



NEWSLETTER

SEPTEMBER 2006

Despite statisticians' reputation for "making up" numbers, they have an important role in auditing data, making sure that the numbers are right. This skill in understanding and evaluating information means statisticians fill an important position for many organisations.

*Dr John Henstridge
Managing Director*

The Value of Information

Today's organisations are data centric. Information capture is easier and storage limits have effectively disappeared, with databases like Customer Relationship Management (CRM) Systems at the heart of the organisation becoming critical assets. But unlike other supposedly more real assets, this one does not appear on the balance sheet and its value is not realised.

One reason is that the value of information depends upon its relevance and quality. Unfortunately when many organisations want to use their data for strategic judgments, the quality is not good enough or the right data has not been collected. This may be discovered when a major upgrade or redesign is carried out – historic data is simply incompatible with the new system. Sometimes the data is needed urgently and vital parts are found to be missing.

Customer name and address data found in a CRM system degrades over time as details change and updates are not made. Names and addresses are often recorded in a free text form with little or no structural checking. This works well *most* of the time – just think of letters you have received with incorrect addresses – since the human postman can often interpret technically wrong details and use other information to overcome problems. Computerised systems are usually less forgiving.

Data Analysis Australia has recognised that a statistical approach is best in assessing just how good such data is. It is also necessary to have a strategic approach to the IT systems and processes to ensure that they assist rather than hinder the maintenance of quality data. For this purpose we have formed an alliance with **Iain Massey and Associates** (www.iainmassey.com), combining analysis with strategic IT planning to provide

an information audit service. Recently collaborative work was done reviewing the "Names Database" central to the Insurance Commission of WA information systems. The project required both a review of the data itself and consideration of the data capture process and the way the data was used.

Data Analysis Australia used a similar approach in working with telecommunications consultants, **Gibson Quai-AAS** (www.gqaas.com.au), in the audit of the Integrated Public Number Database (IPND) for the Australian Communications and Media Authority. Here a major issue was the sheer size of the problem – comparing over 40 million records in the IPND against a reference set of over 30 million. A cluster of high speed computers with clever algorithms was used to do this in days rather than months.

We also use these approaches

internally. For the Perth and Regions Travel Survey, large volumes of imperfect address information are collected in self-completion surveys which have to be converted to reliable geocodes. We developed smart probabilistic algorithms that compare an address being captured with every possible Perth address – almost one million – and determine the best fit using a Bayesian probability criterion that allows for incompleteness, spelling and other errors. A quality index flags if the data is ambiguous, for immediate follow up. This happens in real time as data entry takes place.

As organisations rely more upon large computerised databases, it is essential that they focus on exactly what they are storing in them. For further information on this and particularly on our information audit service, contact Jodie Thompson at Data Analysis Australia.



Statistics plays a major role in sport, typically in measuring performance or determining the likely outcome of a game. In cricket, statistical theory is used to balance one-day games affected by rain delays. This issue's AI article explores the Duckworth-Lewis method used to adjust the required number of runs in rain affected one-day matches.

R	M	B	4s	6s	SR
36	65	39	6	0	92.30
32	55	44	3	0	72.72
77					
6					
33					
35					
22					
0					
6					
14					
261					



See: www.daa.com.au/analyticalideas/cricket.html

Staff Profile



Our resident Metallica fan **Greg Wogan-Browne** joined us in July 2000 as an Information Technology Programmer on a casual basis while completing a Bachelor of Computer and Mathematical Sciences at the University of Western Australia. He became a permanent staff member in early 2003 and is now our Network Co-ordinator.

Greg is responsible for keeping our sophisticated network and IT systems running, one that is designed especially for the specific needs of a statistical and mathematical company. Over the years, he has become more directly involved in consulting on projects like the audit of the IPND for the Australian Communications Media Authority and the design and implementation of several internet surveys. He also contributed to the sophisticated data entry system used by our survey team.

Classic Quote

A knowledge of statistics is like a knowledge of foreign languages or of algebra; it may prove of use at any time under any circumstances.

- **A.L.Bowley (English statistician and economist (1869-1957))**

97 Broadway
Nedlands WA 6009
Telephone: +61 8 9386 3204
Fax: +61 8 9386 3202
Email: daa@daa.com.au
Web: www.daa.com.au

Company News

A strong contingent of Data Analysis Australia consultants attended the 2006 Australian Statistical Conference in Auckland, New Zealand in July. The Conference is a biennial meeting for Australian and international statisticians, providing opportunities to discuss the latest techniques and catch up with colleagues. Our consultants presented conference papers on topics as diverse as travel, gold standards, internet surveys, counting Indigenous preschool enrolments and locating justice facilities – for more details visit our website: www.daa.com.au/presentations/

We congratulate **Joanna Potts** and **Sharon Lau**, recipients of the Data Analysis Australia's Young Statisticians Travel Award. The award provided travel sponsorship to statisticians young in their career to attend the Conference 2006.

John Henstridge was an invited speaker at the Spatial Sciences Institute WA Regional Conference held in Perth in June 2006, speaking about interesting spatial statistics. John also spoke to the Australian Property Institute in August 2006, when he was invited to speak on Mandurah population trends.

In September, **Miriam Maclean** is attending the Australian Psychological Society Conference being held in Auckland, New Zealand. Miriam was invited to attend and present a poster as part of the Australian Psychological Society Prize for being the top graduating psychology student in her year.

Pamela Gant was awarded a Diploma of Spatial Information Services in June. During her study, Pam has been building her GIS role at Data Analysis Australia and handling many of the technical processes involved in maintaining spatial data, geocoding and preparing maps.

Scott Brown, who joined us in March as a Consultant Mathematician, has had his PhD thesis accepted. It won't be long before we have another Doctor in the house!

We bid farewell and good luck to **April Rutkay**, who has moved to Canberra. April joined us as a graduate and provided a valuable contribution to many projects over the last three years.

Congratulation to **Cheryl Praeger**, who has been elected to the Executive Committee of the International Mathematical Union, the international umbrella organisation for Mathematics, both research and education.

