

特集にあたって

森 裕一（岡山理科大学）

「百聞は一見に如かず」、英語で“A picture is worth a thousand words.”（1枚の絵は1,000の言葉に値する）…言葉で説明されるより、実物や写真、イラストを見る方が素早く正確に理解できる…情報の「可視化」を表す最適なことわざである。「可視化」以外にも、視覚化、Data/Information Visualization, Visual Analytics, Graphic Communication, 見える化といった表現があり、細かい違いはあっても、いずれも物理的な特徴のみならず、目に見えない抽象的な情報の理解に大きな役割を果たすことを示すものである。

この「可視化」は、データ分析の世界にはなくてはならないものである。探索的データ分析の提唱者である Tukey も“The greatest value of a picture is when it forces us to notice what we never expected to see.”(J. W. Tukey, *Exploratory Data Analysis*. Addison-Wesley, 1977) と、その大切さを強調している。これらのことに異議を唱える人はいないと思われるが、基本的な表やグラフはあまりにも身近なため、これだけしか見ていないと、その効用や効果的な利用について意識することは少ないかもしれない。しかし、データが複雑で大規模になっている今日、その様相を把握する際、可視化の恩恵を受けていない人は誰もいないであろう。さらに、様相の把握だけでなく、新しい可視化手法が開発され、これまでわからなかった新たな知見が得られるなど、可視化の進展と分析手法の進化はともにあるといっても過言ではない。そして、これらの進展・進化は、計算機の能力の向上に支えられている。計算速度だけでなく、対話的なインタフェースが容易に手に入るようになったことが大きい。ハイライティングやブラッシングに加え、パラメータを自由に変更して描画し直したり、地理情報とリンクさせて空間的な特徴を提示したり、さらにはそれを時間の変化とともに動かすといった可視化手法がわれわれの考察力を飛躍的に向上させた。

本特集は、そのような観点から、五つの話題を提供してもらうことにした。

前半の2編は、可視化の工夫についてである。

藤原美佳氏（岡山理科大学）は、拡張型星座グラフ

と拡張型漢字グラフを紹介する。前者では、既存の星座グラフの星の位置（データ点）を新たな計算方法で求める工夫が施され、後者では、既存の漢字グラフとレーダーチャートを組み合わせた改良が行われる。これにより、数十年前のグラフがサンプルの傾向やクラスタを把握しやすいグラフへと拡張される。

中野純司氏（中央大学）は、統計グラフィックスの作成手順を統計解析システム R を例に示し、大規模データの集約的シンボリックデータを可視化する新しいプロット chestnut plot を例にその開発について紹介する。集約的シンボリックデータの可視化をどう実現するか、詳細な考察のために動的・対話的機能をどう実装するかなどが示される。

後半の3編は可視化を活用した分析事例である。当該分野の問題解決にどのような可視化を用意したか、可視化でどのように問題が解決されるかが示される。

久保田貴文氏（多摩大学）は、ある都市のコロナ禍前後の自殺率の地理的変動を可視化・分析するためのシステムを R の Shiny パッケージを用いて開発し、地域ごとの自殺リスクを明らかにする事例を紹介する。データの平滑化や偶然変動の抑制を取り入れ、生成 AI も活用した情報提供の仕組みとその有用性が示される。

藤野友和氏（福岡女子大学）は、大学の学務データの分析、いわゆる IR(Institutional Research) における可視化の活用事例を紹介する。この実現のために、IR 体制を整備したうえで、科目の評価指標を開発し、それを基にカリキュラムの可視化が行われる。教員へのフィードバックにも可視化が有効であることが示される。

最後の地道正行氏（関西学院大学）は、財務諸表に関する大規模データを前処理し、ラングリングしたものを可視化することにより、得られた知見から統計モデリングを行うという、可視化に基づいた探索的データ分析の事例を紹介する。結果から動的に文書生成することによる再現性の検討も行われる。

以上より、可視化ではどのような工夫を行えばよいか、可視化を活用した分析はどうあるべきでどのような効果をもたらすかを理解いただければ幸いです。みなさんの「一見」の一助になりますように…