

高校生を対象とした初学者向け データサイエンス実習の設計と実践

大場 保幸^{†1}

キーワード：高校データサイエンス教育，オープンデータ活用，生成AI（Pythonコード生成），探究的なデータ解釈

1. はじめに

現在、日本の教育現場では、スーパーサイエンスハイスクール（SSH）やDXハイスクールの指定により、ICT環境の整備が急速に進んでいる。SSHでは統計やデータ分析を用いた探究活動が先行して行われており、DXハイスクールにおいては補助金や「情報II」、「数学B・C」の拡充を通じて、データ活用能力の育成が期待されている。

しかし、高校現場における「データサイエンス（DS）教育」の実態は、いまだ限定的であるのが現状である。多くの場合、Excel等の操作スキルの習得に終始してしまい、データの背後にある意味を深く考える「探究」の段階まで至らないことが多い。

本研究の目的は、生徒がDS実習を行う際の具体的な参考となる説明案および実践案を作成することである。特に、生徒がデータサイエンスを「自分事」として捉え、技術的な障壁を下げつつ、データが直感と異なる結果を示した際にどのように思考を深めるべきか、そのプロセスを体験させる実習を設計した。

2. 実習の教材設計とツールの役割分担

実習の題材として、生徒にとって身近な「静岡県の農業（きゅうりの分析）」を選択した。使用データは、生産現場の状況を反映した以下のオープンデータである。

- ・収穫量・作付面積：ジャパンクロップスより取得
 - ・気象データ：気象庁より取得（日照時間、気温、降水量）
- 分析における工夫として、単なる「収穫量」の比較ではなく、収穫量を面積で割った「単位面積当たりの収穫量（t/ha）」を算出させるステップを導入した。これにより、生産の「効率性」というDS的な視点を持たせることができた。

また、生徒の技術的障壁を解消するため、ツールの役割を明確に分担させた。データの収集と基本的な加工には使い慣れた表計算ソフトを用い、相関分析や高度な可視化には、生成AI（Gemini等）を活用して作成したPythonコードを用いることとした（図1）。これにより、プログラミングスキルの習得そのものを目的とするのではなく、分析結果の解釈に時間を割くことが可能となった。

```
# データの前処理: 年と必要な値（年の値）を抽出
# 気象データは「年」と「年の値」カラムを使うと仮定
def clean_weather_data(df, col_name):
    # 年と年の値（合計や平均）を抽出
    # 実際のCSV構造に合わせてカラム名を調整してください
    temp_df = df[['年', '年の値']].copy()
    temp_df.columns = ['年', col_name]
    return temp_df

df_rain_clean = clean_weather_data(df_rain, 'Total_Rain_mm')
df_sun_clean = clean_weather_data(df_sun, 'Total_Sun_h')
df_temp_clean = clean_weather_data(df_temp, 'Avg_Temp_C')

# きゅうりデータの整理（数値型への変換など）
df_cucumber = df_cucumber[['年', '単位面積当たりの収穫量(t/ha)']]
df_cucumber.columns = ['年', 'Yield_per_ha']

# データの結合
merged_df = df_cucumber.merge(df_rain_clean, on='年', how='inner')
merged_df = merged_df.merge(df_sun_clean, on='年', how='inner')
merged_df = merged_df.merge(df_temp_clean, on='年', how='inner')

# 相関行列の計算
correlation_matrix = merged_df.corr()
```

図1 生成AIが作成したソースコードの一部

3. 実践プロセス：仮説の立案と検証

実習の導入として、まず生徒に仮説を立てさせる。多くの生徒は「日照時間が長いほど光合成が促進され、収穫量は増えるはずだ（正の相関）」という直感的な仮説を立てることが予想される。しかし、実際に静岡県のデータを用いて相関分析を行った結果、驚くべきことに日照時間と単位面積当たりの収穫量の間には「-0.58」という「負の相関」が見られた（図2,表1）。

通常、教育現場ではこれを「分析の失敗」や「予想外のエラー」として片付けてしまいがちであるが、本実践ではここを「指導の転換点」と位置づけた。なぜ直感と異なる結果が出たのかという「なぜ？」の問い合わせ立てることで、生徒をDSの真の醍醐味へと繋げる工夫を行った。

4. 多角的な分析とドメイン知識との接続

生徒がこの矛盾を解消するために行ったのは、データの深掘りと考察である。日照時間単独ではなく、平均気温や降水量のデータを組み合わせた多角的な視点での検討を促した。

考察のプロセスでは、以下のようない仮説が導き出された。

- ・日照が多いということは、猛暑である可能性が高い。その結果、気温が植物の生育限界を超える温度（高温障害）

^{†1} 静岡県立浜北西高等学校

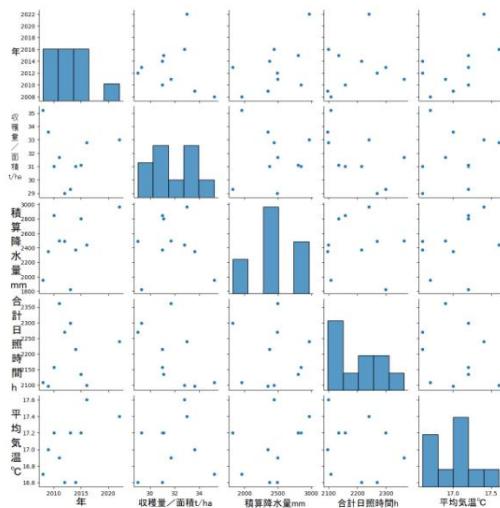


図 2 変数間の散布図およびヒストグラム

表1 相関係数マトリックス

静岡県における単位面積当たりのきゅうり収穫量と気象条件					
	年	収穫量(t/ha)	降水量(mm)	日照時間(h)	平均気温(°C)
年	1.00	-0.10	0.51	0.16	0.55
収穫量(t/ha)	-0.10	1.00	-0.05	-0.58	0.12
降水量(mm)	0.51	-0.05	1.00	-0.05	0.34
日照時間(h)	0.16	-0.58	-0.05	1.00	-0.24
平均気温(°C)	0.55	0.12	0.34	-0.24	1.00

になっていないか。

・静岡県の農業特性である「施設園芸（ビニールハウス）」を考慮すると、外気温の上昇がハウス内の温度管理に顕著な影響を与え、結果として収穫量を押し下げているのではないか。

このプロセスを通じて、データは単なる数字の羅列ではなく、「農業の現場」というドメイン知識と組み合わされて初めて価値を持つということを生徒は学ぶこととなる。

5. 教育的示唆：解釈を教えるマインドセット

本実践を通じて得られた、DS 教育における核心は以下の 3 点である。

第一に、「計算」ではなく「解釈」を教えることである。計算やグラフ化といった作業は、今や AI やソフトウェアが肩代わりできる。人間が担うべき役割は、「なぜその数字になったのか」という背景を想像し、物語を紡ぐことにある。

第二に、AI との適切な付き合い方を促すことである。コードを書かせること自体を目的化せず、自分の仮説を検証するための「道具」として AI を使いこなす姿勢を強調した。

第三に、「外れ値」や「逆の相関」を大切にすることである。予想通りにいかないデータこそが新しい発見の入り口であるというマインドセットを、教員のアドバイスを通じて生徒に伝えた。

6. 成果と今後の課題

本実践の成果として、表計算ソフトの操作に習熟していない生徒であっても、AI の補助を得ることで短時間のうちに「高度な分析結果を解釈する」ステージまで到達できることが示された。

今後の課題としては、生成 AI が output するコードの正確性を判断するための最低限のリテラシー教育が必要である点が挙げられる。今後は、農業以外の分野（観光や交通など）への横断的な展開を図るとともに、教員の役割を「操作の指導者」から、生徒と一緒に悩む「探究の伴走者」へとシフトさせていくことが重要である。

7. まとめ

2003 年の教科「情報」新設から現在に至るまで、日本の情報教育は大きな転換点を迎えている。

共通点は、一貫して「情報の科学的な理解」「情報活用の実践力」「情報社会に参画する態度」の三柱を基盤としている点だ。2003 年当時から、単なる操作スキルの習得にとどまらず、情報を主体的・批判的に活用し、問題解決に繋げる能力の育成を目指してきた。

一方、相違点は教育の「深化」と「目的」にある。かつての教育は、PC の基本操作やリテラシーといった「情報社会への適応」に主眼が置かれていた。しかし、現在のデータサイエンス人材育成では、プログラミングや統計的手法を用いてビッグデータを解析し、新たな価値を創出する「社会の変革者」の育成へと高度化している。

特に現行の「情報 I」では、AI やデータの収集・分析が必修化され、数学との連携による統計的探究が重視されている。大学入学共通テストへの採用も相まって、全ての生徒がデータサイエンスの基礎を備える体制へと進化した。2025 年現在は、ICT を単に「使う」段階から、データを武器に「新たな知見を導き出す」段階へと、教育の質的転換が明確に図られている。

本実践が、これからデータサイエンス教育における一つのモデルケースとなり、若い世代がデータサイエンスを活用して社会活動が大きく進展することを期待する。

謝辞

本研究の実践にあたり、生成 AI を活用したデータ分析の手法について多大なる御指導・御助言をいただいた、静岡県立浜北西高等学校の原弘忠教諭に深く感謝の意を表します。

参考文献

- [1] ジャパンクロップス（静岡県のきゅうり収穫量）：
<https://japancrops.com/> (2025 年 12 月 15 日閲覧)
- [2] 気象庁（過去の気象データ検索：静岡県の日照時間、降水量、平均気温）：
<https://www.jma.go.jp/jma/menu/menureport.html> (2025 年 12 月 15 日閲覧)