# Contents



# Editorial

In 1973, Francis Anscombe (Anscombe, 1973) published a fascinating simulated dataset containing four pairs of variables. On calculating the sample correlation coefficient or fitting a least squares line of best fit the results appear similar, however, anyone who had first conducted a visual exploration of the data would be aware that the datasets had very different properties indeed. The simulation process was later generalised by Chatterjee and Firat (2007). Our first article this issue addresses a related issue. Despite the existence of techniques such as smoothing functions, dichotomising continuous variables remains a much used technique. We're delighted to present an article featuring five datasets generated by Oliver Kuss. These have the property of giving identical two-way tables of counts when suitably dichotomised, but yet clearly have very different properties. Hopefully these data will be widely used to illustrate caveats around dichotomising variables. Ordinary linear regression plays a big part in this issue, with an article by Bittner looking at limitations of predictions and some interesting work by Quinino, Reis Bessegato using distributional characteristics of the coefficient of determination – results that don't feature in many introductory text books yet which may provide additional insights into linear regression. We are also fortunate to have an article by Falk and Lann on weighted means which provides a very thorough introduction to this key concept – a concept that often requires revisiting at higher levels of education despite learners thinking they already know everything they need to know about the topic.

The potential for sport to motivate classroom activities is exemplified with articles by Oke Gerke on collecting football cards (which could actually include non-football cards) as well as an article by Shiwalker and Deshpande on the cricket world cup (for readers in the US cricket is a bat and ball game somewhat akin to baseball).

Finally, given growing interest in online learning we are pleased to host an article specifically concerning learning statistics online by Susan Gazioglu and as usual present a stimulating statistical diversions column from Peter Petocz and Eric Sowey.


Anscombe, F.J. (1973). Graphs in Statistical Analysis. *American Statistician*, **27**(1), 17–21.

Chatterjee, Sangit and Firat, Aykut. (2007). Generating Data with Identical Statistics but Dissimilar Graphics: A Follow up to the Anscombe Dataset. *American Statistician*, **61**(3): 248–54.