Cloud Native Ocean Data Analysis and Visualisation School — Feedback Form

25 responses

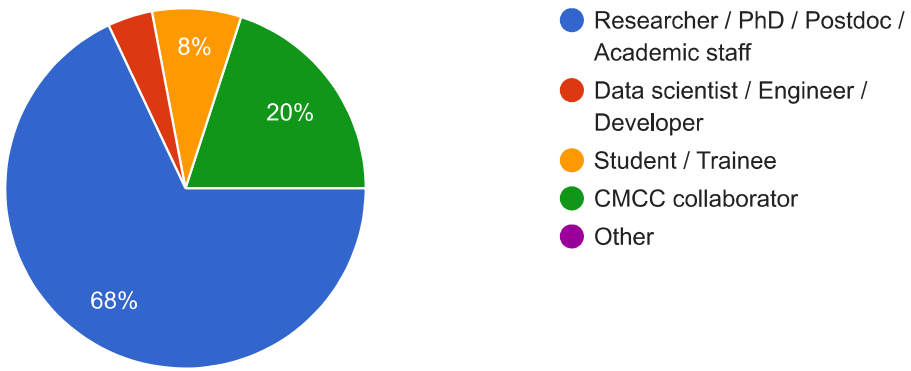
Publish analytics

A. About You

Role

Copy

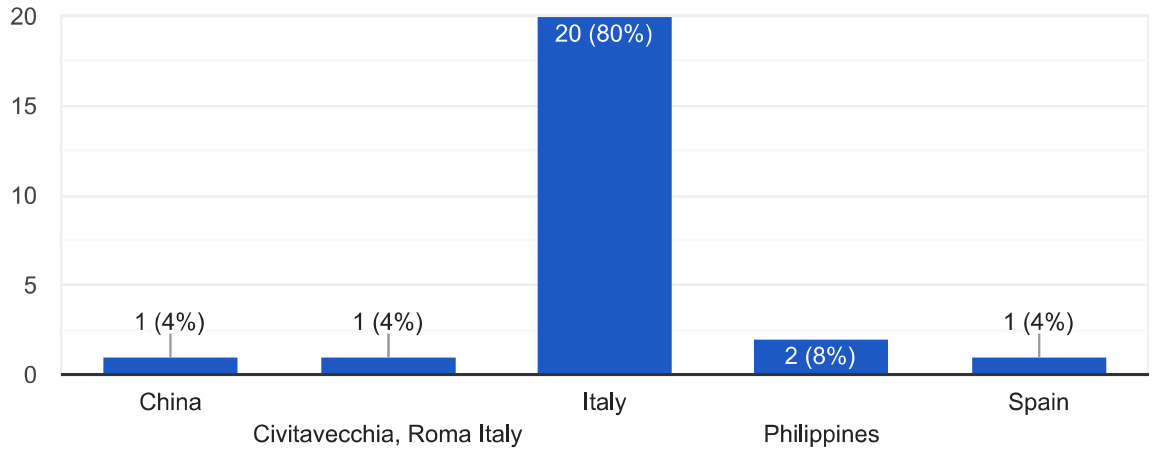
25 responses



Residence country

Copy

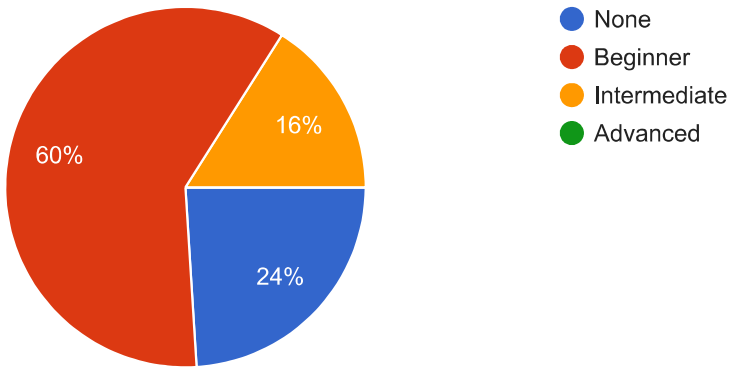
25 responses



Prior experience using cloud for scientific work

Copy

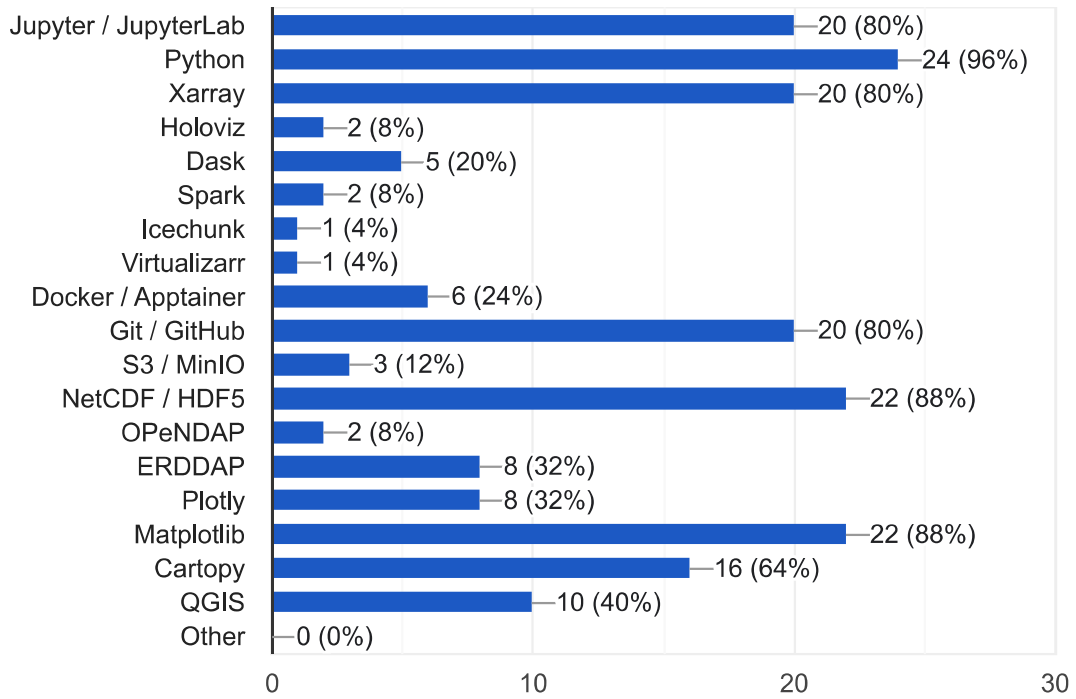
25 responses



Tools you used before today (select all that apply)

Copy

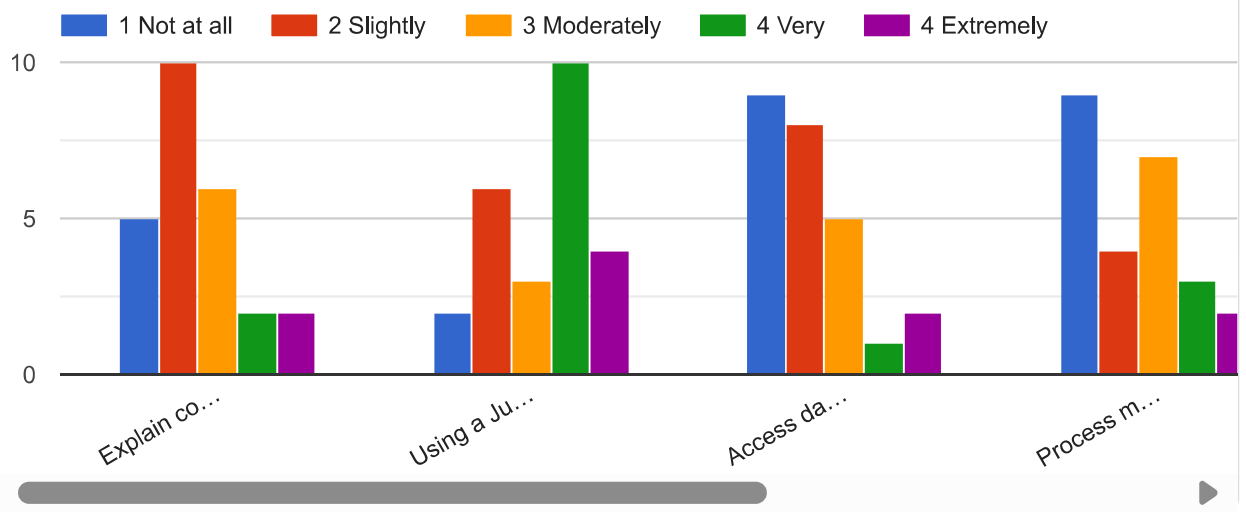
25 responses



B. Learning Outcomes (Retrospective)

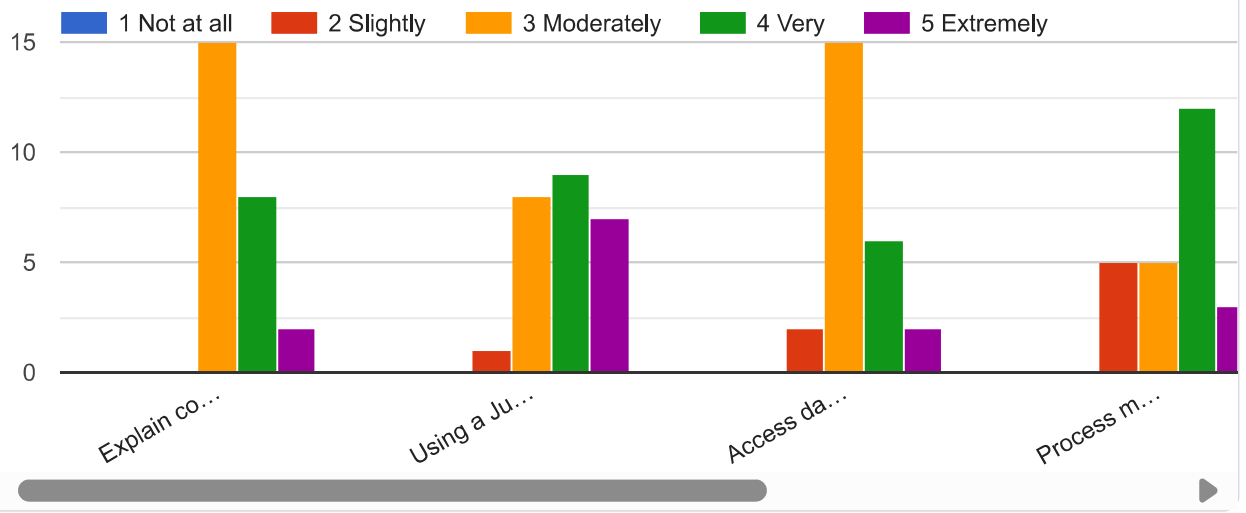
Confidence BEFORE the school

Copy



Confidence NOW

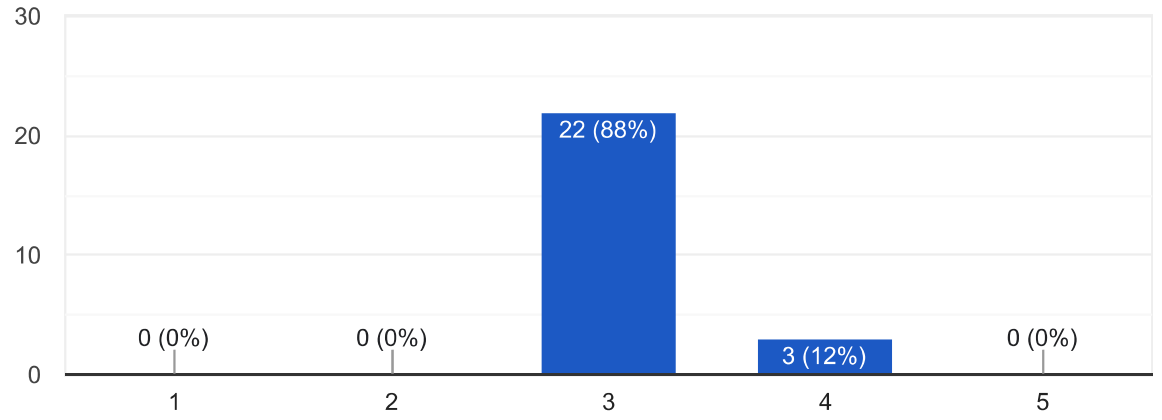
Copy



Overall difficulty level

Copy

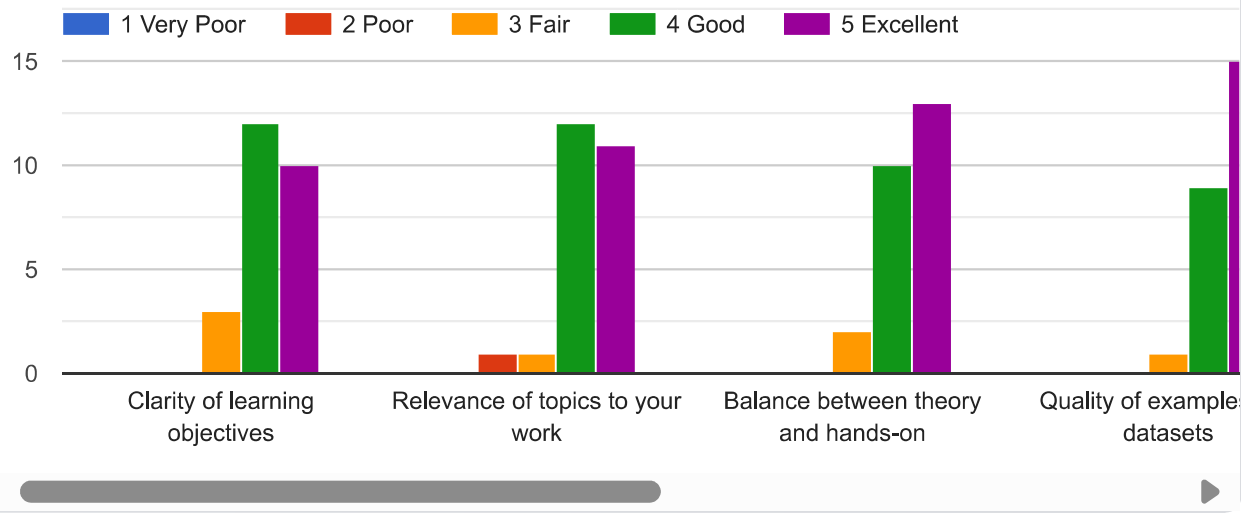
25 responses



C. Content & Instruction

Please rate the following (Quality 1–5)

Copy



What was the most valuable concept or technique you learned?

19 responses

I now understand much better how parallelisation works. Previously, I used it but had missed some concepts. I also find it valuable to explore other options for data storage and access.

perform data analysis completely in the xarray environment, first approach to parallelize the computation using dask

A good introduction to the topic of cloud

Dask usability was really interesting

Use the Pangeo to download and calculate instead of personal computing resources.

dask parallelization

chunk methodologies

Razz

Interactive visualization dashboard using the HoloViz visualization tool (hvPlot)

how to use zarr format to speed up cloud computation

Parallelisation and chunking

The concept of remote data access/storage

I think that the most valuable tools that I learned are VirtualiZarr and Icechunk for accessing multifile datasets and hvplot for data visualization

Probably parallelization with Dask and chunking with Zarr and Virtualizarr

Zarr file strcuture, parallel cloud computing with dask

How to efficiently read/access large data and generating interactive visualizations

I think it's more on parallel ocmputing by using Dask, it is my first time using it and at first quite intimidated since I don't know how to use it

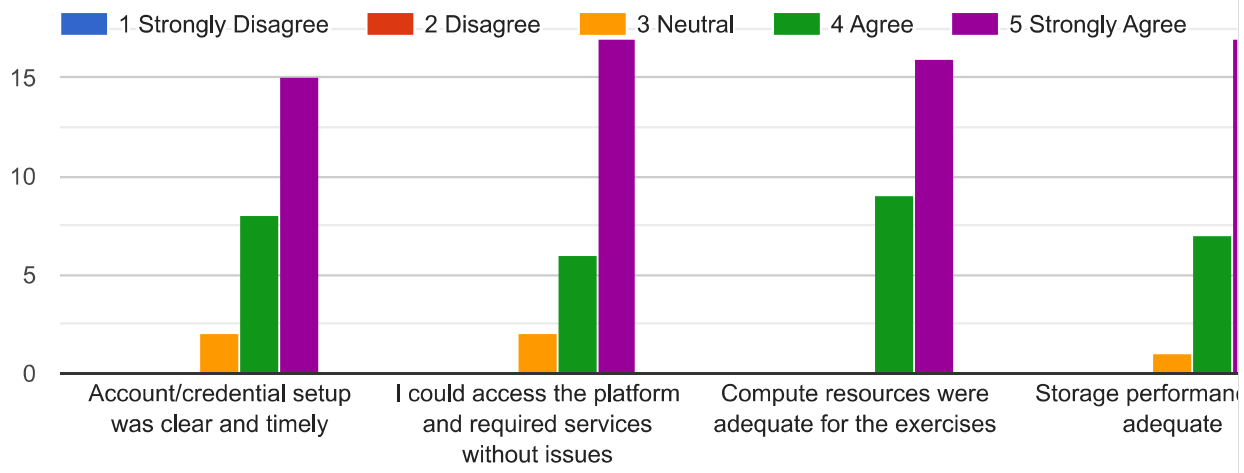
parallelization and lazy management of large dataset

Using dask and chunking, many good practices, and how to use zarr and icechunks to compute over the cloud and avoid recomputing of metadata

D. Hands-on Labs & Platform

Please indicate your agreement (1–5)

Copy



What significant issues did you encounter? Were these issues solved during the class?

14 responses

None

The Copernicus dataset I used for my project (the only one useful for during my daily work) were not suitable to practice on dask parallelization.

No at all, Jupyter was used to crash for huge-data applications

not enough memory in the cloud environment.

creation a bucket on MinIO, but the issue is cause for a setup of the MinIO

too long explanation time risks to loose attention by the students

/

The parallelization of the dataset with Dask was not feasible because of the restriction on Copernicus

Jupyter notebook not always working properly, especially when creating new cells

I had some issues replicating the Virtualizarr + Icechunk for a different dataset. The issue was solved with the help of the instructor.

I was not able to load relatively large datasets from Copernicus Marine Service with Dask and that prevented me from being able to create a realistic project, but that it's not something that could be solved within the class context, so that's fine! :)

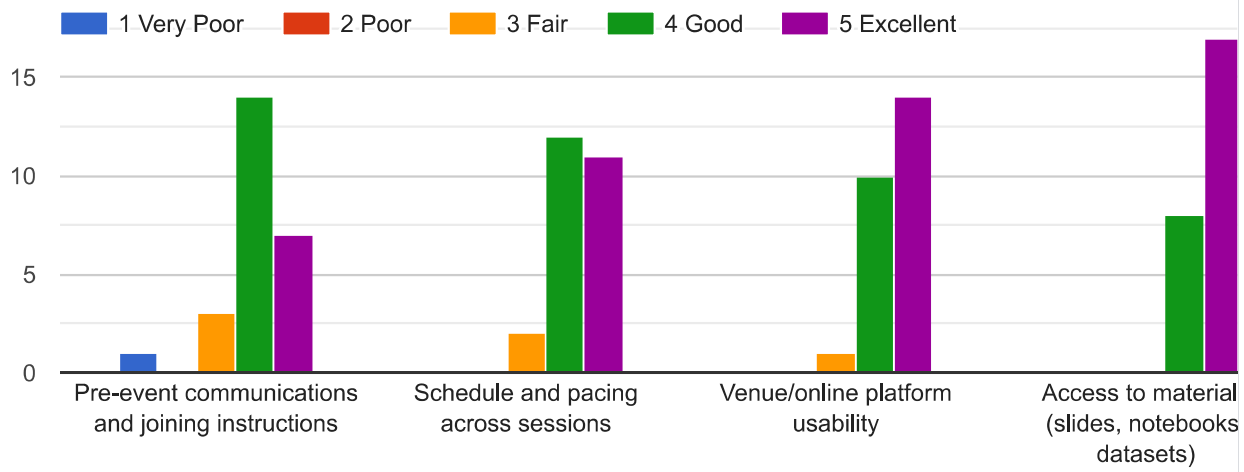
The only tool presented during the course that I did not manage to use is dask. I should go through the material provided again to clear that out.

upload data on the cloud from the HPC, alignement of data with different dimensions

F. Organization & Logistics

Please rate (Quality 1–5)

Copy



Any suggestions to improve logistics or schedule?

10 responses

None - I really think the organization was efficient and flexible

the classroom assignments could be on a unified project rather than scattered topics to encourage teamwork

Working in small groups would support class to be more productive and competitive

morning coffee break too close to the lunch time

no

/

Provide better info about the school BEFORE it starts

Everything was fine on my side

The invitation could be sent a bit more in advance, but not critical

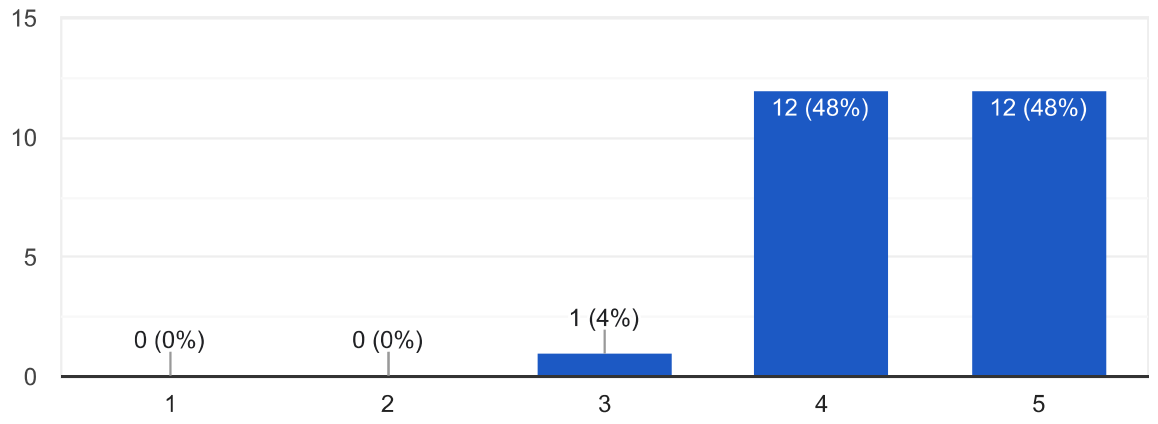
I think the school was well organized.

G. Overall Impact

Overall satisfaction

Copy

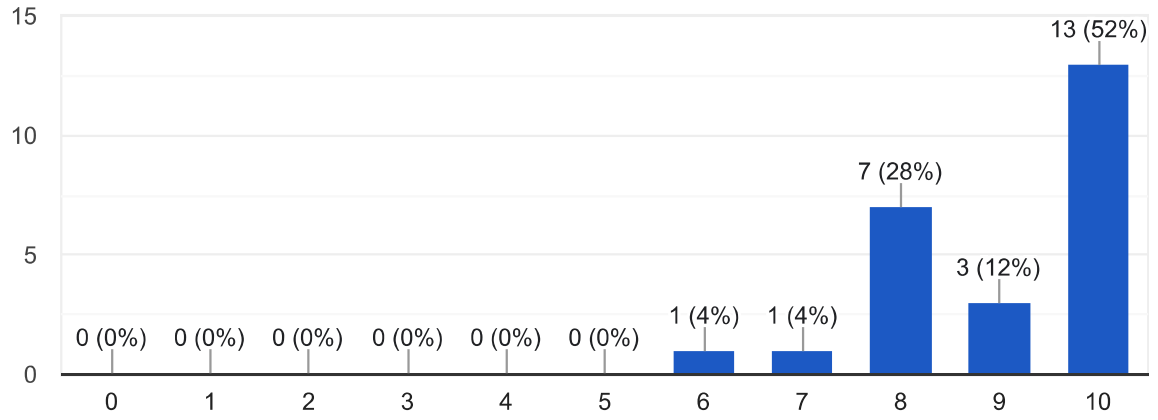
25 responses



How likely are you to recommend this school to a colleague?

Copy

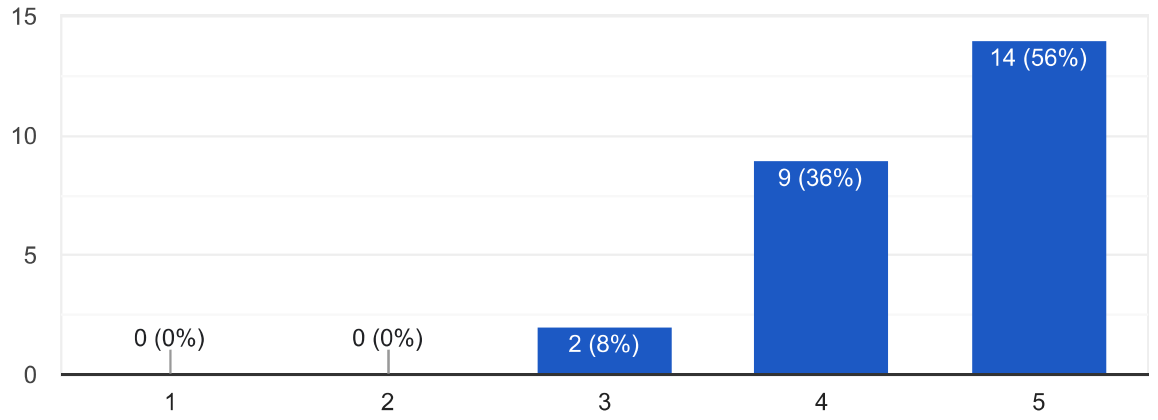
25 responses



Time well spent: the value matched the time I invested

Copy

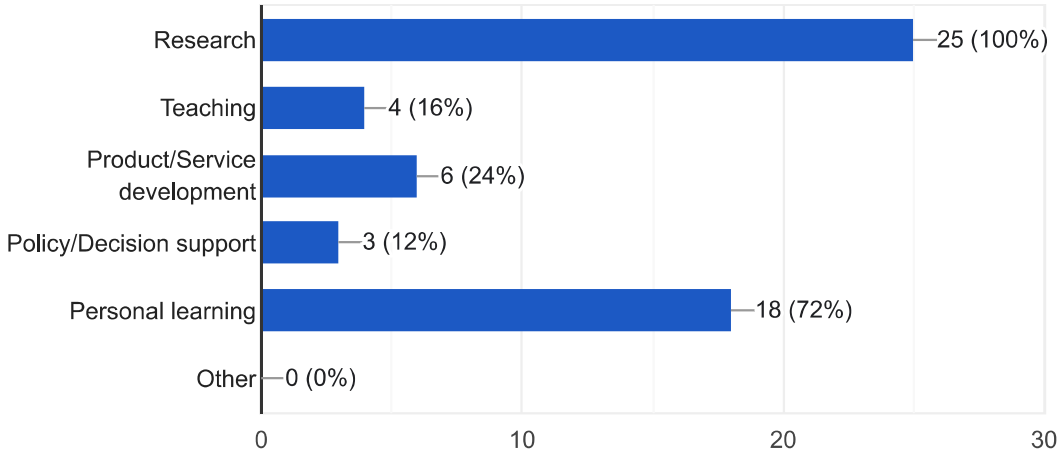
25 responses



Intended application (select all)

Copy

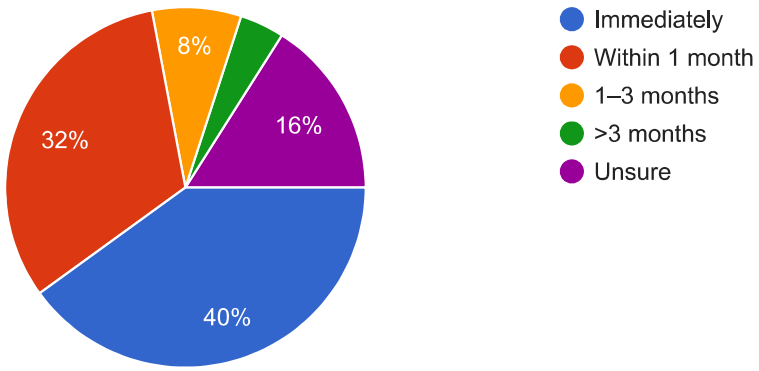
25 responses



Expected time to apply

Copy

25 responses



What will you do differently as a result of this school?

16 responses

Some plotting tuning. Checking the status and memory usage when parallelising.

I work mainly with model output, so I'll store my data better to make the dataset flexible and more efficiently accessible. Next time I have a chance, I'll explore accessing large datasets in the cloud in more depth.

ready to delve into ML

Working on small groups for the final project would help participants feel more productive and competitive

optimize how data is stored and acessed

Data Analisys storage

Apply the knowledge gained to my research

the way I read large dataset and store model output

I haven't planned to use large datasets in the future. But I can now handle them better if I ever need to.

I think that the first thing I'm going to apply is the hvplot tool for plotting. Then I think I'll continue to learn about best practices and other tools from Pangeo community

I will add data parallelization with Dask and fetching with Zarr to my workflow

Loading and procesing data at greater speed

More likely to implement cloud computing and storage, use zarr extension in my work with datasets (if possible)

Reading large data in efficient way issued from the ocean models used in our Research team (mostly NEMO and SHYFEM). Creating easy to use diagnostic/validation tool to assess the output of the models. Inclusion of some of the visualization techniques presented in the workflow of some our software products.

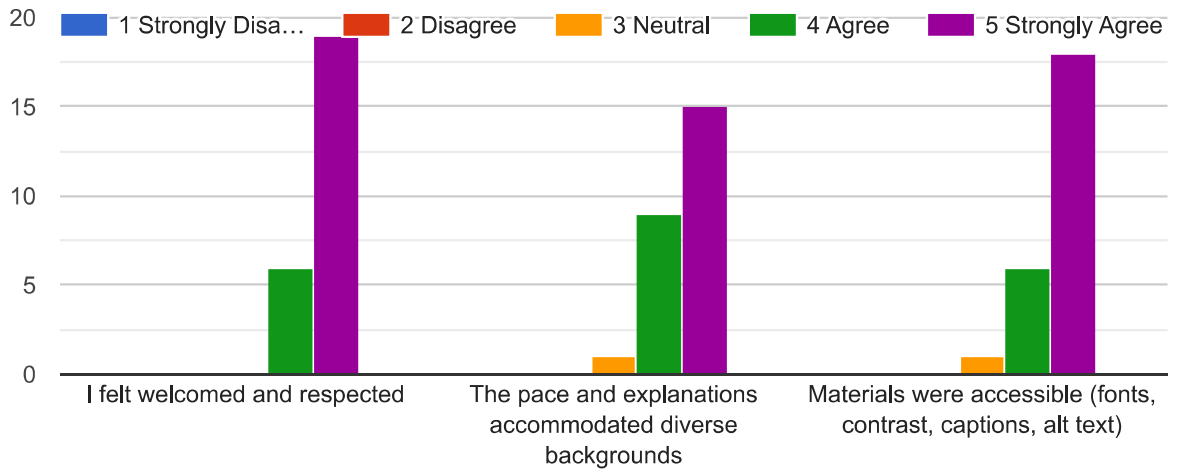
It's more on using reelvant packages and libraries especially hvplot for data visualization

management of large dataset and visualization

H. Inclusivity & Accessibility (Optional)

Please indicate your agreement (1–5)

Copy



Accessibility needs we could better support next time? (optional)

1 response

Perhaps the pace could be adapted to enable better following the exercises.

J. Final Comments & Permissions

Anything else you'd like to tell us?

6 responses

I actually enjoyed as well and the reason I still am not that confident is because of the time difference. Sometimes I miss some because it's too late here in the Philippines. Other than that, I really enjoyed how Richard explained and it is not that formal, but not too casual either. I wish I could understand more, this is actually my first time doing something in the cloud

Many thanks!

No

Thank you for the opportunity to be a participant

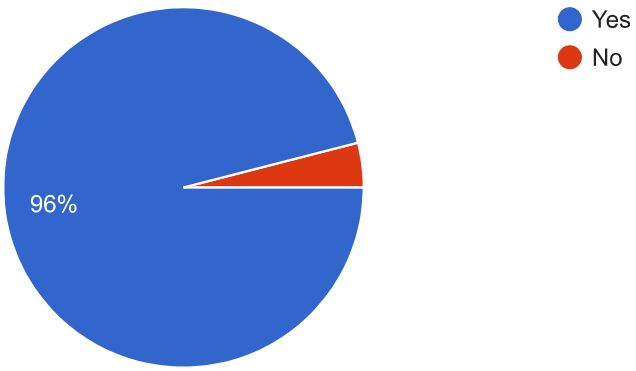
Thank you for organizing such a great school!

I think Richard Signell has an incredible amount of knowledge and succeeded to transmit it in an efficient way.

May we quote your feedback (anonymously) in reports/website?

Copy

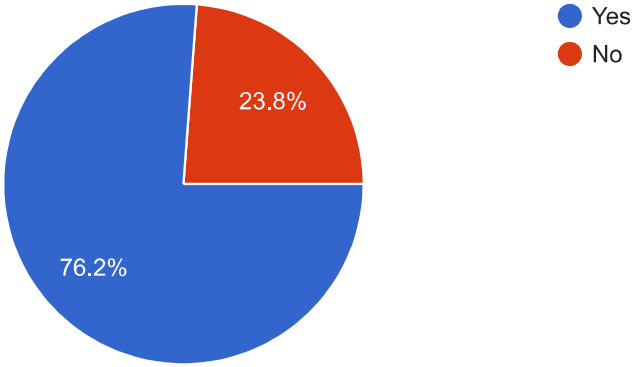
25 responses



May we contact you for a brief follow-up?

Copy

21 responses



Email for follow-up (optional)

14 responses

albert.martinez@cmcc.it

federica.borile@cmcc.it

ampagador@msi.upd.edu.ph

rafael.menezes@cmcc.it

francesco.benfenati@cmcc.it

nicolas.gonzalez@cmcc.it

beniamino.tartufoli@cmcc.it

jacopo.dallaglio2@unibo.it

serena.lezzi@cmcc.it

dmitry.kondrik@cmcc.it

igor.atake@cmcc.it

rodrigo.vicente@cmcc.it

egtorculas@msi.upd.edu.ph

muhammad.saad@cmcc.it

Consent

Copy

25 responses