

# General Brochure



Copyright © Red Centre Software, 2005-2007. All rights reserved.

## The Problem

Despite much development energy and financial investment by many companies across a wide range of fields, the majority of desktop end-user software systems for the processing and analysis of numeric or coded data continue to have many inadequacies. Whereas it is easy to make any tool look good in a controlled demonstration using clean and elementary data, coping with the relentless pressures of the real world is entirely a different matter.

## The Ruby Solution

*Ruby* is an advanced full-featured graphical cross tabulator with time series extensions.

Ruby's management of unlimited complexity and dynamic data is, as far as possible, seamless and simple while never compromising power and flexibility. Ruby addresses the needs of the most demanding professionals by implementing a wide variety of processing and analytical features within an intuitive GUI and infinitely scalable architecture.

## Ruby in Action

With Ruby, you can

- Process raw data into meaningful variables
- Tabulate and chart complex data
- Track changes over time
- Work interactively to pursue hypotheses
- Assemble detailed reports comprising hundreds, even thousands, of interdependent tables and charts, with data drawn from multiple sources
- Seamlessly and transparently reprocess complex report sets under different segment filters or weighting schema.
- Automate common or complex processes using scripts
- Export reports to MS Office, including auto-update of waves
- Work on the desktop, or via internet, intranet, LAN or WAN

Ruby can be used in many different fields. Some examples are

- Surveys
- Science
- Finance
- Textual Analysis
- Business Intelligence

- Athletic Performance

## Surveys

- Cross tabulate anything by anything
- Percentage anything on anything
- Filter anything by anything
- Weight anything on anything
- Reorganise anything, anytime, in any way
- Interact anywhere
- Automate everywhere

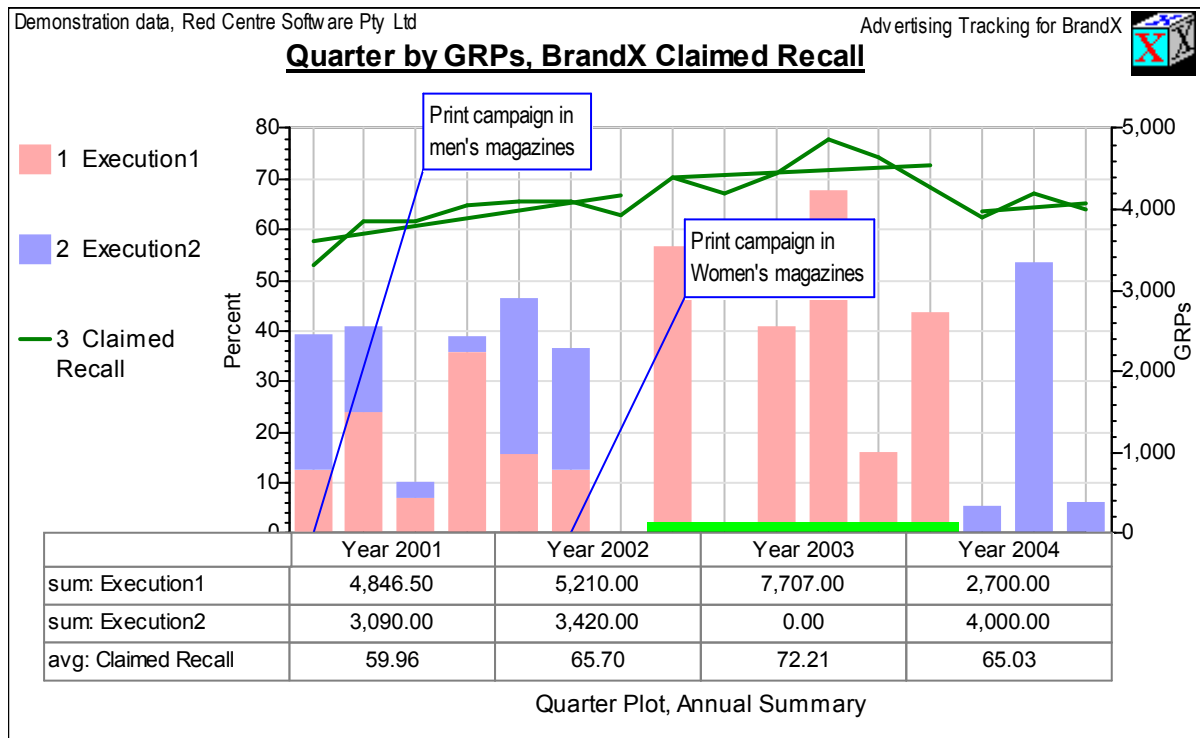
Ruby provides total job management from the point of clean source data through to final tables and charts. This includes all processing of case data into and out of new variables, and all operations required to deliver anything from low level reports to top-line analytics, either interactively or automatically.

All the thorny day-to-day issues of continuous tracking and longitudinal studies (such as miscodes, new codes or variables, changed labels and descriptions, etc) are neatly accommodated by Ruby's many features and sophisticated proprietary architecture.

**Top: Gender**  
**Side: Buy Regularly**

| Frequencies<br>Column Percents<br>Row Percents |        | Gender |       |        |        | North  |       |        | South  |       |        |
|--|--------|--------|-------|--------|--------|--------|-------|--------|--------|-------|--------|
|  |        | Cases  | Male  | Female | Total  | Gender |       |        | Gender |       |        |
|  |        |        |       |        |        | Cases  | Male  | Female | Cases  | Male  | Female |
| Buy Regularly                                  | Cases  | 16,743 | 4,985 | 5,015  | 10,000 | 8,454  | 2,542 | 2,524  | 8,289  | 2,443 | 2,491  |
|  |        |        | 30%   | 30%    | 60%    |        | 30%   | 30%    |        | 29%   | 30%    |
|  | BrandX | 8,886  | 4,441 | 4,445  | 8,886  | 4,503  | 2,263 | 2,240  | 4,383  | 2,178 | 2,205  |
|  |        | 53%    | 89%   | 89%    | 89%    | 53%    | 89%   | 89%    | 53%    | 89%   | 89%    |
|  |        |        | 50%   | 50%    | 100%   |        | 50%   | 50%    |        | 50%   | 50%    |
|  | BrandY | 4,173  | 2,037 | 2,136  | 4,173  | 2,067  | 1,014 | 1,053  | 2,106  | 1,023 | 1,083  |
|  |        | 25%    | 41%   | 43%    | 42%    | 24%    | 40%   | 42%    | 25%    | 42%   | 43%    |
|  |        |        | 49%   | 51%    | 100%   |        | 49%   | 51%    |        | 49%   | 51%    |
|  | BrandZ | 3,684  | 1,871 | 1,813  | 3,684  | 1,884  | 937   | 947    | 1,800  | 934   | 866    |
|  |        | 22%    | 38%   | 36%    | 37%    | 22%    | 37%   | 38%    | 22%    | 38%   | 35%    |
|  |        |        | 51%   | 49%    | 100%   |        | 50%   | 50%    |        | 52%   | 48%    |

This table shows a multi-response variable on the side, with the single-response Gender on top, both unfiltered and filtered to the regions North and South. Row and column percents are calculated simultaneously on the number of brands (16,743) and the number of purchasers (10,000), showing that 89% of respondents buy BrandX 53% of the time, etc, and that differences across gender and region are negligible.



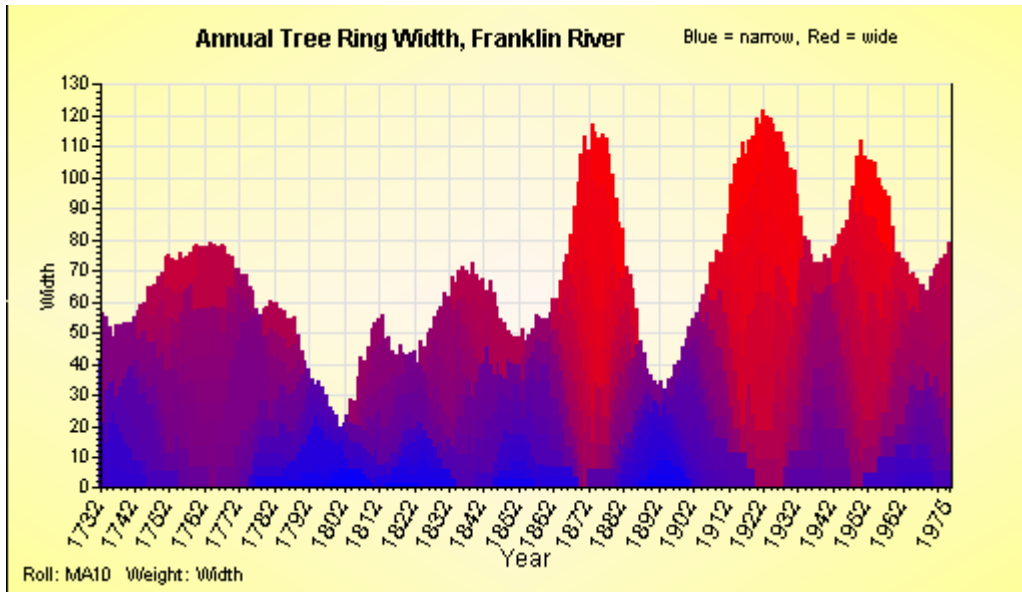
*This chart shows a plot of survey data for Claimed Recall ('Do you recall any advertising recently for BrandX?') against the external data for Gross Rating Points. The resolution is quarterly, with annual sums for GRPs and annual average percentage for Claimed Recall. The series has piece-wise trending to isolate Execution 1 for 2002 Qrt3 to 2004 Qrt1. The x axis codes for 2001 Qrt1 and 2002 Qrt2 are linked to call-outs advising of additional advertising in the print media.*

## Science

- Handle huge data sets
- Plot tens of thousands of points
- Plot thousands of series on a single chart
- Large-scale matrix arithmetic
- Aggregate and re-aggregate at different resolutions
- Unlimited smoothing
- Easy to prepare data from any source
- In-built programming languages for power users

Scientific data comes in all shapes and sizes and formats. Often the numbers are in ASCII files, arranged in columns or rows or lists, space formatted or delimited, and with many different ways of representing missing data. Such files can be manipulated and imported using Ruby's rich automation application programming interface (API).

Once in Ruby, huge data sets can be analysed interactively by varying smoothing algorithms and parameters, aggregating and reaggregating, proportioning, applying statistics, performing matrix operations, conditionally selecting data regions, summarising at coarser resolutions, filtering out data, and so on.

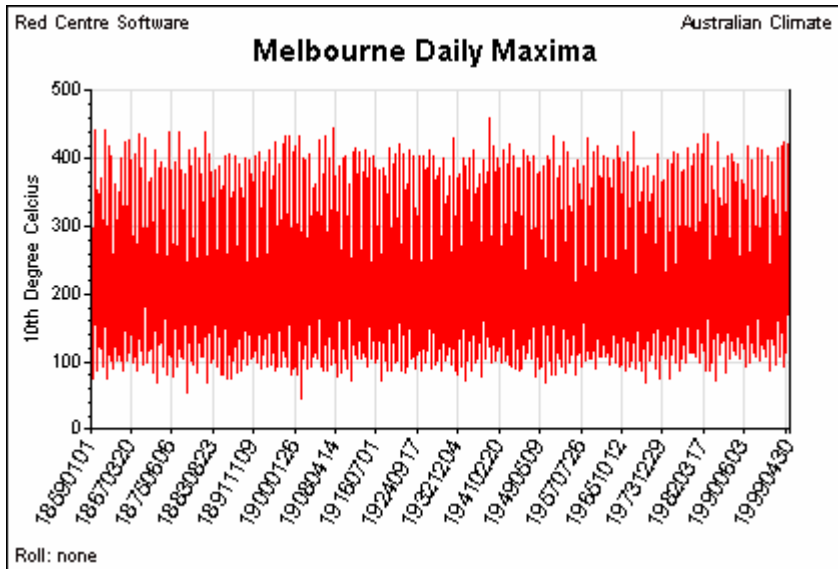


*A spectral plot of annualised tree ring widths on a ten year moving average. The data is for a single Huon Pine in the Franklin River region of Western Tasmania. Wide years, high amplitude red, indicate optimal growing conditions. Low amplitude blue indicates drought conditions.*

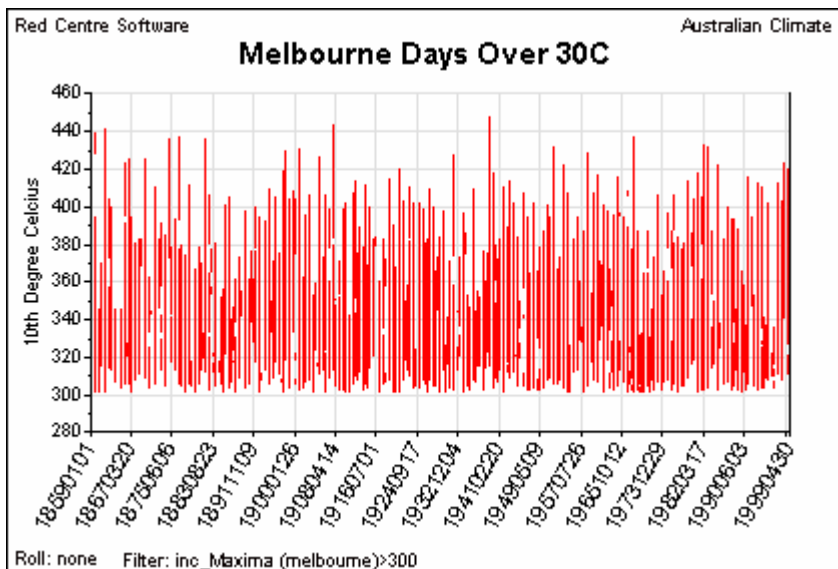
Ruby makes finding the underlying structure of an extremely long time series exceptionally easy. The following sequence of charts shows

- 1) raw plot of daily temperature maxima for Melbourne from 1858
- 2) filtered to days over 30 degrees
- 3) with the gaps removed
- 4) rolled at 3652 calendar days (equals ten years), zoomed on Y1, and trended.

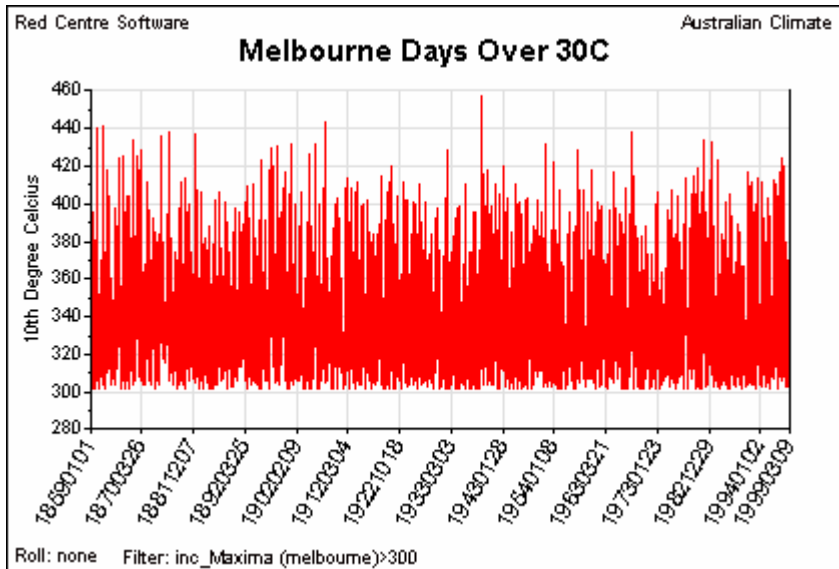
Getting from one chart to the next requires a few mouse clicks at most.



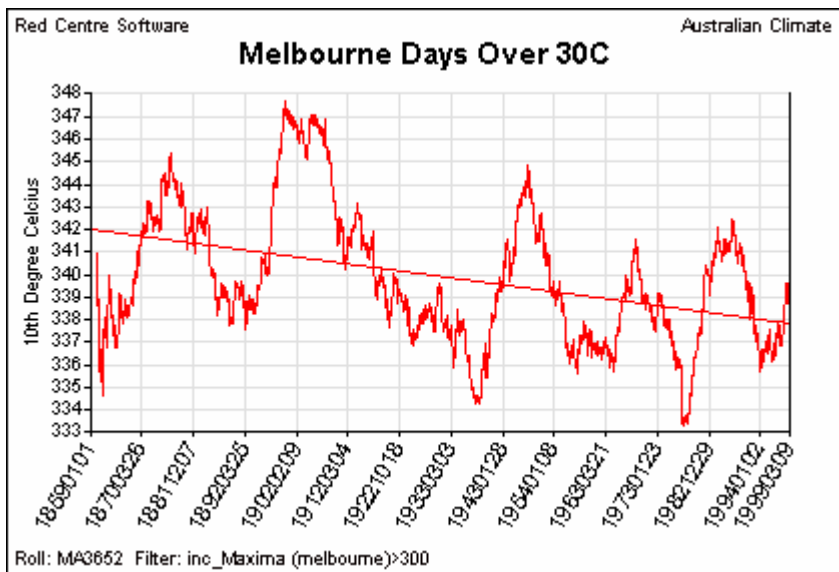
Daily maxima from 1 January 1859 to 30 April 1999. Plotting at daily resolution shows the outliers otherwise hidden by weekly or monthly averages. The lowest maximum, at 4.7C, was in 1901, and highest maximum, at 44.8C, was in 1939.



Observations for days under 30C have been filtered out. The actual filter expression is displayed as a footer.



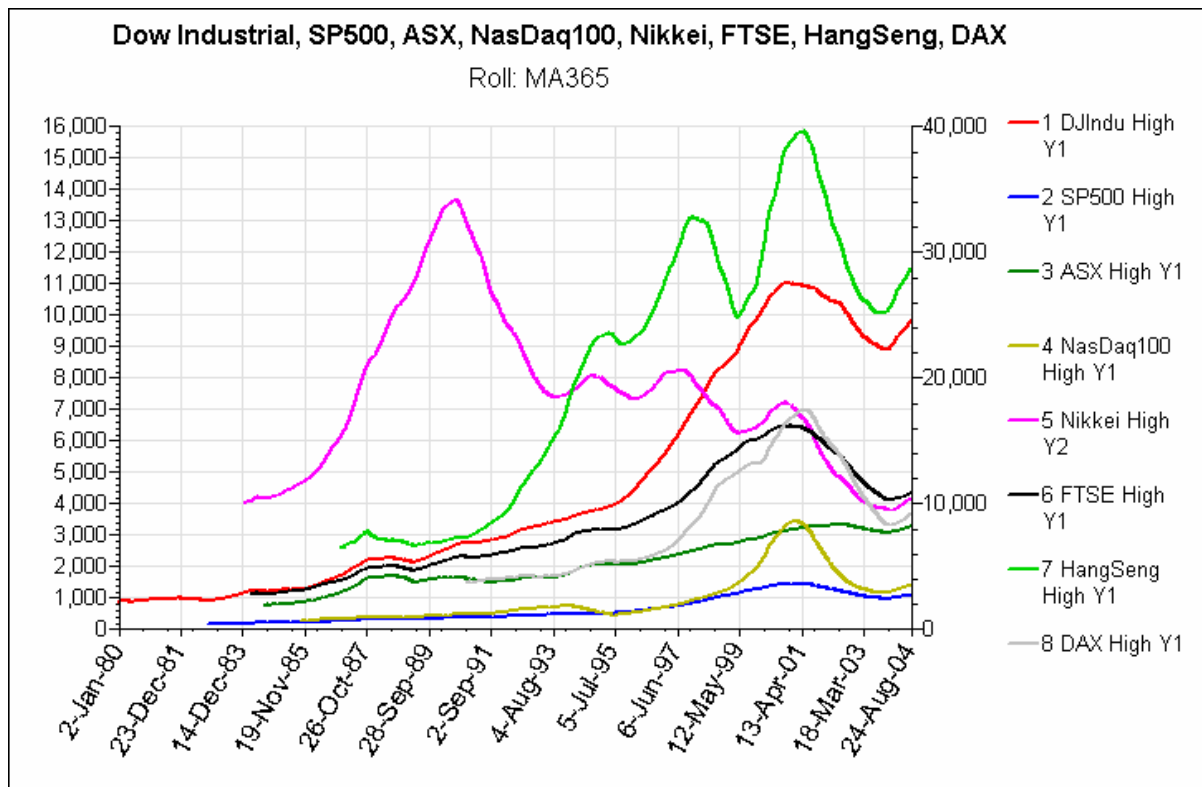
Removing all the gaps for days under 30C reduces the number of plotted points from 51,254 (number of days from 1/1/1859 to 30/4/1999) to 4,063.



Applying a moving average of 3652 days (= 10 years) and reducing the range of the Y1 axis to increase amplitude shows a clear pattern of cyclic peaks and a downward trend of nearly 0.5C over the 140 years. The trend line is calculated against unsmoothed data to avoid distortion. The smoothing is properly date-aware, so that each point is the average of days over 30C for the preceding 10 years (and not the preceding 3,652 plotted points).

## Financial

- Easily handles 100 years of daily data
- Missing data recognised
- Plot multiple market indices on the same chart
- Access and show any item of a database of external events with market impacts
- Calculate and plot expressions
- Zoom in and out
- Piece-wise regressions



A line series chart of the major stock market indices, shown simultaneously at daily resolution since 1980 and on a 365 day roll. The Nikkei is plotted on Y2, the other indices on Y1. This chart shows that the ASX is least affected by global trends.

To see this unrolled and with all out-of-field periods displayed takes a single click on a popup menu (to unroll) and a single click on the status bar (to unhide empty points).

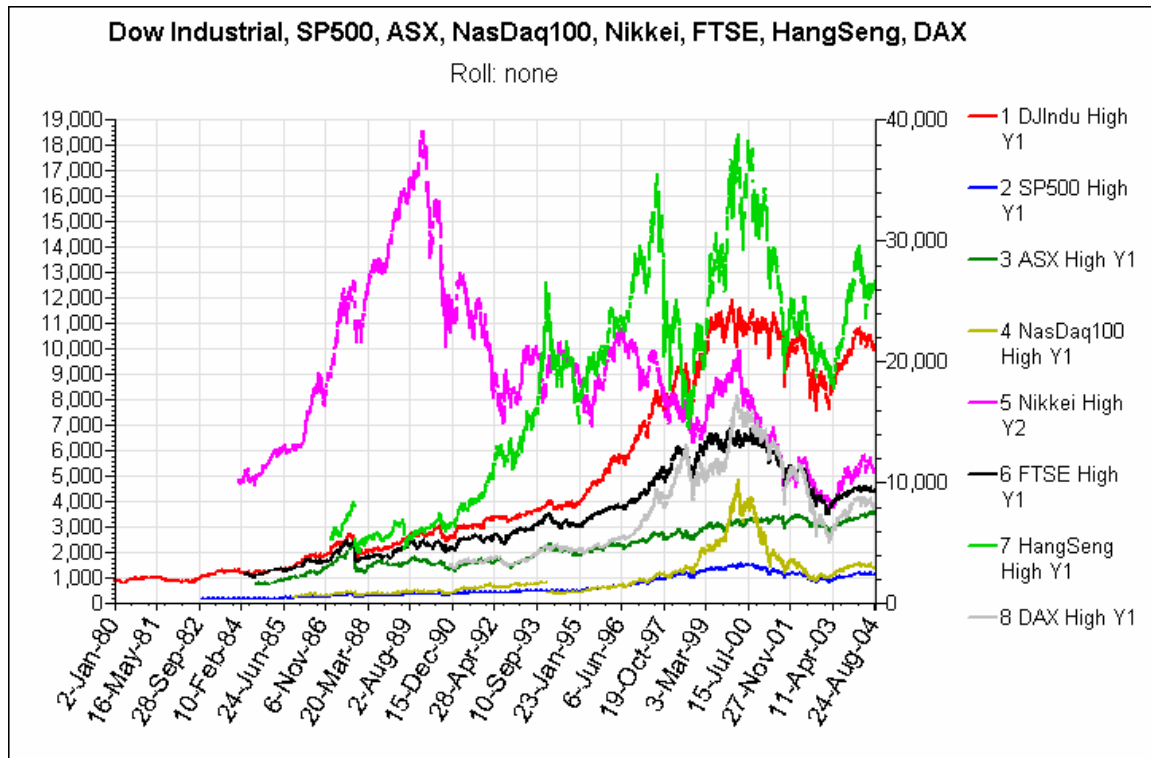
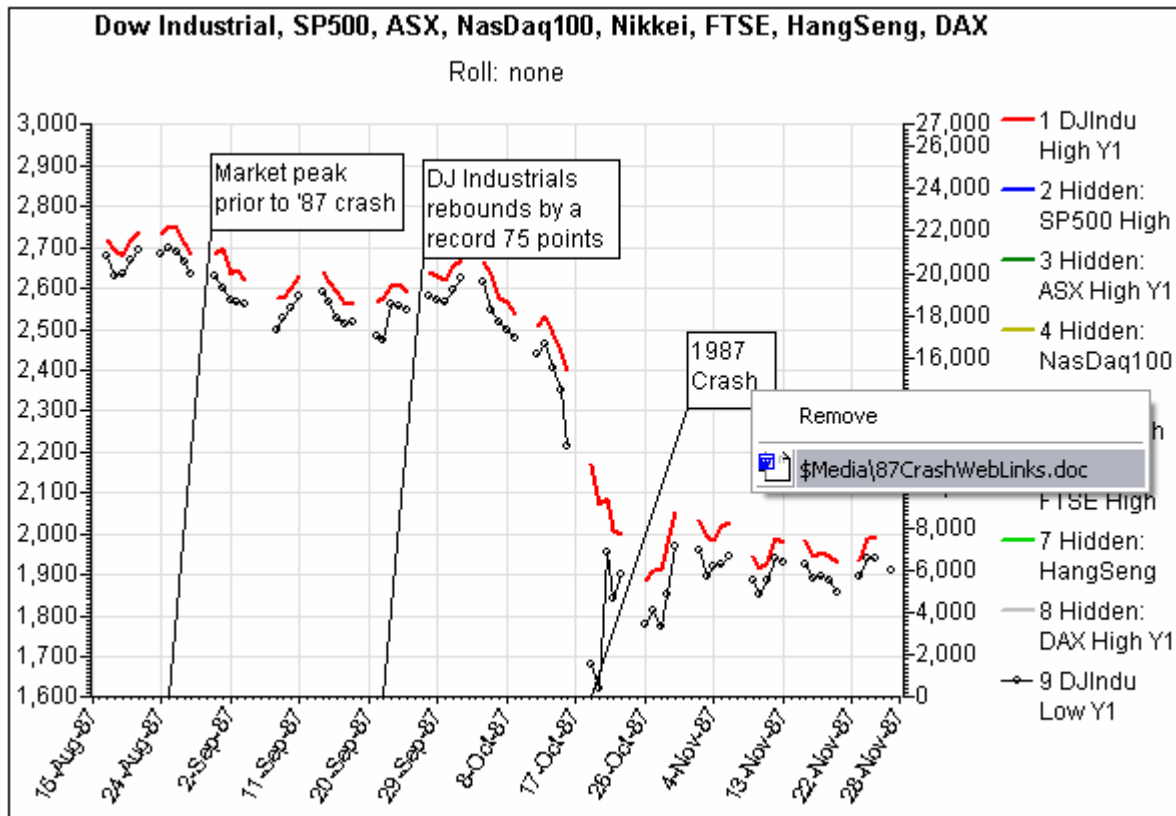


Chart shows all indices unrolled and spaced out across the real calendar. Different indices are out of field at different times.

The entire Highs data set, at daily resolution, is contained in the above chart. To see any bit of it up close, such as the Dow Jones Industrials during the 1987 crash, you can solo the series (that is, hide all others) with a right click and then zoom in with a simple drag. There are some entries in the Events database which would be nice to see too. These are added from the toolbar with a few clicks. Any annotation to an event can carry any number of file attachments. The crash event itself has an attached document of live web links which can be accessed by a right click and then activated with a left click. And for high/low comparison, adding a series for the Dow Jones Industrials Daily Lows takes a right click and a few left clicks. The result, prepared in just the several seconds it takes to do the mouse work, is



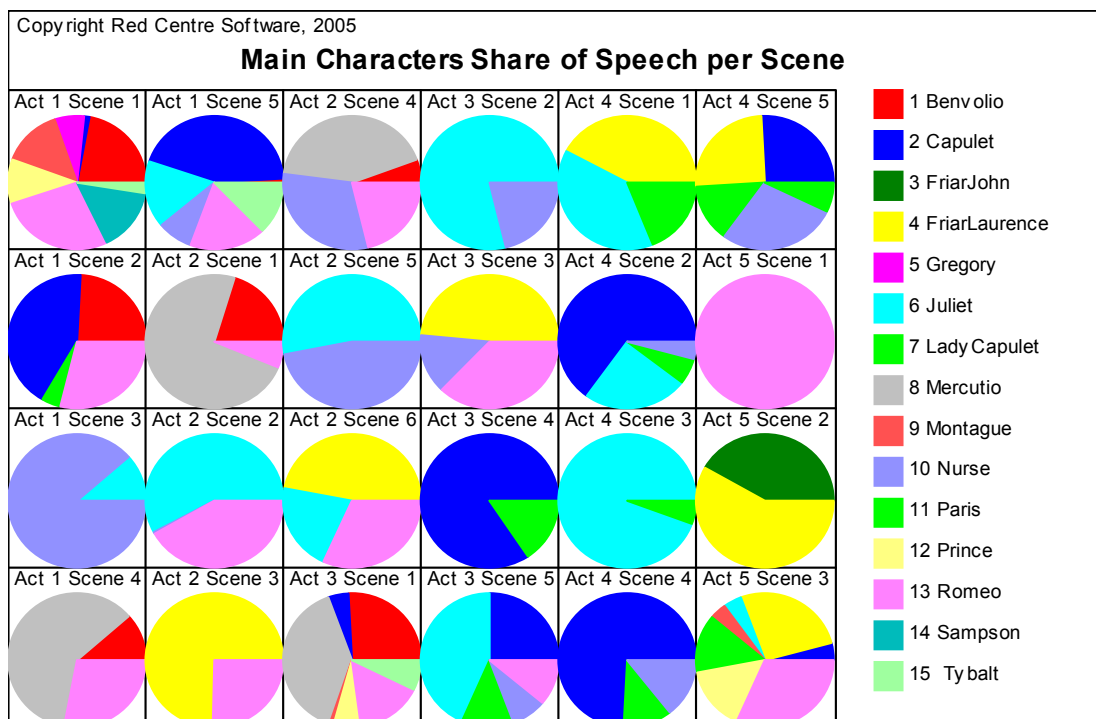
A zoomed view of the Dow Jones Industrials, other series hidden, the Low series added, and with some key events displayed. The crash event has a file attachment which contains live web links to interesting sites which discuss the event in detail.

The techniques for interacting with absolute data such as indices are quite different to survey series – for example percentaging is rarely relevant, and plotting at coarser resolutions such as week or quarter requires on-the-fly averaging. This means that any Market Research job which requires access to such data (temperatures for soft drinks or ice creams or air-conditioner sales, warehouse withdrawals, EFTPOS and scanner data, stock indices for financial services, etc) can seamlessly mix and match survey and external data at will.

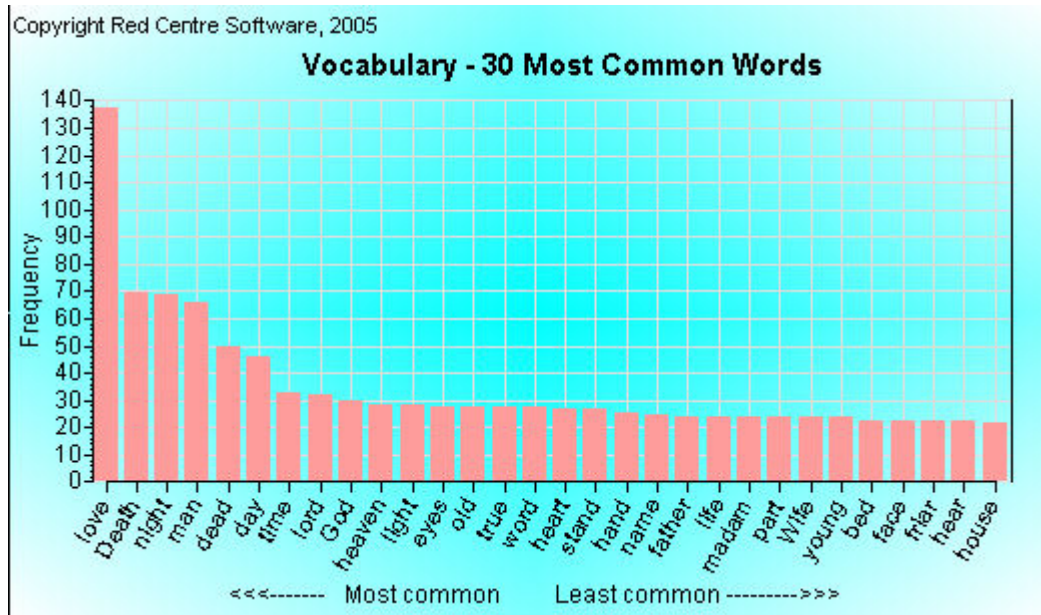
## Texts

- Analyse entire texts such as a play or novel
- Analyse submissions to an enquiry or collection of regional reports
- Track word patterns, recurrent themes, symbols, images, ideas
- Calculate statistics such as share of speech per scene per character
- Plot frequency distributions for any recurrent element

Ruby can be used for many different types of textual analysis. By making each word a case, and recording document and position within document against each, large scale tables can be run across entire sets at a resolution of each individual word. This makes it easy to list and sort all words, net them into groups and look for clusters, index each word to line number or page number, and so on. For historic texts, authorship can be tested by plotting distributions of particles, articles, or other parts of speech or text artefacts such as punctuation.



Above: Multi-pie chart showing the proportional distribution of words across main characters in each scene of Romeo and Juliet.



Frequency plot of most common words shows the emphasis on love and death/dead, day and night. Shows a thematic balance between love (137 occurrences) and death (69+50 dead =129 occurrences).

## Business Intelligence

- Analyse time sheets, employee charge-outs
- Analyse sales by any available criteria (region, personnel, product etc)
- Trawl warehoused data for hidden patterns and indicators
- Manage customer relationships
- Mix and match sales or orders or deliveries against stock or commodity prices

As a modern mariner navigating today's turbulent business world might be heard to say – "Data data everywhere, nor any drop to drink." Most businesses have no shortage of data – it is collected automatically by scanners, entered into thousands of spreadsheets, generated by electronic sales and internet orders, and so on. But making sense of it all is not easy. Who are the best sales reps? In terms of turn over? In terms of profit? Which product lines are declining or improving over time? Do these movements correlate with other data? The questions are endless, but without an effective tool for analysis, answers can be few.

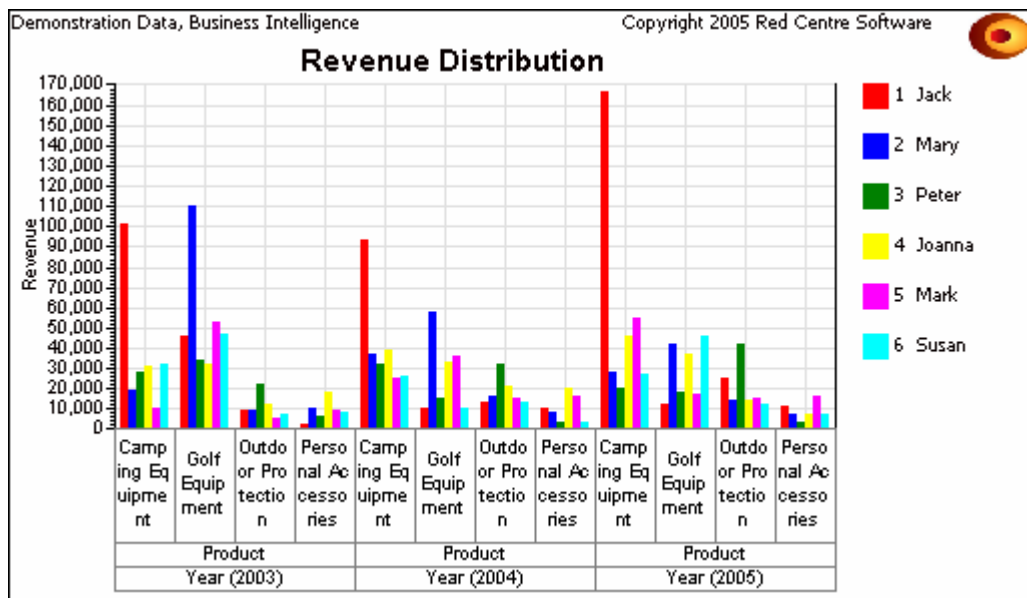
**Top: Year**

Side: SalesRep

SideSort: col 1 decreasing

| Frequencies |              | Year    |           |         |         |
|-------------|--------------|---------|-----------|---------|---------|
|             |              | Total   | 2003      | 2004    | 2005    |
| SalesRep    | Total        | 3,079   | 996       | 1,025   | 1,058   |
|             | Jack         | 714     | 228       | 205     | 281     |
|             | Peter        | 624     | 176       | 219     | 229     |
|             | Joanna       | 472     | 164       | 172     | 136     |
|             | Mark         | 457     | 148       | 159     | 150     |
|             | Mary         | 453     | 158       | 154     | 141     |
|             | Susan        | 359     | 122       | 116     | 121     |
|             | Margin (All) | Total   | 2,224,143 | 872,301 | 670,826 |
| SalesRep    | Jack         | 555,808 | 211,528   | 141,080 | 203,199 |
|             | Mary         | 459,276 | 202,922   | 141,048 | 115,306 |
|             | Peter        | 336,600 | 119,225   | 113,090 | 104,284 |
|             | Joanna       | 305,830 | 123,831   | 92,977  | 89,021  |
|             | Mark         | 297,481 | 103,051   | 109,757 | 84,673  |
|             | Susan        | 269,149 | 111,743   | 72,873  | 84,533  |

Composite table shows the number of orders processed by each sales rep, and the profit margins generated, by year and in decending order. Mary processes the second lowest number of orders (453) but generates the second highest profit margin (\$459,276).



Revenue distribution across four product groups, three years and each sales rep, as cluster bars. This entire chart can be switched to show either margins or costs instead with a few clicks.

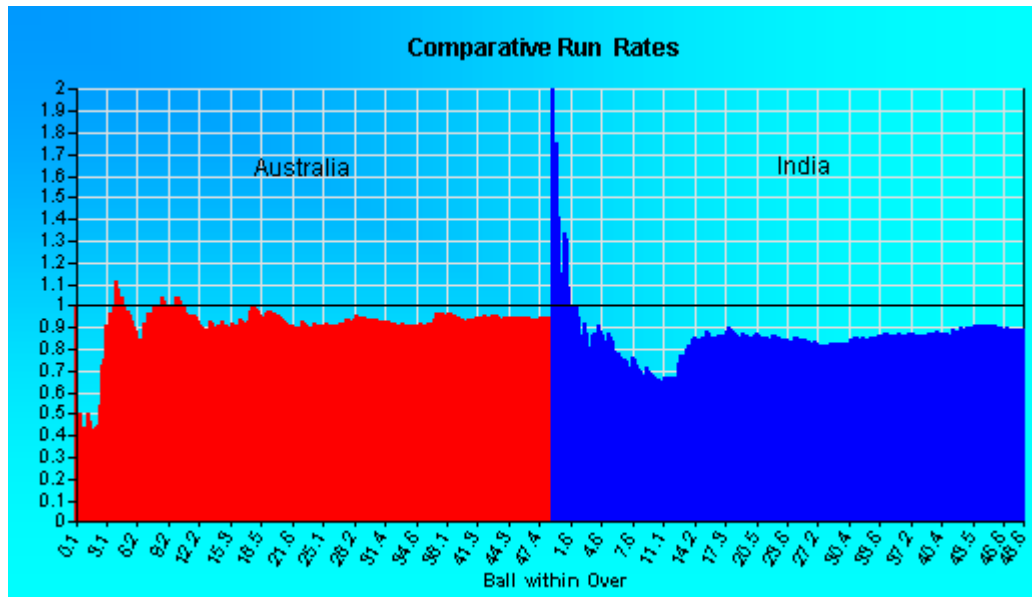
## Athletic Performance

To get an edge, every strength and weakness must be identified, every interaction explored. More important than any particular score are the underlying patterns and tendencies which predispose the outcomes. Understand these, and you will be well placed to exploit your advantages and capitalise on your opponent's weaknesses.

### Ordinal Ball Analysis

| Row Percents          | Ball  |     |     |     |     |     |     |    |    |
|-----------------------|-------|-----|-----|-----|-----|-----|-----|----|----|
|                       | Total | 1   | 2   | 3   | 4   | 5   | 6   | 7  | 8  |
| Batter (Ganguly)      | 82    | 12  | 6   | 10  | 15  | 18  | 19  | 2  |    |
| Batter (Agarkar)      | 1     |     |     |     | 1   |     |     |    |    |
| Batter (Bangar )      | 3     |     | 2   |     |     |     | 1   |    |    |
| Batter (Dravid )      | 16    | 2   | 3   | 1   | 3   | 6   | 1   |    |    |
| Batter (Kumble )      | 5     |     | 1   |     |     | 4   |     |    |    |
| Batter (Laxman )      | 16    | 1   | 4   | 2   | 2   | 6   | 1   |    |    |
| Batter (Pathan )      | 3     | 2   |     |     |     |     | 1   |    |    |
| Batter (Sehwag )      | 35    | 10  | 2   | 4   | 8   | 8   | 1   | 2  |    |
| Batter (Tendulkar )   | 63    | 4   | 13  | 9   | 13  | 14  | 6   | 4  |    |
| Batter (Singh Yuvraj) | 25    | 5   | 3   | 1   | 7   | 1   | 8   |    |    |
| Batter (Ponting )     | 18    |     | 3   | 5   |     | 6   |     | 4  |    |
| Batter (Gilchrist )   | 34    | 9   | 2   | 9   | 5   |     | 4   | 4  | 1  |
| Batter (Bevan )       | 1     |     |     |     |     |     | 1   |    |    |
| Batter (Bichel )      | 1     |     |     |     |     |     | 1   |    |    |
| Batter (Clarke )      | 63    | 7   | 5   | 18  | 19  | 4   | 4   | 6  |    |
| Batter (Gillespie)    | 8     |     | 2   | 1   | 3   |     | 2   |    |    |
| Batter (Harvey )      | 28    | 3   | 5   | 3   |     | 5   | 10  | 2  |    |
| Batter (Hayden)       | 20    | 4   |     | 3   | 6   | 1   | 6   |    |    |
| Batter (Symonds )     | 88    | 13  | 7   | 22  | 15  | 8   | 14  | 9  |    |
| Percent of Total Runs | 510   | 14% | 11% | 17% | 19% | 16% | 16% | 6% | 0% |
| Ordinal Ball          | 613   | 98  | 98  | 98  | 98  | 97  | 97  | 24 | 3  |

Batter performance from a one-day international cricket game. Table of Ordinal Ball by Runs, each row filtered to the batter, shows whether a batter does better at the start, middle or end of each over. Ganguly scores most runs at the end (balls 4, 5 and 6). Symonds dominates on ball 3. Clarke does best on balls 3 and 4, etc. The percentage row indicates an overall tendency to score more on balls 3 and 4, at 17% and 19% respectively.



Run rates for both innings, calculated as the cumulative score divided by the number of balls. Australia started poorly, but sustained almost a run per ball thereafter. India started well in the first few overs, collapsed during overs 4 to 11, soon recovered, but insufficiently to win.

## Major Features

- Import from SSS, Quantum, any card image, Surveycraft, SPSS, SAS, RDBs, etc
- Create and apply your own weighting matrices at will
- Process millions of cases in a few seconds
- See ALL the data, including being able to do tables of raw verbatims
- Create complex multi-layered tables with just a few mouse clicks, such as Behaviour within Age within Gender within Region
- Handle an unlimited number of variables and codes
- Use any type of data – categorical, quantitative, single-response, multi-response, incremented, hierarchic, textual, or any mix thereof
- Cross tabulate and chart external data such as GRPs or warehouse withdrawals or financial indices or temperatures against survey variables
- Manipulate rows and columns arithmetically and statistically in many ways
- Independently specify the variable, code, matrix expression, weight and/or filter for each row and column
- Reorganise or hide or remove or add rows and columns at any time
- Create new variables by mapping or constructing codes under any mix of boolean or arithmetical conditions, including new hierarchic variables to any depth and with any number of nodes
- Chart tables, or make tables of charts
- Analyse either graphically or numerically
- Slice, dice and drill tables and charts by interactive filtering

- Dynamically interact with generated reports by summing rows or columns, showing/hiding frequencies, sorting on any key, re-percentaging on different bases, even basing one table on a row or column of another table altogether
- Handle time series data from daily through to yearly and beyond, with full recognition of the real calendar at every point for intelligent handling of out-of-field
- Interactively or conditionally select periods for statistical comparisons against benchmarks or for piecewise linear regressions
- Select different rolling algorithms for control over roll-ins and out-of-field periods
- Handle tens of thousand plottable points (eg daily from 1950) as easily as 52 weeks
- Access an in-built Events/Attachments data base which allows any type of file (multi media, graphics, AV, spreadsheets, live web links etc) to be attached to any code of any variable, such as advertising executions to a week code, pack shots to a brand code, or anything else you like
- Update continuous tracking jobs in seconds or a few minutes at most
- Produce presentation quality tables and charts, with full user control over fonts and all layout details.
- Organise all of your related documents (Word, PowerPoint, web pages etc) as part of the job, editable, and accessed and viewed from the Table of Contents in the same way as a regular Ruby report
- Export slideshows of tables/charts to MSOffice applications as data, graphics or as recreated MSGraphs
- Choose whether to work on-line or with local data
- Use VBScript or JavaScript to automate complex or repetitive tasks or to integrate inputs and outputs with other applications
- Do the sorts of things you would otherwise have to specify in Quantum or Surveycraft or SPSS, but using a Windows GUI instead of arcane and archaic command syntax

## **Who we are**

Red Centre Software is a consultancy which specialises in developing applications for the processing and analysis of survey data, or of any data which can be represented numerically.

The RCS team has deep experience in these fields, and in the IT world at large, having supplied products and services since the early 1990s to many multi-national research and service companies and their local and global clients.

Our major areas of expertise include

- Complex surveys
- Dynamic data
- Hierarchic data
- Statistics
- Time series, calendars
- Geographic mapping
- Technical documentation

Collectively, our qualifications are in mathematics, computer science, astronomy, meteorology, statistics, education and languages and literature.