

# ***SurfStat.australia*: a Statistics Textbook for the Web**

[Keith B G Dear](#)

*National Centre for Epidemiology and Population Health, Australian National University, Australia*

## **Abstract**

*SurfStat* is a web site comprising the complete text of an introductory service course in Statistics, but with several added interactive features. It demonstrates the concept that existing text-based teaching material (course notes) can be renovated and enlivened for web-based delivery without massive commitment of programming effort. The *SurfStat* project thus contrasts with projects elsewhere, which showcase very polished material but after years of development are far from delivering a complete course. The web technology is limited to what can be handled by an unadorned browser, including use of frames, JavaScript and CGI scripts for indexing but without video, sound, *Shockwave* or other multimedia plug-ins. *SurfStat* is freely accessible at <http://www.anu.edu.au/nceph/surfstat/>.

## **Content**

The text of *SurfStat* is presently structured into five chapters, on data and statistics, probability, design, inference and control charts. Many of the examples are drawn from a business background to suit the original students who were mainly enrolled in degrees in Business or Commerce. Example analyses include *Minitab* output, but there is no necessity to teach using any particular package.



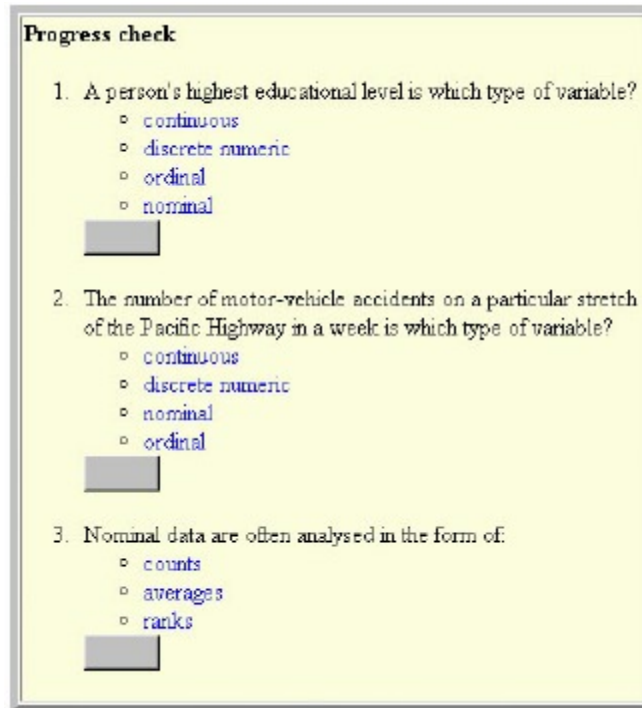
**Figure 1. The *SurfStat* homepage**

The chapters are subdivided into a total of about 60 HTML pages, each short enough to allow rapid loading and comfortable on-screen reading without too much scrolling. A navigation tool with up/down/left/right arrows makes it easy to move between the pages, though some visitors to the site have mistaken this tool for mere decoration! The simple design of *SurfStat* ensures that the chapters can be reorganised very easily and new material added. As the project progresses, we find we are repeatedly subdividing the material, tending towards many small pages each covering only a few concepts.

The chief interactive embellishments are the progress checks, the glossary and the statistical tables. We are also experimenting with interactive tools for exploration of statistical concepts, written in Java. Paul Velleman's remarkable product *ActivStats*<sup>1</sup> shows what is possible in this direction.

## **Progress checks**

Each page (in principle - this is work in progress) ends with a "progress check" consisting of two to four self-marking multiple choice questions, that provide readers with immediate reinforcement of the material on that page. The progress checks are implemented in JavaScript.



**Figure 2. Interactive progress checks on each page**

Other text elements include a detailed contents page linking to all the main text pages, a page of links to other statistics teaching web sites and a page listing some conventional introductory textbooks.

## **The glossary**

It was an early intention, still not fully realised, to provide *SurfStat* with a glossary intimately linked with the text in both directions. The text includes a number of "hotwords" which if clicked cause their definition to appear in a separate glossary window, while terms defined in the glossary are linked back to relevant pages of the text. The existence of links back to the full text permits the glossary definitions themselves to be very terse, to encourage students to browse. The difficulty in fully achieving a rich network of definitions and explanations is not the technology, which is very simple, but rather the effort of manually inserting hundreds of links. Many of the definitions are linked to the glossary provided with the *STEPS*<sup>2</sup> project in the UK.

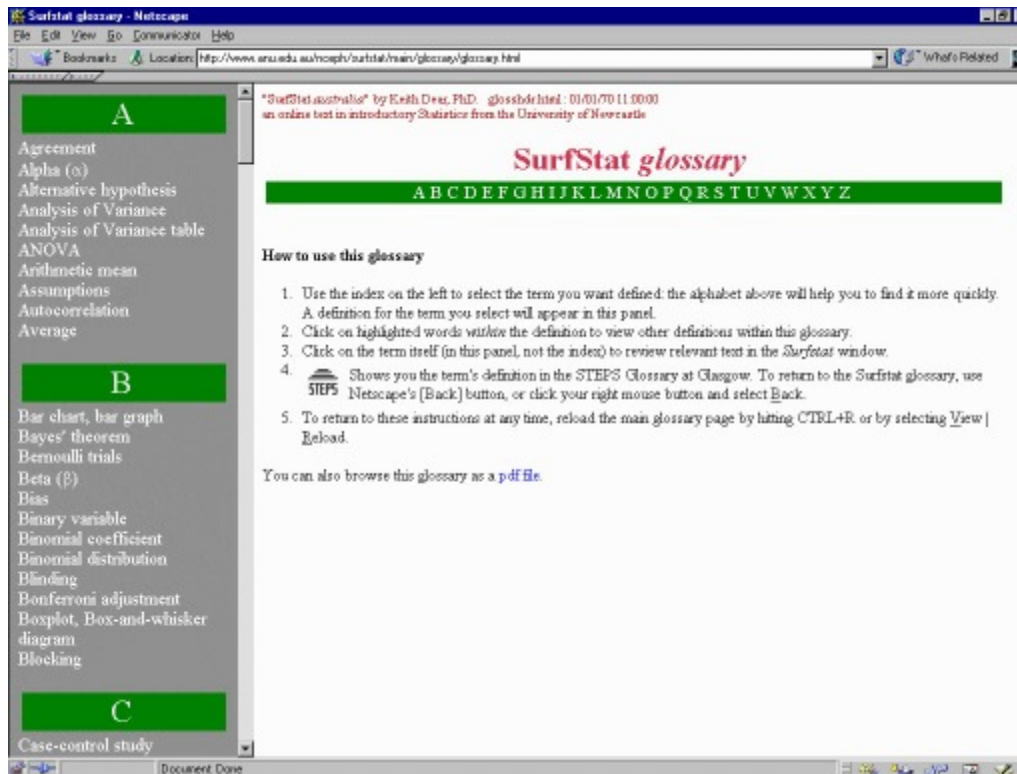


Figure 3. The *SurfStat* glossary is linked to the text

## Statistical tables

Most introductory Statistics texts include tables. In a computer based text it was natural to replace these with distribution calculators. We have so far provided normal, t, chi-squared and binomial calculators. We no longer waste time in our courses teaching students to interpolate values from printed tables. Since we cannot provide computers in examinations for large classes, we compromise by attaching to the examination paper lists of quantiles that include all values that candidates should need to answer the questions, plus a few more.

We decided not to provide direct links from *SurfStat* into any statistics package. To do so adds little convenience if one is already using a windowing environment, and assumes too much about the configuration of the user's computer.

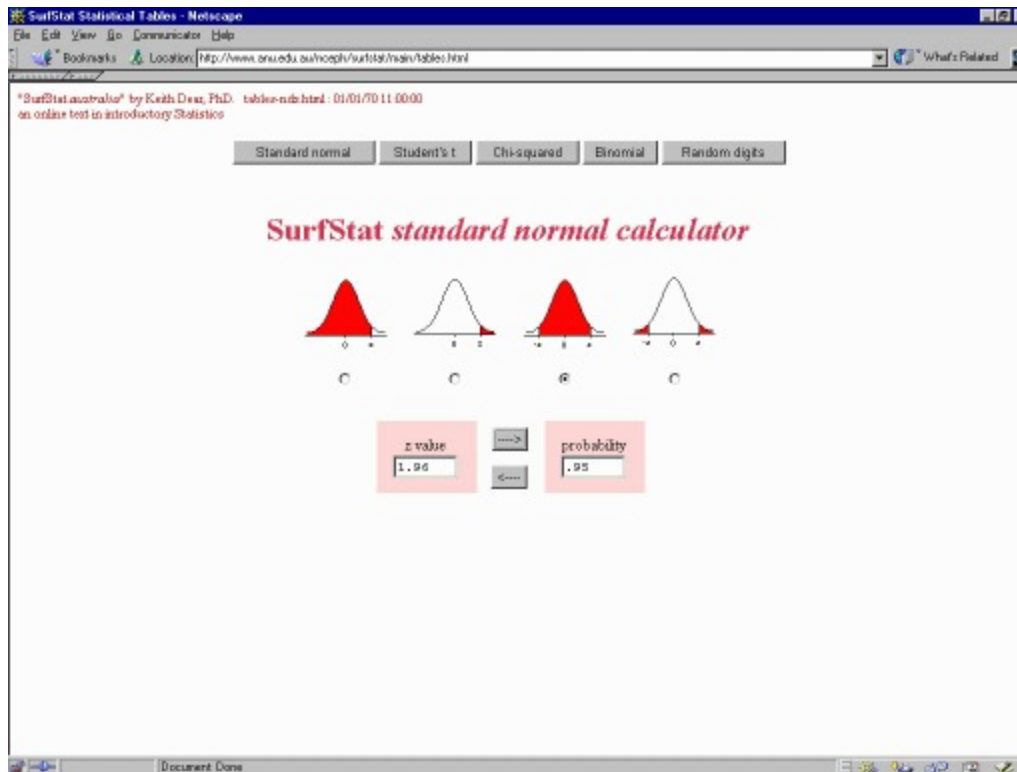


Figure 4. *SurfStat* includes statistical tables in JavaScript

## Use in undergraduate Statistics classes

The model we have used for our undergraduate classes is to have students access *SurfStat* via separate course-specific web-based study guides, which give learning objectives for each fortnightly module and provide a reading list in the form of HTML links to *SurfStat* pages. In this way the web resource itself can expand beyond the requirements for any one course, and can become something between a book and a bookshelf. Our present direction is to extend *SurfStat* from its origins in business statistics towards epidemiology and biostatistics. *SurfStat* has been licensed for use in university courses in several countries, and is also much accessed by individuals searching the web for help with statistical problems.

## Computer literacy training

If students are to learn at the computer, they must be able to use one fluently enough to attend to the message without being distracted by the medium. "Point and click" must be as natural as turning a page. *SurfStat* was first provided to first year students, and we are still seeing surprisingly few who have this level of experience and confidence with computers. As well as lacking skills, there is also a problem of insecurity - many students feel that what appears on the screen is not theirs until it is printed and in their hand, even though they know rationally that what they have just read will still be there tomorrow. In our course evaluations in 1998, 70% had a computer at home, but only 30% had Internet access, so most students accessed *SurfStat* using the University's general-access PC laboratories or the campus libraries.

With the assistance of the University Library staff, we dedicated the first two weeks of a 14-week semester to familiarisation with *MS-Windows*, *Netscape*, Internet resources including *SurfStat* and email. Ideally the training would take place separately from academic teaching, but the issue of computer literacy is one that universities generally have yet to come to grips with, and meanwhile the training is arranged ad hoc and usually within class time. I believe that if we require students to put significant effort into learning these skills within a regular course, then their achievement should be reflected in the course assessment. Our experience suggests that without it, many students fail to take the computer training seriously.

## Evaluation studies

We first made substantial use of *SurfStat* in delivering STAT101 at The University of Newcastle in March (June 1997). Evaluation took two forms - focus group discussions and a formal questionnaire. The focus group included students and several relevant experts: a librarian, an instructional designer and a multimedia programmer. Feedback was positive on the technology of *SurfStat* - students liked the glossary, the statistical "tables" and the few Java applets - and suggested that the place where an accessible teacher was most needed was not in delivering the theoretical material through conventional formal lectures, but in the practical sessions using the statistics package.

Some practical problems also surfaced: for example while students found it easy to send email queries while reading *SurfStat*, it was not so convenient for them to collect my replies from their accounts on the University mailserver. Many other comments were to do with practical problems such as poor access, whether on-campus or by dialling in from home, and problems with print queues. The IT infrastructure provided by one's university is crucial to successfully implementing computer based teaching.

Although we did not set any particular text (other than *SurfStat*) we did encourage students to browse the library shelves and read widely about topics that gave them problems. However many students reported having used no other text. One student wrote "I think it's a good idea using the computer, as it forces people (like me) to confront their fears of computers/technology".

### Web authoring techniques

Mathematical expressions and equations are displayed as GIF images, created using the Microsoft equation editor in *Word* and captured off the PC screen using the *LView* graphics editor. Very little technical content can be displayed directly in HTML. A  $\sigma^2$  can be achieved using `<FONT face="symbol">s</FONT><SUP>2</SUP>`, but even so simple an expression as  $\bar{x}$  requires an in-line image. Line diagrams were created as bitmaps using *Microsoft Paint* and converted to GIFs using *LView*. We did not use an HTML editor, preferring to edit the pages directly in a text editor to retain complete control. *SurfStat* is hosted by a Unix web server with Perl support for CGI

scripts, and we used Perl also for housekeeping maintenance of the many HTML source files.

An early decision was to use frames. Although many web teaching experts recommend against their use, we found an index frame too useful a feature to pass up. The main *SurfStat* window, the tables window and the glossary window all use an index frame to ensure that important resources are always available with a single click, whatever may be displayed in the principal frame.

*SurfStat* uses JavaScript to mark the progress check questions, to drive the navigation tool and to display glossary definitions. The statistical tables, or distribution calculators, are programmed entirely in JavaScript. Unlike the full Java language, JavaScript is easy to learn, requires no compilation and no separate files to manage, and adds almost nothing to the time that pages take to load from the server.

## Future development of SurfStat

The benefits of moving towards computer based learning are probably more educational than directly financial. There seems little potential for saving money by using computers to reduce staff input to teaching. A resource like *SurfStat* can only partially replace direct contact with a teacher, and at least as much time is spent developing and managing the resource as is gained from reduced contact hours. The benefits I hope to see *SurfStat* deliver lie more in attracting students to courses that use it, because of the flexibility it offers them in time and place of study and because of the enhanced quality of their learning experience. In the modern competitive environment of tertiary teaching, attracting students is itself a financial imperative.

## Acknowledgments

The development of *SurfStat* was supported by grants from The University of Newcastle in Australia, and in 1996 by the Australian Government's Committee for the Advancement of University Teaching (CAUT).

## References

1. <http://www.activstats.com/ActivStats/> Paul Velleman's CD-based resource *ActivStats*, from Data Description, Inc.
2. <http://www.stats.gla.ac.uk/steps/> The STEPS consortium: "STatistical Education through Problem Solving".
3. <http://davidmlane.com/hyperstat/> David Lane's HyperStat homepage lists many statistics teaching sites.

Keith B G Dear  
National Centre for Epidemiology and Population Health  
Australian National University  
Canberra, ACT 0200  
Australia  
[keith.dear@anu.edu.au](mailto:keith.dear@anu.edu.au)