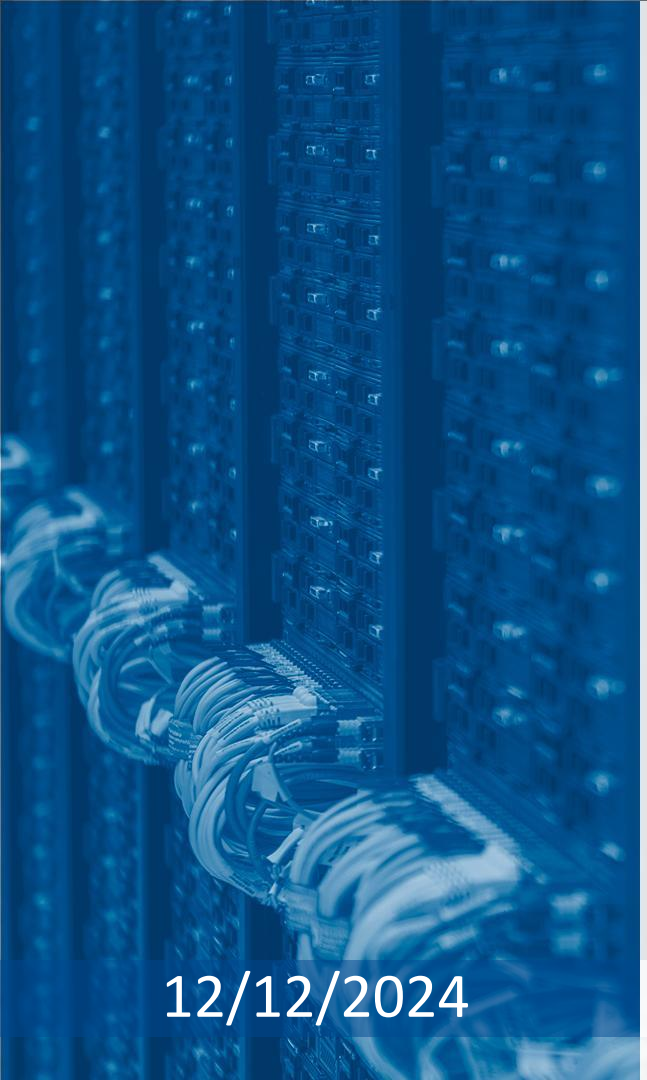


Barcelona Supercomputing Center
Centro Nacional de Supercomputación

R user meeting

12/12/2024

Victòria Agudetse, Ariadna Batalla, Theertha Kariyathan

Agenda

1. Ice-breaker: package ‘christmas’
2. News
    - startR
    - s2dv
    - CSTools
    - CSIndicators
    - esviz
    - SUNSET
3. Survey results
4. Q&A

Ice-breaker:
R package ‘christmas’

Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

R Package ‘christmas’

christmas is a package that **creates different Christmas cards.** Some are cute, some are funny, some don’t really work that well. Let’s take a look…

```
library(christmas)
xmascaganer(year = 2025, language = "catalan", seed = 1)
xmasepitree(year = 2025)
xmashealth(year = 2025)
xmashelix(year = 2025)
xmasneon()
xmaspicasso(year = 2025, language = "catalan")
xmasregression(year = 2025, language = "english")
```

You can see all the functions in the package documentation. Happy Holidays!

startR

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Improve error message about missing dimensions

Start() returns an error message if any of the specified dimensions are not found in some of the requested netCDF files:

```
> Could not find the dimension 'lat' in the file. (...)
```

If there was more than one dimension missing, the error message used to look like this:

```
> Could not find the dimension 'latlon' in the file. (...)
```

Now, it looks like this:

```
> Could not find the dimension(s) 'lat', 'lon' in the file. (...)
```

MR: https://earth.bsc.es/gitlab/es/startR/-/merge_requests/243

status: in master

Possible deprecation of ecFlow as workflow manager

startR can be used to load heavy data in chunks and apply a user-defined function to each chunk. It can use a workflow manager (**Autosubmit** or **ecFlow**) to execute the chunks in parallel on an HPC.

Currently, both workflow managers are supported, but we are considering deprecating ecFlow, because:

1. We have internal support for Autosubmit
2. Newer ecFlow versions seem to give some problems in the BSC-ES Hub
3. It will be time-consuming to do new developments for two different cases

Possible deprecation of ecFlow as workflow manager

The decision is not final yet, and ecFlow will continue working as it is **at least for as long as we have the workstations** and possibly longer if we can easily solve the issues in the Hub.

Steps to do before deprecation:

1. Solve some minor bugs in the implementation of Autosubmit
2. Update the practical guide
3. Update all use cases
4. **What else do you need to support you in this transition?**

s2dv

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

RandomWalk bugfix

Bugfix in .RandomWalkTest() when “A_better” or “N.eff” is NA.

```
if (isFALSE(N.eff)){N.eff <- length(skill_A)}

   if (!is.na(output$score) & N.eff >= A_better) {
     p.val <- binom.test(x = A_better, n = floor(N.eff), p = 0.5,
                         conf.level = 1 - alpha,
                         alternative = test.type)$p.value
```

```
Error in if (!is.na(output$score) & N.eff >= A_better) { :
  missing value where TRUE/FALSE needed
```

Comparing an NA in an if() clause gives an error

```
> if (NA >= 1) {print ("hello")}
Error in if (NA >= 1) { : missing value where TRUE/FALSE needed
```

RandomWalk bugfix

Bugfix in .RandomWalkTest() when “A_better” or “N.eff” is NA.

```
if (isFALSE(N.eff)){N.eff <- length(skill_A)}

   if (!is.na(output$score) & N.eff >= A_better) {
   if (!is.na(output$score) & isTRUE(N.eff >= A_better)) {
     p.val <- binom.test(x = A_better, n = floor(N.eff), p = 0.5,
                         conf.level = 1 - alpha,
                         alternative = test.type)$p.value
```

Now if N.eff or A_better is NA the condition returns as FALSE.

MR: https://earth.bsc.es/gitlab/es/s2dv/-/merge_requests/197
status: in master

CSTools

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

New methods to .make.eval.train.dexes()

- Nuria added new cross-validation methods to the function .make.eval.train.dexes().
- The function creates training indices for different cross-validation methods.
- The new methods are tested and unit tests were added.
- Next we will focus of simplifying the function and making it visible to the user with its own documentation.

issue: https://earth.bsc.es/gitlab/external/cstools/-/issues/150

status: in branch dev-cross-indices

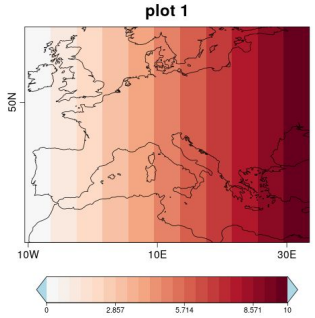
esviz

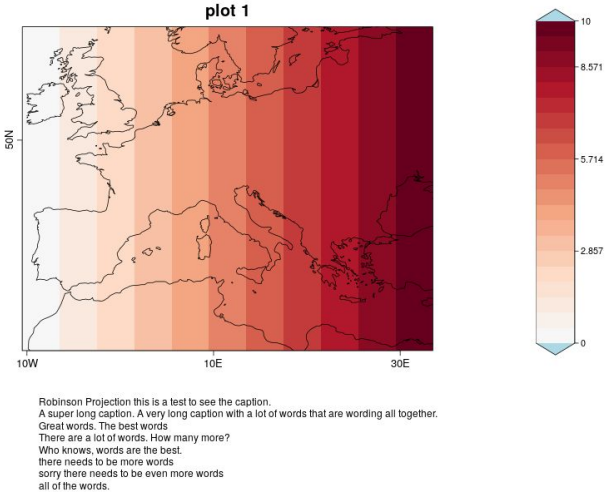
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

New parameters: caption + vertical colorbar

- **caption**: a character string of the caption located at left-bottom of the plot. Use "\n" to indicate a new line.
- **caption_size**: scale factor for the figure caption. Default is 0.8.

- **vertical**: TRUE/FALSE for vertical/horizontal colour bar. Default is FALSE. Parameters 'width' and 'height' might need to be modified to accommodate the vertical colour bar.

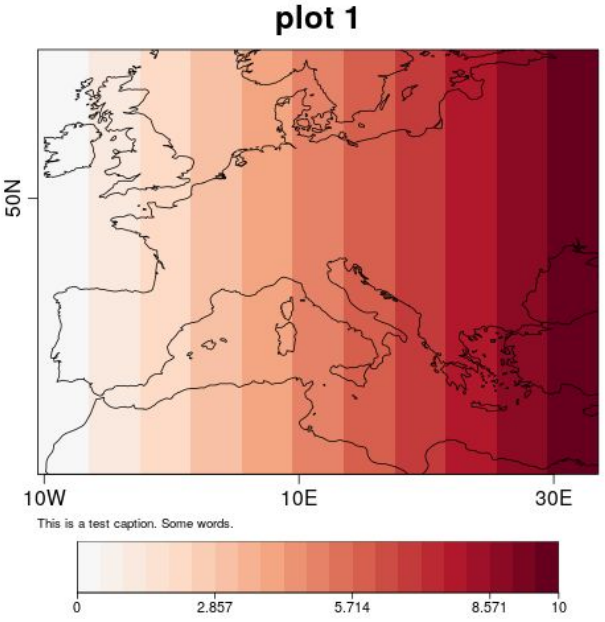


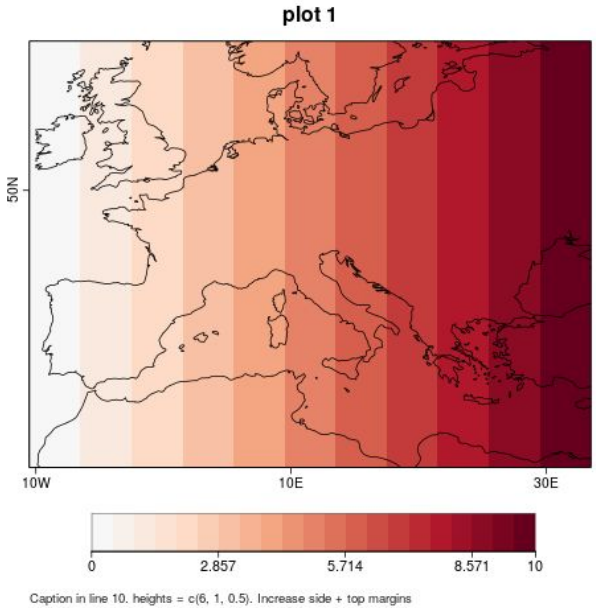


New parameters: caption + vertical colorbar

Question:

- When the color bar is horizontal, should the caption be placed above or below it?





SUNSET

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Launching jobs on CTE-AMD and MN5

The launch_SUNSET.sh script can now be used to launch jobs on CTE-AMD and MN5.

1. **Without Autosubmit:** Your output directory 'Output_dir' should be in **/gpfs/.** You can launch the script directly from the HPC machine.
2. **With Autosubmit:** Your output directory 'Output_dir' should be in **/esarchive/**, and you should launch the script from the Hub or a workstation. A temporary directory will be created in your /gpfs/scratch/ automatically. When all the jobs finish, the outputs will be transferred back to **esarchive** and the temporary files will be cleaned.

We are working on a hands-on tutorial to show the step-by-step process.

issue: https://earth.bsc.es/gitlab/es/sunset/-/issues/149
status: in master

Downscaling module updates

There have been some updates from the Intlr() function from CSDownscale:

1. In the `lr_method = "9nn"` option, Principal Component Analysis (PCA) is applied. A bug related to this application was identified in the function and has been corrected.
2. If `exp` ("hindcast") and `exp_cor` ("forecast") have the same number of members, each member is now downscaled individually, so that the downscaling equations derived from each hindcast member are applied to the same forecast member.

MR: https://earth.bsc.es/gitlab/es/sunset/-/merge_requests/174

status: in master

Documentation improvements

We have added some new information to the SUNSET README.md file detailing the first steps to using the tool: https://earth.bsc.es/gitlab/es/sunset#getting-started-with-sunset

Remember that, additionally, we have:

- The wiki
- Step-by-step use cases
- A recipe template

We are working on more improvements. If you have any suggestions or miss anything, please comment in this issue: Documentation: Errors, improvements and changes

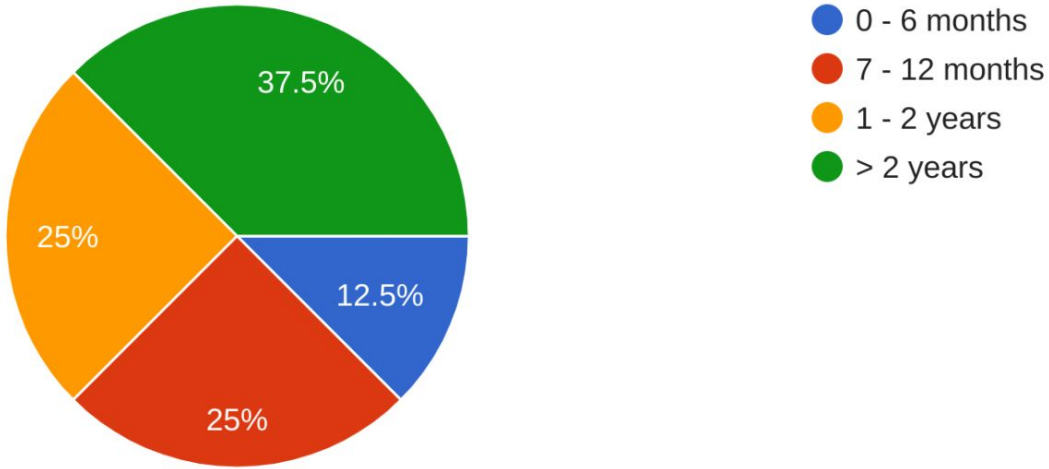
R user meeting survey

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

R user meeting survey

0. How long have you been in the BSC R community

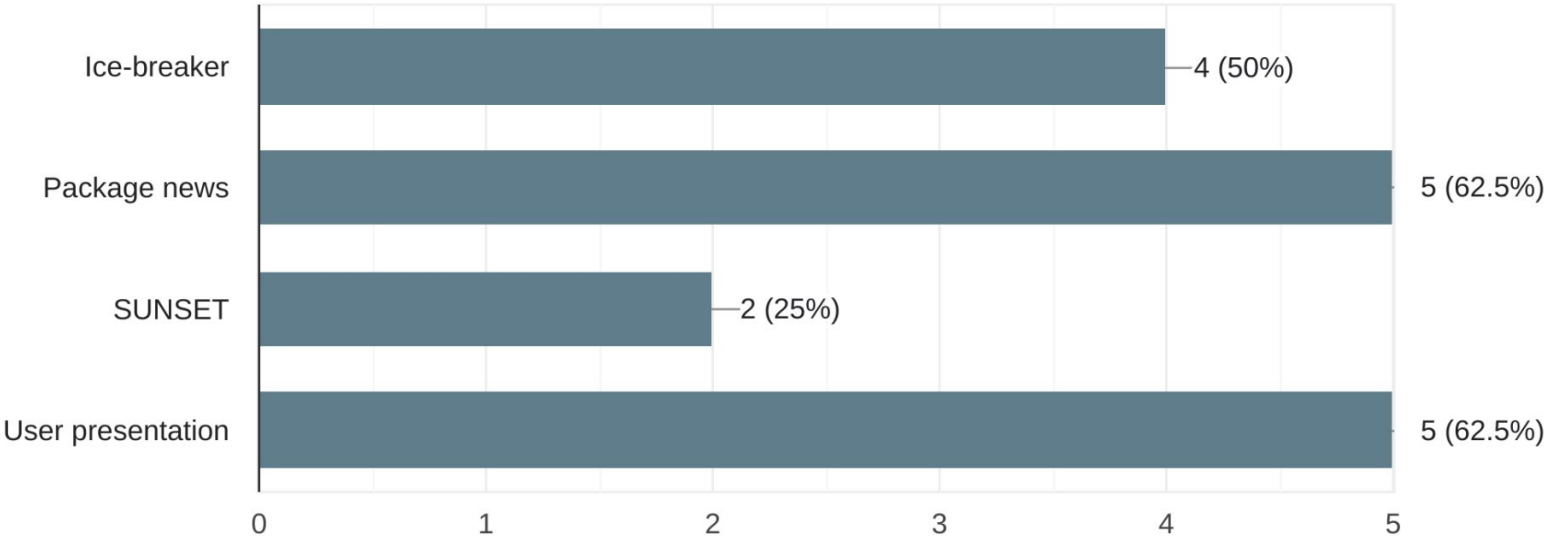
8 responses



R user meeting survey

1. Which parts of the user meeting do you find useful?

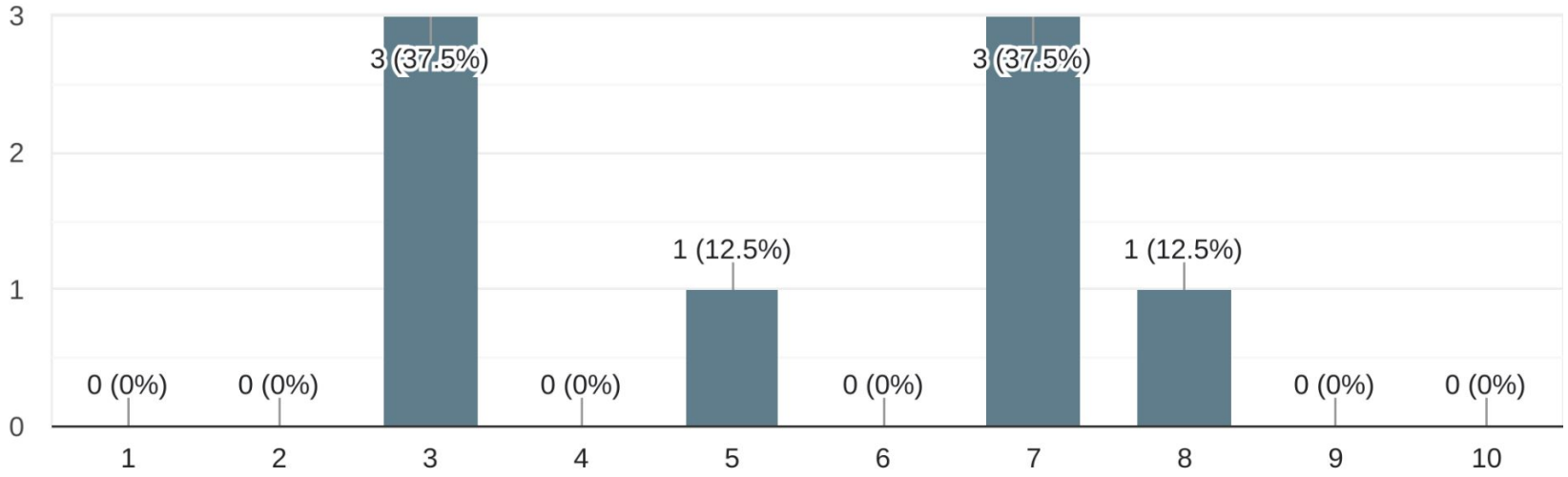
8 responses



R user meeting survey

2. What percentage of the PACKAGE NEWS can you follow usually?

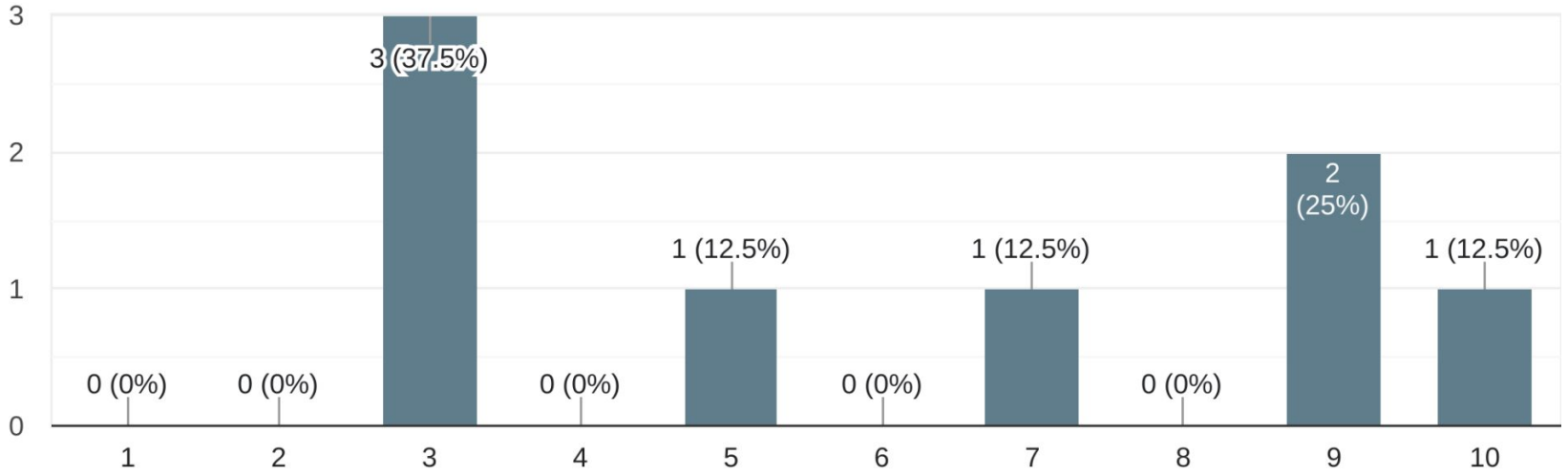
8 responses



R user meeting survey

3. What percentage of the SUNSET UPDATE can you follow usually?

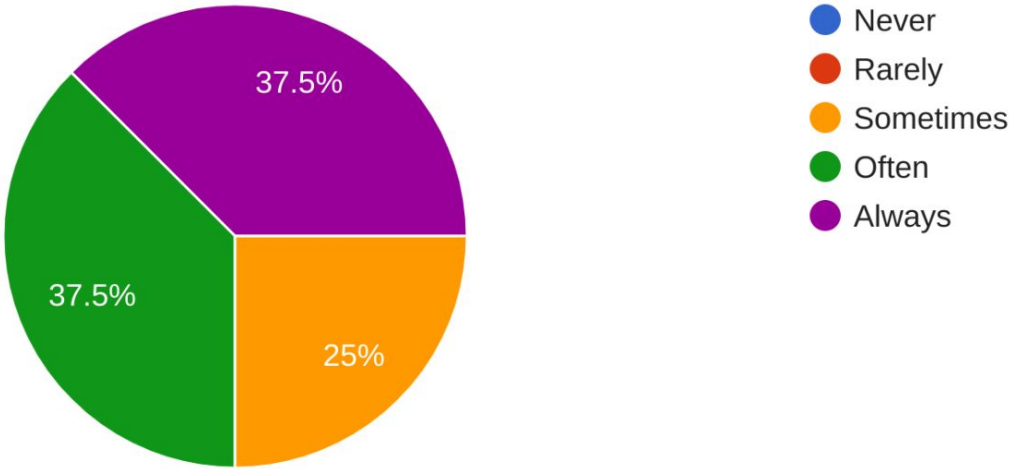
8 responses



R user meeting survey

4. How often do you attend R user meetings?

8 responses



R user meeting survey

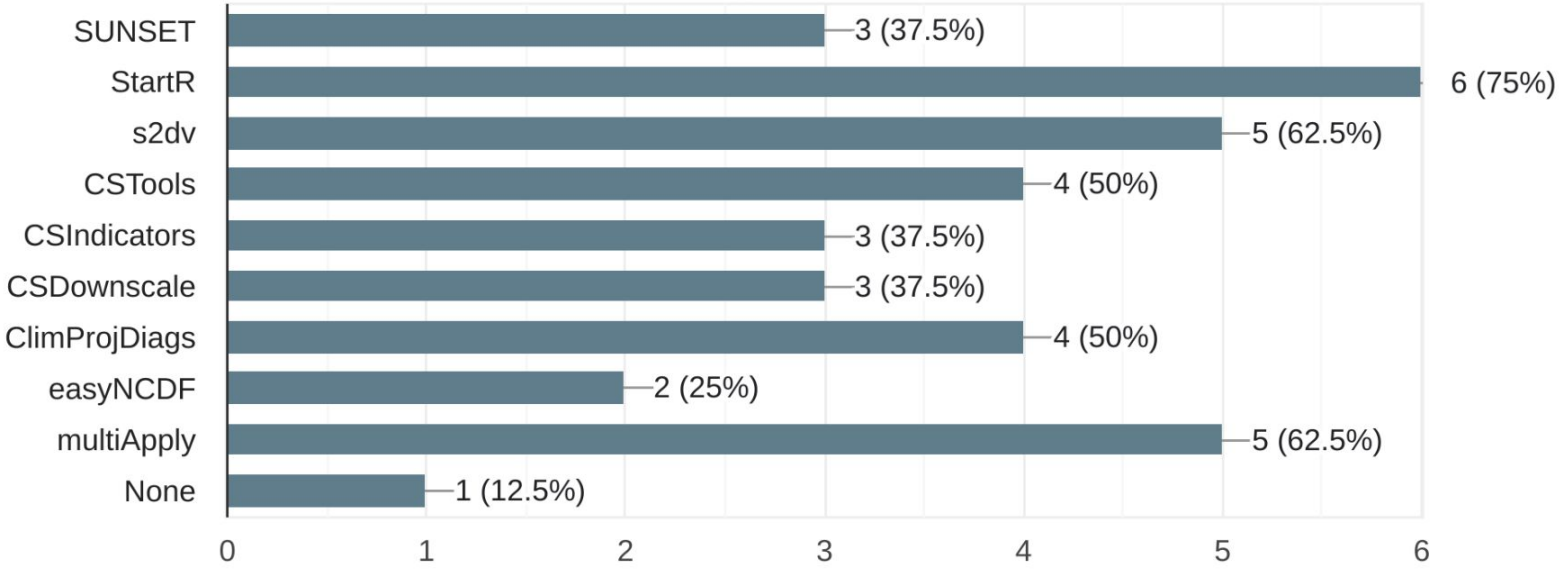
5. We are considering replacing some R user meetings with R support hours, where the R Dev Team books a meeting room and the users can come durin...r meeting to be replaced with an R support hour?

8 responses

R user meeting survey

6. Which internal packages/tools do you use? Note that if you only use the package within SUNSET, you should only select SUNSET.

8 responses

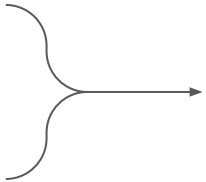


R user meeting survey

7. What things did you find difficult when you started using R tools in our department? Is there any missing information that would be nice to have?

Difficulties:

- Difficulty knowing which what the packages do and how to use them
- Difficulty using MultiApply, startR, SUNSET
- Configuring RStudio/VSCode in Nord3 and MN5



Nice to have:

- Documentation,
- exercises / use cases
- R packages guide / summary

How can we make the information more easy to access?

R user meeting survey

8. Do you have any other suggestions/requests/feedback?

- Help in improvement of functions
- Showing and sharing codes during the User presentation, with some to-do and not-to-do personal feedback of the presenter.

Thanks for joining

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Next meeting: February 6th