

栄養士養成課程におけるデータサイエンス教育の試み

稲 益 智 子

美作大学・美作大学短期大学部紀要（通巻第67号抜刷）

栄養士養成課程におけるデータサイエンス教育の試み

Implementation of Data Science Education: a Case in the Department of Nutrition at Mimasaka Junior College

稲 益 智 子^{1)†}

要 旨

高等教育機関における数理データサイエンス科目の必修化にともない、美作大学・美作大学短期大学部においても、2020年度より試験的にリテラシーレベルのデータサイエンス教育プログラムを導入することとなった。本稿では、地元の企業との連携により実施した、美作大学短期大学部栄養学科におけるデータサイエンス入門にあたる授業の実践例を紹介し、授業終了時に毎回実施したアンケートから明らかになった課題を報告する。

キーワード：データサイエンス、栄養士養成課程、データ分析

1. 背 景

1.1 研究の目的

データサイエンス科目の導入が本格化したのは、令和元年の「AI戦略2019」¹以降だが、2017年には滋賀大学、2018年には横浜市立大学にデータサイエンス学部が設置され、この分野で諸外国に後れを取る日本においても、近年多くの大学でデータサイエンス教育が行われるようになった²。

美作大学・美作大学短期大学部（以後、本学とする）では、2020年度後期より、試験的にデータサイエンスの入門科目にあたる授業を開始した。

情報科学や数理統計学を専門としない学生を対象としたデータサイエンス教育は、新たな試みであり、経験も教材も限られている³⁻⁴。そこで本報告は、本学における第一期データサイエンス教育の学修効果を、授業終了時に毎回実施したアンケートを用いて分析し、本学におけるデータサイエンス教育の課題と展望を明らかにすることを目的とした。

1.2 授業の概要

データサイエンス科目の基礎となるのは情報科学と統計学である。本学栄養学科における、情報科学と統計学にあたる科目は、2018年まで「教養・基礎教育科目」に分類されていた「情報リテラシー科目」の「情報リテラシーI」と「情報リテラシーII」および、「専門教育科目」に分類されていた「統計処理論」の計3科目であり、これらはいずれも選択科目として提供されていた。

データサイエンス教育を導入するにあたり、「統計処理論」を「調査と統計」と名称変更し、2020年以降は、これと情報リテラシー2科目の計3科目を、「教養・基礎教育科目」「データサイエンス科目」と定めた。この変更により、学生は上記3科目の中から必ず1科目以上を履修することが卒業要件となった。本報告が取り上げる2020年後期は、この変更の過渡期にあたる。

「情報リテラシーI」は、多様なレベルのICT技術を身につけて入学してくる新入生に向け、本学のICT環境への導入から、エクセルの初歩的操作を含む基本的なPC操作を教えるもので、新入生全員のICTレベル

[†]責任著者

¹⁾ 美作大学・短期大学部 学修・学術情報センター

の底上げを目的としている。この授業は、例年新入生全員が履修するが、実践的なデータ分析は扱わない。

栄養士養成課程においては、エクセルによる栄養価計算等、表計算ソフトの基本操作は不可欠である。そこで、「情報リテラシーII」では、「情報リテラシーI」のエクセルの内容から一步進んだ、初中級レベルのエクセル操作を扱う。これには基本的な関数やピボットテーブルを用いたデータ集計が含まれる。

一方で、「調査と統計」と名称変更された統計科目では、入門レベルの統計学を扱い、実際にデータを収集し、エクセルを用いた回帰分析、 χ^2 検定、対応のない2標本t検定などの比較的易しいデータ分析も行う。

「情報リテラシーI」1科目のみを履修する学生も少なくないが、例年、半数以上の学生が「情報リテラシーII」も履修する。一方、「調査と統計」は数学的要素を多く含むため、数学に苦手意識を持つ学生に敬遠される傾向がある。履修者は全体の1/3程度である。

このような背景から、2020年後期以降は「情報リテラシーII」の中心的内容をデータ分析に変更し、「調査と統計」を履修しなかった場合でも、より多くの学生が基本的なデータ分析を経験することができるよう意図した。

本学の入学試験においては、いずれの学科においても数学は必須ではなく、数学に苦手意識を持つ学生が多い。特に専門高校から入学してくる学生の中には、そもそも数学I（「データの分析」を含む）に該当する数学科目を履修していない者もいる。こうした学生が無理なく継続できるよう、「情報リテラシーII」では、可能な限り数学的要素をそぎ落とした。

さらに、本学の学生が参画しているお弁当メニューの開発に関するデータを、提携企業から提供を受け、企業の担当者にも授業に参加して頂くことで、学生の当事者意識を高め、課題への取り組みへのモチベーションを向上できるよう工夫した。

後半のデータ分析を円滑に進めるため、全15回の授業のうち、最初の9回をデータ分析に必要なエクセル操作の習得と強化にあてた。データ分析によく使われ

る関数については、繰り返し4回の授業を費やして演習を行った。また、フィルターやピボットテーブルの使い方に2回の授業をあてた。

残り6回のうち前半3回は、データ編集の必要がない演習用データを用い、比較的小規模なアンケート結果の集計を行った。これに先立ち、与えられたデータから調査票をデザインし、追加すべき質問項目がないかも、クラス全体で議論した。

最後の3回は、企業の担当者を招き、提供を受けた大規模データを加工せずに学生に渡し、データセットの理解からデータクリーニング、データ編集、表やグラフの作成、簡単な統計分析までを扱った。

企業データの分析結果をWordによるレポートおよびPowerPointによるスライドにまとめたものを提出させ、成績評価を行った。表1に授業の概要を示す。なお、表中の区分は、1はエクセル操作、2は小規模データ分析、3が大規模データ分析である。

表1 「情報リテラシーII」の授業概要

回	区分	内 容
1	1	基本操作（データの移動、複写、挿入、削除）
2	1	オートフィル、MAX、MIN、COUNT関数ほか
3	1	DATE、DATEDIF、LEFT、RIGHT、TRIM関数ほか
4	1	VLOOKUP、IF、AND、OR、ROUND関数ほか
5	1	MID、LEN、PROPER、REPLACE関数ほか、表の編集
6	1	条件付き書式、入力規制
7	1	検索と置換、ハイパーリンク
8	1	データベースの取り扱い、並べ替え、抽出法
9	1	ピボットテーブル、度数分布とヒストグラム
10	2	調査票の作成、データセットのデザイン
11	2	関数によるデータ分析(COUNTIF、SUMIF関数ほか)
12	2	分析結果のレポート、絶対参照・相対参照・複合参照
13	3	ピボットテーブルによるデータ分析
14	3	データ編集(SUMIF、VLOOKUP関数ほか)
15	3	分析結果のレポート発表ほか

1.3 履修者の概要

本学栄養学科1年生の、2020年後期時点における休学者を除く在籍者は33名であり、うち54.5%にあたる18名が「情報リテラシーII」を履修した。うち1名

はベトナムからの留学生である。全員が前学期に履修した「情報リテラシーI」の成績の平均値は、履修者の75.5に対し履修しなかった学生が71.3、入学時に実施した基礎学力テストの平均値も、履修者の9.3に対し、履修しなかった学生は7.7と、いずれも統計的に有意な差は認められなかったものの、履修者の方が若干高かった。また、多くの学生が既に「情報リテラシーI」の単位を履修していたため、「情報リテラシーII」は必ずしも卒業に必要ではない。従って、その中で「情報リテラシーII」を履修した学生は、ICTスキルを身につけることに強い意欲を持っていた可能性はある。とは言え、上述の基礎学力テストで、学科平均の半分にも満たなかった、学力的に不安のある学生が2名、当該授業よりも易しい内容を扱う「情報リテラシーI」で単位が取得できなかった、パソコン操作に不安のある学生も1名含まれていた。同年前期に実施された「調査と統計」では、「情報リテラシーII」で扱うデータ分析よりも高度な統計分析を行うが、「調査と統計」を履修した学生は、「情報リテラシーII」の履修者18名中8名（うち1名は不合格）であった。

15回の授業は、いずれの回も出席率が88.9%から100%と概ね高かった。エクセルの操作に伴う課題も、提出率は80%から100%と高く、最終レポートも全員が提出し、全員が単位を取得した。

2. 方法

全15回の授業の最後に毎回、教員の説明の明瞭さ、内容の理解度、授業の進度、授業の難易度、授業で学んだ内容の有用性の5項目について、学生に評価してもらった。進度と難易度は5件法、ほかは4件法のリッカート尺度による回答とした。評価は本学が採用する学習管理システムWebClassのアンケート機能を用いて実施した。

エクセル操作を中心に行った第1回から第9回の授業を「エクセル操作」、データの個数100の小規模デー

タを分析した第10回～第12回を「小規模データ」、データの個数47,062の大規模データを分析した第13回～第15回を「大規模データ」と、3つのセッションに区分し、回答を集計した。

また、明瞭さ、理解度、進度、難易度、有用性の5項目について、各学生の出席した授業回分の回答を合計し、出席回数で除したものを、「各項目に対する満足度」、5項目の平均値を「満足度」と定義した。さらに、各項目に対する満足度の値が75%に満たないものを、満足度の低い項目とし、各学生の満足度の低い項目数を集計した。

最後に、履修者のICTスキルに近似する前期の「情報リテラシーI」の成績をもとに履修者をA（80点以上）、B（70～79点）、C（60～69点）、D（60点未満）の4群に分け、ICTスキルの違いによる「各項目に対する満足度」の違いを示した。さらに「満足度」の平均値を成績A群とBCD群の2群で比較した。

リッカート尺度で測定した項目間の相関分析にはSpearmanの順位相関係数を算出し、2群における連続変数の平均値の比較には対応のある2標本t検定を行った。有意水準は5%とし、統計処理にはSTATA SE 16を使用した。

3. 結果

3.1 授業評価5項目

授業評価のデータから、1) 教員の説明の明瞭さ、

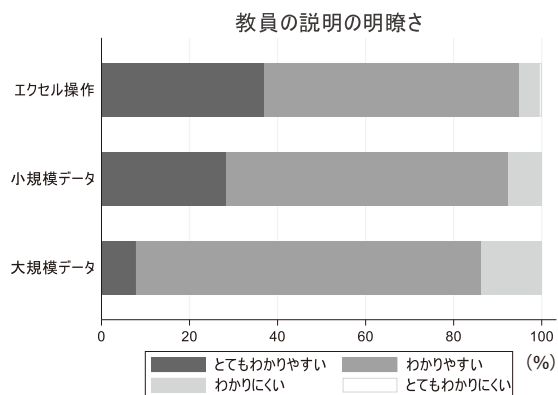


図1 各セッションにおける明瞭さの評価

*言語分野（国語能力）10問と非言語分野（数的処理能力）10問からなる本学独自の試験（20点満点）で、2020年度から新入生を含む全学生を対象に実施している。

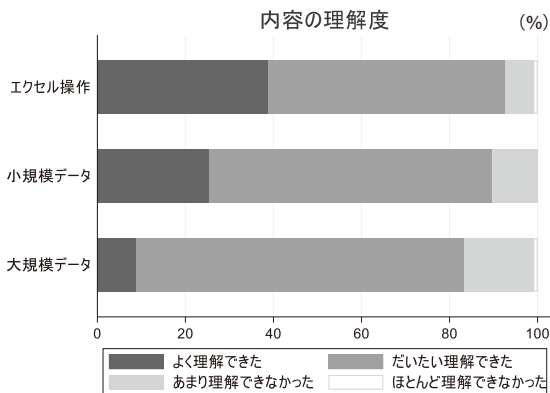


図2 各セッションにおける理解度の評価

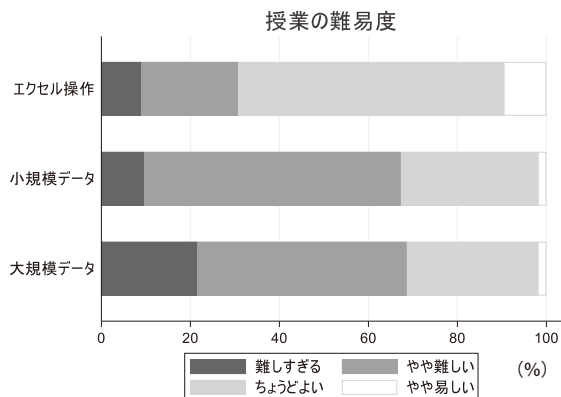


図4 各セッションにおける難易度の評価

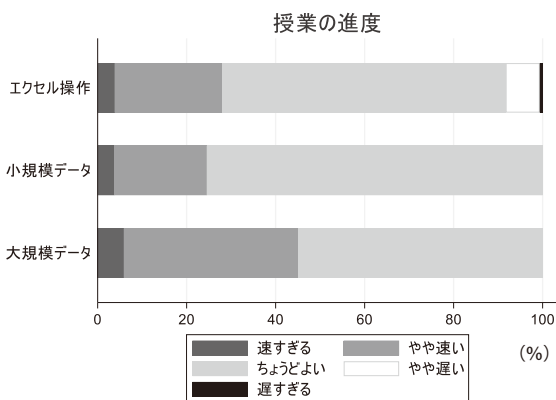


図3 各セッションにおける進捗の評価

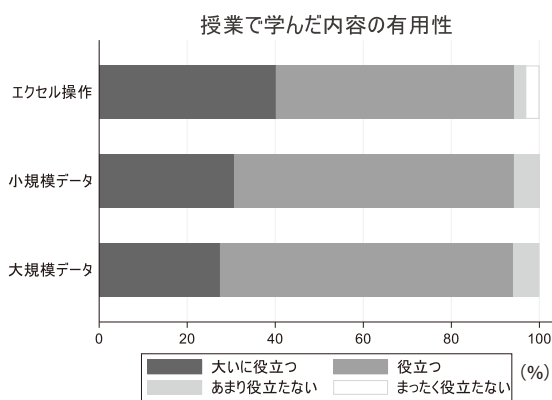


図5 各セッションにおける有用性の評価

2) 内容の理解度、3) 授業の進捗、4) 授業の難易度、5) 授業で学んだ内容の有用性を図1～5に示す。

「教員の説明の明瞭さ」では、前半のエクセル操作の授業については、「とてもわかりやすい」と「わかりやすい」が全体の95%に及んだ一方で、エクセル操作を扱った9回の授業の中で、「とてもわかりにくい」という回答も1例あった。小規模データを扱った3回も、わかりやすいという回答が92.5%と高かったが、大規模データの3回でこの割合が86.2%まで低下した。特に、「とてもわかりやすい」という回答が大幅に減り、代わりに「わかりやすい」に評価が移った。

「内容の理解度」は「教員の説明の明瞭さ」との間の順位相関係数が0.7552 ($p < 0.01$) と、概ねと連動するが、「よく理解できた」と「理解できた」を合わせ

た回答の割合は、「教員の説明の明瞭さ」の回答と比べ若干低かった。また、「ほとんど理解できなかった」という回答が、エクセル操作の9回のなかで2例、大規模データの3回のなかで1例見られた。

「授業の進捗」は、いずれのセッションにおいても半数以上が「ちょうどよい」という回答であった。エクセル操作の回では、「やや遅い」「遅すぎる」という回答も少数あった。データ分析に入ると、授業の進捗が遅いという回答はなくなり、特に大規模データの回で、「やや速い」「速すぎる」という回答が半数近くを占めた。

「授業の難易度」が、最も見解が分かれた項目である。過半数が「ちょうどよい」と感じたエクセル操作のセッションにおいても、「難しすぎる」という回答

も1割弱あり、それぞれの授業回をみても、多くの回で「難しい」と「易しい」が混在した。データ分析のセッションに移ると「易しい」と感じる学生はほぼいなくなった（大規模データのセッションで「やや易しい」と回答している1例は、自由記載欄に「わかりにくく時間がかかった」と入力しており、回答の選択ミスと考えられる）。

授業で学んだ知識や技術が、学業や将来の仕事に役立つかどうかを問うた「授業で学んだ内容の有用性」は、いずれのセッションも「大いに役立つ」と「役立つ」を合わせた回答が95%程度を占めた。エクセル操作のセッションで「まったく役立たない」とした回答が4例あったが、いずれも同一の学生である。この学生は15回の授業を通じて授業の有用性を見出すことが出来ず、特に「難しすぎる」と回答した回に「まったく役立たない」と回答する傾向があった。

授業評価5項目のいずれも、VLOOKUP関数やIF関数のネストを扱った第4回、煩雑なデータ編集を扱った第14回、授業で身につけた知識と技術を駆使して独力でデータ分析をしなければならなかった第15回で低迷した。

3.2 授業評価項目の満足度

授業評価の5項目において、満足度が75%を下回った学生が最も多かったのは、「授業の難易度」で、18名中7名の満足度が低かった。この中には、エクセル操作のセッションでの「易しすぎる」とした不満も含まれるが、概ね「難しい」とした評価によるものである。

「教員の説明の明瞭さ」と「内容の理解度」とについては、それぞれ5名の満足度が75%を下回った。前者の内訳は、前期の成績が振るわなかった学生が2名、それ以外の学生が3名、後者は前期の成績が振るわなかった学生が3名、それ以外の学生が2名であった。

進度に関しては、前期の成績が振るわなかった学生1名の満足度が75%を下回ったほかは、概ね満足度は高かった。

「授業で学んだ内容の有用性」は、1名の満足度が

†合計点を割合に直し、小数で表示している。

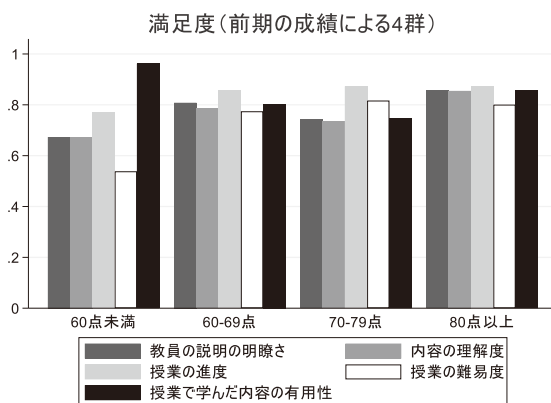


図6 ICTスキル別の授業満足度†

46.7%と特出して低かったほかは、やはり概ね満足度は高かった。

3.3 満足度の低い項目数

授業評価全5項目のうち、満足度が75%に満たなかった項目が一つもなかった履修者が18名中9名と最も多く、満足度の低い項目数の平均値は1.11項目であった。その中で、4項目で満足度が低かった学生が1名いた。この学生は、前期の「情報リテラシーI」および「調査と統計」のいずれも単位を取得できておらず、授業を「説明がわかりづらく理解できず、授業の難易度は高く、進度も速い」と受け止めていたことがわかる。その反面、扱う内容の有用性は、ほかのどの学生よりも強く感じていた。

3.4 前期の成績による4群の満足度の特徴

基本的なPC操作のみを扱う前期の「情報リテラシーI」の成績は、履修者が入学までに身につけてきたICTスキルに近似する。そこで、「情報リテラシーI」の成績をもとに、履修者を4群に分け、各群における満足度5項目を図6に示した。

ICTスキルが最も高いと思われる「80点以上」のグループにおいては、いずれの項目も高い満足度を得られている一方、その他のグループでは、進度以外の項目の満足度が概して低い。人数が1名である60点未満の群は、3.3で挙げた満足度の低い項目数が最も多かった学生と一致するが、この学生が授業を非常に難解と捉え、理解に苦しんでいたことがわかる。

対応のない2標本t検定の結果、前期の成績が80点以上の学生群と80点未満の学生群で、満足度5項目の平均値はそれぞれ0.85と0.788と、統計的に有意ではない($p=0.0749$)ものの、若干80点未満の群の満足度が低い傾向にあった。

4. 考察

本稿では、本学のデータサイエンス教育プログラムとしては初めて開講した2020年後期の「情報リテラシーII」について、エクセル操作を習いたいだけの学生から、簡単な統計の知識を持つ学生まで、多様なレベルの学生が履修した授業の3つのセッションごとに、学生の満足度を分析した。

難易度が急激に上がるデータ分析の回に入る時期に、学生の不満や反発が出ることを予想していたが、9割以上の学生が全15回の授業を通じて、扱った内容の有用性を高く評価していた。こうした意味では、授業の目的であった、履修者に初歩的なデータ分析の経験をさせるという試みは、概ね成功したと言える。

後半のデータ分析に入り、授業の進度を速いと感じた学生が少なくなかったことから、エクセル操作9回、データ分析6回という配分は、今後、若干変更する必要があるかもしれない。

また、今回は初の試みとして、「情報リテラシーII」の授業開始以降に企業との連携を図り、大規模データ分析の回の2週間ほど前にデータを入手したため、企業データの分析に必要なエクセルスキルを、それまでの授業で重点的に教えることができなかった。今後は、エクセル操作で取り扱う内容を、データ分析に必要な技術に特化するなどの工夫をし、無駄と学生の負担を省く努力も必要だろう。

授業の難易度については、履修者に予習復習の必要性を認識させ、反復練習を促すためにも、多くの学生にとって「やや難しい」レベルを目指すのが、大学教育としては妥当だと考えるが、「難しすぎる」と回答している学生については、個別に質問に応じるなどの対応が必要になる。本稿で紹介した栄養学科の「情報リテラシーII」は、例年履修者が20名程度であり、こ

うした個別対応がしやすいが、今後、同様の授業を人数の多い学科で実施する場合は、授業について行けない学生のフォローアップが大きな課題となるだろう。

前期の成績による満足度の違いの分析から、学生のICTレベルの差が課題として浮き彫りになった。これは情報リテラシー科目に共通する、授業運営上の最大の課題である。典型的な例として、前半のエクセル関数の部分で躓き、やる気を失ってしまった学生が、後半の授業について行けるかどうか心配された。

実際、今回の履修者の中に、前半で意欲を失ってしまったように見えた学生がいたが、この学生は「調査と統計」を履修しており、後半のデータ分析に入ると、授業内容が「調査と統計」の知識と結びつき、予想外に力を発揮するようになった。最終レポートは、単純集計の域を出ないものの、よく練られており、結果的に良い成績を収めることができた。毎回の授業の感想で吐露していた不平不満や、満足度の低さを考慮すれば、データサイエンス科目の履修の組み合わせと学生自身の努力が良い結果に結びついた例と言える。

今回、学生の大半がエクセルに苦手意識を持っていたため、エクセルの強化に9回もの授業を費やし、段階を追って、小規模データ、大規模データと進めていった結果、多くの学生が無理なくデータ分析、レポート作成と進めることができた。一方で、入学前に相当エクセルの知識と技術を身につけていた学生にとっては、エクセル操作の授業は退屈であったに違いない。

初期の回の進度が遅いと不満を持っていた学生がいたが、この学生は、後半の論理的思考や分析を非常に苦手としており、結果、後半で躓き、ほぼ満点だった「情報リテラシーI」のような良い成績を収めることができなかった。もしも「情報リテラシーII」が、エクセル操作よりも分析に時間をかける性質の授業であったなら、この学生は毎回の授業に危機感を持って臨み、多くの新しい知識と技術を学び、問題なく良い成績を収めていたかもしれない。当該学生は、分析に対する苦手意識から、「調査と統計」の履修を回避していたが、こうした学生こそ、「調査と統計」を履修することで、エクセルの技術を武器に、苦手な分析を克服で

きたかもしれない。こうした事例を、履修相談の際などに紹介する意義は高いだろう。

今回データサイエンス科目として定めた情報リテラシー科目および統計科目は、3科目のうち1科目で卒業要件を満たすため、3科目すべてを履修する学生は少ないが、今回、3科目すべてを履修した学生が8名いた。履修者の多くの最終レポートが単純集計にとどまった中、3科目履修した学生のうち2名は、「調査と統計」で学習した内容を駆使し、それぞれ相関分析、回帰分析を行っており、データサイエンス科目の組み合わせの効果の可能性を示した。

エクセル操作のスピードが速く、授業中に時間を持って余してしまう学生には、エクセル操作の応用課題を与えているが、手応えを考慮すれば、こうした空き時間に取り組むべき課題は、データ分析の応用課題が望ましい。現状、統計学の知識が全くない学生にデータ分析の応用課題を与えることはできないが、ICTスキルの高い学生には、こうした背景を説明した上で、積極的に統計科目の履修を促したいものである。同時に、ICTスキル別に、異なるデータサイエンス科目の履修選択が可能になるよう、柔軟なカリキュラムを提案していくことも必要であろう。

本研究は、本学におけるデータサイエンス教育活動の一環として実施した。開示すべき利益相反はない。本稿で紹介した事例が、同様にデータサイエンス教育の導入に試行錯誤する高等教育機関の参考になれば幸いである。

謝 辞

大規模データ分析のセッションのために、貴重なデータをご提供頂いた上に、授業にもご参加頂き、学生の学習意欲を大いに刺激して下さい、株式会社マルイの植松様、辻様に、心よりお礼申し上げます。また、「難しい」と悲鳴を上げながらも、最後まで授業に食らいつき、毎回几帳面に授業の感想書き、本研究のために情報として提供して下さい、本学短期大学部栄養学科の当時1年生の皆さんに感謝の意を表しま

す。

参考文献

1. 統合イノベーション戦略推進会議「AI戦略2019」令和元年6月11日
https://www.kantei.go.jp/jp/singi/ai_senryaku/pdf/aistratagy2019.pdf [参照2021年10月16日]
2. 竹村 彰通、和泉 志津恵、齋藤 邦彦、姫野 哲人、松井 秀俊、伊達 平和「データサイエンス教育の滋賀大学モデル」統計数理(2018) 66 (1) , 63-78
3. 富山 清升、庄野 宏「数理データサイエンス教育を鹿児島大学の全学必修分野として導入した経緯」鹿児島大学総合教育機構紀要(2021) 4, 101-116
4. 藤澤 修平、吉田 秀典、林 敏浩「香川大学における数理・データサイエンス教育の検討:一基礎科目を踏まえた応用科目への展開一」日本科学教育学会研究会研究報告(2021) 35 (6) , 1-4