

高等学校「情報I」教科書および傍用問題集でのデータサイエンスの扱いについて

成瀬 浩健^{1,a)} 辰己 丈夫²

概要：2022年度からはじまる高等学校の科目「情報I」にむけて、各高校の情報科教員は具体的な授業設計に取り掛かり出しているところだろう。これまで「社会と情報」しか授業して来なかった教員の多くは、新たに取り組むことになるプログラミング等の指導に不安を感じているようだ。その一助となるよう多くの研究も行われ、指導プランなども開発されている。しかし、その多くがプログラミングに関するもので、データサイエンス分野に関するものが少ない。そこで、データサイエンス分野の指導案や教材を考えていく参考となるよう、「情報I」の教科書及び傍用問題集で、データサイエンス分野の扱いについてどのように扱われているか調べた。また、アクティブ・ラーニングを用いた授業展開を求められているので、その実習の扱いを中心に比較整理した。

キーワード：情報I, データサイエンス, データの活用と表現, データベース, データの分析

1. はじめに

高等学校では、あと数ヶ月で新しい「情報I」の授業がはじまる。「情報I」では、多くの高校で現在履修されている「社会と情報」で扱われていない、コンピュータの演算の仕組みとしての論理回路、プログラミング、そしてデータサイエンス（「情報I」ではデータの活用と表現）が必修となっている。

これまでこれらの項目を教えてきていない教員にとって、新たな教材を準備することは大きな負担となってくる。その中で、プログラミングの指導は話題になり研修の機会も多かったのに対して、データサイエンスをどのように教えていくかはほとんど話題になっていないと感じる。そこで、「情報I」の教科書および傍用問題集で、データサイエンスがどのように扱われているかについてまとめた。

2. 学習指導要領の記述と教科書

まず、新たな学習指導要領等でのデータサイエンスに関する記述部分を整理したのち、データベースに関する部分と、データの分析に関する部分に分けて、解説と実習内容（実習につながる記載）を示すことにする。

2 内容

(4) 情報通信ネットワークとデータの活用

ア 次のような知識及び技能を身に付けること。

(イ) データを蓄積、管理、提供する方法、情報通信ネットワークを介して情報システムがサービスを提供する仕組みと特徴について理解すること。

(ウ) データを表現、蓄積するための表し方とデータを収集、整理、分析する方法について理解し技能を身に付けること。

イ 次のような思考力、判断力、表現力等を身に付けること。

(ウ) データの収集、整理、分析及び結果の表現の方法を適切に選択し、実行し、評価し、改善すること。

図1 学習指導要領 [1] でのデータサイエンスに関する内容

2.1 学習指導要領でのデータサイエンスに関する記述

平成30年公示の学習指導要領 [1] の「情報I」において、データサイエンスに関する部分の記述は図1のようになっている。

また、学習指導要領解説情報編 [2] において、授業で扱う内容について、より詳細に示されている。

「ア (イ) データを蓄積、管理、提供する方法、情報通信ネットワークを介して情報システムがサービスを提供する仕組みと特徴について理解すること」では、情報システムにおけるデータの位置付けや、データベースについて理解することが求められている。データベースのしくみや、それを運用、管理するためのデータベース管理システムについても学ぶ。

データモデルとしては、リレーショナルデータベースのような構造化されたものだけでなく、多様かつ大量のデー

¹ 京都女子高等学校
Kyoto Girls' High School

² 放送大学
The Open University of Japan

a) naruseh@kyoto-wu.ac.jp

データを扱うもの、構造化されていないものについても触れることなどを理解させるよう求めている。

「**ア (ウ)** データを表現、蓄積するための表し方と、データを収集、整理、分析する方法について理解し技能を身につけること」では、データをファイルとして蓄積するための様々な形式、データ処理の流れについて理解する。データの形式では、表形式をはじめとして、時系列データ、ネットワーク型、キーバリュ形式といった様々な形式のデータを扱う。

名義尺度、順序尺度、間隔尺度、比例尺度などのデータの尺度水準の違い、文字情報として得られる「質的データ」と数値情報として得られる「量的データ」などの扱い方の違いを理解するようにする。

データの分析で、基礎的な分析及び可視化の方法、多量のテキストから有用な情報を取り出すテキストマイニングの基礎やその方法を理解する。テキストマイニングの学習として、単語の出現頻度を調べタグクラウドを作らせ、単語の重要度や他の単語との関係性を捉える学習活動などが考えられる、などと示されている。

「**イ (ウ)** データの収集、整理、分析及び結果の表現の方法を適切に選択し、実行し、評価し改善すること」では、分析の目的に応じた方法を選択、処理する力。その結果について多面的な可視化を行うことにより、データに含まれる傾向を見いだす力を養う。データの傾向に関して評価するために、客観的な指標を基に判断する力、生徒自身の考えを基にした適正な解釈を行う力を養う。

また、データ整理の際に表計算ソフトウェアだけでなく統計ソフトウェアで扱うことも考えられる、としている。

さらに文部科学省が補足する資料として提示した「平成30年改訂の高等学校学習指導要領に関するQ&A <情報に関すること>」[3]において、上記内容や、「数学I」のデータ分析の、分散、標準偏差、散布図、相関、仮説検定など関連づけて扱うことが述べられている。

2.2 「情報I」の教科書

文部科学省の令和4年度使用教科書目録[4]によると、情報Iの教科書は以下の通り6社による12種類である。各教科書は、便宜上記号(A~F)をつけて表すことにする。

東京書籍

- A-1 新編情報I
- A-2 情報I Step Forward!

実教出版

- B-1 高校情報I Python
- B-2 高校情報I JavaScript
- B-3 最新情報I
- B-4 図説情報I

開隆堂

- C 実践 情報I

数研出版

- D-1 高等学校 情報I
- D-2 情報I Next

日本文教出版

- E-1 情報I
- E-2a 情報I 図解と実習 図解編
- E-2b " 実習編 (合冊)

第一学習社

- F 高等学校 情報I

B-1とB-2は書かれている内容は基本的に同じで、プログラミングの単元で使っている言語が異っているだけなので、1種類と扱うことにする。E-2aとE-2bは綴じ込んだ形の合冊となっていて、1種類と扱うことにする。よって、6社11種類の教科書について、データサイエンスの扱いについて内容を比較する。

3. 各教科書でのデータベースの扱い

各教科書で、どのようにデータベースが取り上げられているか表1にまとめた。

3.1 データベース分野の解説の扱い

各教科書ともに、以下の内容の解説を扱っている。DBMS、データモデルやデータの形式(リレーション型だけでなく、階層型やネットワーク型、キーバリュ型に触れる)、量的データ・質的データ、尺度について、全教科書で取り上げている。

また、データベースに関して説明だけで実習を伴わないものが1種類、ソフトウェアを使わず印刷されたテーブルのみで実習しているものが1種類ある。

3.2 表計算ソフト(Excel)を用いた実習

多くの教科書(3社6種類)でデータベースの実習にデータベースソフトウェアではなく、表計算ソフトウェアを用いていて、単一の表として抽出や並べ替えに取り組んでいる。表計算ソフトの使い方のページでの実習として取り上げているものも1社2種類ある。

3.3 MS-Accessを用いた実習

2014年に長瀧らが行った調査[5]では、「情報の科学」の教科書においてリレーショナル型データベースを学習するソフトとして、マイクロソフト社のアクセスを用いた実習が取り上げられている。しかし、今回「情報I」の各教科書を見て、データベース実習でMS-Accessを使うことを意識したものは無かった。

ただし、巻末資料でMS-Accessの使い方の解説をしているものが1種類あった。

表 1 教科書と傍用問題集でのデータベース実習の扱い

教科書	実習ツール	実習内容, 例題	解説
A-1, A-2	表計算ソフト, sAccess	抽出, 並べ替え, 追加, 結合, 選択, 射影	DBMS, データモデル
B-1, B-2	表計算ソフト, (発展)Python または JavaScript	抽出, 並べ替え, (発展) キー・バリュー型	DBMS, 結合, 選択, 射影
B-3	表計算ソフト	並べ替え, 抽出	DBMS, データモデル
B-4	表計算ソフト	並べ替え, 抽出	DBMS, データモデル, 結合, 選択, 射影
C	オープンデータ	結合	データモデル, MS-Access の使い方
D-1	表計算ソフト, 印刷された テーブル	並べ替え, 抽出, 結合, 選択, 射影	DBMS, データモデル
D-2	表計算ソフト	並べ替え, 抽出	DBMS, データモデル
E-1	印刷されたテーブル	結合, 選択, 射影	DBMS, データモデル
E-2a, E-2b	sAccess	選択, 射影, 結合	DBMS, データモデル
F	—	—	DBMS, データモデル, 射影, 選択, 結合

3.4 sAccess を用いた実習

sAccess[6] とは, 東京書籍 (A) の執筆者の一人である兼宗進氏らの開発したもので, オンラインで RDB の学習ができるツールである。

4. 各教科書でのデータの分析の扱い

各教科書で, どのようにデータの分析が取り上げられているか表 2 にまとめた。

4.1 データの分析分野の解説の扱い

各教科書ともに, 以下の内容について実習を伴わない解説を行なっている。

量的データ・質的データ, 尺度水準, テキストマイニングについて, 全教科書で取り上げている。また, オープンデータ, ビッグデータといった項目もほとんどの教科書 (6 社 10 種類) で取り上げられている。

学習指導要領解説では, 基礎的な分析と可視化が求められているが, どの教科書もその扱いは解説だけでなく実習を伴うようになっている。

4.2 表計算ソフト (Excel) を用いた実習

全ての教科書で, データの分析に表計算ソフトを使用していた。その実習内容は, 多くの教科書で, データをグラフで表すこと (5 社 7 種類) や分散・標準偏差を求める (4 社 7 種類), 散布図で表す (5 社 8 種類) というものを扱っている。どこまでを基礎的な分析とするか, 各教科書の編集方針の違いのようだ。発展的と思えるものに, クロス集計や回帰分析や検定を扱っており, 発展的内容の扱いには差があった。

4.3 統計ソフトを用いた実習

「情報 I」で求められている基礎的な分析は, 表計算ソフトですべて対応可能である。しかし, 「情報 II」に向けて学習を発展させるためには, 統計ソフトを使った実習を

行ってもよいのではないだろうか。

4.4 テキストマイニングの扱い

学習指導要領やその解説では, テキストマイニングの実習にも触れられているが, 実習を扱っていたのは 3 社 4 種類のみであった。テキストマイニングのツールを実習用の端末にインストールすることの困難さが, 扱われていない理由であろう。

5. 実際の授業プラン作成に向けて

各教科書でのデータベースやデータ分析の実習について, データベースや統計処理の専用のソフトを使わずに表計算ソフトで済ませているのは, 学校によって実習環境が異なっているあるいは実習環境の構築が困難である, といった理由があるからだろう。そのように, すべて表計算ソフトで代用してしまうのではなく, 実際に社会で用いられている環境に近い実習が望まれる。

そういった実習の実現に向けて, 筆者が「情報 I」に向けて試行的に行った授業を振り返り, 都高情研が学指指導要領を基に検討した授業時間数 [7] も参考にしながら, より現実的な授業プランを, 実際にこの単元の授業実施に間に合うよう 2022 年夏までに提案したいと考える。

5.1 データベースの実習に向けて

多くの教科書で扱っている, 並べ替えと抽出だけの実習ではデータベースに対するイメージを理解するには不十分だろう。例えば, リレーショナルデータベースのリレーションシップについて, 図表で示されるだけでなく, 実際にデータベースソフトで扱ってみてこそ理解が深まると感じている。クエリについても, 実際に操作してみないとちゃんとした理解には至らないだろう。

筆者はこれまで自宅でも復習できるよう, オープンソースの LibreOffice Base[8] を活用してきた。ただ, このソフトはコンピュータにインストールして使うものなので, 近

表 2 教科書と傍用問題集でのデータの分析の扱い

教科書	実習ツール	実習内容, 例題	解説
A-1, A-2	表計算ソフト	クロス集計, 散布図・相関	テキストマイニング
B-1, B-2	表計算ソフト	グラフ, クロス集計, 分布, 検定, 回帰分析	テキストマイニング
B-3	表計算ソフト	グラフ, クロス集計, 散布図・相関, 回帰分析	テキストマイニング
B-4	表計算ソフト	散布図・相関	テキストマイニング
C	表計算ソフト, テキストマイニングのソフト	グラフ, 分散・標準偏差, 散布図・相関, テキストマイニング	テキストマイニング
D-1, D-2	表計算ソフト	グラフ, 分散・標準偏差, クロス集計, 散布図・相関	テキストマイニング
E-1	表計算ソフト	グラフ, 散布図・相関, 回帰分析, クロス集計, 検定, テキストマイニング	—
E-2a, E-2b	表計算ソフト, テキストマイニングの Web サイト	分散・標準偏差, グラフ, 散布図・相関, テキストマイニング	クロス集計
F	表計算ソフト, テキストマイニングの Web サイト	グラフ, 分散・標準偏差, 散布図・相関, 代表値, クロス集計, テキストマイニング	データマイニング

年の実習環境を考えると, 学校での実習の続きを自宅で継続するような授業展開には不向きだと感じている. sAccess のようなオンライン教材で, Google クラウドで管理できるようなツールが望まれる.

5.2 データの分析の実習に向けて

どの教科書もエクセルを使うように書かれている. しかし, 多くの学校で Chromebook の採用が進んでいるようだし, 授業管理でよく使われてきている Google クラウドで管理できる Google のスプレッドシート [9] で, グラフ, 分散・標準偏差あたりの基礎的な分析まで実習できると考える.

その上で, 「情報 II」への継続も意識し, 回帰分析, クロス集計, 検定といった分析を, オープンソースの R, あるいはそのオンライン版である R-Studio Cloud[10] の利用を提案したい.

5.3 質の高い理解や大学入試に向けて

確かな学力の定着のため, アクティブ・ラーニングを意識した実習が必須だろう. 教師側が提示した練習課題をそのまま書き写した写経のような実習にならないようにしないといけない. そのための工夫のひとつとしてループリックがある.

教師の評価基準であるとともに, 生徒が自ら工夫して改良できる目標設定をしておけば, 生徒は反復して思考を重ね, 工夫することで理解が深まるだろう.

また, 実習で理解を深めた後, その理解 (知識や思考) が大学入試に向けて定着し続けているか, ということも気になる. 受験期まで学習の継続が有効と考えるが, 「情報 I」は 1 年時の履修でかつデータサイエンスにかけることが出来る授業時間は 6 週分くらいだ. 大学入試を考えると, 「情報 II」の履修あるいは別途補習なども必要となるかもしれない. どのような方策が取れるかは, 各校の事情によ

るため今後の研究対象とはしない.

「情報 I」でのデータサイエンス分野の実習について, 今回整理したことを踏まえ, 今後具体的な授業プランを提案したい.

参考文献

- [1] 文部科学省: 高等学校学習指導要領 (平成 30 年告示), 2018.
- [2] 文部科学省: 高等学校学習指導要領 (平成 30 年告示) 解説情報編, 開隆堂出版株式会社 (2019).
- [3] 文部科学省: 平成 30 年改訂の高等学校学習指導要領に関する Q&A <情報に関すること>, 入手先 (https://www.mext.go.jp/a_menu/shotou/new-cs/qa/1422454.htm) (2019.11.18 参照 2022.01.06).
- [4] 文部科学省: 高等学校用教科書目録 (令和 4 年度使用), 入手先 (https://www.mext.go.jp/a_menu/shotou/kyoukasho/mext_00002.html) (2021.04.23 参照 2022.01.06).
- [5] 長瀧寛之, 中野由章, 野部緑, 兼宗進: データベース操作の学習が可能なオンライン学習教材の提案, 情報処理学会論文誌, Vol.55 (1), pp.2-15 (2014).
- [6] 兼宗進: sAccess, 入手先 (<http://saccess.eplang.jp>) (参照 2022.01.06).
- [7] 中山享司 (都高情研): 次期学習指導要領「情報 I」年間指導計画とその具体案, 入手先 (https://www.zenkojoken.jp/wp-content/uploads/2019/07/wakayama_3a.pdf) (参照 2022.01.10).
- [8] The Document Foundation: LibreOffice Base, 入手先 (<https://ja.libreoffice.org/discover/base/>) (参照 2022.01.10).
- [9] Google.com: Google ドキュメントについて Google スプレッドシート, 入手先 (<https://www.google.com/intl/ja-jp/sheets/about/>) (参照 2022.01.10).
- [10] RStudio, PBC: RStudio Cloud, 入手先 (<https://rstudio.cloud/>) (参照 2022.01.10).