SCRUM Club: use open source software for your research

When it comes to open access, most people's minds go directly to the availability of research articles for everybody. But there's more to it, for instance the software that researchers use to analyse their data. The Scientific Computing and Research @ UM Club (SCRUM Club) feels that using a licensed programme like SPSS always shuts out a group of people. That's why they promote and teach openly and freely available software and tools for scientific computing and research at their biweekly meetings.

Most Dutch universities teach students how to use SPSS for their statistical analyses. The university pays for the license and everyone within the organization can use it for free. "But that access ends once someone leaves the university, unless they pay the high fee themselves," says Wolfgang Viechtbauer, associate professor of Methodology and Statistics at the Department of Psychiatry and Neuropsychology and one of the founders of the SCRUM Club. Anybody who is not at a university may have difficulties accessing SPSS files.

Open source software doesn't have that problem and there are other perks as well, says co-founder Sander van Kuijk, clinical epidemiologist. "In SPSS, users typically use their mouse to carry out their analyses. But say I want to run new data in the same way I ran some of my old data. I can't remember exactly what I clicked on six months before. With R (the most-used open source statistical analysis programme, ed.) you type your commands and save the script. You can see exactly what you did – as can everybody else, which makes it more transparent."

Using a command line instead of an interface where you just have to click on what you want the programme to do, can be intimidating at first, admits Viechtbauer. You're programming instead of just using the software. "But it's a hurdle that can be overcome. As you start mastering it, you become much more efficient. You can automate tasks. For instance, a colleague of mine had to transfer his work into another format. If you do that by hand, it's 4-5 hours of tedious and error-prone copying and pasting. When you use the command line to programme a task, it will take you about an hour the first time and only seconds after that." An extra benefit is that employers will be more impressed by programming skills than with the standard SPSS skills, according to Viechtbauer. "Students are already aware of this," says Van Kuijk. "And I see more and more people outside the statistical world switching to R."

People who'd like to learn more about R or other scientific computing programmes like reference managers, can join the SCRUM Club. "Even if it's just for a few meetings. We post the topics on our website," says Viechtbauer. "If you are enthusiastic about a programme and want to teach others about it, that's also possible," adds Van Kuijk. "Every meeting someone else takes the lead. We founded it, but that doesn't mean we determine what happens."

Some universities have officially switched to open source software, but Viechtbauer doesn't see Maastricht University making that change anytime soon. "SPSS has a user-friendly interface, they are worried that students will feel overwhelmed if they have to use a more complicated programme, especially since a lot of them already struggle with statistics. My question then is: how come other universities were able to do it? It's good to push students a bit. Besides, switching to R when you already know SPSS is much harder than learning R from day one."

Cleo Freriks
For more information and the meeting schedule,
see https://wviechtb.github.io/scrum-club/

In rugby the term 'scrum' is used for players packing closely together with their heads down. At the SCRUM Club members put their heads together to learn about open source software for scientific computing
Photo: commons.wikipedia.org