

データ活用の基礎教育 ～社会人教育における事例から～

立石 亨

公共システム政策研究所

ttateisi@asahi.email.ne.jp

著者が近年社会人向けに講義を続けてきた「データ活用の基礎」の内容について報告する。

当教科は、データにもとづく分析や報告を様々な場面で求められる行政職員に対して、「データそのものあるいはデータを使って成された表現を分析し、理解する能力」と「データを用いて事実や意見を的確に表現する能力」とを養う講座の入門編として企画されたものである。講座の内容とともに、受講者の反応がどのような点に集まっていたかを報告する。

キーワード：情報科教育，データ活用，社会人教育，統計以前

1. はじめに

筆者はながらく国，地方公共団体といった公共分野における情報技術の適用に「利用側」を支援する立場から携わってきた。そこでは、情報を取り扱う道具としてのコンピュータやネットワークなどの技術全般についての利用者へのアドバイスや計画づくりの支援などとともに、情報にもとづいて分析・判断し、情報をもとにした「事実」を伝え、その事実にもとづいて予測する活動が必要とされている。例えば「(データや数理にもとづく)科学的行政」という言葉は筆者の経験のごく初期から耳にした言葉であり、かつ今もキーワードになり得ることに鑑みると、これを支える知識や人材は、なお今日でも育成が必要とされていることは間違いが無い。

本稿は、こうした認識のもと、社会人一般に行ってきたデータ活用の基礎教育の構成と受講者の反応について報告する。

2. 社会人におけるデータリテラシーの課題

リテラシーという言葉が示す「読み書きの能力」をデータについて適用すると「データそのものあるいはデータを使って成された表現を分析し、理解する能力」と「データを用いて事実や意見を的確に表現する能力」と整理することができる。

試みに平成29年告示の「小学校学習指導要路解説」の算数編をひもとけば、そこには「データ活用」教育のための3つの狙いと2つの「身につけるべき能力」が示されている^①。特に後者の「目的に応じてデータを収集、分類整理し、結果を適切に表現すること」および「統計データの特徴を読み取り判断すること」という内容には、社会人が身につけるべきデータ活用に

関する素養が投影されており、実際、これまで筆者が社会人に対して行ってきたデータ分析・活用の講座で求められてきた「育成目標」にも通じるものがあると考えられる。

3. データ活用教育の具体的方針

以下に社会人向けデータ活用入門教育の方針を示す。

- ① 世に行われるデータ分析を正しく理解できるようになること
 - 1) 対象データの性質を正しく理解できること
 - 2) 分析方法の用途や制限が理解できること
 - 3) 分析結果の主張が理解できること
 - 4) 対象データの扱い方、適用されている手法の適切さ、分析の主張の正しさが判別できること
- ② データを使って伝えたいメッセージを正しく表現できるようになること
 - 1) 伝えたいことを明確に認識できること
 - 2) 最も効果的かつ誤解を生まない表現方法が選択できること
- ③ より高度なデータ活用の意義を理解できるようになること
 - 1) より高度な現状分析手法についての導入
 - 2) 確からしさとは何であるかについての導入
 - 3) 将来予測の手法についての導入

4. 研修内容

前章の方針に基づいて、具体的な研修内容を次の通り構成している。なお、方針の③については、リテラシー向上のための入門としては、その入り口を示すことにとどめ、①②の中で触れることとしている。

表1 研修内容の構成

大項目	中項目	小項目
1 データの読み方	1-1 データの諸元の重要性	1-1-1 調査対象の差異 1-1-2 カテゴリーの差異 1-1-3 時間的性質の差異
	1-2 実値と比率	1-2-1 分母の選択 1-2-2 差異の読み取り 1-2-3 変化の読み取り 1-2-4 複数の変化群の分析
	1-3 相関と因果	1-3-1 疑似相関 1-3-2 因果を見つけ出す
2 データの見せ方	2-1 好例と悪例	2-1-1 円グラフのわかりづらさ 2-1-2 軸への注目 2-1-3 色の選択
	2-2 グラフの種類と適用方針	2-2-1 棒グラフと度数分布 2-2-2 箱ひげ図 2-2-3 折れ線グラフ 2-2-4 泡グラフ 2-2-5 散布図 2-2-6 ヒートマップ 2-2-7 円グラフ 2-2-8 レーダーチャート
	2-3 デザインとしての見せ方	2-3-1 インフォグラフィックス

具体的には各小項目を次の通りに研修内容に結びつけている。

1-1-1 調査対象の差異

調査手法や取得対象の定義によってデータの持つ意味が異なる例を示し、データを得た調査・実験手法の把握の必要性を学ぶ。

1-1-2 カテゴリーの差異

得られた時期や集計された観点によって、データが持つカテゴリーの定義が変化する例を示し、データカテゴリーの定義の把握の必要性を学ぶ。

1-1-3 時間的性質の差異

フローとストックという、データの持つ時間的性質を把握する必要があることを学ぶ。

1-2-1 分母の選択

異なる集団間での指標比較の際に「集団の規模の差異」を取り除く手段として「比率」を用いるという導入と同時に、比率における「分母の選び方」の留意点について学ぶ。

1-2-2 差異の読み取り

実値の推移や集団間の比較において、実値同士の比較以外にも、基準(基準時点や基準集団等)を定めて、これとの比較を行うことで、より差異に注目した読み取りや表現が可能になることを学ぶ。

1-2-3 変化の読み取り

時系列での指標の表現において、実値の他に「基準時点からの変化」や「一定期間前の時点からの移動変化」があること、ならびに適用の考え方について学ぶ。

1-2-4 複数の変化群の分析

時系列で変化する二つの指標について、双方を二軸

に配した「軌跡」を表現する手法について学ぶ。

1-3-1 疑似相関

相関があるように見える二つの指標には必ずしも因果関係が無いことを、過去の実例から学ぶ。

1-3-2 因果を見つけ出す

前項のより進んだテーマとして、因果関係を統計的に推論する方法の存在と、その基本的な考え方を学ぶ。

2-1-1 円グラフの問題点

実例をもとに、一般的なグラフのひとつとして数多く利用されている円グラフが、変形等の手段によってミスリードにつながる表現になり得ることを学ぶ。

2-1-2 軸への注目

目盛りの一部分省略や、二種類のY軸等、棒グラフや折れ線グラフで利用されることのある記法がもたらす「解釈の難しさ」について理解するとともに、適切に使うための考え方を学ぶ。

2-1-3 色の選択

グラフの線や面を色で塗り分ける場合の色使いについて、見やすさ、表現する内容との関連等の基本的な考え方を学ぶ。

2-2-1～2-2-8

各種のグラフ表現の性格と適用場面について学ぶ。

2-3-1 インフォグラフィックス

より進んだ表現方法として、グラフ表現をアイコンやピクトなどのイメージと組み合わせて、さらに直感的な理解を促進する技法について学ぶ。

5. おわりに

本稿の内容で、これまでに数百名の社会人を対象に講義を行ってきた。講義後の評価における反応の大多数は「データ分析・表現の手段には、分析側の主張ないしは意図が必ず含まれていることを認識できた」という点にあり、これに次ぐものが「グラフ表現の種類と適用場面の考え方を認識できた」というものだった。手法や表現の学習と、その適用範囲や場面を知ることには、他者によってなされた分析の表現を正しく理解することにもつながり、初学であるか否かを問わず重要なデータ活用の学びになると認識している次第である。今後は具体的方針の③に示した、より高度なデータ活用(検定、予測等)の意義に関する内容を充実させることとしたい。

参考文献

- (1) 小学校学習指導要領(平成29年告示)解説 算数編, pp. 67 (2018).