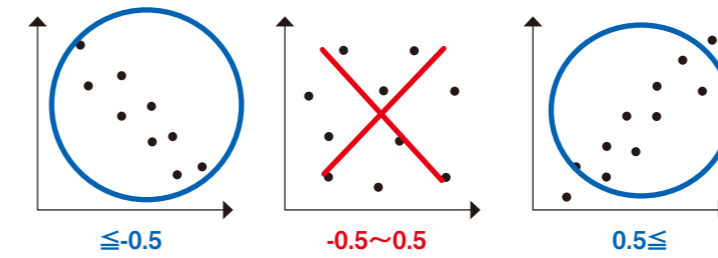
実際の都道府県データを用いて、相関係数の数値が持つ意味を理解する演習を行います。例えば、**家賃と鉄道駅数などのデータを比較し、相関係数を算出**します。これにより、相関係数を用いて2つのデータの相関を数値的に判断する方法を学び、数値が示す関係性を解釈する力を養います。



出典：「SSDSE-市区町村(SSDSE-A)」(独立行政法人統計センター) <https://s.nikkei.com/4dHUU4t>より抜粋して作成

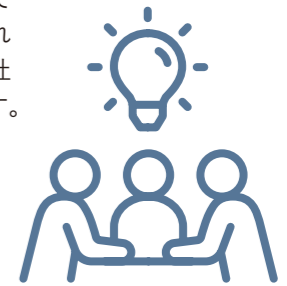
展開2：個人ワーク

各自で都道府県データを分析し、相関係数が0.5以上または-0.5以下の組み合わせを探します。その上で、**なぜ強い相関が出るのかを考察**します。散布図を作成し、データの関係性を視覚的に確認することで、相関の背後にある要因や傾向を深く理解します。



まとめ：ディスカッション

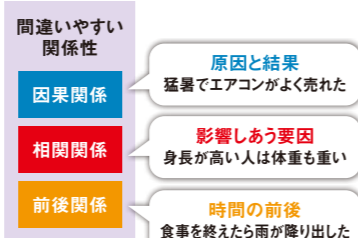
個人ワークの結果をグループ内で共有し、**考察内容について相互にフィードバック**を行います。さらに、日経電子版の記事検索を使って「相関係数」に関する情報をリサーチし、実際にどのような2つの要素間の相関が議論されているかを調べます。これにより、相関係数の実社会での応用例を学びます。



授業展開のポイント

相関係数の概念の導入と興味喚起

相関係数を導入する際には、日常生活での具体的な例を用いて生徒の興味を喚起することが重要です。例えば、気温とアイスクリームの売上、身長と体重など、身近なデータを示すと理解が深まり、主体的な学習意欲を高めることができます。注意点は、相関関係と因果関係を混同させないことです。



相関係数の計算方法とツールの活用

相関係数の計算では、表計算ソフトを活用し、実際にデータを入力して算出するプロセスを体験させることが効果的です。操作手順を明確に示し、生徒同士でサポートしあうように促します。計算結果を散布図と比較し、数値とグラフの関係性を理解させることで、相関係数の意味を深めることができます。

データを用いた実践的演習

実際の都道府県データを使った演習では、生徒自身に興味のあるデータ項目を選ばせることで主体的な学習を促進します。相関係数の数値が持つ意味を理解し、なぜそのような相関が得られたのかを考察させます。注意点は、データの信頼性や偏りについても考慮するよう指導することです。また、結果を生徒同士でプレゼンテーションさせることで表現力も養えます。

ディスカッションを通じた相関係数の理解

個人ワークの結果をグループ内で共有し、相互にフィードバックを行うことで多角的な視点を養います。特に、相関関係が必ずしも因果関係を示すわけではないことを強調することが重要です。生徒が安易に因果関係を結論付けられないよう指導し、データ分析の限界や注意点についても考えさせます。また、日経電子版を活用して、実社会での相関係数の応用例を紹介し、学習内容の定着を図ります。

2-4 単回帰分析の授業展開

本節の目標

データの形式や分析目的に応じた可視化方法を選択し、相関係数や相関関係・因果関係、交絡因子を理解する手がかりをつかむ。Python プログラムでの単回帰分析を通じてデータの関係を数式で表し、データの予測に活用できる力を養うことを目指す。

評価規準

知識・技能: 単回帰分析の手法を理解し、Python を用いて回帰式を求め、予測値を計算できる。

思考・判断・表現: データの相関関係を考察し、適切な分析方法を選択し、結果を明確に表現できる。

主体的に学習に取り組む態度: データ分析に積極的に取り組み、他の指標を探究し、学習を振り返ることができる。

準備

Google Colaboratory: プログラミング言語のPython を、Web ブラウザから直接記述・実行できるサービス
<https://s.nikkei.com/3U96Bdu>

授業の組み立て

導入 (10分): 散布図と相関係数を復習し、それが予測につながることを説明する。単回帰分析の概要を示し、興味を持たせる。

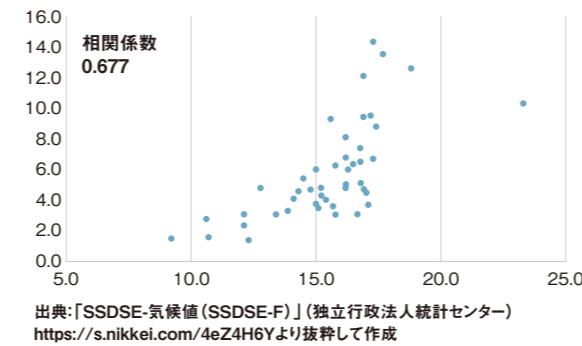
展開1 (40分): Python で為替レートと日経平均株価をグラフ化し、単回帰分析を実施する。

展開2 (35分): データの関係を数式で表し、予測を行う演習を行う。

まとめ (15分): 日経電子版で、日経平均株価と単回帰分析できそうな他の指標を探す。見つけた指標についてグループでディスカッションし、手法の応用を考える。

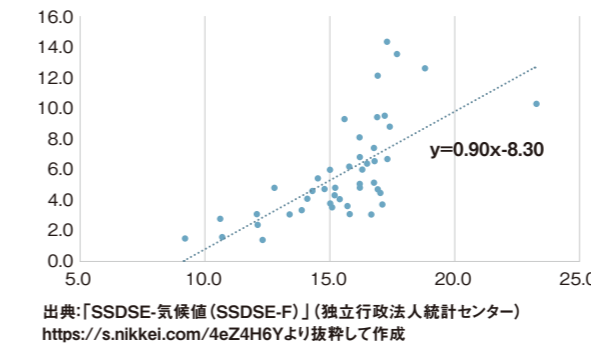
導入1: 散布図と相関係数の復習

まず、**前回までに学習した散布図と相関係数について復習します。** 前回までのワークを例示して、散布図がどのようにデータの分布や相関関係を視覚化するかを確認します。また、相関係数がデータ間の関係性を数値で表す指標であることを再確認し、これらがデータの予測につながるの見通しを立てます。



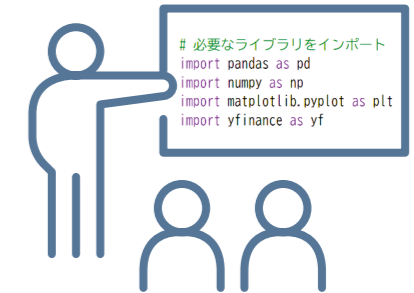
導入2: 単回帰分析の概念説明

単回帰分析の基本的な概念を説明します。**相関関係のある二つの変数を用いて、一方の変数から他方の変数を予測する手法であることを示します。** 回帰直線の方程式や回帰係数の意味を紹介し、データの関係を数式で表現することで予測が可能になることを伝えます。



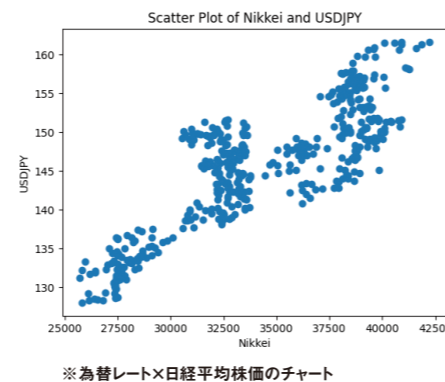
展開1: Pythonでのデータ可視化

Google Colaboratory を用いて、為替レートと日経平均株価のデータを取得し、**Pythonで読み込む演習を行います。** データの前処理や必要なライブラリのインポート方法を一緒に実施します。取得したデータを使って散布図を作成し、データ間の相関関係を視覚的に確認します。



展開2: 単回帰分析の実施と解釈

Python で単回帰分析を実行し、回帰直線の方程式を求めます。**得られた回帰係数や切片の解釈方法を説明し、モデルを用いて為替レートから日経平均株価を予測する方法を示します。** さらに、決定係数からモデルの適合度について説明し、結果の評価に必要なポイントを理解します。



まとめ: 他の指標の検討とディスカッション

日経電子版を活用して、**為替レート以外で日経平均株価と関連性がありそうな指標をリサーチします。** 見つけた指標についてグループでディスカッションし、それらを用いた単回帰分析の可能性を考えます。モデルを変更することで予測結果がどう変化するかを議論し、単回帰分析の応用範囲を理解します。



授業展開のポイント

散布図と相関係数の復習、単回帰分析の導入

まず、前回までに学習した散布図と相関係数を復習します。これらがデータ間の相関関係を視覚化し、数値で表す手段であることを確認し、それがデータの予測につながる見通しを立てます。次に、単回帰分析の基本的な概念を説明します。一方の変数から他方の変数を予測する手法であり、回帰直線の方程式や回帰係数の意味を紹介します。データの関係を数式で表現することで、予測が可能になることを理解します。

Pythonでのデータ可視化と単回帰分析の実施

Google Colaboratory を用いて、為替レートと日経平均株価のデータを取得し、Python で読み込む演習を行います。必要なライブラリのインポートやデータの前処理方法を解説します。取得したデータを使って散布図を作成し、データ間の相関関係を視覚的に確認します。その後、Python で単回帰分析を実行し、回帰直線の方程式を求めます。得られた回帰係数や切片の解釈方法を説明し、為替レートから日経平均株価を予測する方法を示します。

単回帰分析の結果の評価と解釈

単回帰分析の結果について、決定係数の統計指標を用いてモデルの適合度を評価し、妥当性や予測精度について考察します。相関関係と因果関係の違いを復習し、交絡因子の存在がないか、結果をグループで共有して相互に意見交換します。これにより、データ分析における注意点を理解し、より正確な予測を行うための視点を養います。

他の指標の検討とディスカッション

日経電子版を活用して、為替レート以外で日経平均株価と関連性がありそうな指標を各自でリサーチします。見つけた指標についてグループでディスカッションし、それらを用いた単回帰分析の可能性を考えます。モデルを変更することで予測結果がどのように変化するかを議論し、単回帰分析の応用範囲や限界について理解を深めます。これにより、重回帰分析への発展や他の分析手法への興味を喚起します。

3章

データを使った探究

- 3-1 GIS の活用
- 3-2 統計データの活用
- 3-3 地域課題の発見・考察

3章

データを使った探究

本章では、GIS(地理情報システム)であるjSTAT MAPとe-Statの社会統計データの基本操作を学び、統計データを地理的に可視化する意義と活用方法を理解することを目指す。生徒はe-Statから都道府県や市区町村のデータを取得し、jSTAT MAP上で地図データと属性データを組み合わせることで、地理的なパターンや傾向を視覚的に把握し、社会課題への応用力を養う。防災や都市計画などの具体例を通じ、GISと統計データの組み合わせが問題解決にどう貢献するかを実感させる。データサイエンスの基礎的な知識と多角的な解釈力を身につけ、データに基づいた考察力や問題解決力を身につける。

	時間・概要	目標	学習活動	評価	留意点
3-1 GISの活用	時間：50分 概要： GISの基本知識を学び、総務省統計局が提供するjSTAT MAPを活用して、地図データと属性データを組み合わせた統計データの可視化と分析を行う。空間的思考を養い、GISの多様な活用方法を理解することを目指す。	GISであるjSTAT MAPの基本的な機能を身につけ、統計データの地理的可視化の意義や役割を理解することを目指す。 地図データと属性データを活用して、様々な分野での分析や解釈に応用できる力を養う。空間的思考を育成し、WebGISの基本機能や使用時の注意点を理解することで、データサイエンスの基礎を深める。	導入では、GISであるjSTAT MAPの基本的な使い方を学ぶ。展開では、地図データ(地理データ)と属性データ(統計データ)の違いや役割を理解する。jSTAT MAPを使用して、統計データを地図上に可視化する方法を習得する。まとめでは、可視化したデータから空間的なパターンや傾向から考察を行う。	知識・技能： GIS(jSTAT MAP)の基本機能を理解し、基本的な操作ができる。 思考・判断・表現： 可視化されたデータを読み取り、地理的特徴や傾向を考察し、自分の考えを明確に表現できる。 主体的な学習態度： 課題解決に向けてGISの活用方法を積極的に探究する姿勢を持っている。	生徒がjSTAT MAPの操作に不慣れである可能性が高いため、基本操作を丁寧に指導する必要がある。また、データの解釈においては、多面的な考察が重要であることを強調し、単純な結論に飛びつかないよう注意を促す。
3-2 統計データの活用	時間：100分(2時限) 概要： 統計データの基本概念と重要性を理解し、e-Statを活用して都道府県や市区町村のデータを取得・操作するスキルを身につけることを目指す。データの可視化と分析を通じて、社会問題や地域課題の解決にデータを活用できる力を養う。	統計データの基本概念とその重要性を理解し、e-Statを活用してデータを取得・操作するスキルを身につけることを目指す。 e-Statや日経電子版の経済指標ダッシュボードでデータをグラフ化するなど可視化を行い、データの特性を分析・解釈する力を養う。 これにより、生徒は実際の社会問題や地域課題に対してデータを活用し、現実世界の問題解決に貢献できる能力を培う。	はじめに、e-Statの概要と基本的な使い方を演習を通して学習する。次に、取得したデータを表計算ソフトで整形・加工する方法を学ぶ。その後、データの平均値や中央値を計算したり、ダッシュボード機能でグラフ化してデータの傾向を分析する。最後に、分析結果をグループで共有し、データの解釈と社会への応用について議論する。	知識・技能： e-Statを使ってデータを適切に取得し、基本的な統計手法で分析できる。 思考・判断・表現： データの特性を理解し、分析結果から妥当な結論を導き、それを的確に表現できる。 主体的な学習態度： 自ら課題や疑問を見つけ、データの活用と分析に積極的に取り組み、学習を深める姿勢を示す。	生徒がデータの取得や操作に不慣れである可能性があるため、e-Statや表計算ソフトの基本操作を丁寧に指導する必要がある。また、データの解釈においては、単純な結論に飛びつかず、多面的な視点で考察することを促す。
3-3 地域課題の発見・考察	時間：100分(2時限) 概要： e-Statから多様な統計データを検索・取得し、jSTAT MAPに読み込んでエリア作成とプロット作成を組み合わせることで、地域データの可視化を行う。これにより、統計データから地域課題を発見・考察し、データに基づく問題解決の方法と注意点を理解することを目指す。	jSTAT MAPを用いて、e-Statの社会統計データを整形して読み込み、 エリア作成とプロット作成を組み合わせることで地域課題の発見や考察を行う方法と注意点を理解することを目指す。 多様な統計データを検索・取得し、地域データを可視化・分析・解釈する力を養う。これにより、データに基づいた考察力や問題解決に向けたアイデアを創出する能力を身につける。	はじめに、統計データが社会問題の解決に役立つことを具体例で示し、興味を喚起する。その後、e-StatからjSTAT MAPへのデータ取り込みを演習し、使い方を理解する。生徒は興味のある地域の統計データを取得し、エリア作成とプロット作成を組み合わせることで可視化を行う。まとめでは、グループで結果を共有し、地域課題の発見と考察を行う。	知識・技能： e-Statからデータを整形し、jSTAT MAPへ取り込むことができる。エリア作成とプロット作成を用いて地域データを可視化できる。 思考・判断・表現： データを分析・解釈し、地域課題の原因や解決策を論理的に考察・表現できる。 主体的な学習態度： 統計データの活用に関心を持ち、積極的にデータを探索・分析する姿勢を示すことができる。	生徒がデータの整形やjSTAT MAPへの取り込みに困難を感じないように、手順を丁寧に指導する必要がある。また、データの分析・解釈においては、単純な結論に飛びつかず、多角的な視点で考察するように注意を促す。

授業展開のポイント

GISと統計データの基本理解と操作

GIS(地理情報システム)の基本概念を理解し、jSTAT MAPの操作を通じて地理データと統計データの組み合わせを学ぶ。GISの活用が防災や都市計画など多様な分野で役立つ事例を取り上げ、地理的なパターンや傾向を可視化する意義を生徒に伝える。地理データ(位置情報)と属性データ(統計情報)の違いを理解し、データの視覚的な表現方法の基礎を養うことで、データに基づいた意思決定力を高める。

e-Statを用いたデータ取得と整形の技術習得

e-Statから地域の統計データを効率的に検索・取得する手順を学び、データ形式や統計情報の多様性を把握する。さらに、取得データの整形・加工を表計算ソフトで実践し、分析に適した形に加工する技術を習得する。データの欠損値や不要項目の整理などの前処理が正確な分析に不可欠であることを理解させる。生徒は自らのテーマに沿ったデータを選定し、データの整形から分析へのステップを体験することで、データサイエンスの基礎を確立する。

データの可視化と多角的な分析力の育成

整形したデータを基に、jSTAT MAPの統計グラフ機能を用いて視覚化を行う。色分けやプロット作成といった手法を用い、データの分布や密度を視覚的に示すことで、情報の傾向や異常値を発見する練習を行う。データを地図上に配置し、地理的・社会的要因を多面的に考察することの重要性を理解させる。これにより、生徒は日経電子版からデータの裏にある要因をリサーチし、批判的かつ客観的に分析するスキルを身につける。

データに基づく問題発見と解決策の考察

可視化されたデータを基に、グループで地域課題を発見し、データに基づいた解決策を考察する。日経電子版などを活用して関連するニュースや事例を調べ、データの社会的意義や応用例を学ぶ。生徒は自身の分析結果を共有し、意見を出し合い、社会的背景や現実世界との関連を見つけ出すことで、データに基づく合理的な判断力を養う。これにより、データに基づいた課題解決力とプレゼンテーションスキルが向上する。

3-1 GISの活用の授業展開

本節の目標

GIS (jSTAT MAP) の基本知識を身につけ、統計利用における意義や役割を理解し、地図データと属性データを活用して様々な分野での分析や解釈に活用できるようになることを目指す。空間的思考を養い、WebGISの基本機能や注意点を理解する。

評価規準

知識・技能: GIS (jSTAT MAP) の基本的な機能を理解し、操作できる。

思考・判断・表現: 視覚化されたデータを読み取り、地理的特徴や傾向を考察し、自分の考えを明確に表現できる。

主体的に学習に取り組む態度: 課題解決に向けてGISの活用方法を積極的に探究する姿勢を持つ。

準備

jSTAT MAP: 総務省統計局が提供している、誰でも使える地理情報システム <https://s.nikkei.com/3C10lws>

授業の組み立て

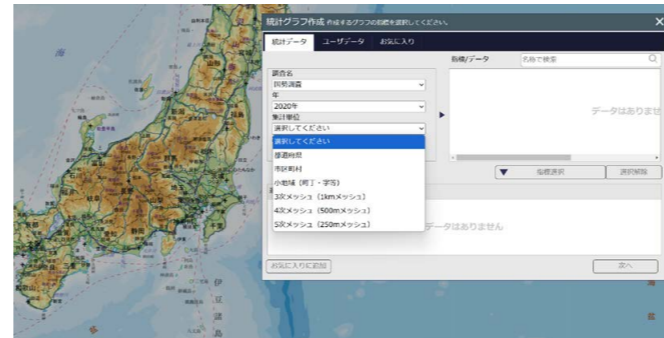
導入 (5分): 統計の利用におけるGISの意義や役割を説明する。GISが防災やマーケティングなど多様な分野で活用されている事例を紹介し、小地域統計をGISで活用する場合の利点から興味の喚起につなげる。

展開 (40分): GISで使用する地図データと属性データについて説明・演習を行う。地図データや統計データを反映する機能を紹介します。国勢調査などの社会統計データの活用について学ぶ。空間的リテラシーとGISの概要を理解し、統計データをGISに読み込ませ、概観するスキルを身につける。

まとめ (5分): 数値である社会統計データを地図上に反映したときに、どのような傾向が表れるかを読み解くためには、多面的に考察する必要があることを伝える。

導入: GISとjSTAT MAPの基本理解

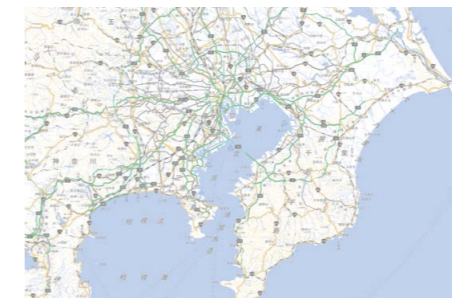
GIS (地理情報システム) の基本概念と、統計データを地理的に可視化する意義を学びます。**GISを利用することで、データの空間的なパターンや傾向を把握できることを説明します。**さらに、総務省統計局が提供するjSTAT MAPの概要と基本的な操作方法を紹介し、生徒が自ら地図を作成する基礎を身につけることを目指します。



jSTAT MAPの概要はこちら <https://s.nikkei.com/3NGX1ul>
jSTAT MAPのマニュアルはこちら <https://s.nikkei.com/3YDzbc1>

展開1: 地理データと統計データの理解

地図データ (地理データ) と属性データ (統計データ) の違いと、それぞれの役割を説明します。**地図データが場所や位置情報を示し、属性データがその場所に関連する統計情報を持つことを理解します。**jSTAT MAPを使用して、地図データに統計データを重ね合わせる方法を実践し、地理情報と統計情報の連携の重要性を学びます。



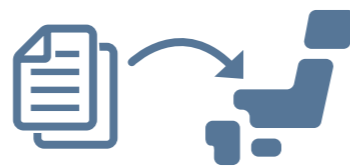
出典:「政府統計の総合窓口(e-Stat)」, jSTAT MAPより作成

都道府県	人口 (A)		
	総数	男	女
北海道	14,047,594	6,898,388	7,149,206
千代田区	66,680	33,637	33,043
中央区	169,179	80,931	88,248
港区	260,486	123,410	137,076
新宿区	349,385	174,822	174,563
文京区	240,069	115,483	124,586
台東区	211,444	108,588	102,856
墨田区	272,085	134,787	137,298
江東区	524,310	258,015	266,295
豊島区	422,488	208,688	213,800
目黒区	288,088	135,820	152,268
大田区	748,081	372,464	375,617
世田谷区	943,664	445,592	498,072
渋谷区	243,883	117,907	125,976
中野区	344,880	172,525	172,355
杉並区	591,108	284,301	306,807
豊田県	301,599	151,020	150,579
光市	355,213	176,289	178,924
厚木市	217,475	107,683	109,792
横浜区	584,483	286,179	298,304
津島区	752,608	361,770	390,838
岡谷市	695,043	347,408	347,635
岡谷市	453,093	225,758	227,335
江戸川区	697,932	351,327	346,605
八王子市	579,355	291,307	288,048
立川市	183,581	91,191	92,390

出典:「政府統計の総合窓口(e-Stat)」, 令和2年国勢調査 / 人口等基本集計(総務省)より作成

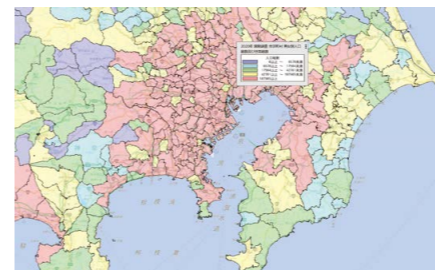
展開2: 統計データの読み込みと管理

地理データと統計データの基本的な違いと役割を学びます。地理データは地球上の特定の位置情報であり、地図上に表現できることを説明します。一方、統計データは人口や経済などの数値情報で、**地理データと組み合わせることで地域ごとの特徴を明らかにできます。**jSTAT MAPを使用して、地理データに統計データを重ね合わせる体験を通じ、データの関連性とGISの基本的な活用方法を理解することを目指します。



まとめ: データの可視化と分析

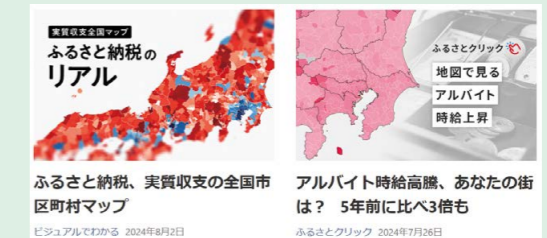
jSTAT MAPの基本機能を使って、統計データを地図上に可視化する方法を学びます。具体的には、**色分けやグラフ表示を用いてデータの分布を視覚的に表現します。**また、可視化されたデータから空間的な傾向や関連性を解釈し、地理的な要因がデータに与える影響を考察する力を養います。



出典:「政府統計の総合窓口(e-Stat)」, jSTAT MAPにて令和2年国勢調査データを用いて作成

日経電子版活用のワンポイントアドバイス ビジュアルデータ

日経電子版のビジュアルデータでは様々なデータを可視化した事例を確認できます。



出典:日経電子版 ビジュアルデータ <https://www.nikkei.com/special/vdata/archive>

授業展開のポイント

GISとjSTAT MAPの基本理解

GIS (地理情報システム) の基本概念を理解し、統計データを地理的に可視化する意義を把握します。GISを活用することで、データの空間的なパターンや傾向を視覚的に分析し、問題解決や意思決定に役立てることができることを示します。防災やマーケティング、都市計画など、GISが様々な分野で活用されている事例を紹介し、その重要性を強調します。さらに、総務省統計局が提供するjSTAT MAPの概要と基本操作方法を学び、生徒が自ら地図を作成し、統計データを可視化する基礎を身につけることを目指します。これにより、生徒は空間的思考を養い、WebGISの基本機能や注意点を理解し、データサイエンスの基礎を深めることができます。

地理データと統計データの理解

地図データ (地理データ) と属性データ (統計データ) の違いと役割を理解します。地図データは場所や位置情報を示し、属性データはその場所に関連する人口や経済などの統計情報を持つことを説明します。具体的な例として、都道府県別の人口データを地図上に表示し、地域間の人口分布を視覚的に理解する方法を示します。jSTAT MAPを使用して、地図データに統計データを重ね合わせる方法を実践し、地理情報と統計情報の連携がデータ分析において重要であることを理解させます。これにより、生徒はデータの空間的な関連性を把握し、より深い分析や考察が可能になることを学びます。

統計データの読み込みと管理

社会統計データの基本的な読み方を学び、jSTAT MAPでの可視化に必要なスキルを習得します。国勢調査などの統計データがどのように収集・整理されているかを理解し、都道府県や市区町村などの地域区分がデータに与える意味を把握します。具体的なデータの例を用いて、jSTAT MAPに統計データを正しく読み込む手順を実践します。データの形式や項目名に注意し、正確に可視化するためのポイントを学びます。これにより、生徒は社会統計データを効果的に活用し、jSTAT MAPでのデータ可視化を通じてデータサイエンスの基礎を深めることができます。

データに基づいた意思決定のための 分析力を高める

jSTAT MAPの基本機能を活用し、統計データを地図上に可視化する方法を学びます。色分けやグラフ表示、リッチレポートなどの手法を用いて、データの分布や密度を視覚的に表現する技術を習得します。具体的な事例として、人口密度や高齢者割合、所得水準などのデータを地図上に表示し、地域ごとの特徴や傾向を分析します。また、可視化されたデータから空間的なパターンや関連性を読み取り、地理的要因や社会的要因がデータに与える影響を考察する力を養います。これにより、生徒はデータを多面的に解釈し、データに基づいた意思決定を行うための分析力を高めます。

3-2 統計データの活用の授業展開

本節の目標

統計データの基本概念と重要性を理解し、e-Statを活用して都道府県や市区町村のデータを取得・操作するスキルを身につける。具体的には、e-Statで取得したデータを用いてグラフにするなど可視化を行い、データの特性を分析・解釈する力を養う。これにより、生徒は実際の社会問題や地域課題に対してデータを活用し、現実世界の問題解決に貢献できる能力を培う。

評価規準

知識・技能：e-Statを使って都道府県や市区町村のデータを適切に取得し、基本的な統計手法で分析できる。

思考・判断・表現：データの特性を理解し、分析結果から妥当な結論を導き、それを的確に表現できる。

主体的に学習に取り組む態度：自ら課題や疑問を見つけ、データの活用と分析に積極的かつ継続的に取り組み、学習を深める姿勢を示す。

準備

e-Stat：総務省統計局が提供している、各府省が公表する統計データを一つにまとめたポータルサイト
<https://s.nikkei.com/4fngkFj>

Excelなどの表計算ソフト

授業の組み立て

導入 (5分)：統計データの基本概念と重要性を説明し、統計が社会で果たす役割を理解させ、興味の喚起につなげる。

展開 (70分)：e-Statの使い方を演習し、都道府県・市区町村データの取得方法を学ぶ。対象のデータはe-Stat上でグラフなどの可視化を行い、データの特性について議論する。生徒にデータの傾向を分析させる。

まとめ (20分)：実際の社会問題に関連するデータ分析例を示し、データの解釈力を養う。学習内容を振り返り、統計データの活用が問題解決に役立つことを考察する。

展開1：e-Statの紹介とデータ取得の基本

最初に、e-Statの概要と基本的な使い方を学習します。具体的には、都道府県や市区町村の統計データをどのように検索し、ダウンロードするかを実演します。**データの形式や種類、利用可能な統計情報の多様性についても解説します。**生徒は自身の興味やテーマに沿ったデータを選定し、データ取得の基本スキルを習得します。e-Statの使い方を理解し、データの取り扱いに慣れることを目指します。



e-Statのマニュアルはこちら <https://s.nikkei.com/3YEGRBf>

展開2：データの整形と加工

取得したデータを分析に適した形に整形・加工する方法を学びます。具体的には、**表計算ソフトを用いて不要な項目の削除や必要なデータの整理などを行います。**データの欠損値についても紹介します。この段階で生徒は、データの前処理の重要性を理解し、分析に必要なデータ準備のスキルを身につけます。

就業状態	地域コード	地域	時間軸(年)	時間軸(年/年齢階級)	15~64歳	15~24歳	25~34歳	35~44歳	45~54歳	55~64歳	
15歳以上人口	00000	全国	2E+09	2000年	***	1617	1872	1586	1937	1643	
15歳以上人口	00000	全国	2E+09	2001年	***	1573	1907	1576	1953	1615	
15歳以上人口	00000	全国	2E+09	2002年	***	1529	1896	1595	1894	1662	
15歳以上人口	00000	全国	2E+09	2003年	***	1493	1883	1625	1808	1731	
15歳以上人口	00000	全国	2E+09	2004年	***	1457	1862	1650	1731	1812	
15歳以上人口	00000	全国	2.01E+09	2005年	***	1420	1836	1679	1664	1863	
15歳以上人口	00000	全国	2.01E+09	2006年	***	1389	1797	1715	1614	1880	
15歳以上人口	00000	全国	2.01E+09	2007年		8312	1358	1726	1754	1584	1890
15歳以上人口	00000	全国	2.01E+09	2008年		8243	1332	1674	1792	1563	1881
15歳以上人口	00000	全国	2.01E+09	2009年		8165	1307	1622	1823	1556	1857
15歳以上人口	00000	全国	2.01E+09	2010年		8113	1281	1574	1836	1561	1861

グレーは不要な項目、赤は欠損値

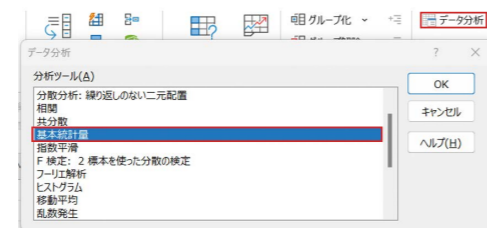
出典：「政府統計の総合窓口(e-Stat)」、労働力調査(総務省) 就業状態、年齢階級別15歳以上人口(2000年～) より作成

展開3：データの分析

e-Statのデータを用いて基本的な統計分析を行います。平均値や中央値などを計算し、データの傾向を把握します。例えば都道府県別の世帯数の平均値や中央値を示します。中央値よりも平均値の値が大きいことを示し、**日本の人口が都市に集中する傾向にあることなどを説明します。**ここで、データから有益な情報を得る力を養います。

都道府県別世帯数(単位：千)	
平均	1155.532
標準誤差	185.8347
中央値 (メジアン)	706
最頻値 (モード)	398
標準偏差	1274.019
分散	1623124
尖度	6.720235
歪度	2.444031
範囲	6358
最小	221
最大	6579
合計	54310
データの個数	47

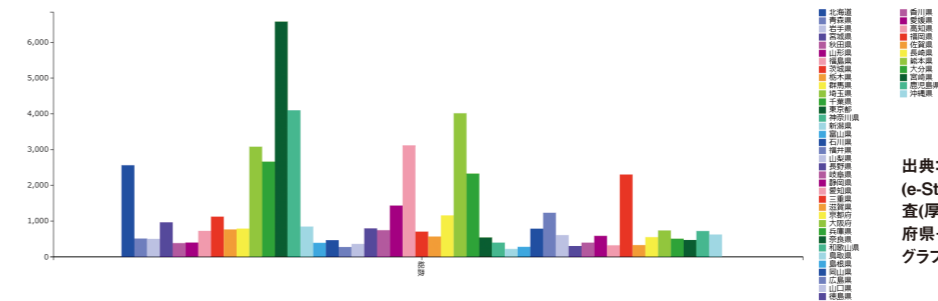
出典：「政府統計の総合窓口(e-Stat)」、令和4年国民生活基礎調査(厚生労働省) 世帯数、市都道府県-21大都市(再掲)別 より作成



Excelの「データ分析」機能を使うと基本統計量が簡単に出来ます。
参考：総務省統計局 なるほど統計学園
<https://s.nikkei.com/481Y3pm>

展開4：データの可視化

分析を行ったデータを用いて、e-Statのグラフ機能を使ってデータを可視化する方法を学びます。これにより、**生徒はデータの分析結果を視覚的に表現するスキルを身につけます。**ここでは、データの特徴を視覚的に理解する力を養うことを目指します。



出典：「政府統計の総合窓口(e-Stat)」、令和4年国民生活基礎調査(厚生労働省) 世帯数、市都道府県-21大都市(再掲)別をe-Statのグラフ機能で加工して作成

まとめ：結果の解釈と発表

分析した結果をどのように解釈するか、グループでディスカッションを行います。また、日経電子版を活用して、国勢調査などのデータが社会でどのように活用されているか、記事のリサーチを行います。これにより、**データの社会的意義や実際の応用例を理解し、自身の分析に活かすことができます。**ここでは、データに基づく合理的な判断力と表現力を養うことが目的です。

授業展開のポイント

e-Statの基本理解とデータ取得スキルの習得

e-Statの概要と基本操作を理解し、統計データの取得方法を学びます。都道府県や市区町村のデータを検索・ダウンロードする手順を実践し、データ形式や統計情報の多様性を把握します。生徒は自身の興味に基づいたデータを選定し、データ取得の基本スキルを身につけます。これにより、データの取り扱いに慣れ、データサイエンスの基礎を築くことができます。

データの整形と加工の技術習得

取得したデータを分析しやすい形に整形・加工する方法を学びます。表計算ソフトを用いて不要な項目の削除や必要なデータの整理を行います。生徒はデータ前処理の重要性を理解し、分析に必要なデータ準備のスキルを身につけます。これにより、正確なデータ分析の基盤を作ることができます。

データ分析と可視化の実践

整形したデータを用いて基本的な統計分析を行います。平均値や中央値の計算を通じてデータの傾向を把握し、e-Statのグラフ機能で可視化する方法を学びます。生徒は分析結果を視覚的に表現するスキルを身につけ、データから有益な情報を抽出し理解する力を養います。これにより、データの洞察力を高めることができます。

結果の解釈と社会への応用理解

分析結果を正確に解釈し、グループディスカッションで共有します。日経電子版を活用して、国勢調査などのデータが社会でどのように活用されているかを調査します。生徒はデータの社会的意義や応用例を理解し、自身の分析に反映させます。これにより、データに基づく合理的な判断力と表現力を養い、問題解決に活かすことができます。

3-3 地域課題の発見・考察の授業展開

本節の目標

jSTAT MAP を用いて e-Stat の社会統計データを整形して読み込み、エリア・プロット作成を組み合わせることで地域課題の発見や考察を行う方法と注意点を理解する。e-Stat を用いて多様な統計データを検索・取得し、jSTAT MAP で地域データの可視化を行い、得られたデータから意味を読み解き、分析・解釈する力を養う。これにより、データに基づいた考察力や問題解決に向けたアイデアを創出する力を身につけることを目指す。

評価規準

知識・技能：e-Stat からのデータを整形し、jSTAT MAP へ取り込むことができる。エリア・プロット作成から分析を用いて地域データを可視化できる。

思考・判断・表現：データを分析・解釈し、地域課題の原因や解決策を論理的に考察・表現できる。

主体的に学習に取り組む態度：統計データの活用に関心を持ち、自ら積極的にデータを探索・分析し、学習に取り組む態度を示すことができる。

準備

jSTAT MAP：総務省統計局が提供している、誰でも使える地理情報システム <https://s.nikkei.com/3C10lws>

e-Stat：総務省統計局が提供している、各府省が公表する統計データを一つにまとめたポータルサイト <https://s.nikkei.com/4fngkFj>

授業の組み立て

導入 (10分)：統計データが日常生活や社会問題の解決に役立つことを具体例で示し、生徒の興味を引き出す。

展開 (75分)：e-Stat から jSTAT MAP へのデータの取り込みを演習して使い方を理解する。その後、生徒が興味ある地域の統計データを検索・取得し、jSTAT MAP でのエリアとプロット作成を組み合わせ、地域データの可視化について探究的な視点でワークを行う。

まとめ (20分)：グループで結果を共有し、お互いのデータから地域課題の発見と考察を行う。日経電子版を活用して関連記事をリサーチする。学習内容を振り返り、統計データの活用が問題解決に役立つことを考察する。

導入：社会統計データの探索と新たな発見

前回の授業で学んだ e-Stat の基本操作を復習しながら、まだ見たことのない社会統計データの種類を探索します。生徒は e-Stat を用いて、自身の興味や関心に合わせた新しい統計データやテーマを検索します。これにより、統計データの多様性を理解し、新たな視点でデータを捉える力を養うことを目指します。

e-Stat のマニュアルはこちら <https://s.nikkei.com/3YEGRbF>



地域	人口	世帯数	人口1000人あたり	世帯数1000人あたり	人口1000人あたり	世帯数1000人あたり
全国	125,814,000	47,200,000	125,814	47,200	125,814	47,200
北海道	5,219,000	1,895,000	5,219	1,895	5,219	1,895
東北	10,610,000	3,750,000	10,610	3,750	10,610	3,750
関東	46,940,000	16,500,000	46,940	16,500	46,940	16,500
中部	22,810,000	8,100,000	22,810	8,100	22,810	8,100
近畿	21,100,000	7,400,000	21,100	7,400	21,100	7,400
中国	12,100,000	4,100,000	12,100	4,100	12,100	4,100
四国	3,500,000	1,100,000	3,500	1,100	3,500	1,100
九州	12,500,000	4,300,000	12,500	4,300	12,500	4,300

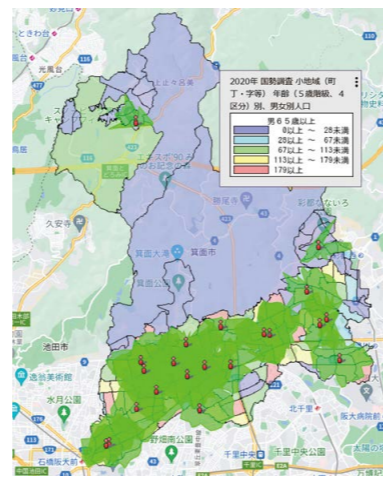
出典：「政府統計の総合窓口(e-Stat)」、平成30年住宅・土地統計調査(総務省) 住宅の構造等に関する集計を加工して作成

展開2：プロット分析の活用とデータ間の関連性の発見

エリア分析に加えて、プロット分析の機能を学習します。生徒は地域のオープンデータを活用してプロットし、偏りや傾向を探ります。例えば、AED の設置箇所や、避難所の分布などのデータを選び、地図上でその関係性を視覚化します。これにより、データ間の関連性を発見し、深い分析力を養うことを目指します。

jSTAT MAP のデータインポート、プロットはこちら <https://s.nikkei.com/4e4CV8u> (10分40秒～)

出典：「政府統計の総合窓口(e-Stat)」、jSTAT MAP に令和2年国勢調査データ、大阪防災ネット 避難所一覧をプロットして作成



まとめ：地域課題の発見とデータに基づく考察

エリア分析とプロット分析で得られた結果を基に、地域課題の発見と考察を行います。生徒はグループでデータを共有し、地域が抱える社会的・経済的課題について議論します。さらに、日経電子版の地域記事から関連情報をリサーチし、データと現実の問題を結びつけます。これにより、データと実情に基づいた問題解決のアイデアを創出する力を身につけます。



授業展開のポイント

e-Stat の基本操作を復習し、新たな統計データの検索方法を学ぶ

e-Stat の基本操作を復習することで、既存の知識を確認し、新たな統計データの検索方法を学習します。具体的には、e-Stat の検索機能を活用して、目的の統計情報を効率的に探し出す方法を習得します。また、検索結果の絞り込みや条件設定のテクニックを学ぶことで、必要なデータを迅速に取得できるようになります。e-Stat を活用した効果的なデータ探索能力を高めることが、この学習ポイントの目的です。

jSTAT MAP にデータをとり込む方法を習得する

データの形式変換やインポート手順を学び、jSTAT MAP に取り込む方法を習得します。具体的には、e-Stat から取得した CSV データを整形し、jSTAT MAP が認識できる形式に整える手順を理解します。取り込みに必要なデータの調整など、前処理の重要性と方法を学びます。これにより、データの互換性を確保し、分析に必要なデータを正しく取り込むスキルを身につけることを目指します。

jSTAT MAP で地理的なデータの傾向を視覚的に理解する

jSTAT MAP のエリア作成機能を用いて、地理的なデータの傾向を視覚的に理解します。具体的には、地図上に統計データを重ね合わせることで、地域ごとの数値の違いや分布を直感的に把握する方法を学びます。また、グラフ表示などを用いた表現方法を習得し、情報を効果的に伝えるスキルを磨きます。これにより、データの空間的な特徴や地域間の比較を容易に行うことが可能となることを理解します。

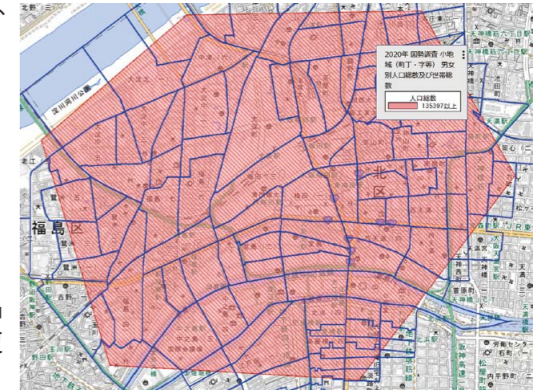
プロット作成の機能を学習し、データ間の関連性を発見する

プロット作成の機能を学習し、データ間の関連性を発見します。具体的には、エリア作成に重ねたプロットを可視化することで、地域間での偏りや傾向を見つけ出す方法を理解します。データの可視化を通じて、どのような社会背景が存在するかなど、日経電子版の地域記事からリサーチを行います。これにより、データと実情に基づいた論理的な考察と問題解決のためのアイデアを生み出す力の育成を目指します。

展開1：jSTAT MAP へのデータ取り込みとエリア分析の実践

取得した新しい統計データを jSTAT MAP に取り込み、エリア分析を行います。データの形式変換やインポート手順を学び、地図上で地域ごとのデータを可視化します。生徒は自分が選んだデータを用いて、特定の地域の統計情報を地図上に表現します。これにより、地理的なデータの傾向を視覚的に理解し、エリア分析の基礎を習得します。

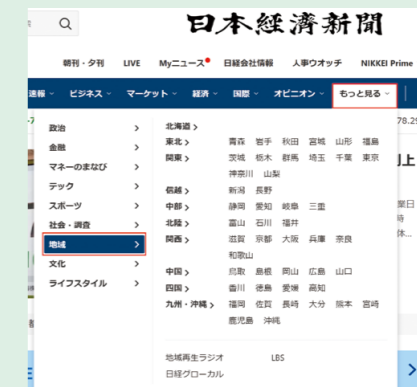
jSTAT MAP のエリア作成方法は [こちら](https://s.nikkei.com/48NsdC) <https://s.nikkei.com/48NsdC> (5分30秒～)



出典：「政府統計の総合窓口(e-Stat)」、jSTAT MAP に令和2年国勢調査データを用いて作成

日経電子版活用のワンポイントアドバイス 地域別の情報

日経電子版のトップページで「もっと見る」→「地域」を選択すると、各地域のニュースを調べることができます。



出典：日経電子版トップページより

4章

データの分析・解釈

- 4-1 重回帰分析
- 4-2 データの整形
- 4-3 探究課題：GIS ×統計分析
- 4-4 検証と分類

4章 データの分析・解釈

本章の目標は、高校生が「情報II」で学ぶデータサイエンスにおける統計分析やデータ可視化の実践的なスキルを、探究的な学びを通して身につけることである。重回帰分析や分類などの統計手法を理解し、さらにGISを活用したデータの地理的可視化を通じて、多面的な分析能力を養うことを目指す。これにより、現実社会の課題を客観的かつ論理的に捉え、データを活用した意思決定や課題解決の力を育成する。また、データの取得から分析、解釈、応用に至るプロセスを探究的に学習し、データリテラシーを深化させる。

	時間・概要	目標	学習活動	評価	留意点
4-1 重回帰分析	時間：150分（3時限） 概要： 重回帰分析の基本を実践的に学ぶ。単回帰分析との違いを理解し、複数の説明変数を活用して予測精度を向上させる手法を習得する。実データを用いてモデルを構築・評価し、統計的な考察を深める。	複数の説明変数を用いた重回帰分析を学び、統計分析の基礎を身につける。表計算ソフトの統計分析ツールを活用して 実データからモデルを構築し、重回帰式の解釈や適合度評価の方法を理解する 。これにより、統計手法を活用した実践的な課題解決力を高める。	単回帰分析と重回帰分析の違いを例を用いて学ぶ。サンプルデータを用いて補正R2や有意性の判定基準となるp値を算出し、e-Statから得た統計データでモデルを構築する。結果を解釈し、改善方法を検討する。最後に統計的視点で考察を行う。	知識・技能： 重回帰分析の基礎概念を理解し、表計算ソフトの統計分析ツールを用いて分析できる。 思考・判断・表現： データを根拠に考察・説明し、結果の改善策を提案できる。 主体的な学習態度： データ分析に積極的に取り組み、考察を深める意欲を示している。	単回帰分析と重回帰分析の基礎を明確に伝える。補正R2や有意性の判定基準となるp値の解釈を丁寧に行い、分析結果の妥当性を説明する場面を設ける。使用するデータの出典や信頼性にも留意する。
4-2 データの整形	時間：50分 概要： 入手した社会統計データを整形し、WebGISツールへのインポートを実施してデータを可視化する。列名やデータ形式の統一、住所コードや緯度経度の整理を実践し、データを地図上の視点から扱うスキルを身につける。	データ整形の重要性を理解し、表計算ツールを用いた前処理を習得することを目標とする。 列名や単位の統一、欠損値処理、住所コードや緯度経度の整理 など、GIS対応形式への変換を学び、データから有益な情報を引き出す力を養う。これにより、空間的な視点を活用した多面的なデータ解釈力を高める。	まず、e-Statや市区町村オープンデータからデータを取得し、表計算ソフトで不要な列の削除や形式の整形を行う。その後、住所コードや緯度経度があるものは整理して、WebGISへインポートする。最後に、データ整形の重要性やつまづいた箇所について振り返りを行う。	知識・技能： データ整形の基本概念を理解し、表計算ソフトで前処理を適切に行える。 思考・判断・表現： 整形したデータを用いて情報を可視化し、その有用性を説明できる。 主体的な学習態度： データ処理に積極的に取り組み、試行錯誤を重ねて問題を解決しようとする意欲を示している。	データ整形では列名や単位の統一、欠損値・異常値の処理、住所コードや緯度経度の確認など、一貫したフォーマットを整えることが重要となる。これにより、WebGISへのスムーズなインポートにつながる。正確な分析を行うためには、前準備が必要であることを実践的な学びから理解させる。
4-3 探究課題： GIS× 統計分析	時間：400分（8時限） 概要： jSTAT MAPを活用し、統計データを地図上に可視化して地域課題を探究的に分析する。地理情報と統計分析を組み合わせることで、社会課題の多面的な要因を探究し、仮説を立てて重回帰分析の手法を用いて、統計的な検証を行う方法を学ぶ。	地理情報と統計データを活用し、地域課題の背景を探究する力を育成することを目標とする。具体的には、WebGISツールである、 jSTAT MAPを使ったデータの可視化や重回帰分析を通じて、統計的手法と地理的視点を融合した問題解決の方法を学ぶ 。これにより、データサイエンスの実践的な応用力を高める。	まず、地域課題に関する仮説を立て、jSTAT MAPを活用してデータを可視化する。その後、Googleスプレッドシートを用いてデータを整理し、重回帰分析を行う。最後に、分析結果をポスターやスライドにまとめ、グループで発表とフィードバックを行う。	知識・技能： WebGISツール（jSTAT MAP）を活用し、統計データを地図上に可視化できる。 思考・判断・表現： 重回帰分析を用いて課題を分析し、考察を論理的に表現できる。 主体的な学習態度： 興味を持って課題を設定し、データ分析を通じて探究を深める姿勢を示している。	入手するデータの信頼性はあるか、また、データの整形時に誤ってデータを書き換えたり、消去したりしていないかなどを確認し、分析の前処理が重要であることを理解させる。また、地理的な特性を考慮する視点を指導する。分析後にデータにおける誤解や偏りがいないかチェックを行う。
4-4 検証と分類	時間：100分（2時限） 概要： 多重共線性と分類の基本的な概念を学ぶ。データ分析の信頼性や妥当性を確認する方法を理解し、分類を活用してデータをカテゴリに分ける考え方を習得する。	多重共線性の影響と分類の基礎を理解することを目指す。 説明変数間の相関が分析結果に与える影響を考察し、回帰と分類の違いを学ぶ 。さらに、分類データに関連した内容を日経電子版からリサーチして、実社会の課題解決に活かす視点を身につける。	まず、多重共線性の概念と影響を具体例で学び、探究課題で使用したデータを基に影響を議論する。次に、分類についてクイズやワークを通して学ぶ。例えば、架空データを用いて分類基準を設定し、結果を共有する。最後に、分類結果の応用可能性を考察する。	知識・技能： 多重共線性や分類の基本概念を理解し、提供されたデータを分析できる。 思考・判断・表現： データの特徴を考察し、分析結果を論理的に説明する能力を評価する。 主体的な学習態度： データ分析やディスカッションに積極的に取り組み、探究心を示している。	多重共線性の影響を直感的に理解できる例を用いる。分類と回帰の違いを明確に説明し、データ分析が実社会の課題解決にどのように役立つかを議論の中で強調する。

授業展開のポイント

単回帰から重回帰へのステップアップ

単回帰分析から発展して、重回帰分析で複数の要因を使った分析の考え方を学ぶ。例えば、売上高を価格や広告費、天候などいろいろな要因から予測することで、一つの要因だけでは分からなかった結果を見つける。これにより、データ分析の可能性が広がることを実感する。

データ整形の大切さを知る

データをそのまま使うのではなく、必要な情報を整理したり形をそろえたりする準備が大切であることを学ぶ。例えば、e-Statから取ったデータを使いやすく加工し、正確な分析ができるようにする。この活動を通じて、データを整えることの重要性を理解する。

地図とデータを合わせた新しい視点

WebGIS（地理情報システム）ツールである、jSTAT MAPを使ってデータを地図に重ね、地域ごとの特徴を目で見て分かるようにする。例えば、人口密度と公共施設の分布を地図に表し、「どこに課題があるのか」を見つける。データを地図と組み合わせることで、新しい視点からの分析ができることを学ぶ。

データの正しさを検証し、活用する力を育む

データの中にある問題を見つけたり、データをグループに分けたりする方法を学ぶ。例えば、多重共線性（データの中で似たような要素があること）を考えながら分析を進める。さらに、分類の技術を使って、データから課題解決に役立つヒントを得る力を育てる。

4-1 重回帰分析

本節の目標

単回帰分析を拡張することで重回帰分析の概念を理解し、複数の説明変数による予測精度向上を体験する。表計算ソフトの統計分析ツールを活用し、実データからモデルを構築して重回帰式を解釈し、モデルの適合度や予測精度の評価・改善手法を習得する。

評価規準

知識・技能: 重回帰分析の基本概念・用語を理解し、表計算ツールでモデルを構築し、当てはまりのよさ・予測精度を確認できる。

思考・判断・表現: 得られたモデルについての考察や改善策を考え、データの根拠に基づき適切に説明できる能力がある。

主体的に学習に取り組む態度: データに主体的かつ積極的に関わり、考察・改良を継続的に行う姿勢がある。

準備

Google スプレッドシート: 表計算ツール
<https://s.nikkei.com/3YIAw4C>

e-Stat: 総務省統計局が提供している、各府省が公表する統計データを一つにまとめたポータルサイト
<https://s.nikkei.com/4fngkFj>

未インストールの場合は

XLMiner Analysis ToolPak:

スプレッドシートの分析ツール

<https://s.nikkei.com/4f3Q3LR>

※XLMiner Analysis ToolPak の利用が難しい場合は、Excel などの表計算ソフトで代用できる

授業の組み立て

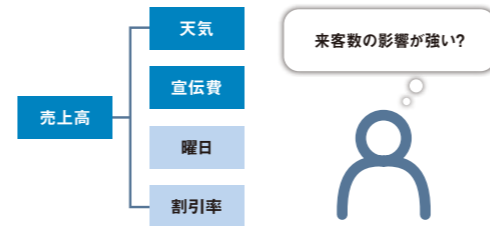
導入 (10分): 単回帰分析と重回帰分析の違いを示し、説明変数と目的変数の関係性を推定するための統計手法であることを例を挙げて理解させる。

展開 (125分): サンプルデータで練習した後、e-Stat の実データを用いて重回帰モデルを構築し、偏回帰係数・決定係数に基づく当てはまりの良さを確認し、考察を行いモデル改善を試みる。

まとめ (15分): モデルの解釈上の留意点や実社会応用場面について、ディスカッションを通じて考えることで、生徒の理解を深める。

導入: データの背景理解と目的明確化

これまでに学んだ単回帰分析を復習し、説明変数が一つのみでは予測精度や分析の材料が限定的になりやすいことを確認します。例えば、売上高を価格、宣伝費、天気など複数の要因で考えると、単一変数では見えなかった関係が明らかになります。このように、**複数の説明変数を同時に取り扱う重回帰分析を導入することで、より多面的な分析が可能になります。** 重回帰分析の目的と必要性を明確にすることで、これから学ぶ内容に納得感をもって取り組みます。



展開2: 重回帰モデルの構築

選定した説明変数を用いて重回帰モデルを構築します。表計算ソフトの統計分析ツールを使い、回帰分析機能を実行します。この際、**偏回帰係数や決定係数といった評価指標が得られ、説明変数の重みやモデル当てはまりの程度が確認できます。** また、有意性の指標であるp値からは、各変数が結果へ及ぼす影響について数値で確認できます。実際の操作を通して、複数の変数を考慮するメリットを実感し、重回帰分析がデータに基づく予測・要因分析の有効な手法であることを理解できます。

回帰統計		自由度	変動	分散	観測された分散比	有意 F	
重相関 R	0.98143438	回帰	3	448.1422271	149.3807424	375.3017498	7.59766E-31
重決定 R2	0.963213443	残差	43	17.11521975	0.398028366		
補正 R2	0.960646939	合計	46	465.2574468			
標準誤差	0.630894893						
観測数	47						

	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	17.414395840	4.727810214	3.683395706	0.000638985	7.879857862	26.948933818	7.879857862	26.948933818
緯度	-1.419641350	0.065423958	-21.699105320	0.000000000	-1.551581335	-1.287701365	-1.551581335	-1.287701365
経度	0.278178174	0.047760760	5.824408436	0.000000658	0.181859422	0.374496927	0.181859422	0.374496927
標高	-0.010381241	0.001284168	-8.084019445	0.000000000	-0.012971013	-0.007791469	-0.012971013	-0.007791469

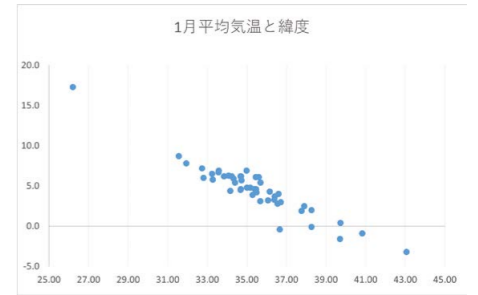
出典:「SSDSE-気候値 (SSDSE-F)」(独立行政法人統計センター) <https://s.nikkei.com/4eZ4H6Y> より抜粋して作成

展開1: データの準備と可視化

分析に用いるデータを整理し、可視化を通して理解を深めます。e-Stat などから得た実データを表計算ソフトで前処理し、不要な項目を除去したり、形式を統一したりして整理します。続いて、各説明変数の分布や相関関係をグラフ化します。これにより、**どの変数が有用か、どの組み合わせが予測精度向上につながるかを直感的に把握できます。** この可視化段階で、データの特徴や傾向を明確にし、モデル構築に向けた土台を築きます。

地域コード	都道府県	市	月・年	平均気温
R01100	北海道	札幌市	01月	-5.3
R02001	青森県	青森市	01月	-6.9
R03001	岩手県	盛岡市	01月	-11.4
R04100	宮城県	仙台市	01月	-2.5
R05001	秋田県	秋田市	01月	6.4
R06001	山形県	山形市	01月	-6.1
R07001	福島県	福島市	01月	-1.5
R08001	茨城県	水戸市	01月	-3.3
R09001	栃木県	宇都宮市	01月	2.8
R10001	群馬県	前橋市	01月	3.7
R11001	埼玉県	さいたま市	01月	4.5
R12100	千葉県	千葉市	01月	6.1
R13100	東京都	東京都港区	01月	5.4
R14100	神奈川県	横浜市中区	01月	6.1
R15100	新潟県	新潟市	01月	2.5
R16001	富山県	富山市	01月	3.0
R17001	石川県	金沢市	01月	4.0
R18001	福井県	福井市	01月	3.3
R19001	山梨県	山梨市	01月	3.1
R20001	長野県	長野市	01月	-5.4
R21001	岐阜県	岐阜市	01月	4.4

緯度	経度	標高
43.00	141.30	17.4
40.82	140.77	2.8
38.70	141.12	188.7
38.20	140.90	38.3
39.72	140.10	6.3
38.20	140.30	182.5
37.30	140.47	67.4
36.38	140.47	29.0
36.55	139.87	118.4
36.41	139.00	112.1
36.20	139.38	300.0
35.80	140.10	3.5
35.69	139.70	25.2
35.44	139.40	38.1
37.48	139.02	41.3
36.71	137.20	8.6
36.59	136.60	5.7
36.00	136.20	8.8
35.07	138.50	27.8
36.55	138.10	418.2
35.40	136.70	12.7



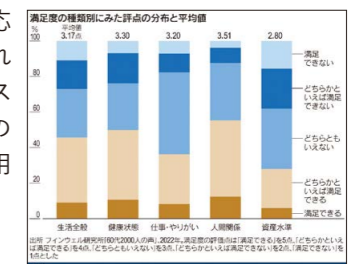
出典:「SSDSE-気候値 (SSDSE-F)」(独立行政法人統計センター) <https://s.nikkei.com/4eZ4H6Y> より抜粋して作成

展開3: 結果の解釈と考察

モデル構築後は、得られた重回帰式を用いて結果を解釈します。各説明変数が目的変数に与える影響の強弱や正・負の関係、単回帰にはなかった要因の組み合わせによる相乗効果などを考えます。**予測精度が高まったかを確認し、なぜそのような結果が得られたのかを探ることで、モデルの有用性と限界を理解できます。** こうした考察を通じて、生徒は分析結果を批判的に評価し、改善の方向性を見いだせます。

まとめ: モデル改善と応用例の提示

重回帰分析の考え方を実践的な学びに結びつけ、データサイエンスへの関心と理解を深めます。まず、モデルをより良くする方法を検討します。**作成したモデルについて、説明変数の追加・削除、変数変換といった改善策を試し、予測精度向上を図ります。** また、実社会での応用例を考えるため、日経電子版などを用いて重回帰分析が実際に活用されているケースを調べ、多面的な問題解決への応用可能性を確認します。これらについて、グループディスカッションを行い、モデルの改善方法や実社会での応用について整理を行います。



出典:60代ライフの「幸せ」に必要な資産額とは、日本経済新聞、2022-4-22、日経電子版、<https://s.nikkei.com/3WmCwX4>

授業展開のポイント

複数の要因を考慮する意義の理解

単回帰分析では、一つの要因から得られる視点は限定的になりがちですが、重回帰分析では複数の説明変数を同時に扱うことで、現象をより多面的に捉えることができます。例えば、売上高予測を価格や広告費、気温、天気など複数の要因で考えると、より現実に即したモデルが構築できます。このような多変量解析は、複雑な現実世界の課題に対して、より説明力を持つ予測や解釈を可能にします。これを理解することで、分析の幅が広がり、データサイエンスのメリットを体感できるようになります。

データ準備と可視化の習慣化

効果的な分析を行うためには、まず用いるデータを丁寧に準備することが大切です。欠損値の処理や不要な項目の削除、数値形式の統一などの前処理を行い、その上で表計算ソフトの機能を活用して可視化を行いましょう。視覚的な表現を通じて、説明変数同士の相関や分布の偏りが直感的に把握できます。こうした前処理と可視化を習慣化することで、分析に用いる変数の選定が容易になり、モデル構築のためのベースが身につきます。

評価指標と操作スキルの習得

重回帰分析を深く理解するには、偏回帰係数や決定係数といった評価指標や有意性の指標となるp値などについて正しく解釈できることが欠かせません。これらはモデルの当てはまりや各変数の寄与度や有意性を数値で示す重要な手がかりです。また、表計算ソフトの統計分析ツールを使いこなす操作スキルを身につけることで、データから結果を迅速かつ正確に導き出せるようになります。評価指標の理解とツール操作は、データに基づく判断や説得力ある説明に役立ち、分析全体の質を高める基礎となります。

モデル解釈と改善への探究心

得られた分析結果をただ受け入れるのではなく、なぜそのような結果が得られたのかを考え、必要に応じて改善を図る姿勢が重要です。例えば、予測精度が思うように上がらない場合、説明変数を追加・削除するなどの処理を試してみよう。また、実社会での応用例や関連する研究事例を参照することで、分析の方向性が明確になります。こうした探究心を持つことで、モデルの信頼性と妥当性が向上し、より納得感のあるデータ分析へとつながっていきます。

4-2 データの整形

本節の目標

e-Stat や市区町村オープンデータから社会統計データを収集し、列名や型の統一、住所コードや緯度経度などの整形・前処理を行い、WebGIS へのインポートに対応可能な形式のデータ作成スキルの習得を目指す。これにより、地図上に社会統計データを表現できるようになり、空間的な視点を用いたデータの多面的な解釈が可能となる。

評価規準

知識・技能: オープンデータの整形手順(列名・型統一、住所コード・緯度経度付与)を理解し、GIS インポートに対応したデータセットを作成できる。

思考・判断・表現: GIS インポートや重回帰分析を視野に入れた前処理を適切に行うために必要となるデータと不要なデータの判断をすることができる。

主体的に学習に取り組む態度: データ整形・前処理という、地道な作業に積極的に取り組み、継続的に改良を試みる姿勢がある。

準備

Google スプレッドシート: 表計算ツール
<https://s.nikkei.com/3YIAw4C>

※Excel などの表計算ソフトで代用できる

e-Stat: 総務省統計局が提供している、各府省が公表する統計データを一つにまとめたポータルサイト
<https://s.nikkei.com/4fngkFj>

jSTAT MAP: 総務省統計局が提供している、誰でも使える地理情報システム
<https://s.nikkei.com/3C10lws>

授業の組み立て

導入 (5分): e-Stat など得た社会統計データを提示し、地図上での可視化や実社会での活用例を示すことで、生徒の興味・関心を喚起する。

展開 (40分): 列名・単位の統一、住所コード・緯度経度付与など、実際に前処理作業を行い、GIS対応データ(CSV)の整形手順を身につけた上で、インポート処理でWebGISに反映させる実践を行う。

まとめ (5分): オープンデータも整形することで有用な情報となるなど、「データ」を「情報」にすることの有用性について話題を提供し、実践からのつながりを深める。

導入: データ取得・確認

e-Stat や自治体オープンデータサイトから必要な社会統計データをダウンロードする方法を身につけるとともに、データの前処理時の確認ポイントを理解します。e-Stat や自治体オープンデータサイトから必要な社会統計データをダウンロードし、表計算ソフトやプログラムで開き、ヘッダー行や列名、データ型や単位、欠損値の有無などを丁寧に確認します。また、解析したいデータの期間、注釈情報も把握しておくことで、後の整形作業が円滑になります。

e-Statのマニュアルはこちら <https://s.nikkei.com/3YEGRbF>

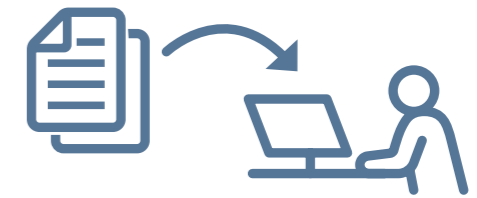
第15表 主目的地(10区分)、宿泊の有無(2区分)、居住地(10区分)別延べ旅行者数【国内旅行】

主目的地	延べ旅行者数		延べ旅行者数		延べ旅行者数		延べ旅行者数		延べ旅行者数	
	千入	千入	千入	千入	千入	千入	千入	千入	千入	
国内旅行	154,108	9,577	13,293	44,839	13,064	21,030				
国内旅行	7,086	6,032	161	484	29	33				
国内旅行	9,578	377	6,318	1,812	229	136				
国内旅行	61,642	1,703	5,728	31,588	6,100	6,749				
国内旅行	6,562	37	381	1,990	2,460	1,143				
国内旅行	20,705	397	268	4,269	2,893	9,143				
国内旅行	23,451	570	283	2,530	1,082	3,278				
国内旅行	8,465	115	69	739	66	139				
国内旅行	4,627	122	29	412	122	124				
国内旅行	11,282	206	48	906	82	240				
国内旅行	709	19	7	109	0	46				

出典:「政府統計の総合窓口(e-Stat)」、旅行・観光消費動向調査/旅行・観光消費動向調査(速報)2024年 7~9月期を加工して作成

展開 1: データクレンジング・標準化

取得データから不要な行や列、注釈情報を除去し、解析に必要な情報のみを抽出します。さらに、解析対象の列名やデータ型を数字は数字型・文字は文字型へ統一し、解析したいデータの期間や単位を揃えることで、比較・分析がしやすいデータに整形します。また、データに欠損値がある場合は、平均値や最頻値などの値で補完することを検討します。これにより、後続のGISインポートや回帰分析が円滑に進められます。



展開 2: 地理情報付与・整合化

地図上にデータを表現するには、住所コードなど地理的情報が必要になります。住所コードは左から2桁目までが都道府県コード、3桁目から5桁目が区市町村コードなどのように、各桁に意味が付与されているため、必要なコードのみに整形します。位置情報が紐づいていないデータは、位置情報を付与する必要があります。住所コードのほか緯度・経度情報などを用いることもできます。これにより、各データがWebGISの地図上の位置との紐づけが可能となり、可視化可能なデータとなります。必要に応じて、今後用いるWebGISの座標系や投影法を事前確認し、地図にデータを表現した際にも正確に理解できるように準備します。

総務省HP 全国地方公共団体コード
<https://www.soumu.go.jp/denshijiti/code.html>

まとめ: GIS対応形式・分析準備

整形済みのデータをjSTAT MAP対応形式(CSV)で書き出し、jSTAT MAPにインポートします。jSTAT MAP上での表示や操作を行うことで、地図上にデータを表現する方法を身につけるとともに、地図上に表現されたデータを用いて、インポートしたデータの特徴や傾向を視覚的に確認します。また、重回帰分析を見据えて、説明変数や目的変数が適切に整形されているか再確認の必要性も説明します。

就業状態	地域コード	地域	時間軸(年)	時間軸(年)	年齢階級	15~64歳	15~24歳	25~34歳	35~44歳	45~54歳	55~64歳
15歳以上人口	00000	全国	2E+09	2000年		***	1617	1872	1586	1937	1643
15歳以上人口	00000	全国	2E+09	2001年		***	1573	1907	1576	1953	1615
15歳以上人口	00000	全国	2E+09	2002年		***	1529	1896	1595	1894	1662
15歳以上人口	00000	全国	2E+09	2003年		***	1493	1883	1625	1808	1731
15歳以上人口	00000	全国	2E+09	2004年		***	1457	1862	1650	1731	1812
15歳以上人口	00000	全国	2.01E+09	2005年		***	1420	1836	1679	1664	1863
15歳以上人口	00000	全国	2.01E+09	2006年		***	1389	1797	1715	1614	1880
15歳以上人口	00000	全国	2.01E+09	2007年		8312	1358	1726	1754	1584	1890
15歳以上人口	00000	全国	2.01E+09	2008年		8243	1332	1674	1792	1563	1881
15歳以上人口	00000	全国	2.01E+09	2009年		8165	1307	1622	1823	1556	1857
15歳以上人口	00000	全国	2.01E+09	2010年		8113	1281	1574	1836	1561	1861

出典:「政府統計の総合窓口(e-Stat)」、労働力調査(総務省) 就業状態,年齢階級別15歳以上人口(2000年~) より作成

授業展開のポイント

データの取得と理解を深める

e-Stat や自治体オープンデータサイトから必要な社会統計データを取得し、その構造を把握させます。ダウンロードしたファイルを表計算ツールやプログラムで開き、見出し行や列名、データ型、欠損値の有無など、ポイントとなる部分を丁寧に確認します。解析対象となるデータの期間を整理し、目的に合った情報だけを抽出できるスキルを身につけます。初期段階でデータの特徴を正しく理解することが、後続の整形工程や分析作業を円滑に進めるための基盤となることを理解させ、動機づけを行いながら指導します。

データを整理し、扱いやすくする

取得したデータから不要な行や列を取り除き、必要な情報のみを抽出させた上で、列名やデータ型、単位などを統一します。こうすることで後から読み取りやすく、比較・分析しやすいデータを作成します。こうしたデータクレンジングは、WebGISへインポートする際や重回帰分析でも欠かせない処理になることを説明し、作業の動機づけを行います。データクレンジングは地道な作業であるものの、質の高いデータセットを得ることが的確な分析を支える鍵となることを説明します。

地理情報を付加して空間的な意味を捉える

住所コードや緯度・経度といった地理情報を適切に付与することで、データを地図上に表現できるようにします。自治体名と住所コードを対応させるとともに、WebGISの座標系や投影法を確認しながら正確な位置情報を整えます。これにより地域間比較や空間的な傾向の把握が容易になります。地理情報が付与されたデータは、WebGISを利用することで視覚的に理解しやすくなり、空間的な視点から多面的な分析や考察を行う基盤を築くことにつながります。

GIS対応形式で可視化し、分析への橋渡しを行う

最後に、整形を終えたデータをjSTAT MAP対応形式(CSV)で出力し、jSTAT MAPへインポートします。地図上で可視化された情報は、単なる数値以上の意味を持ち、社会的・地域的文脈を踏まえた理解を促すことに気づくよう指導します。また、重回帰分析に用いる説明変数や目的変数が適切に整形されている必要性を説明することで、重回帰分析の準備にもつながります。空間的なデータ活用と統計的な分析手法の組み合わせにより、多面的な解釈が可能となり、より効果的な問題解決や意思決定に資することを実感させます。

4-3 探究課題:GIS×統計分析

本節の目標

GIS (地理情報システム)と統計分析を組み合わせて、地域における社会課題を多角的に分析し、その背景や要因を探究する力を養うことを目標とする。具体的には、WebGIS ツール「jSTAT MAP」を用いてデータを可視化し、重回帰分析を通じて複数の変数間の関係性を明らかにする。これにより、地理的データの解釈と統計的思考を身につけ、現実社会の課題解決に向けたデータサイエンスの活用方法を学ぶ。

評価規準

知識・技能: WebGIS ツールである jSTAT MAP を用いて地理情報を正確に可視化し、統計データを効果的に重ね合わせることができ、さらに統計分析ツール (XLMiner Analysis ToolPak 等) を活用して重回帰分析を実施し、目的変数と説明変数の関係性を適切に解析することができる。

思考・判断・表現: 重回帰分析で得られたデータをもとに、統計的指標である補正 R2 や p 値に基づいてデータの傾向や相関関係を考察し、地理的および統計的な視点を統合して課題や傾向を明確に表現することができる。

主体的に学習に取り組む態度: 自ら興味を持った地域や社会課題を設定し、データの収集・整理・分析に積極的に取り組み、データの限界や偏りを意識しながら批判的思考をもって信頼性を評価し、探究的な学習を推進しようとする姿勢を示している。

準備

Google スプレッドシート: 表計算ツール

<https://s.nikkei.com/3YIAw4C>

e-Stat: 総務省統計局が提供している、各府省が公表する統計データの一つにまとめたポータルサイト

<https://s.nikkei.com/4fngkFj>

jSTAT MAP: 総務省統計局が提供している、誰でも使える地理情報システム

<https://s.nikkei.com/3C10lws>

未インストールの場合は

XLMiner Analysis ToolPak: スプレッドシートの分析ツール

<https://s.nikkei.com/4f3Q3LR>

※ XLMiner Analysis ToolPak の利用が難しい場合は、Excel などの表計算ソフトで代用できる

授業の組み立て

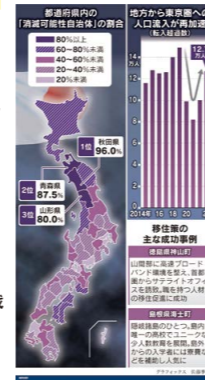
導入 (15分): WebGIS を使った地図データの可視化を体験し、「地理情報を用いたデータ分析はどのような課題解決につながるのか?」と問いかける。jSTAT MAP を活用し、人口密度と公共施設数を重ね合わせた地図を提示して、地理データの可視化が生む新しい視点や気づきを共有する。

展開 (350分): jSTAT MAP を使って自分たちが興味を持つ地域課題を設定し、関連データを収集する。例えば、「人口密度が高い地域ほど公共施設が少ないのではないか」という仮説を立て、jSTAT MAP 上にデータを可視化することで、地理的特徴を視覚的に捉える。次に、収集したデータをスプレッドシートで整理し、XLMiner Analysis ToolPak を使った重回帰分析を実施する。データ分析の結果から、地域課題の背景にある要因を検証し、分析結果の考察を深める。最終的に、分析結果をポスターやスライドにまとめ、成果物を作成する。

まとめ (35分): 探究活動の成果をグループ内で発表し、メンバーからのフィードバックを受ける。相互評価を通じて、データ分析を用いた問題解決の視点を広げるとともに、GIS と統計分析の実社会での応用可能性について議論する。

導入:探究テーマをリサーチする

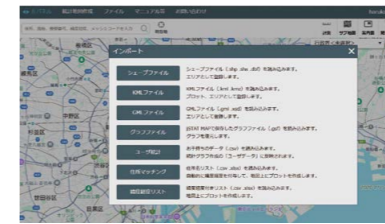
日経電子版の記事やニュースを活用して、地域社会や日常生活の中で気になる課題をリサーチします。例えば、「人口減少による空き家問題」や「観光地での交通渋滞問題」など、興味を持てるテーマをリサーチします。この活動を通じて、身近な課題への関心を高め、データを活用した探究の第一歩を踏み出します。



出典:移住者獲得、「暮い合い」の消費戦 東京集中巡り対立、日本経済新聞社、2024-10-13、日経電子版。
<https://s.nikkei.com/4h2ULes>

展開1:データの準備と可視化


e-Stat や市区町村オープンデータなどの社会統計データを活用して、jSTAT MAP にインポートし、地図上で課題の特徴を可視化します。例えば、交通事故件数やAEDの分布を地図上に重ねることで、地理的な傾向や特徴を直感的に理解します。ここでは、地図を使ったデータ可視化から仮説を立てる準備を行います。



出典:「政府統計の総合窓口(e-Stat)」、jSTAT MAPより作成

展開3:重回帰分析で関係性を検証する

Google スプレッドシートとXLMiner Analysis ToolPak (もしくはExcel などの表計算ソフト) を使用して、重回帰分析を行います。目的変数を設定し、仮説に関連する説明変数を分析に用いることで、複数の変数間での関係性を検証します。例えば、交通事故件数を目的変数とし、交番の数や降雨量を説明変数としたモデルを作成します。立てた仮説に対して、統計的手法で検証する分析力を確認します。



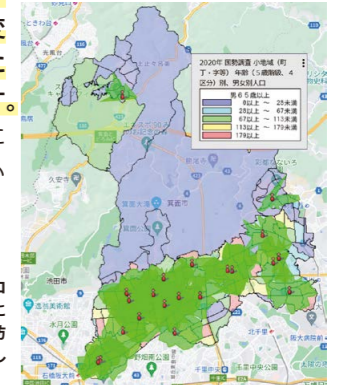
回帰統計	自由度	変動	分散	検定された分散比	有意 F		
重相関 R	0.98143438						
重決定 R ²	0.963233843	回帰	2	448.1422271	149.3807424	375.3017498	7.59766E-31
修正 R ²	0.960460339	残差	43	17.11521975	0.398028366		
標準誤差	6.30894893	合計	45	465.2574468			
観測数	47						

	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	17.414395840	4.727810214	3.683395706	0.000638985	7.879857862	26.948933818	7.879857862	26.948933818
雨量	-1.919821959	0.065423958	-29.099105320	0.000000000	-1.551581335	-1.287701365	-1.551581335	-1.287701365
交番	0.278178374	0.047760750	5.824408436	0.000000056	0.181859422	0.374496927	0.181859422	0.374496927
標高	-0.010381241	0.001284168	-8.084019445	0.000000000	-0.012971013	-0.007791469	-0.012971013	-0.007791469

出典:「SSDSE-気候値(SSDSE-F)」(独立行政法人統計センター) <https://s.nikkei.com/4eZ4H6Y> より作成

展開2:仮説を立てる

可視化したデータを基に、「65歳以上人口が多い地域での避難所が不足しているのではないか」や「観光客の多いエリアほど交通事故が頻発する」といった仮説を立てます。データの傾向を観察し、どの変数が目的とする変数に関連するかを考えます。ここでは、課題解決に向けた仮説設定を行います。



出典:「政府統計の総合窓口(e-Stat)」、jSTAT MAPに2020年国勢調査データ、大阪防災ネットワーク 避難所一覧をプロットして作成

まとめ:成果を発表し考察を深める

重回帰分析の結果を基に、スライドやポスターを作成し、グループ単位やクラス全体で発表します。学習者同士の発表を聞き、フィードバックを行うことで、新たな視点や改善点を得られます。課題解決に向けた提案についての統計データの有用性や分析結果の考察を深めます。



授業展開のポイント

GISを活用した課題発見のプロセス

jSTAT MAP を使い、社会統計データを地図上に重ねて可視化します。このプロセスで、地図上に表れた傾向から、地域がもつ課題を発見することを促します。例えば、「ある地域で空き家率が高い理由」や「公園の数と昼間人口の関係」を探る問いを設定します。これにより、学習者の当事者意識を高め、データ収集への意欲を育みます。

多変量データの整備と重回帰分析の実践

収集したデータセットを基に、説明変数と目的変数を設定し、Google スプレッドシートでXLMiner Analysis ToolPak を用いて重回帰分析を実施します。分析結果から、各変数が目的変数に与える影響を数値で確認し、統計指標 (例: 補正 R2、p 値) の解釈を活用します。これまでの学習から統計分析の基礎的な手法を実践します。

結果の可視化と考察

重回帰分析の結果を基に、jSTAT MAP 上で影響の大きい変数を強調表示するなど、視覚的に示します。この可視化結果を基に、分析結果の妥当性を議論し、仮説と結論の整合性を確認します。さらに、データの限界や疑似相関、交絡要因についても考察し、批判的思考力を養います。

探究活動の成果発表と社会的応用の議論

分析結果をスライドやポスターにまとめ、クラス内で発表する活動を行います。他の学習者のフィードバックを通じて、自身の探究についてのリフレクションにつなげます。また、日経電子版などの信頼性の高い情報源を用い、実社会での GIS や重回帰分析の応用例 (例: 都市計画、災害対策) についてリサーチし、その重要性と可能性についてディスカッションします。

4-4 検証と分類

本節の目標

多重共線性と分類の基本的な概念を理解し、データ分析において信頼性や妥当性を確保する方法を学ぶ。探究課題を振り返り、多重共線性の影響を考慮した上で分析する批判的思考力を養う。また、分類の手法を活用してデータを効果的にカテゴリ分けし、その結果を基に課題解決や意思決定に結びつける視点を身につける。これにより、回帰と分類を適切に使い分ける力と、データサイエンスの実社会での価値を理解する態度を育成する。

評価規準

知識・技能：多重共線性や分類の基本概念を理解し、提供されたデータを基に、説明変数間の関係性を適切に判断できる。また、分類の特徴を活かしてデータを効果的にカテゴリ分けし、分析の基礎となる視点を身につけている。

思考・判断・表現：多重共線性の影響や分類結果について、データの傾向や特徴を考察し、具体的な事例や統計的視点を交えて課題や仮説を明確に表現することができる。さらに、議論や発表を通じて他者の意見を取り入れ、分析結果の妥当性を論理的に説明できる。

主体的に学習に取り組む態度：多重共線性や分類の応用可能性を探る過程で、自ら積極的にデータのパターンや課題を見出し、グループディスカッションに主体的に参加している。データ分析の重要性や限界を意識しながら、自分たちの探究活動を深める意欲を示している。

準備

3章で作成したデータサイエンス探究課題のスライドやポスターデータ
日経電子版 <https://www.nikkei.com/>

授業の組み立て

導入 (10分)：多重共線性の概念を紹介。例として「人口密度と公共交通利用者数」を挙げ、説明変数間の相関が分析に与える影響を説明する。クイズ形式で「どの変数間で多重共線性が起きそうか」を考えさせ、興味を喚起する。

展開1 (40分)：これまで取り組んだ探究課題を振り返り、使用したデータの中で多重共線性が含まれていそうな変数を議論。各グループで疑わしい変数の組み合わせを考え、分析結果への影響を共有する。教員が多重共線性への対処方法を補足する。

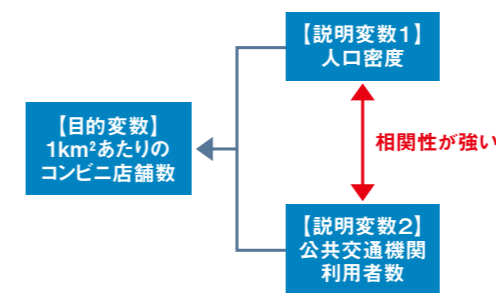
展開2 (35分)：回帰と分類の違いを説明。例として「明日の天気を予測する(回帰)」と「晴れか雨かを分類する(分類)」などを挙げ、分類の特徴を理解させる。ワークとして、架空のデータを基に「駅を主要駅、中規模駅、小規模駅に分類する基準は何か？」などをグループで考え、結果を共有する。

まとめ (15分)：各グループが議論結果を発表。多重共線性と分類の理解を整理し、探究課題についての授業を締めくくる。

導入:多重共線性の理解

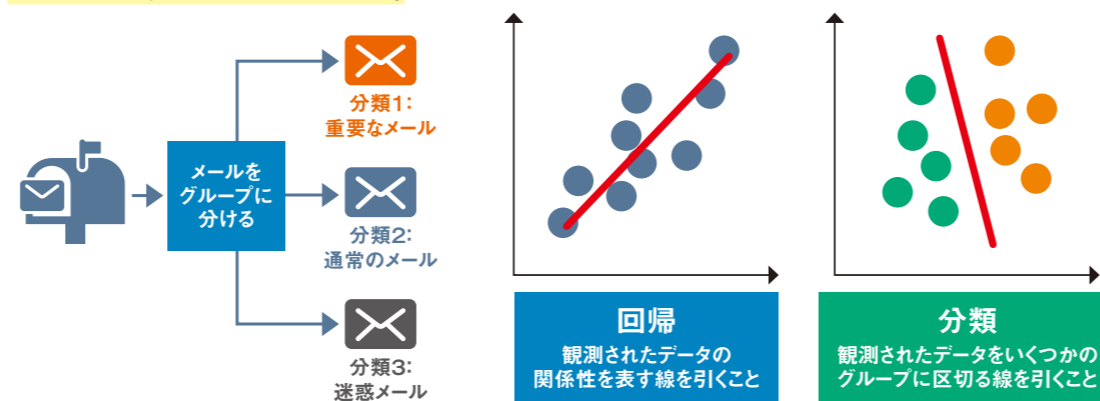
多重共線性は、説明変数同士が強く相関している場合に発生し、回帰分析の結果が不安定になる要因であることを学びます。「人口密度」と「公共交通利用者数」などの具体例を挙げながら、多重共線性が分析結果に与える影響をクイズ形式で考えさせます。この活動を通じて、多重共線性の基本的な概念を直感的に理解します。

多重共線性の一例



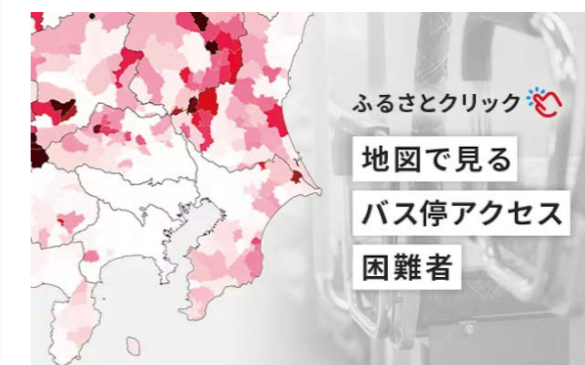
展開2:分類の特徴を学ぶ

分類は、データをカテゴリに振り分ける手法であることを学びます。「明日の気温を予測する(回帰)」と「明日の気温が今日より高いか低いかを分類する(分類)」の例を挙げ、回帰との違いを理解します。その後、グループで与えられたデータを基に、(例: 駅乗降客数から、「主要駅」「中規模駅」「小規模駅」を見る)に分類する基準を検討し、分類結果を共有します。



まとめ:分類結果の応用と考察

分類結果を基に、課題解決や意思決定に役立てる方法をグループで議論します。具体的には、「分類された駅の中で公共交通機関の追加が必要な地域をどう決めるか」などの問いを考え、分類の応用可能性を検討します。この活動を通じて、分類手法が実社会の課題解決にどのように貢献できるかを理解します。



出典:ふるさとクリック 地図で見るバス停アクセス困難者.日本経済新聞社.2024-6-14.日経電子版.<https://s.nikkei.com/3WouWgu>

授業展開のポイント

データ分析の信頼性を支える視点を養う

多重共線性や分類の特徴を学ぶことで、データの信頼性や分析の精度を向上させる方法を理解します。単にデータを用いるだけでなく、変数間の関係性やカテゴリ分けの妥当性を考慮する力を育て、より深い分析を可能にする基礎を築きます。

データを基にした問題解決力を高める

多重共線性や分類の考え方を応用し、実際の探究課題に適用することで、データを基にした課題発見と解決策の提案力を養います。これにより、生徒はデータの特徴を客観的に捉え、課題を論理的に整理する力を身につけます。

回帰と分類の違いと使い分けを理解する

回帰と分類の手法を比較し、それぞれが異なる目的を持つことを学びます。数値的な予測を行う回帰と、カテゴリごとの特徴を把握する分類を適切に使い分ける視点を育て、データ分析の幅を広げます。

実社会に結びつくデータサイエンスの価値を実感する

探究課題やグループディスカッションを通じて、データ分析が社会課題の解決にどのように役立つかを体験的に学びます。例えば、地域ごとの公共交通の最適化や消費行動の分類によるマーケティング改善など、実社会での応用例を日経電子版からリサーチして、考えを深める手助けをします。

おことわり

本授業案は授業実践の事例を紹介するものであり、教育的効果や記載内容の正確性・妥当性等は保証していません。

本授業案に関するご意見やご質問等は、受け付けておりません。ご了承ください。

本授業案の利用によって生じる不利益等について、著者および日本経済新聞社では一切の責任を負いません。

本授業案の一部もしくは全部を無許可で転載、複製、複写、転用することを禁止します。

本授業案は下記執筆者と日本経済新聞社で製作をしています。著作権等の権利は執筆者および日本経済新聞社に帰属します。

執筆者



西出 新也

関西学院千里国際中等部・高等部 技術科・情報科主任

インテリアアドバイザーからシステムエンジニアを経て現職。

テクノロジーの力で学習者のクリエイティビティとモチベーションを高める学びのデザインを探索する。

業務と並行し、神戸大学大学院 博士後期課程にてHCI（ヒューマンコンピュータインタラクション）分野の研究に従事する。

・ Apple Distinguished Educator

・ Adobe Creative Educator (ACE) Innovator

日経電子版 for Education



「日経電子版 for Education」は電子版の記事を生徒が読みこなしながら自ら問いを立て、答えを導き出す授業や、教科書の学びを深める調べ学習の資料として活用されています。また本プログラムを支援するパートナー企業の強みを生かし、中高生にとっての新しい学びの創出も進めています。

<https://www.nikkei.com/promotion/education/>