

# 高等学校情報科における通年型データサイエンス教育の取り組み

## —大阪教育学附属高等学校 WWL 事業「データサイエンス基礎」を例に—

若杉 祥太

大阪教育大学理数情報教育系

wakasugi-s83@ex.osaka-kyoiku.ac.jp

増田 高行

大阪教育大学附属高等学校平野校舎

masuda\_sjb@yahoo.co.jp

AI やビッグデータの活用が進む現在、研究機関のみならずビジネスの現場においてのデータサイエンス人材の育成が急務である。また、学校教育では、平成 30 年告示高等学校学習指導要領情報科にデータサイエンスに関する学習内容が言及されるなどデータサイエンス教育に注目が高まりつつある。

本稿では文科省 WWL の指定を受けた大阪教育大学附属高等学校 平野校舎での通年型のデータサイエンス教育の取り組みとしての「データサイエンス基礎」の授業実践や成果の概要を報告する。

### 1. はじめに

我が国では様々な機関や大学、企業を中心に Society5.0 の実現に向けてデータサイエンスの習得・活用が可能な人材育成が進められている。大阪教育大学では 2020 年度より文科省の WWL コンソーシアム構築支援事業「これからの新しい社会をリードし、世界で活躍できるイノベティブなグローバルリーダー」の指定を受け、同大学附属高等学校の平野校舎（以下、平野校舎と略記）と池田校舎に通年型のデータサイエンス教育の取り組みとし「データサイエンス基礎」を設定した。

本稿では、平野校舎における「データサイエンス基礎」の授業における単元や学習内容を中心とした授業実践と成果の概要を報告する。

### 2. データサイエンス基礎

授業実践の対象は全 1 年生 120 名、期間は 2021 年度の 1 年間とした。本授業実践は、高等学校の必修教科である情報科 2 単位のうち 1 単位を学校設定科目とし代替設置する（授業担当：増田）。平成 30 年度告示情報科学習指導要領による位置づけとしては、情報 I 「情報通信ネットワークとデータの活用」の分野において、データサイエンスで欠かせない「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」に対応する。

授業の目標は、①課題解決やイノベーションをするための統計やエクセルが利用できること、②データ処理や機械学習の理解のための Python のプログラミングを知ること、③データと AI に関する倫理について理解を深めることとした。表 1 に、学習する単元・学習項目・学習内容を示す。

授業では、一人一台の windows PC を使い Microsoft Excel を活用した。情報科教科書は用いず、配布プリントと副読本として「はじめての AI

（優しく知りたい先端科学シリーズ 6）、土屋誠司、創元社）、「問題解決ができる！武器としてのデータ活用術 高校生・大学生・ビジネスパーソンのためのサバイバルスキル、柏木吉基、翔泳社」を用いた。

授業方法は、既有知識とつなげる導入（講義）と一人一人の体験的な学習（演習）を組み合わせることを基本とした。また、生徒の文理を問わず、データサイエンスに関わる話題を幅広く扱い基礎的な知識や技能の習得とで実際のデータサイエンスの経験を通じて思考・判断・表現力、主体的に取り組む態度を養うようにした。

### 3. 授業実践の成果

授業実践の前後でアンケート調査を実施した（平野校舎：有効数 99 名）。調査項目は、授業の目標①を評価するために、平成 30 年告示情報科学習指導要領やその解説、教員用研修資料を参考に作成した統計的な知識（10 項目）、情報学的技能（9 項目）、問題解決に向けた流れ（10 項目）の 3 つのカテゴリの計 29 項目とした（5 件法）。

事前事後調査別の平均値と標準偏差および t 検定の結果（対応有、両側、有意水準 5%）、事前調査では、いわゆる平均値や中央値など現在に至るまで馴染み深い統計的な知識の平均値は高いものの、PC 操作が伴う情報学的な技能や問題解決に向けたデータの整理・変換・分析・可視化の平均点が低かった。事後調査の結果では、全項目の平均値の向上が認められた ( $p < 0.01$ )。とりわけ、情報学的技能に関する平均値は大きく向上する結果となった。また、同様に問題解決に向けたデータの整理・変換・分析・可視化の平均値が大きく向上した。

自由記述調査では、データサイエンスのイメージや学習歴、知っていること、役立つことを授業

実践の事前事後に自由記述形式で実施した。テキスト分析には、kh-coder3 を用いた。その結果、事前調査では、データサイエンスに関する学習歴有（あてはまる）は全体の 18.8%であった。共起関係を明らかにするため、Jaccard の類似性測度を用いて共起ネットワークを作成した。その結果、事前のデータサイエンスのイメージでは「パソコン-使う」、「データ-活用」「堅苦しい（難しい）」という語のつながりが見られた。事後のデータサイエンスが役立つことでは、「データ-分析」、「答え-導き出せる-技術-必要」などの語のつながりが見られた。この結果からも、単なる前述の平均値の向上が認められただけでなく、データサイエンスの学習を通して本来の意味でのデータサイ

エンスの習得につながっていることが示唆された。

#### 4. おわりに

本稿では、文科省の WWL 指定を受け平野校舎で実施される通年型のデータサイエンス教育の取り組みである「データサイエンス基礎」の授業実践と成果の概要を報告した。本稿では紙面の都合上、成果の概要を述べるに留まったが、今後更なる分析と考察を行う。また、今回の実践の成果と課題を真摯に受け止め、課題の改善と探究的統計プロセスを加味した 2022 年度版の「データサイエンス基礎」として再度授業実践を行い、社会活動における意思決定や問題解決に役立てられるデータサイエンス教育を探究する。

表 1 「データサイエンス基礎」の概要（単元・学習項目・学習内容）

単元	学習項目	学習内容
【1】データサイエンスの導入 ・DSの意義と歴史 ・人工知能と機械学習 ・DSのよさ	データサイエンス（教育）の位置づけ	Society5.0、WWL、統計と機械学習の概要、ビッグデータ、高校の他教科との関係
	AIと歴史（概要）	身近にあるAI事例、AI、AI史
	データの使い道（意思決定と機械学習、自動化など）	問題解決思考（PDCA）、データを根拠とすること、データ活用サービス・アプリ
【2】データサイエンスの基礎 ・データ ・グラフ ・表計算ソフトの活用① ・記述統計 ・倫理	データ、オープンデータ	身の周りのデータ、データの定義、質と量、尺度水準
	Excelの使い方と集計	セルとシート、条件付き書式など
	グラフ作成、分布の可視化	棒グラフ、円グラフ、折れ線グラフ、帯グラフ、箱ひげ図
	グラフに関わる注意点	グラフの印象、グラフの使い分け
	数値計算	基本演算、べき乗、セル参照、SUMなど
	分布と中心化傾向	度数分布とヒストグラム、平均値、中央値、最頻値、代表値の使い分け、外れ値
【3】データサイエンスの応用（統計編） ・問題解決とデータ分析 ・アンケート調査 ・表計算ソフトの活用②	データの取得と取扱いに関する現代的な倫理課題	世論調査、オプトアウト、人種差別などのデータに関わる社会的な事例、ELSI
	散布度	分散、標準偏差、偏差値、四分位数、箱ひげ図、代表値・標準偏差・偏差値の数値比較等
	相関分析と回帰分析	散布図、相関係数、回帰直線、相関と因果、相関係数や散布図の比較、予測等
【4】データサイエンスの応用（AI編） ・プログラミング ・人工知能	クロス分析	連関、クロス集計表、ピボットテーブル、φ係数、YuleのQ、アンケート分析、欠損値
	プログラミングとPython	Python、簡単なプログラミング
	Pythonでのグラフ作成、数値計算	基本演算、Numpy、Matplotlibの利用
	Pythonでのデータ処理	CSVファイルの処理
	AIと人間の在り方	AIの事例、これからの職業とAI、IA