Vol.2021-CE-160 No.7
2021/6/5

高等学校データサイエンスにおける導入教材の教育効果と有効性

林宏樹 [1]

概要：データサイエンス（以後 DS）の重要性が高まる中，いままで DS 教育の実践経験がない高等学校の現場は困惑している．そのような中で，一般社団法人データサイエンティスト協会では DS 導入に適した教材を開発している．本研究では，本教材を使った具体的な授業を提案し，高等学校における授業実践によって本教材の有効性を検証することを目的とする．本発表では，教材の内容を紹介するとともに，高校生を対象に実践した授業展開を，生徒の思考調査から教育効果を分析した結果について述べる．

キーワード：データサイエンス教育，データサイエンス教材，情報教育，情報Ⅰ，問題解決 [**]

Educational effects and effectiveness of the teaching material for introduction of data science in high schools

HIROKI HAYASHI[†1]

1. はじめに

Society5.0 と呼ばれる時代の到来が社会や生活を大きく変えていくとの予測がされる中，様々な情報を見極め，知識の概念的な理解を実現し，情報を再構築するなどして新たな価値につなげることができる人材の育成が求められている[1].

令和４年度より実施される学習指導要領[2] [3]では，「情報Ⅰ」に情報通信ネットワークとデータの活用，「情報Ⅱ」に情報とデータサイエンス（以下，DS）が位置付けられることとなった．また，「数学Ⅰ」にデータの分析（仮説検定の考え方を含む），「数学 B」に統計的な推測が位置付けられている．高等学校の授業内容に数理・DS に関する内容が多く含まれることとなり，DS 教育の重要性が高まっている．

一方，昨今，ビッグデータからビジネスに活用する知見を引き出す中核人材として「データサイエンティスト」に注目が集まっている．そのような背景から，高等学校では，DS の知識・技能を身に付けるだけでなく，データ活用を楽しめるような授業作りする必要がある．

令和 2 年度，DS 教育を実践する高等学校の教員で設立した「JDSSP 高等学校データサイエンス教育研究会」[4]において，データサイエンティスト協会の方からデータサイエンス教育に活用できる教材（以下，本 DS 教材）が提供された．

一般社団法人データサイエンティスト協会[5]は，新しい職種であるデータサイエンティストに必要となるスキル・知識を定義し，育成のカリキュラム作成，評価制度の構築など，高度 IT 人材の育成と業界の健全な発展への貢献，啓蒙活動を行っている．その活動の中で，本 DS 教材が開発された．本研究は，1 年間 DS 教育を受けた高校生に対して本 DS 教材による授業を行い，生徒の思考調査を行うことで，本教材の高等学校での有効性を検証することを目的とする．

2. 本 DS 教材について

2.1 教材の概要

一般社団法人データサイエンティスト協会によって，DS 教育の導入を目的として開発された本教材の名称は,「職業体験プログラム・データサイエンティスト・コンビニ出店計画編」である．

学習者がコンビニ会社のデータサイエンティストという立場となり，どこにコンビニを設置するのかを決めるという体験ワークである．３つの候補地から 1 か所に決定するプロセスから，データで決定する体験，複数のデータの中から必要なデータを取捨選択する体験，課題解決する体験を実践するというストーリーである．決定した候補地を発表するまでの一連の流れが体験できる構成となっている．

提供された DS 教材は，授業用のスライド教材，候補地決定のためのデータのエクセルファイル，スライド教材のシナリオが記されたエクセルファイルの３点である．シナリオのエクセルファイルに準じて授業を進めれば，初めて授業を実施する人でも授業実践できるように工夫されている．なお，シナリオ通りに実施すれば，85 分間で完結する内容構成である．

2.2 スライド教材の内容

本スライド教材を構成するスライドと，そこでの学習活動を以下に示す．また，典型的なスライドの例を図 1 から図 4 に示す．

1 兵庫県立姫路西高等学校，帝京大学大学院
Hyogo Prefecture Himeji Nishi Senior High School , Teikyo university.

ⓒ2021 Information Processing Society of Japan

スライド１（図 1）
体験ワークのテーマについて説明する．
スライド２
地図からコンビニを設置する場所を考えさせる．
スライド３
解決のために必要なデータを考えさせる．
スライド４
必要なデータの集め方を考えさせる．
スライド５（図 2）
候補地と既存店に関するデータを提示する．
スライド６
データを活用して候補地を決定させる．
スライド７（図 3）
既存店に関する散布図を提示する．
スライド８
散布図からデータの読み方の事例を指導する．
スライド９
どの要因をもとに候補地を選んだのか考えさせる．
スライド１０
決定した候補地を発表するフォーマットを提示する．
スライド１１（図 4）
データ活用の決定プロセスの一連の流れを復習する．
スライド１２
回帰分析法を用いた手法を提示する．

2.3 本 DS 教材で用意されているデータについて

スライド５（図 2）における生徒に与えるデータである候補地のデータと既存のデータについて説明する．候補地，既存のコンビニのデータは，A 店前交通量，B 視界性評価，C 間口の広さ，D 店舗面積，E 店長の年齢，F 駅からの距離，G 駐車場台数，H 周辺人口，I 競合店舗数，J 平均日販が用意されている．なお，各データ間の相関係数は表 1 の通りであるが，教材中には相関係数は示していない．

図 1　スライド 1
Figure1 Slide 1

図 2　スライド 5
Figure2 Slide 5

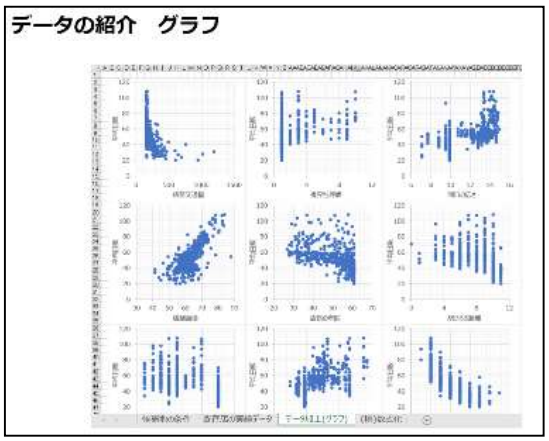

図 3　スライド 7
Figure3 Slide 7

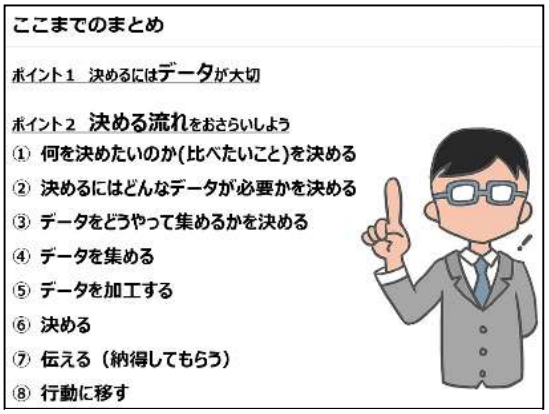

図 4　スライド 11
Figure4 Slide 11

ⓒ2021 Information Processing Society of Japan

表１　各変数間の相関係数の一覧表
Table 1　List of correlation coefficients between each variable

	B	C	D	E	F	G	H	I	J
A	-0.19	-0.41	-0.21	0.35	0.38	0.55	-0.29	0.44	-0.45
B		0.52	0.30	-0.56	-0.66	-0.32	0.38	-0.40	0.40
C			0.39	-0.64	-0.69	-0.52	0.39	-0.61	0.60
D				-0.25	-0.21	-0.12	0.28	-0.62	0.71
E					0.73	0.42	-0.27	0.61	-0.49
F						0.44	-0.24	0.50	-0.35
G							-0.38	0.43	-0.39
H								-0.36	0.52
I									-0.73

2.4 教材の授業展開の提案

統計的探究プロセスには,国際的な枠組みとして PPDAC サイクルがある．Problem（問題をとらえる），Plan（計画を立てる），Data（データを集める），Analysis（分析する），Conclusion（結果を読み取る）の５つのフェーズを繰り返しながら問題を解決するステップを定義したものである[6]．本教材は使い方の工夫をすることで，様々な段階の生徒に対応できる教材である．以下に，高等学校における本教材を用いた授業展開例として，PPDAC サイクルのステップのうちのどの部分を重点的に体験させるかによって，入門編，標準編，発展編の 3 つの例を提案する．

提案１　（入門編）
結論（Conclusion）を身に付けるための授業
データサイエンス教育の導入として，スライドの１～１２を教員のファシリテートに基づいて順序立てて授業を展開していく．生徒には可視化した状態のデータを与え，グラフの特徴の読み取り方，見方・考え方を重点的に指導していく授業展開．

提案２　（標準編）
Analysis（分析する）・結論（Conclusion）を身に付けるための授業
スライド８を除いた１～１２で授業を実施していく．スライド７は生徒に提示するだけで，生徒には散布図行列を表計算ソフトで作成する作業をさせる．そして，散布図行列を個人やグループで考え，結論を見出させる授業展開．

提案３　（発展編）
Plan（計画を立てる）・分析（Analysis）・結論（Conclusion）を身に付けるための授業
スライド７・８を除いた１～１２で授業を実施していく．与えられたデータから，可視化の方法，データ解析手法，結果の考察を生徒に考えさせ,結論を見出させる授業展開．

なお，すべての授業展開において，データから結論を見出す体験ができる．また，具体的な技能として，グラフからデータの特徴を見抜く「データ解析力」，表計算ソフトでデータを可視化する「データ研磨力」，結果に対して「考察力」など，どの力を身に付けることを目的とするのか工夫できる．発展的には,「情報Ⅰ」における単回帰分析法,「情報Ⅱ」における重回帰分析法の指導も可能であると考えられる．

3. 本教材を用いた実践授業

3.1 授業の概要

2020 年，データサイエンス教育を学校全体に実装するカリキュラム開発校として，文部科学省のスーパーサイエンスハイスクールに兵庫県立姫路西高等学校（以後,姫路西）が採択された．著者らは，本プロジェクトにおいてカリキュラム開発を行うと共に，データサイエンスを基盤とした課題研究を実施するために，1 年次においてデータサイエンス教育を実施した[7]．

本教材を用いた授業は，2020 年度の 1 年間，そのデータサイエンス教育カリキュラムで学んだ国際理学科 1 年（39 名）に対して，2021 年 3 月 18 日に実施した．

今回は,2.4 で述べた提案３（発展編）によって実施した．

3.2 生徒による候補地の選定結果

授業展開に沿って，生徒が選定結果を Microsoft の Forms に入力するスタイルで授業を進めた．授業進行にともない，与えた情報によって，生徒による候補地の選定結果は次のようになった．

(1)　スライド２における地図のみによる候補地の選択
　候補地①4 名　候補地②29 名　候補地③6 名
(2)　スライド６における候補地・既存店のデータを与えた場合の候補の選定
　候補地①12 名　候補地②26 名　候補地③1 名

3.3 生徒の回答

生徒が候補地を選定した理由を以下に示す．

(1)　スライド２，３における生徒の回答

候補地①～③を選定した 39 名全生徒が大通りの道の幅から交通量の多さを推測したり，駅から近いので顧客が多くなるという地理的な要素を挙げている．

また，競合店舗数（まわりのコンビニ以外の店舗の種類）に関するデータ，交通量などの顧客が増加するための要素となるデータを必要と考えた生徒が多かった．

それに対して，立地面積や駐車場に関する情報というコンビニの設立に関する要素を必要と考えた生徒は 2 名いた．そして，既存のコンビニのデータが必要であると考えた生徒は 2 名いた．

ⓒ2021 Information Processing Society of Japan

(2)　スライド９における生徒の回答

まず，ほとんどの生徒が各変数間の相関係数を求めており，相関係数が高い要素を用いて結論を導くことが多かった．

候補地②を選定した 26 名の生徒のうち 16 名は，I 競合店舗数を選定した要素に挙げている．I と J の相関係数は絶対値が最も大きい - 0.73 である．

候補地①を選定した 12 名の生徒のうち 10 名が，D 店舗面積を選定した要素に選んでいる．D と J の相関係数は 0.71 であり，正の相関の中では最も大きい値である．

そして，39 名中，18 名の生徒が複数の要素を組み合わせて結論を導いた．

4.　生徒の変容や意識調査・感想

4.1　授業内における生徒の回答から見受けられる変容

(1)　スライド２，３の生徒回答に対する分析

コンビニ設立の候補地を決定するために，必要なデータを生徒に考えさせたとき，地理的な観点から顧客の増減に係わるデータ収集に着目する生徒が多かった．

その中で，立地面積や駐車場立地など土地に関するデータが必要であると考えた生徒が存在したことは, PPDAC サイクルにおける P（研究計画を立てる）部分において，多面的に課題を判断する力が養われていたと判断できる．

また，今回のコンビニ設立の目的である「売り上げ」との関係をきちんと念頭に置き，既存のコンビニのデータが必要であると判断した生徒が存在したことも姫路西の DS 教育の１つの成果であるといえる．

(2)　スライド９の生徒回答に対する分析

データを与えた後，データから意思決定をするプロセスとして，単にグラフの特徴から結論を出すのではなく，相関係数の活用，単回帰分析の活用をほとんど生徒が実施していた．データ解析において，自然に相関係数を求める習慣や単回帰分析を利用している生徒の様子は DS 教育の成果の１つである．また，単回帰分析を複数行い，その関係性から結論を見出す生徒が 46%おり，データを順序立てて，論理的に結論を見出す様子が見られた．

そして，重回帰分析法を用いた予測値を利用している生徒がおり，重回帰分析を習得し，活用した生徒も見受けられた．

4.2　本授業による生徒が実感した身についた力

姫路西では，統計的探究プロセスの PPDAC サイクルを「課題定式力」,「データ収集力」,「データ研磨力」,「データ解析力」,「考察力」の 5 つの力を意識させながら，生徒に指導をしている．

Problem（問題をとらえる），Plan（計画を立てる）の 2 つのステップで「課題定式力」, Data（データを集める）で「データ収集力」，Analysis（分析する）で「データ研磨力」と「データ解析力」の 2 つの力，Conclusion（結果を読み取る）で「考察力」を身に付けるものとして対応付けている．

そこで，本教材を用いた授業において，これらのうち，どの５つの力が身についたのか意識調査を行った．その結果，データ研磨力 9 名，データ解析力 21 名，考察力 9 名であった．

つまり，本教材を用いることで，統計的探究プロセスの PPDAC サイクルで分類すると，分析（Analysis）30 名，結論（Conclusion）9 名のステップに関して成長したと生徒は実感していることがわかる．

4.3　授業後の生徒の感想

本授業を受けて，「楽しかった」と答えた生徒が 97%，データサイエンティストという職業に関する調査では,「職業にしたい」と回答した生徒が 5%,「職業の選択肢の１つにする」と回答した生徒が 67%であった．

生徒の感想として，データ研磨やデータ解析などの技能的な部分で「難しい」と回答した生徒は 12 名であったが，データサイエンティストの「必要性」については 20 名が必要であると回答している．また，「楽しい」「やりがいがある」という意欲的な回答は 6 名であり，「AI にできないこと」「人間だからできること」という回答が 6 名であった．

5.　本 DS 教材の有効性に関する考察

本 DS 教材を用いた授業展開の検討と発展編の授業実践を通して，本 DS 教材は以下のような点で有効であると考えられた．

(1)　P（計画を立てる）のステップに有効である．

第一に，スライド３によって，統計的探究を実践する上での P（計画を立てる）のステップとして必要なデータを考察させ，スライド５によって用意されたデータと生徒の考えた必要なデータを比較することで，研究計画の立て方を学ぶことができた．

また，1 年間のカリキュラムの中の課題研究では生徒だけでデータ収集から行うことを経験しているが，今回の授業実践では既存店のデータが収集されている状態であった．それでも，与えられたデータに既存店の間口の広さ等の量的データや視界性評価等の質的データがあることで，生徒たちは，収集する必要のあるデータの視点・考え方，どのようなデータが必要であるのかなど，必要なデータ収集の参考になったようである．

(2)　A（分析）のステップに有効である．

データ解析において，様々なアプローチが考えられるため，生徒が身に付けた技能を駆使し，試行錯誤しながらデータから結論を導く体験ができた．また，単回帰分析，重回帰分析など，教科「情報」における内容を盛り込むこと

ⓒ2021 Information Processing Society of Japan

ができる．

(3) C（結果を読み取る）のステップに有効である．

最終的に候補地の選定を行うため，結果を読み取る体験ができた．また，クラス内で同じデータで分析しても，結論が異なる場合，どのような解析結果から結論を見出したかを討議する場面もあった．

(4) DS への興味・関心を喚起するために有効である．

データサイエンスの重要性が言われており，スーパーサイエンスハイスクールによって DS 教育を受ける機会を得た生徒であるが，現実的に社会で必要であるという実感を味わえることは少ない．しかしながら，本教材によって，高校生の生活範囲でも想像でき，データサイエンスを実践できることで生徒の意欲は喚起された．また，既存データをみて，何を説明変数にすればよいのか考える際，現実の実態がイメージできるため，生徒の主観的な考えと客観的なデータを結びつける経験ができ，データサイエンスの本質を味わうことができた．

(5) 高校生に職業としてのデータサイエンティストの重要性を意識させるのに有効である．

生徒の授業後アンケートでも，データサイエンティストが必要であることを実感した生徒は多く，候補地を決定するという解決策を必ず見出す難しさを痛感した生徒も多い．体験を通して，高校生にとってデータサイエンティストの価値が高まり，データを活用することの難しさから学ぶ意欲へとつながった．

6. おわりに

本 DS 教材を活用した高等学校での授業実践は，PPDAC サイクルにおける計画（Plan），分析（Analysis），結論（Conclusion）のステップの向上を目的とした教材として有効であった．また，高校生にデータサイエンスへの興味関心をもつきっかけとなる教材としても有効であることがわかった．

本研究では，本教材を用いた授業を姫路西高等学校における１年間の DS 教育を受けた生徒に実施したが，今後は DS 教育の入門として使用し，スタートアップでの役割としてどのような有効性があるのか検証したい．また，１つ１つの技能の定着は重要であるが，DS ってどんなことをするのかを仮想的に体験し，興味関心を喚起することでその後の DS 教育の効果を高めることを目的に使用することも検討したい．

高校生がデータサイエンスを身近に感じ，データを活用して意思決定するプロセスを楽しめる教材開発はデータサイエンスの垣根を取り払い，未来のデータサイエンティストの育成に寄与する．楽しく，かつ，効果的な教材を活用することで，質の高いデータサイエンスの素養を育むことにつながる．そして，教員側はデータサイエンスで身につく生徒の変容を見定めて実践することで教育効果が高まるであろう．

なお，本 DS 教材は「高等学校データサイエンス教育研究会」ホームページから問い合わせることで入手できる．今後，データサイエンティスト協会ホームページで公開される予定である．

謝辞

本実践は一般社団法人データサイエンティスト協会の教材の提供によって実施できた．教材および，教材の使い方のシナリオの提供などをご協力頂いた皆様に，謹んで感謝の意を表する．また，本論文をまとめるにあたって，ご助言いただいた帝京大学渡辺博芳教授に感謝する．

参考文献

[1] 文部科学省（2018）『高等学校学習指導要領 解説 総合的な探究の時間編』
[2] 文部科学省（2018）『高等学校学習指導要領 解説 情報編』
[3] 文部科学省（2018）『高等学校学習指導要領 解説 理数編』
[4] JDSSP 高等学校データサイエンス教育研究会 https://ds-education.com/
[5] 一般社団法人データサイエンティスト協会 https://www.datascientist.or.jp/
[6] 深澤弘美・櫻井尚子・和泉志津恵（2018）「統計的探究プロセスとその評価」，統計数理研究所，第 66 巻 第 1 号 15-36
[7] 林宏樹・笹嶋宗彦・大里隆也（2020）「高等学校におけるデータサイエンス教育のカリキュラム開発と実践，評価基準の作成」，情報処理学会，SSS2020 情報教育シンポジウム論文集

ⓒ2021 Information Processing Society of Japan