Visualisation driven analytics—novel 'point and click' tools for general and population statistical analysis

Daniel Barnett[1], **Tom Elliott**[1], Andrew Sporle[1], and Chris Wild[1]

[1] The University of Auckland

The utility of data resources is often limited by technical barriers to accessing and analysing the available data. In response to this need, we have developed two software tools that reduce the cost, security, and skill restrictions affecting the use of data. Both tools are developed using the R software platform.

The first software tool is 'iNZight'—an R-based freeware package that provides a mouse-driven user interface to a wide range of R packages for analysis and graphics. It is standalone software that enables data to be easily explored visually as well as offering a wide range of statistical analyses. The interface has pull down menus and contextual controls that recognise data types, empowering novice users to explore their data while the software guides them. 'iNZight', along with its Shiny-based online version 'iNZight Lite', are routinely used in the maths syllabus in New Zealand. The package itself not only produces tabular, graphical, and map output, but can also work with data from complex survey designs. For interested users, 'iNZight' generates an R code script allowing users to replicate their results directly from R, both for learning and reproducibility.

The second tool is a population statistics package that can work in an online or standalone environment to create and compare age-standardised rates (with confidence intervals) for population outcomes by user-defined selections of population and reference (standard) population. It also can compare and graph population outcomes over time, including absolute and relative rate differences with statistically robust confidence intervals. All results are exportable into other formats. This software works with both count data and microdata (unit record file). The data behind the tools is not visible and the software can sit on the user's computer, meaning that control of the data can be maintained.

These tools provide easily accessible and simple means to produce robust statistical analysis and visualisations with minimal training but still provide robust analytic output.