Ross Ihaka

ADRIAN MALLOCH

Accidental Success of an Anarchist

The numbers man who creates order from chaos.

West Auckland statistician Ross Ihaka says he's "an anarchist from way back". And that would seem to be a damn fine trait: the free data-analysis software Ihaka and his Canadian co-creator let loose on the internet in 1996 has taken the world by storm.

It's called R. Anyone can download it, crunch numbers with it and extend it, as long as they share the underpinning infrastructure, or source code. R's relative simplicity has made it must-have software, used on a daily basis by companies such as Shell, and Merck Sharp & Dohme, and universities such as Cambridge and Harvard. If you use Google, it touches your life too – the company uses R to find and analyse patterns in user searches.

Some IT companies estimate more than one million people use the

software regularly; R has spawned more than 40 books, which is rock-god territory in academic terms.

This hasn't made millions for Ihaka, 55, an associate professor in the University of Auckland's statistics department, or his US-based colleague Robert Gentleman. But they do have groupies of a sort. R – named for the first letter of their forenames – has drawn together thousands of collaborators who have designed and shared more than 2500 free, bespoke R add-ons, including software that analyses speech patterns, a program that studies the human genome and even an R Sudoku game.

R was born in the early 1990s after Ihaka and Gentleman realised their students needed simpler data-analysis software than what was then available. Although they weren't programmers, they "cobbled together" something to

help their students and put it on the net.

They didn't expect R to have a life outside the university, but it took off.

In 2008, Ihaka was awarded New Zealand's top science innovation award, the Pickering Medal. And last August, he and Gentleman went to Canada to receive an inaugural $9000 American Statistical Association award for innovation.

Ihaka is working on a new version of the software which he calls "daughter of R"; he hopes it will be a thousand times faster. There's a need for speed, he says, as our capacity to wrangle data into something meaningful is exceeded by the astonishing rate at which we're collecting it.

He points to a new radio telescope project between New Zealand and Australia that harvests five petabytes of data a day. Five whats? "A petabyte is 1000 terabytes, a terabyte is 1000 megabytes, and a megabyte is the size of a novel – *War and Peace* rather than a Mills & Boon."

If you see anarchism as the impulse to just do it, Ihaka embodies it — and then some. And who knows where he'll go next: "I'm only an academic until I decide what I want to be when I grow up."

JULIE MIDDLETON