

So, why is research essential in today's climate? If you can't measure it, you can't manage it. And if you can't manage it, you can't fix it.

### What statistical analysis tools and database software have you previously used? What are your favourites and why?

The pre-eminent statistical analysis tool we use is the ASTEROID software user interface. This is an easy-to-use tool with minimal experience and enables users, including ourselves and our clients, to interrogate the Roy Morgan Single Source consumer database, arguably the best consumer database of its kind in the world. Another favourite we use at Roy Morgan is Tableau which is an easy-to-use statistical analysis tool that functions in an attractive user interface.

We predominantly use Tableau as a data analysis and visualisation tool in our dayto-day work. The data fed into Tableau is prepared using Google Big Query. Although not strictly a statistical analysis tool, another favourite we use is Python programming language. Python's

simplicity, versatility, rich ecosystem and active community make it a great tool for beginners and experts alike.

# What are some best practices for data cleaning? What are the steps you take?

We put in place a number of checks to identify poorquality responses, seeking to eliminate what is known as speeding and pattern (also known as straight-lining). Parameters for such checks are specific to each survey, but for example, we seek to identify responses that have been completed too quickly (which indicates a lack of consideration for answers, and the likelihood that the respondent is completing answers at random), or where the respondent is using the same response continually. We will, for example, identify repeated and excessive use of the 'don't know' and 'prefer not to say' answer options, and remove them. We also oversample to allow

headroom for the removal of problematic records while maintaining the required overall sample size – this is undertaken within a rigorous quality assurance process.

#### Describe your most complex data project from start to finish. What were the most difficult challenges, and how did you handle them?

Our most complex recent data projects are probably those that link survey data and machine-based data, device movement data and then new survey data. However, the most interesting project throughout the pandemic is "Taking the Pulse of the Nation" – known as the "TTPN". The dynamic nature of the survey means the questions are constantly changing on

#### In your opinion, what do you think is the most important message businesses need to know about how data can help their organisation?

We live in the data era – it is central to everything that businesses encounter in the modern hyper-connected and digital world. The world of data grows at an exponential rate that is simply impossible for any one person to fully comprehend – and especially the myriad ways that different sets of data interact and inter-react and shape the wonderful and exciting world that we live in.

Roy Morgan works as a living organism to bring together complex data sets on all the significant industries in Australia.

With this wide and rich understanding of data, we have developed valuable

## "Investing in data is an investment in the future of a business - it shapes its future."

a monthly basis to keep up-to-date with the changing focus of the survey, which leans heavily into the current socio-economic issues.

This project represented a learning curve for team members on how complex data projects can frequently present novel and unforeseen challenges, and require agile responses and an intimate connection with the data being collected, analysed and eventually passed through to key stakeholders. This impacts everyone involved in the project, constantly refining how the nuts and bolts of collecting data in complex segmentation arrangements can translate into an illuminating and richly informative output in the eventual hands of national stakeholders, including media companies.

segmentation tools to distil the maximum value for businesses looking to solve problems and find a competitive and comparative edge on their competitors. For instance, Roy Morgan's Helix Personas incorporates values, beliefs and attitudes, which are the best predictors of consumer behaviour. Businesses, brand owners and media agencies can then optimise marketing spend to effectively message those potential customers and, at the same time, find, and grow new markets. To truly grasp the marketplace in which a business operates, understanding the factors that drive consumer attitudes, behaviour, and spending decisions is vital - and data is central to that understanding.

Investing in data is an investment in the future of a business – it shapes its future.