Abstract

The advent of data science has led to statistics education researchers re-thinking and expanding their ideas about what computational tools to use for teaching statistical modelling. Consequently, the use of computer programming languages such as R have been promoted in textbooks, tasks, and other learning activities to support data science at the high school level. The shift in pedagogical knowledge required for teachers, however, is not restricted to the introduction of code-driven tools and computational thinking to the curriculum. The statistical methods taught need modernising to include digital sources of data, such as APIs, and algorithmic approaches, for example, predictive and classification modelling. Therefore, teaching data science at the high school level using code-driven tools requires consideration of how learning activities can be designed to support the co-development of statistical and computational thinking.

Minimal research exists about the design of tasks that support the introduction of high school students and teachers to the use of code-driven tools for statistical modelling. Hence, the dual objectives of the study were to: (1) explore the observable thinking practices that emerge when learners completed statistical modelling tasks that introduced code-driven tools; and (2) develop a task design framework to introduce code-driven tools and support the integration of statistical and computational thinking. Using a design-based research approach, four structured tasks were developed for teaching statistical modelling at the same time as introducing the programming language R. These four tasks were implemented with high school statistics teachers across four full-day face-to-face professional development workshops.

The study resulted in the development of the Introducing Code-driven Tools (ICT) task design framework, which was produced by identifying key design elements for one task, and refining and re-evaluating design principles and processes through consecutive retrospective analyses of the other three tasks. The findings from this exploratory study indicate that the tasks constructed supported teachers' introduction to code-driven tools and encouraged an integration of statistical and computational thinking. The ICT task design framework contributes to statistics and data science education research by building on previous work in effective pedagogy, and by providing practical guidelines for the introduction of code-driven tools to facilitate the integration of statistical and computational thinking to learn from modern data. Another contribution is the production of two hypothesised frameworks, which provide guidelines for assessing and clarify the integration of statistical and computational thinking.