

「データの活用」を育む表計算ソフトウェア活用学習とその実践

Design and Practice of Learning Data Operation with Spreadsheet Software

鎌田 敏之

愛知教育大学技術教育講座

Toshiyuki Kamada

Department of Technology Education, Aichi University of Education

キーワード：情報に関する技術，表計算ソフトウェア，データの活用

Keywords: Technology of Information Processing, Spreadsheet Software, Data Operation

1. はじめに

中学校技術・家庭科の技術分野における「情報に関する技術」の内容は、平成20年改訂の学習指導要領¹⁾においては、情報活用能力の育成に関して「(2) デジタル作品の設計・制作について、次の事項を指導する」の項目を中心として行われることとなった。基本的な操作については小学校各教科の指導のなかで身に付け、高等学校「情報科」において情報の表現と選択や問題解決の学習において、より現実の問題に即した活用方法を実習していく構造といえる。中学校技術免許の教員養成においては、中学校学習指導要領に対応する知識・技能を教師として実践する能力を身に付ける一方で、高等教育の場にふさわしい、大学の学びに対応した情報活用のありかたを踏まえた授業が求められていると言える。

現代の情報化社会において、情報技術活用の核のひとつに、データの活用がある。企業や行政機関を問わず、経営判断や業務の適切な遂行のためにはデータを適切に収集し、分析することが欠かせない。このことは、加納らによる高等学校「情報科」検定教科書の用語分析²⁾において、語の共起関係のネットワーク図のなかで、「デジタル情報の表現と活用」語群とされるグラフのハブとなる位置に「データ」の語が高い頻度を表す大きな円として現れることからわかる。「データ」の語が情報活用に属するとされるのは、「表現」「表」「グラフ」などと有意な共起関係を示しているためことから理解される。したがって、情報活用教育についての学びにおいても「データの活用」を意識した演習が重要であり、ア

カデミックスキルとして、また教員として業務を高いレベルで遂行する能力を育成するためにも意義があると考えられる。

そこで、本稿では、中学校技術専攻の大学生を対象としたアプリケーションソフトウェア活用の授業において、表計算ソフトウェアを用いた「データ」の分析を行う授業を構想し、実践した結果について報告する。

2. 表計算ソフトウェアとデータ

本節では、教員養成学部における授業を構想するにあたり、学習者である教育学部学生の実態を考慮しつつ、使用するソフトウェアとデータについて考察した事柄について述べる。

2.1. ソフトウェア

コンピュータを用いたデータ処理について学ぶことは、理科系の学問では実験結果の処理を、人文社会系の学問では統計処理などを通して現実の問題を分析することを、効率よく行い、検討を加えるために別の条件による再計算を行うなど、学びの内容を充実させることにもつながる。このとき、データ処理を行う汎用的なアプリケーションとして、パーソナルコンピュータ(PC)とともに普及し、広く社会で使われている表計算ソフトウェアは、入門用として適当なソフトウェアであるといえる。

表計算ソフトウェアは学校教育においても小学校段階から利用されることがあるほど一般的となっているが、その嚆矢は1979年にApple Computer社(当時)の8bitパーソナルコ

ンピュータ Apple II 向けに発売された「VisiCalc」である。Apple II は当時のパーソナルコンピュータとしては高度な描画機能を持っていたが、現在のような高精細のグラフィクス表現までには至っておらず、「VisiCalc」も単色の文字と、文字領域の明暗反転状態で枠を表した表示のみで構成されているが、基本的なアイデアと操作の発明は、そのまま現代の表計算ソフトウェアに引き継がれている。

画面が行と列の格子に分割されており、格子内の領域を「セル」と呼ぶこと、カーソルキーによって（画面外を含む）任意のセルへと操作対象を移動できること、セルにはデータである数値や文字列をその場でタイピングして入力できること、セルは座標（行は 1, 2, 3 の数値で列は A, B, C のアルファベットで表され、その組み合わせを座標とする）を持つこと、セル座標を指定して、他のセル内のデータを参照して計算する「計算式」もセル内に記入できること、そして表に含まれるすべての計算は利用者の特別な指示なく自動的に行われ、参照するデータを書き換えればすべての式の再計算が自動的に行われること、これらすべてを「VisiCalc」は備えていた³⁾。「VisiCalc」はプログラミングの作業なしに、データと計算式をセルに書き込むだけで自在に、任意の帳票を作成し計算までを行うことができるため、家庭や小規模ビジネスにおける経理計算でのパーソナルコンピュータ普及に貢献してきた⁴⁾。コンピュータ活用教育が世界的に広まった 1980 年代⁵⁾には、高等学校数学での教育活用事例も報告されている⁶⁾。

現在広く使われている表計算ソフトウェアとしてマイクロソフト社の Excel があり、少なくとも日本国内においては Windows を搭載した PC と同時に購入することが一般的である。学生個人に授業用ノートブック型 PC の準備を入学時に求めている本学の授業においても、Excel を含むマイクロソフト社のオフィス製品をインストールしていることが同時に求められており、特に追加の出費なく共通の学習環境が実現できるという意味で、Excel を教材として用いることに十分な妥当性はあると考えられる。

2.2. データ処理

1999 年閣議決定された「ミレニアムプロジェクト」に盛り込まれた「教育の情報化」施策以来、2000 年開始の事業を通して全国のすべての学校を対象に急

速に PC 教室が整備されてきた。現在では Excel を用いて簡単な表を作成し、何らかの値を入力して合計や平均を求める手順については高等学校「情報科」の教科書に掲載されており、実際に小・中・高等学校の少なくともいずれかでこのような操作についての体験を経て入学する学生が大多数である。しかし、それを超える内容について具体的知識や技能を持つ学生は、2006 年立野らによる獨協大学新生を対象とした調査⁸⁾ではこれに最大値と最小値を求めるとグラフ化まで、2010 年石崎による福岡県立大学新生を対象とした調査⁹⁾では合計と平均及びグラフの作成までが高い割合で操作可能な範囲と報告されている。データ集計において、合計を求めることと平均値、入力データのグラフ化は無意味ではないが、グラフ化について、石崎は作成後に見た目の変更のできない割合が高いと報告しており、高等学校の実態として、単一の数値データについて、手入力と単純な操作を行うにとどまっている授業の姿が推察される。与えられるデータが生徒にとって操作する意味を感じさせるものかどうかについても疑問である。

アカデミックスキルとしてのデータ処理の側面を考慮するならば、この授業を終えた後に行う大学での学びのなかで、意欲のある学生が、授業その他の場面でふと気付きを持ち、自ら目的を持ってデータ処理に取り組もうと考えられるような水準の内容でなければ意味がない。そこで、構想する授業では、実社会のなかで集められた本物のデータを活用すること、一定の規模を持ち、処理する必要性が明確なデータセットを選択すること、データから何らかの発見があるような処理を行うことを考慮し、理科系人文社会系いずれにも関係のある、統計について学ぶことを大きな枠組みとすることとした。

3. Excel と統計

大学学部生を対象とした統計学入門授業において、Excel による演習を含むための教科書は、少なくとも Windows 98 により家庭や学校にオフィスアプリケーションが普及した時期以後の版が図書検索によって複数確認できる。

PC が入手しやすく、直感的な操作が可能になったことで、授業を理論のみで終わらせることをやめ、手頃な価格の Excel を大学の PC 室で演習する形をとって理解の促進を図る十分な動機があり、出版に至ったものと思われる。

ただ、統計処理を Excel で学ぶことについて本稿では深入りしないが、Excel が持つ関数、特に統計分析に関わる関数及び分析ツールの処理内容が示されていないことから品質に関する疑義がネットではたびたび言及されており、学習教材としての利用についての妥当性を検討する必要はある。しかし、全体的な傾向として、Excel の実装はバージョンごとにそれぞれ異なるが、改善が進められ、判明している実装について計算の信頼性が問題となるケースは限られてきているようであり、研究等で用いられる商用統計パッケージソフトウェアとの用途の違いを説明した上で利用するならば、大きな問題はないと考えた。むしろ注意すべきは、理論に関心の薄い実践側の研究者が執筆した教科書のほうである。実際に、入手した複数の教科書のうち、途中の議論が省略されている部分の証明を試みると、明らかに誤っているものが存在し、講義ノートと思われる Web ページで同様の式を用いた解説も散見される。何らかの原典に含まれる誤りが検証されることなく孫引きされていることが懸念される。また、回帰分析の章で「決定係数 R^2 は相関係数 r の二乗である」という教科書の記述に対して、掲載されているデータでは確かに Excel の計算結果に再現性があったが、「よく言われるこの定義は実際には Excel の採用しない定義であり、回帰式の適合度とされる決定係数の定義は少なくとも 7 つ存在するという論文があるが、適合度の基準を何に置くかという点でなお議論の余地がある」ことを指摘している Web ページ¹⁰⁾が存在する。

4. 構想した授業

上記の考察を踏まえ、「データの活用」を意識させられる表計算ソフトウェア演習授業として、統計について学ぶことを全体の骨格とした、以下のような授業構成を構想した。全 15 回の構成の概要を、表 1 に示す。

それぞれの授業内容について、各時における実施内容と用意した教材について以下述べる。

表 1 構想した授業の概要

時	内 容
1	統計とは
2	数理的な問題解決とは
3	前回の課題解説
4	Excel の基本操作
5	データの抽出と並び替え
6	度数分布表とヒストグラム
7	代表値と散らばりの尺度
8	人口動態推計から未来を考える
9	二次元のデータと関係
10	散布図とクロス集計
11	共分散と相関係数
12	回帰分析
13	最終課題の構想
14	最終課題
15	最終課題

4.1. 第 1 回：統計とは

授業全体の導入として、社会においてデータを活用する具体的事例を示し、学習の意義を理解させる必要があると考えた。そこで、第 1 回では、文献¹¹⁾第 2 章の内容に基いた資料を提示し、統計学の位置づけについて、データの単純集計とデータ分析の違い、データ分析は犯罪容疑者を裁く法廷証拠となるか？ というアメリカでの実際のエピソードについて、統計的仮説検定の限界と、相関と因果の相違について話題提供を行い、議論を行うこととした。

4.2. 第 2 回：数理的な問題解決とは

データを活用するためには、どのようにデータを収集するか、問題を分析に適合したモデルとして表現するなど、数理的な考え方が必要となる。そこで第 2 回では、2005 年にアメリカ CBS で放映された数学者と FBI 捜査官兄弟を主人公としたミステリードラマシリーズ「Numb3rs」のパイロット版のエピソードの DVD を上映し、シナリオや演出に含まれる数学に関わる話題、確率と統計について語られる場面、データ収集手段、データを用いた問題解決のための数理的アプローチなどに注意して記録し、レポート課題としてまとめさせること

とする。上映するエピソードはアメリカで実際に捜査手法として FBI が用いた実話を下敷きにしたものであるが、この点にも言及し、フィクションではあるが現実社会に対応したものであることを意識させる。

4.3. 第3回：前回の解説

提出されたレポートの記述をとりまとめ、エピソードの各場面を振り返りながら、読み取りのできている事柄と、見落とされていた事柄について確認するとともに、核となる問題解決手段がどのように活用されたかについて、文献¹¹⁾の第1章で式を含めて解説されている部分を示す。

4.4. 第4回：Excelの基本操作

大学入学時の平均的スキルと考えられる、表の作成と装飾、数値と関数を用いた合計と平均の計算を含む演習を、5項目程度のシンプルな売上票の作成と、統計データから作成した表へのやや難しい式を含む計算の追記によって行う。データは公開されている政府統計に基づき、都道府県の面積、人口動態、所得を組み合わせるCVS形式として作成しておく。計算として、人口密度と年当たりの人口増加率を計算させることとする。

4.5. 第5回：データの抽出と並び替え

Excelの特徴的な機能のひとつである「オートフィルタ」を用いて、マウス操作で対話的に条件に合致する行のみの表示に制限する手順と、条件をセルに設定し、項目(列)を含めて選択した内容からなる新しい表を作成する手順を演習させる。また、項目を指定して並び替えを行う手順も演習させる。これは第7回で用いる四分位値を求める手順を確認する際に、データを昇順に並べる作業が含まれることから、この時点で実施しておく。

4.6. 第6回：度数分布表とヒストグラム

得られたデータを分析する前に、データの傾向を確認する手順を最初に行うことが一般的であることから、2010年県別人口の分布を対象に、度数分布表を階級数の設定、階級上限と下限の設定から行い、Excelの「分析ツール」アドインを設定して度数の自動集計を行わせる手順について演習する。度数から相対度数、累積度数、累積相対度数の計算を行い、全体をひとつの表とした後、グラフ作成によってヒストグラムを作成する。作成したグラフをどう解釈

するかについても考えさせる。

4.7. 第7回：代表値と散らばりの尺度

データの最小値、最大値、合計値、平均値は、取得したデータ全体から求められ、データの性質を表すという意味で「代表値」と呼ばれることを説明し、データの分布により、中央値と平均値に差が出ることを、解釈においては中央値がより妥当な値となることが多いことを指摘する。これらを2010年県別人口の値を使い、定義とその計算をExcelにより計算する。次に、データの分布に関する代表値として、分布を視覚的に把握しやすい箱ひげ図に用いられる四分位値を紹介し、平均値から分散・標準偏差という別の散らばりの尺度を求められることを述べる。また、学力試験における偏差値やIQ値が標準偏差を用いて定義されていることについても触れる。

4.8. 第8回：人口動態推計から未来を考える

データの集計から分析へと扱う内容が次回から切り替わるので、ここで再び統計を活用する意義について触れる。近年、社会保障と年金負担の問題について少子高齢化に関する議論とともに活発となっているので、国立社会保障・人口問題研究所が作成してWebで公開している、国勢調査による1920年からの人口動態時系列データと2050年までの推計値に基づく年齢別人口分布グラフを用い、年齢と、社会的役割を想像し、将来を見据えて日本社会は何に取り組んでいくべきかを話し合う。

4.9. 第9回：二次元のデータと関係

二次元のデータは、散布図を描くことで、平面的に散らばった点の集合として表現されることを確認し、これらの観測値が、一方の増加に対して他方はどうなるか、という「関係」をみることに役立つことを述べる。関係のうち、正または負の比例の関係を「相関」と呼び、強さで言い表すことを確認する。また、2つの観測値を用いればどんなものであれ散布図を描き、相関をみることができるが、荒唐無稽な2つであっても「相関がある」ように見えること、しかしそれは「因果」を表すとは限らず、混同しないことを、健康や学力などに関する言説か

らいくつか例を挙げ、妥当性を考えさせることで実感させる。演習として、県別年収と人口密度及び人口増加率について散布図を描き、相関の有無や相関を考える対象になりうるかを話し合う。

4.10. 第10回：共分散とクロス集計

相関の大小は二次元的な散らばりであることを確認し、一次元データにおける分散を拡張した共分散の概念について、図解とともに説明する。共分散では異なるデータについて同じ尺度で比較しにくいいため、「相関係数」が使われることを説明する。県別年収と人口密度、人口増加率を使い、相関係数を求める演習を行う。

性別その他、質的な値について散布図は不適当なので、表の行と列に集計した「クロス集計表」にすることを説明する。クロス集計表の作成には Excel の特徴的な機能である「ピボットテーブル」を用いる方法を説明し、県に対応する地方名の項目を作成し、人口密度を適当な値を基準に大・中・小の3つに置き換えたデータによりピボットテーブルの操作について演習する。

4.11. 第11回：回帰分析

散布図にデータの間中部分を通る直線を加えることで、関係が見やすくなることを、グラフ機能で回帰直線を加える項目にチェックを入れて、見せる。この直線がどのようにして求められるかを最小二乗法より順に説明する。データから計算した値が Excel のグラフ設定で表示される直線の式の値とほぼ等しいことを確認する。

4.12. 第12～15回：最終課題

政府統計データから、学習者それぞれが普段感じている問題意識に対し、何らかの分析が可能なものがあるかどうかを探索させ、仮説とそれを説明できそうなデータと分析方法を決めさせる。学校基本調査を例に挙げるが、学校基本調査ではほとんどの項目が「概要」の文書で時系列グラフを掲載しており、ダウンロードしたデータをグラフ化するだけ不足であることを伝え、思考を伴わない、単なる作業で終わることを避ける。レポートライティングの指導を兼ね、仮説に至る問題意識から、どのデータを使用して、どのような分析を行い、得られた結果を評価する構成でレポートを作成するよう指導する。

5. 結果と考察

2015年度の1年生前期に開講された「ソフトウェアI」の授業で、上記構想に基づく授業を行った。受講者数は13名であった。

第1回と第2回では資料や映像に基づき、データに対する関心を持つこと、数理的な考え方についての理解がある程度できることを期待した。これらの授業後のレポート課題の記述から、社会においてデータの活用が役立てられていることについて、理解できたことは確認できたが、題材が犯罪捜査や死刑を含む裁判であったことから、「数字でものごとを決めるのはよくない」という感情的な記述が目立った。

統計は確率に基づくものであり、完全な正解を与えるものではないが、そのことも、「決定的証拠を出さない統計は使うべきではない。人がきちんとした証拠を見つけ出すことが大事だ」のように、シナリオを十分に咀嚼できない段階での、道徳的な観点での見かたを示す記述が目立った。

そのため、確率についての言及が数多いエピソードにもかかわらず、レポートでそれを指摘できたものはほとんどなかった。よって、第3回の解説も、期待した回答がどこにあったかを順に確かめる形の一方的な授業になったことは改善を要する点である。

第4回から7回までのデータの集計操作演習は、提示資料に沿う比較的単純なものであったため、遅れがちな者を教え合いで助ける順調なものであった。第8回は昭和の歴史を学ぶ面、人口減と高齢者の厚い層が明白な推計グラフの解釈という取り組みやすい内容であったことから、統計を用いた予測の有用性については納得できるという発言が多く得られた。

第9回からは分析となり、Excelでは煩雑な操作が増えること、後期の「ソフトウェアII」で扱う予定の統計的仮説検定に向けて数学的な理解を着実にするため、数式の多い板書となったため、難しく感じられている様子が伺われた。特に、ここでたびたび現れる「偏差」という語を最後までのみこめない者が半数程度見られた。「データにはばらつきがある」ことを、

ヒストグラムを見る段階からうまく積み上げていく授業改善が必要と考えられる。

最終課題についても、出題の意図に沿った取り組みに至ったのは13名のうち3名であった。10名は目的を明確化できないまま、データを散布図または棒グラフ化し、計算を行わずに終わってしまっていた。教科書的な基礎演習から自由課題に取り組みさせることはギャップが大きかったことが理由として考えられる。入門段階の最終課題としては、データセットとゴールの組をいくつか用意して、演習内容をどの程度活用してゴールを達成したかを説明させ、合理的な説明が可能かどうかで理解度を判断するほうがより適切なのであろう。

6. まとめ

データの活用という視点から、統計を導入した授業を構想した。しかし、今回を行った結果は、いくつかの活用事例を紹介することはできたが、データを取り扱うことの有効性を十分に実感させられることには至らなかった。今後、理由が比較的明らかな部分の改善を検証するとともに、受動的な操作演習をよりアクティブラーニングに近づける工夫について検討したい。

参考文献

- 1) 加納寛子, 菱田隆彰, 長谷川元洋, ほか: 文部科学省検定教科書高等学校「情報」の用語分析, 日本科学教育学会年会論文集, 第37号, pp.152-155 (2013)
- 2) 文部科学省: 中学校学習指導要領, 2008 (2010一部改訂)
- 3) Burton Grad: The Creation and the Demise of VisiCalc, IEEE Annals of the History of Computing, Vol.29, No.3, pp.20-31 (2007)
- 4) The first Ad of VisiCalc in Byte Magazine, September Issue, pp.51, 1979.
<http://www.americanradiohistory.com/Archive-Byte/70s/Byte-1979-09.pdf>
(2015年9月15日閲覧)
- 5) 岡本敏雄, 高橋三吉, 西野和典: 情報科教育法 第2版, 丸善出版, 2015
- 6) J Coate: Using VisiCalc and DYNAmo to Make Models and Solve Problems in High School Math, Computers in Schools, Vol.3, No.1, pp.75-81

(1986)

- 7) 首相官邸: ミレニアムプロジェクト (新しい千年紀プロジェクト) について, 1999
<http://www.kantei.go.jp/jp/mille/index.html> (2015年9月15日閲覧)
- 8) 立野貴之, 加藤尚吾, 加藤由樹: 高等学校「情報」の履修内容と大学入学後の情報に関する意識や成績との関係の分析, 日本科学教育学会年会論文集, 第22号, pp.222-223 (2006)
- 9) 石崎龍二: 福岡県立大学人間社会学部新入生の入学時のコンピュータスキル調査 (2010年), 福岡県立大学人間社会学部紀要, Vol.19, No.2, pp.99-109 (2011)
- 10) 井口豊: 決定係数 R^2 の違い: Excel, OpenOffice, LibreOffice および統計解析ソフトRを用いて
<http://www.ab.auone-net.jp/~biology/regline/R-Squared.htm>
(2015年9月15日閲覧)
- 11) キース・デブリン, ゲーリー・ローデン (山形浩生, 盛岡桜 共訳): 数学で犯罪を解決する, ダイヤモンド社, 2008

(2015年10月4日受理)