

人流ビッグデータを扱ったプロジェクト型授業の実践報告

公立千歳科学技術大学 理工学部 情報システム工学科 石田雪也

ソフトバンク株式会社 奥野修敬

公立千歳科学技術大学 理工学部 情報システム工学科 曾我聡起

1. はじめに

公立大学である本学は、地域連携センター¹を設置するなど、地域との繋がりを強く意識した教育・研究活動を行っている。特に情報システム工学科においては、地域が抱えるさまざまな課題についてサービス科学・サービス工学の知見により発見し、ユーザービリティエンジニアリングのアプローチで解決を提案している。また、2021年8月には文部科学省より、「数理・データサイエンス・AI教育プログラム（リテラシーレベル）」に認定された。我々は、2020年度から、学生に「平成30年北海道胆振東部地震」の前後の時期を含む人流データを提供し、学生のアイデアに基づく課題の発見と解決の提案について取り組むカリキュラムを開始した。

ビッグデータは、文字通り膨大な量的データであり、そのままでは学生が授業内で操作すると支障が出る場合もある。そこで、我々の授業設計に向けたデータ形式になるように、提供する人流ビッグデータの精査とツールの提供などを、ソフトバンク株式会社の奥野氏と共に行った。本論では第2章で奥野氏はその概略について記している。今回は、情報システム工学科2年生、3年生の科目でビッグデータを用いた授業を実施した。授業の詳細や学生の授業評価に関する考察などを本学の石田氏が第3章、第4章で記している。また、本取り組みを行なった2020年度、2021年度はコロナ禍にあり、学生はオンライン授業や対面授業を繰り返す日々であった。この点についても効果的な講義様式の展開について記しているので参考にしていきたい。第1章と第5章は曾我がまとめた。

2. 人流ビッグデータ

2.1. 人流ビッグデータの活用

学生らが授業で分析した人流ビッグデータは大手通信キャリア子会社の株式会社 Agoop²から購入し、提供を受けた。人流データはスマートフォンのGPS情報をもとに利用者の承諾のもと特定アプリを通じてデータが蓄積されている。同社が提供する人流データには「メッシュ型データ」および「ポイント型データ」の2種類がある。

前者は地図上の経緯度方眼として定められた特定メッシュ内にどれくらいの人口が集まっているかを統計的に示したものであり、特定アプリの利用者数から国内総人口規模に換算した拡大推計という考え方に基づき推計人口を算出している。どのメッシュにどれくらいの人が集まっているかをヒートマップとして可視化するのに適している。

¹ 公立千歳科学技術大学地域連携センター、<https://snc.chitose.ac.jp> (2021.12.25 参照)

² 株式会社 Agoop、<https://www.agoop.co.jp> (2021.12.25 参照)

一方後者はスマホ利用者個人の回遊状況を取得したものであり、個々の行動履歴として日時や位置情報等が蓄積されたものである。表1にポイント型データの形式例を示す。この形式のデータは、他にも性別や年代、推定居住地などの属性も記録されていることから、主に観光業界や小売業界等の人流・出店分析で大きな効果が期待される。いずれのデータも特定アプリから収集されるデータのため、特定通信キャリアに偏ることはなく、また個人情報も特定されないよう加工されており、個人情報保護に関する各法律等に準拠する形で提供されている。

表1 ポイント型データの形式例

No	データタイトル	CSV項目	データ型
1	デイリーID	dailyid	文字列
2	年	year	整数(4桁)
3	月	month	整数(1-2桁)
4	日	day	整数(1-2桁)
5	曜日	dayofweek	整数(1桁)
6	時間	hour	整数(1-2桁)
7	分	minute	整数(1-2桁)
8	緯度	latitude	小数点以下6桁
9	経度	longitude	小数点以下6桁
10	その他、多数	—	—

2.2. ビッグデータの事前処理

今回の授業で主に利用した人流ビッグデータは、2018年9月に発生した北海道胆振東部地震のものであるが、同月の千歳市におけるアプリのMAU (Monthly Active Users) は約22,000である。これらのユーザーから取得されるデータは膨大であるが、特に回遊分析に使用するポイントデータについては、生データをそのまま利用できない。稀に発生するデータ欠損の補完や、回遊の可視化ツールのフォーマットに合わせて大量のデータ加工(事前処理)が一般的に必要である。今回学生が用いたデータは学生が扱いやすいようにデータ加工を行ったものを購入した。

2.3. ビッグデータの可視化

最近では人流ビッグデータを可視化するためのBI (Business Intelligence) ツールが多く登場しているが、本学が利用した可視化ツールは、Mobmap³である。このツールは人流データの可視化・分析に特化したツールであり、東京大学空間情報科学研究センター(CSIS)の柴崎亮介研究室が開発したツールである。図1にMobmapの可視化例を示す。陸上自衛隊演習場内に期間内に入った人を点で示し、それぞれの回遊動向を直線で示している。

³ Mobmap、<https://locationmind.com/products/mobmap/>、LocationMind社(2021.12.25参照)

Mobmap で取り込み可能なデータフォーマットが決まっているため、前項の人流ビッグデータの生データを加工およびフォーマット変換処理を経て回遊動向が可視化できる。属性毎の回遊分析（訪問地や訪問時間、滞在時間等）を細かく分析することで、各種課題および解決ソリューションの検討、プロモーション実施を効果的に行うことができると考えられる。

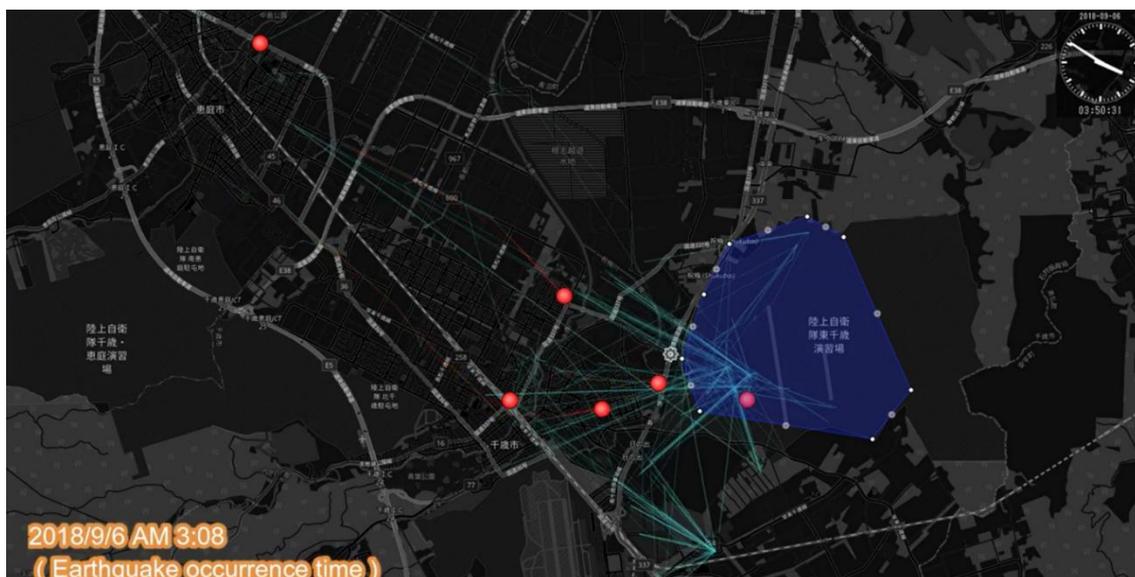


図 1 Mobmap 可視化例（北海道胆振東部地震）

東京大学 CSIS 柴崎研究室 Mobmap

3. ビッグデータを活用したプロジェクト形式授業の実践

3.1. 授業概要

人流ビッグデータを用いた授業は 2020 年度、2021 年度の 2・3 年次の情報システム工学科の必修科目で行った。それぞれの科目、授業形態をまとめたものを表 2 に示す。3 年次開講のシステムデザインプロジェクトは前期の科目で、ハードウェアを行うグループとビッグデータを扱うグループに分けて 7 週ずつ実施する形式で授業を行っている。2 年次開講の情報処理プロジェクトについては、他の実習科目とグループを 2 つにわけ、それぞれ 15 週実施している。

表 2 授業概要

年度	科目名	学年	履修者数と 単位認定者数	授業形態
2020	システムデザインプロジェクト	3	88 人 (83 人)	オンライン
2021	システムデザインプロジェクト	3	95 人 (82 人)	ハイフレックス
2021	情報処理プロジェクト	2	90 人	対面

授業は履修者を半数に分け 44～48 名で実施している。単位認定がされなかった内訳とし

では、2020年度のシステムデザインプロジェクトは全員途中でドロップアウトした学生であり、2021年度のシステムデザインプロジェクトはドロップアウトした学生が9名、授業への参加は行ったが、課題点が合格点に満たなかった学生が4名であった。なお、2021年実施の情報処理プロジェクトについては本稿記載時点では、成績が確定していないため単位認定者数は記載していない。

Covid-19の影響のため、2020年度は完全オンライン形式での授業とし、2021年前期の科目は一部対面、一部オンラインを授業週や、学生の希望に応じて選択させる形式で授業を行った（ハイフレックス型）。2021年情報処理プロジェクトは対面で授業を実施している。

3.2. 3年次科目でのビッグデータの活用授業

2020年度、2021年度の3年次必修科目システムデザインプロジェクトでの7週間の授業実践について述べる。2年間の授業は、1グループ4人程度で人流ビッグデータや他の情報も調査し、それらからわかることを分析し、問題点、解決案を検討させることを目的とした。表3に7週間の授業スケジュールを示す。2020年度と2021年度では授業スケジュールは同じとした。2020年度はオンライン会議アプリ Zoom を活用して授業を行った。2020年度前期は Covid-19 の影響を考え大学での受講を許可しなかったためオンライン上で行ったが、学生の自宅の PC では、Zoom と Microsoft Office、特に Mobmap を同時に活用するための環境が整わない学生が数人いた。そこで2021年度は1週目、4週目を2021年度は、対面形式として実施した。特に4週目はノート PC を持参させ授業を PC 教室で実施し、Mobmap を各自のノート PC で活用させることで、学生の環境でグループワークが実施できるのかについて検討させた。残りの週は学生に PC 教室か自宅等での受講を選ばせる形とした。5、6、7週目は約3割程度が大学で授業を受講した。

授業スケジュールについて示す。まず人流データについての簡単なレクチャーと Excel を活用した簡単なデータ処理、グラフ作成、情報検索を行わせ、分析を行う個人課題をレポートとして課した。利用したデータは、2020年度は、前述のとおり2018年9月に発生した北海道胆振東部地震の前後9月5日、6日、7日のメッシュデータのうち、公立千歳科学技術大学付近のデータを分析させた。2021年度は千歳駅、新千歳空港の周辺のメッシュデータから分析を行わせた。両年度とも2、3週は千歳駅、新千歳空港駅、千歳市民病院、自衛隊東千歳基地周辺のメッシュデータを含めたデータの分析、他の情報調査をもとに、分析、問題点の抽出、解決案の提示を行うグループワークを行い、3週目後半にプレゼンテーションを行わせた。

4週目以降は、グループメンバーを入れ替えポイントデータと Mobmap を用いた分析を中心に行わせ、各自で自由にテーマを決めて、グループワークを行わせる形とした。

グループのメンバーは、前半は学生が個々に3~5名程度で決めさせたが、独自で決められない学生もいたため、その際には教員、SA や TA も加わってグループを決めさせた。ポイントデータを扱う際にはランダムでグループを作成した。

表3 授業内容と2021年度の実施形態

週	概要	2021年度の実施形態
1	概要説明と個人課題実施	対面
2	メッシュデータ GW	学生によって選択
3	メッシュデータ GW・発表	学生によって選択
4	ポイントデータ GW	対面
5	ポイントデータ GW	学生によって選択
6	ポイントデータ GW	学生によって選択
7	ポイントデータ GW・発表	学生によって選択

2021年度に、7週目の発表を終えた時点で学生に振り返りを学内ポータルシステムのアンケート機能を用いて行わせた。振り返りは2021年度の前半グループは振り返りのみとし、2021年の後半グループには、振り返りに加えて、(1)オンライン型か対面型のどちらで行うのがよいか、(2)グループはランダムか任意かとそれぞれの理由について記入させた。

オンラインか対面かについては、回答者39名のうち、22名が対面を希望し、17名がオンラインを希望した。対面を希望した学生からは、「相手の顔を見て、反応を感じることができると話しやすい」、「TAや教員にアドバイスをもらいやすい」といった前向きな回答が多かった。一方で、オンライン型に関して、「オンラインで行うと積極的に参加しない学生が出てくる」といった回答もあった。一方で、オンライン型を希望した学生からは「画面共有がしやすい」「自宅から気軽に授業を受けられる」「通学時間がかからない」などという意見があった。また、いずれの学生からも「最初は対面がよいが、慣れてくればオンラインでもよい」といった意見もあった。

グループについての質問では、39名中24名はランダムが良いと回答し、15名は任意(自由)が良いと回答した。ランダムが良いと回答した学生からは、「就職活動時のために仲良くない学生と行うほうがよい」、「妥協が少なくなる」といった意見があった。一方で「友人同士だと甘えてしまう」、「積極的でない(あまり参加しない)学生とあたるにつらい」、「毎回同じメンバーだと慣れてしまう」などといった意見もあった。任意を希望した学生からは、「自由だと遠慮せずにグループワークができるから」、「授業外に集まりやすい」などの意見があった。

振り返りからは、回答者81名中56名(69%)がビッグデータに関する内容を振り返っていた。ビッグデータの解析についての楽しさ、難しさについての言及が多かった。また、81名中53名(65%)がグループワークに関する記載をしていた。同じデータを扱っても考える視点が異なる点の指摘や、オンラインと対面のそれぞれのメリット、デメリットや積極的でないメンバーがいた場合の対応の難しさ、授業外に集まる必要性などについて言及していた。そのほかに81名中17名(21%)が発表や発表資料の準備などについて振り返りを行っていた。

3.3. 2年次科目でのビッグデータの活用授業

2021年度の2年次必修科目情報処理プロジェクトでの活用実践について述べる。情報処理プロジェクトは、情報システム開発において必要な、分析、調査、要件定義、画面設計、プレゼンテーションなどをグループワークによって学ぶ科目である。3年次科目の振り返りからビッグデータを扱うことが学生にとって好評であった点を考慮し、ビッグデータに触れる機会をさらに増やすために担当教員間で相談の上、2年次科目でビッグデータを扱うこととした。

情報処理プロジェクトの2021年度のスケジュールは、1週目から10週目までは、まず個人で課題を行わせ、グループワークを行うという形をとった。扱った内容として、顧客のニーズ分析やターゲットの策定などについて3週間、要件定義、画面遷移、簡易データベース設計などを活用したシステム提案を7週間かけて行った。本格的なグループワークとしては、8～10週の3週間で要件定義書を作成したうえでシステム提案を行わせるテーマで発表も行わせた。授業はすべて対面で実施し、学生はノートPCを持参のうえ参加させた。グループはすべてランダムに割り振り、グループワーク中はCovid-19の影響も考え、オンライン会議ツールのZoomで画面共有をさせ、机の向きを移動しないまま(平行)のまま行わせた。学内のネットワーク環境としても受講者最大44名での学内無線ネットワーク利用についても、不定期で数人の回線不具合があったものの概ね問題がなかった点も付記しておく。

11～14週目に3年次で行った、人流ビッグデータのメッシュデータを用いた授業を行った。11週目は、個人で2018年9月の胆振東部地震時の千歳駅と新千歳空港駅のデータを用いた分析、調査を行わせた。12～14週目には、千歳各地域の胆振東部地震時のデータに加え、Covid-19の影響分析として、支笏湖温泉付近の2019年から21年の2月のデータ、2019年、20年の8月のデータ、新千歳空港付近の2020年、2021年のゴールデンウィークのデータを自由に分析し、問題点を抽出し、改善案を提示することを目的としたグループワークを実施した。

21グループ中、14グループが支笏湖温泉についてのデータ調査を実施し、6グループが空港からの調査を実施し、胆振東部地震震災時の病院についての調査を1グループが実施した。各班ともメッシュデータの分析を行った後に、例えば他の地域の観光客数の増減や新千歳空港やその他の地域の利用者数などの調査を実施したうえで、問題点、解決案を実施していた。

グループワーク実施後に、アンケート調査を実施した。設問は「これまでの授業で行ってきたものの中で難しかったと思うものに3つだけ、チェックを入れてください。」とし、記名式で実施した。当日の授業参加者全員の83名から回答があった。表4にアンケート調査結果を示す。情報調査、データ分析が難しかったと約9割の学生が回答していた。「ビッグデータを見ても何をどのように使えばよいかわからず、非常に難しかった。」「データ量が大きくなればなるほど、自分が今持っていないデータとつながっていることが多く、情報収集がとても大切になってくると感じた。」など、自由記述でもそれらの内容についての言及が多く見られた。

表 4 グループワーク実施後のアンケート調査

項目	回答数
Excel での表計算・グラフ作成	34
データ分析・解析	73
情報調査	74
発表資料作成・発表	38
グループワーク (役割分担など)	24

N=83

4. 考察

2年間の3年次科目及び2021年度の2年次科目について考察を行う。まず授業の実施形式であるが、Covid-19の環境下で、全面オンライン、ハイフレックス型、対面型とさまざまな授業形式で実施を行ってきたが、学生のPC環境やネットワーク環境などを考えるとPC教室にノートPCを持ち込ませた形での対面形式での授業が望ましいと考えられる。学内のネットワーク環境も40人程度であれば、画面共有を行うためのZoomやWeb検索も同時利用が可能であったため、この人数であれば対面授業が可能となる。もちろん、Covid-19の感染拡大状況に応じてとはなるが、Covid-19以前のように、顔を突き合わせての対面形式でなくても、画面共有を行うことで、講義形式の授業の机の位置(平行)でも対応できることが2021年度の2年次の実証からもわかった。

次に授業概要についてだが、2021年度は新たに2年次科目にビッグデータ分析を行い、問題点をさせるという内容を扱った。アンケートの結果からも3年次の学生同様に、データ分析や情報調査が重要であると認識をしている。このことから2年次、3年次と連動したカリキュラムを行っていくべきかと考える。2022年度は、2年次の経験をもとに、ポイントデータの分析、課題解決を行う授業内容を検討している。

また、2020年度から初年次必修科目として、地域課題プロジェクトという地域の問題点の解決案を考えるグループワークの科目が新設された。2021年度以降の学生は、初年次での経験も踏まえたうえで、グループワーク、地域の課題解決にあたることができる。初年次、2年次、3年次度と段階的にデータの分析や情報調査、課題解決案の検討を行わせることは、今後研究室活動や社会に出る際にも有用であると期待できる。

5. おわりに

千歳市には支笏洞爺国立公園という日本有数の観光地があるが、近年はコロナ禍における観光客の減少などの地域課題がある。また、2018年9月6日には北海道胆振地方中東部を震源地とする最大震度7の「平成30年北海道胆振東部地震」を体験した。この地震では、北海道全域が電源喪失するブラックアウトが起こった。しかし、支笏湖畔では、苫小牧市の王子製紙に電力を供給しており、僅かな時間で電源を回復した。さらに、この地震の前日には、台風21号が北海道日本海側に接近し、停電や農作物などに大きな被害を与え、JR北海

道や新千歳空港を発着する航空機の運航にも多大な影響があった。

近年では、ビッグデータによる課題の発見と解決の提案に注目が集まっている。しかし、上述したような地震の前日に台風による影響があったことなどは、時が経つと人々の記憶からは忘れ去られてしまう。今回の人流データの分析授業において、地震による影響などについては様々な情報からすぐに気がつく学生が多かったが、時間をかけることにより、台風の影響について考察する学生が現れ始めた。学生たちが多角的に課題に取り組んだ成果であると考えている。優れたデータサイエンティストはリベラルアートに立脚する多角的にデータを捉えることで、課題の発見と解決に到達する。その意味で、我々の取り組みは「数理・データサイエンス・AI教育プログラム」が目指すべきゴールに一步近づく試みであったと考えている。

最後に、過去二年間におけるビッグデータを用いた授業を通じて取り組んできたことにより得られたものが二点ある。一点は学生たちの意識を、千歳という土地に向けさせて、課題の発見とその解決に取り組んだことである。公立大学として必要な地元意識を培うことができ、将来的にはこの地で活躍する学生を育成できたと自負している。これは、単に統計の授業を行うだけでは得られない成果だと考える。

もう一点は、コロナ禍にあって、オンライン授業と対面授業の組み合わせにおける知見の習得である。2021年12月現在、欧米をはじめとする世界各地では新たな変異株による感染者の急増という課題に直面している。この先我々は、適材適所にオンラインと対面を組み合わせた教育様式を確立することが求められる。これは、感染症のみならず、さまざまな災害や社会変化に対応する上で必要不可欠な知見となる。一例を挙げると本学のキャンパスは、千歳市から車で10分以上離れた位置にあるが、学生の多くはJR駅から連絡バスを利用している。しかし、JRそのものが災害により運休や遅延が発生することはもちろんだが、今後の社会情勢の変化に伴い、運賃の高騰や路線の廃止などに影響されることもある。こうした不測の事態に際しても、ネットワークなどを用いることにより強靱で柔軟な対応が可能な大学が求められる。

「数理・データサイエンス・AI教育プログラム」を取得した多くの大学が、ビッグデータを用いた様々な工夫を凝らしたカリキュラムを提案している。また、多くの大学がビッグデータに関する専門の学部学科の増設を行っている。我々も、今回のカリキュラムを支えるために今後様々な補強を行う必要があると考えている。授業を通じて「量的データ」と「質的データ」について意識することを学生に伝えた。量的データはトレンドを示し、質的データは意向を示す。双方を組み合わせることにより、「なぜ」という疑問に対する答えが潜んでいる、と。2018年度まで学生たちは千歳市内でフィールドワークを行い、高い満足度の授業を実施してきた。ポストコロナの時代が訪れた際には、フィールドワークを通じて「質的データ」も分析できる授業を再開し、より充実したデータサイエンスのカリキュラムが実現できることを待ち望んでいる。