What country are you from?

colombia

china

germany

usa

us

united states

russia
australia
austria

canada

brazil

greece

japan

united statws

coastal engineering

59

What field do you work in?

geomorphology

data curation

space physics

united states

machine learning

geophysics

economic geology

coastal engineering

data archives

science data management

climate science

researcher

polar meteorology climate

planetary science

climatology

it management

big data

astrophysics

data management

stratigraphy

italy

data

atmospheric science

geochemistry

seismology

remote sensing

libraries

gnss

srismology

bathymetry

metadata

gravimetry

oceanography

emissions diagnostics

open science policy

geology

cfd

data infrastructure

climate and conflict

hydrology

cloud computing

plasma physics

biogeosciences

ocean modeling

infomatic

ecology

computing

59

What software tools do you use?

Integrated Development Environments (IDEs)

5.5

Test Frameworks (pytest, ctest, testthat)

3.2

Version Control Systems (git, svn)

5.9

Collaboration Platforms (github, gitlab)

6.5

Documentation tools (oxygen, sphinx, docutils)

3.5

58
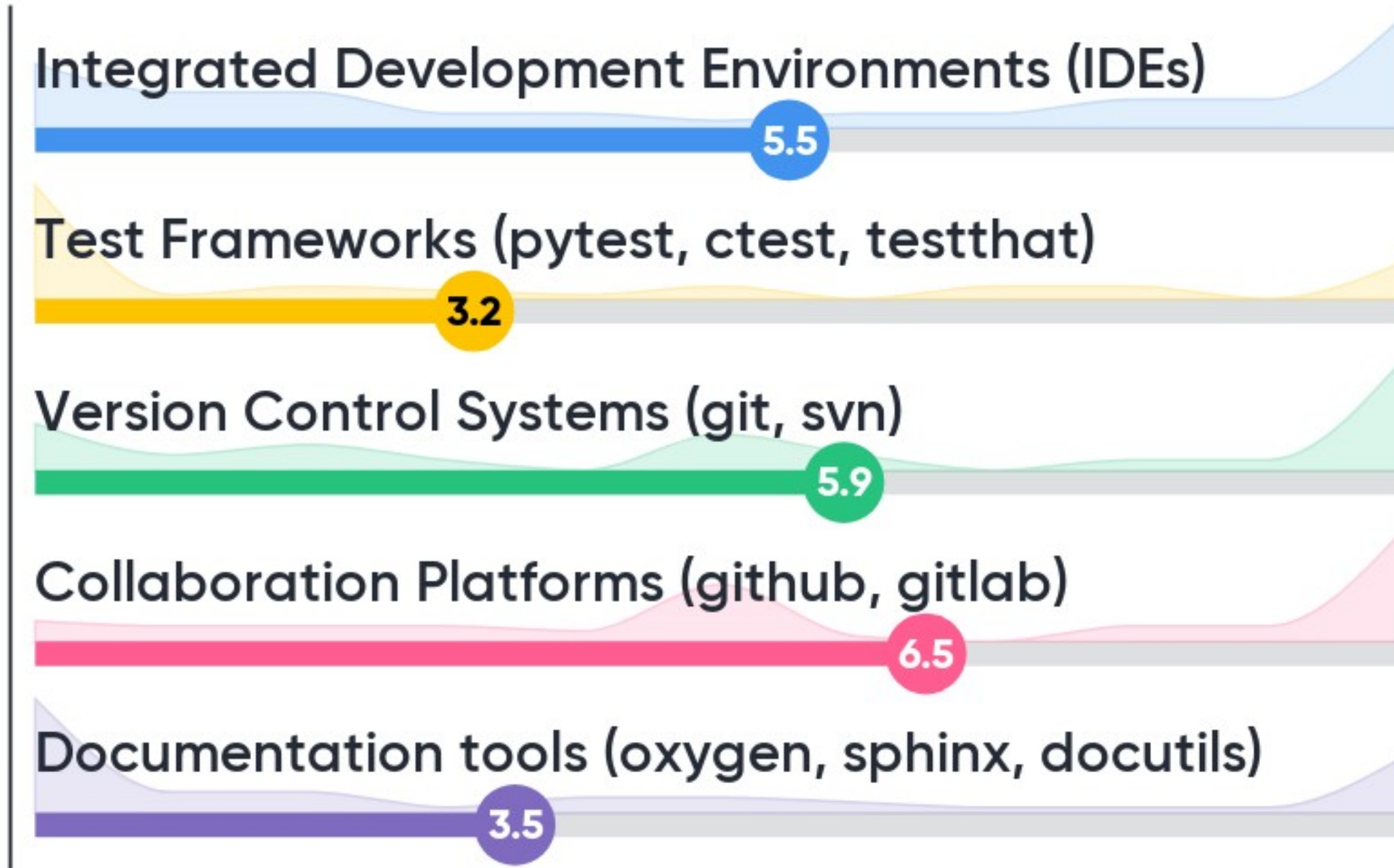
What python packages do you use the most (list as many or as few as you want!)?

55
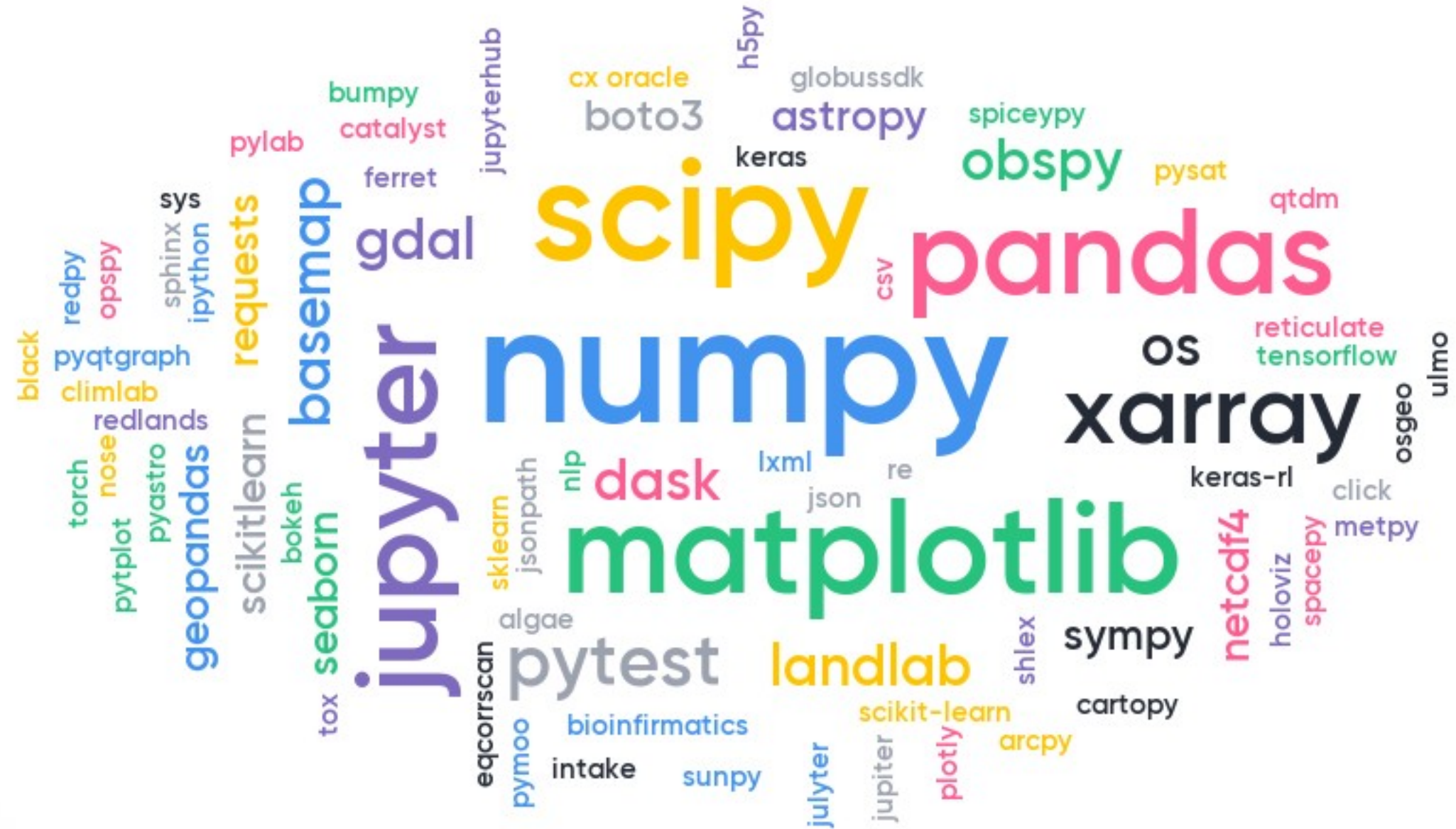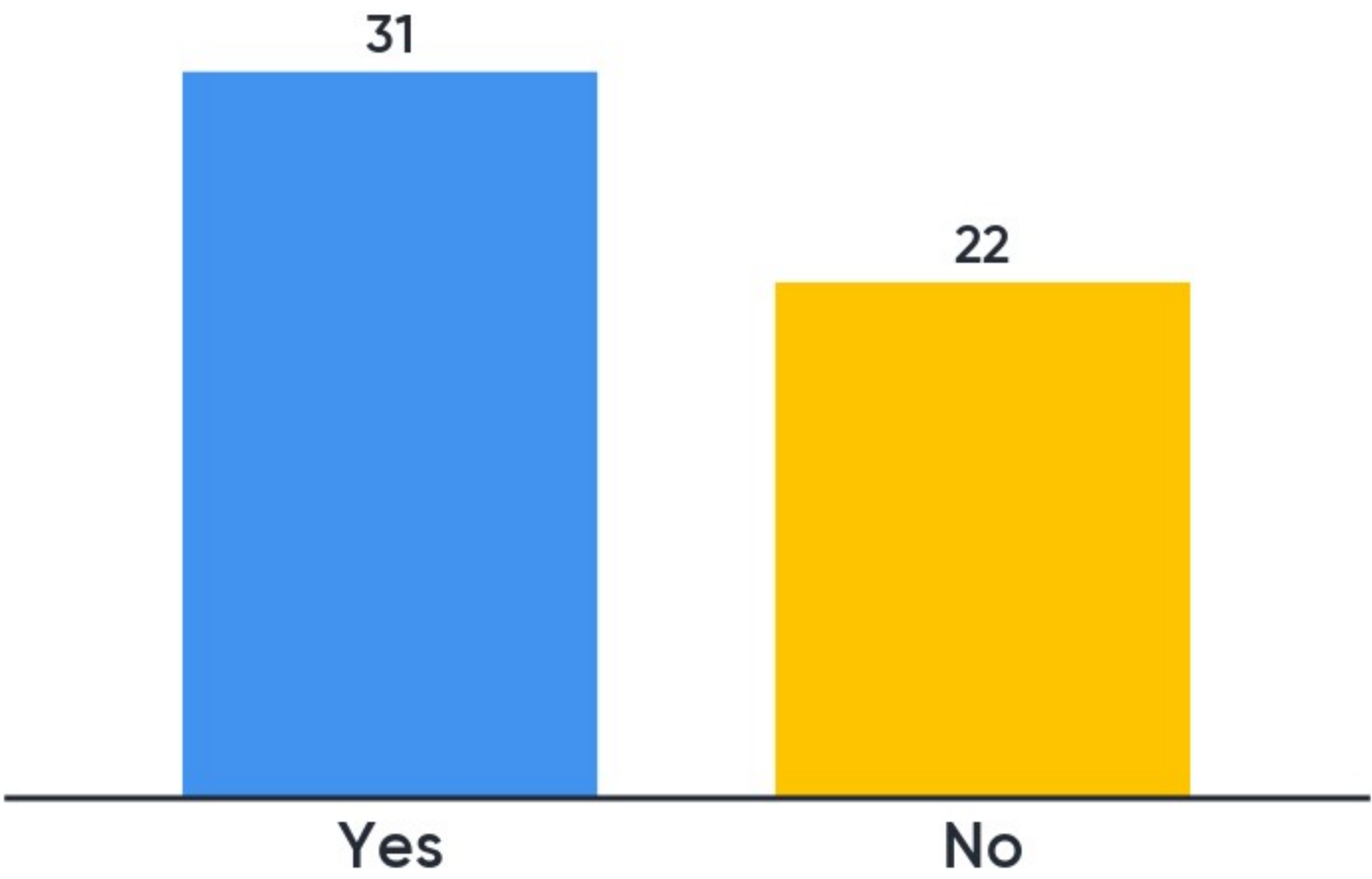
questions about pyopensci?

? 8 questions

6 upvotes

How could pyopensci help you with your science and code and software?

Scientific software review

What are best practices of sharing the code?

Bullet-proof, well-documented packages for Earth science.

Coordinate development activites to reduce parallel efforts and increase collaboration

All the docs on Python packaging. Blogs, tutorials, how tos, videos.

Streamline development of good quality software, socially responsible and easily shareable.

Training about developing tools

Quality assurance and interoperability

\

12

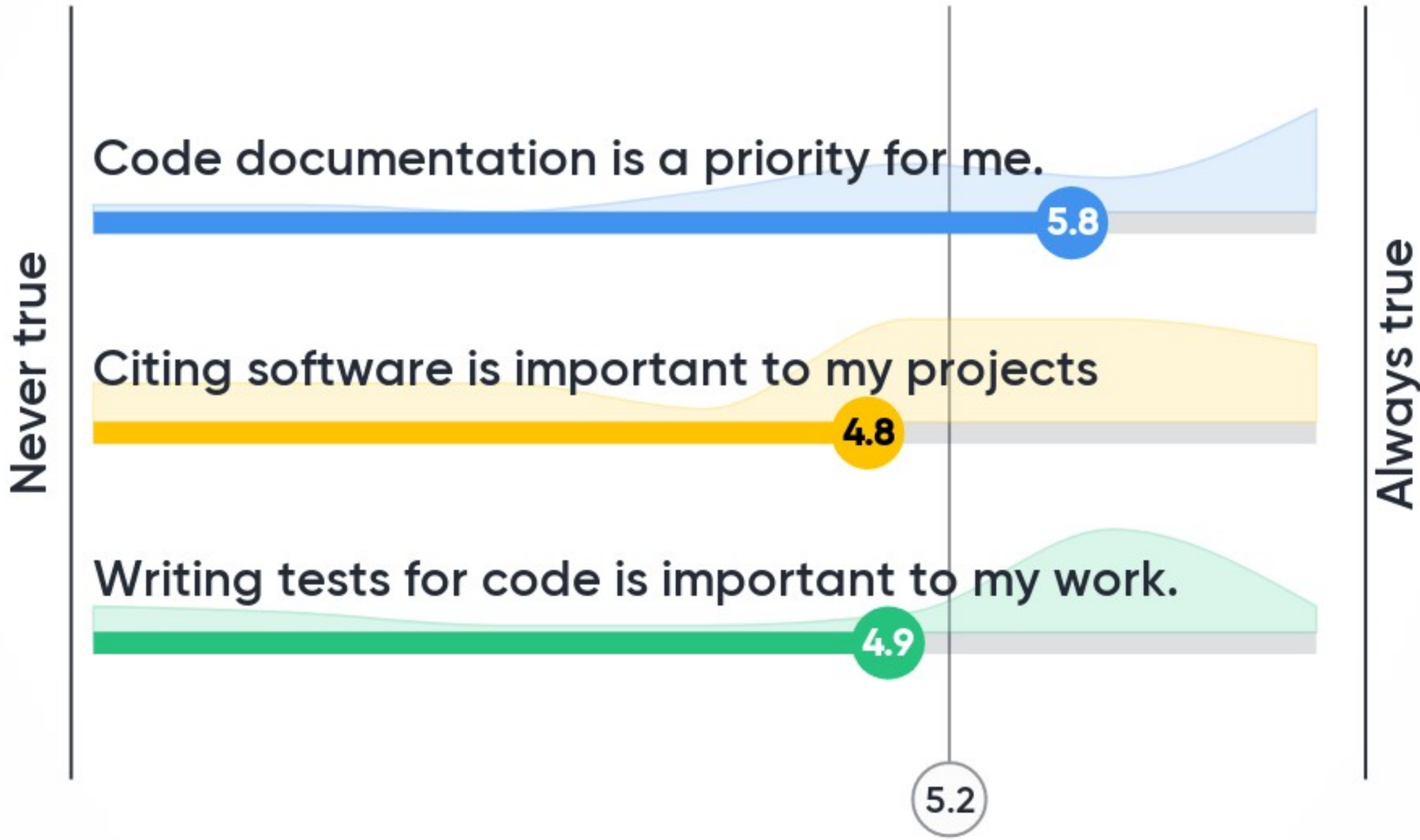# How could pyopensci help you with your science and code and software?

Templates for new packages – – for scientists who aren't well versed in things like testing, CI, docs
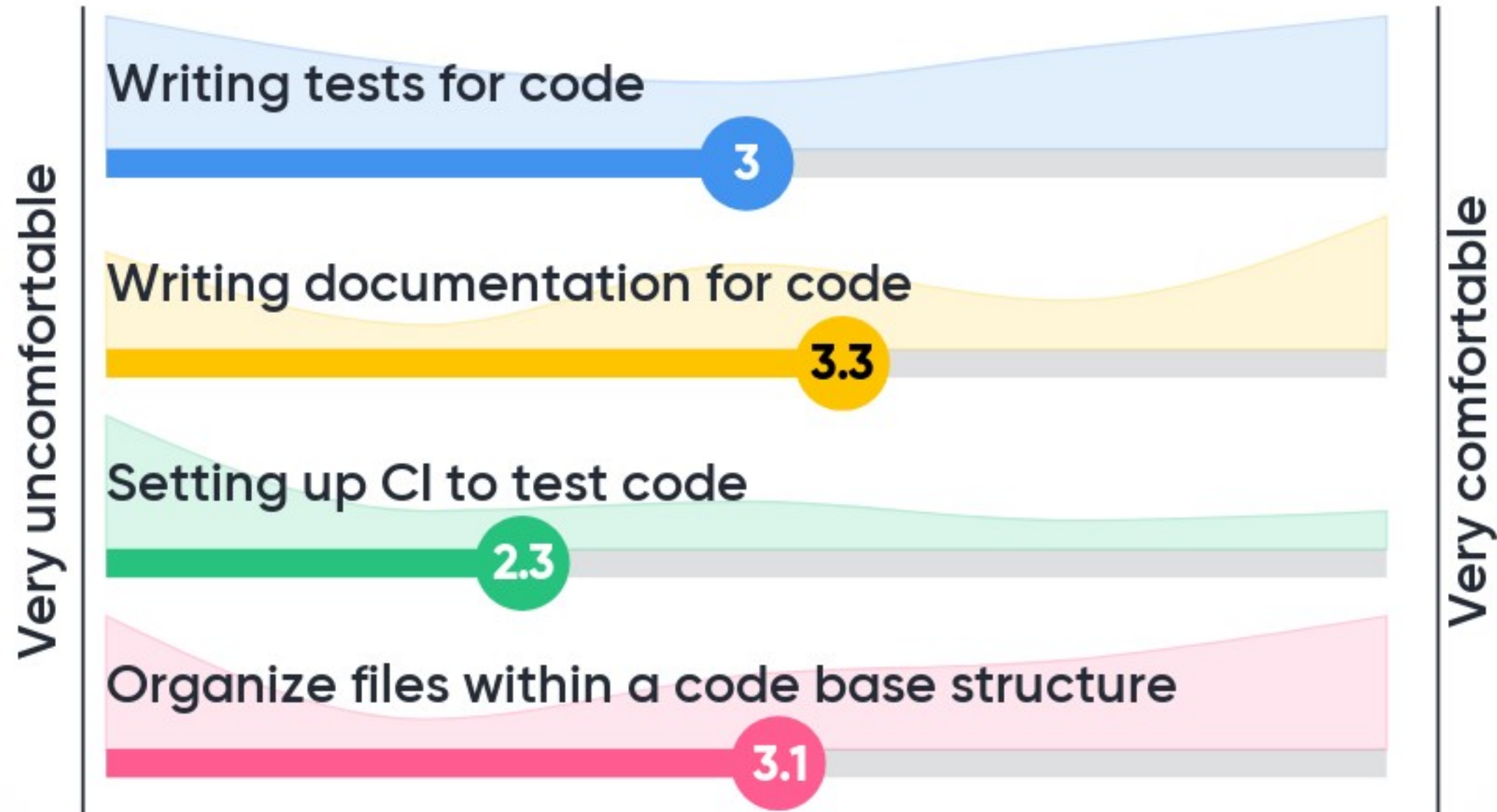
Best practices including PR forms and commit hooks

Training for people who can code for themselves, but want to start developing software for others. Style, documentation, testing, got, etc.

12

Rate the following

Rate your level of comfort for each of the items below.

Very uncomfortable / Very comfortable

Writing tests for code — 3

Writing documentation for code — 3.3

Setting up CI to test code — 2.3

Organize files within a code base structure — 3.1

33

List some of the software or code challenges that you have encountered.

Cache invalidation of intermediate products in large / complex workflows

Unit testing and CI in long-running / complex analyses

Large data sets and unit testing

Documentation

Peer-review

Dependency hell. Unnecessary duplication.

Dealing with backwards incompatibility in frameworks like python in maintaining software over decades

Country matching, accurate zonal extractions,

Which library to use for task x? Keep it simple and functional or go object-oriented?

23

List some of the software or code challenges that you have encountered.

Outdated dependencies

typing/exceptions -- lack of good documentation about return types and exceptions raised. Imagine a days-long run that aborts at the report generation process due to a mishandled type.

Rapidly evolving ecosystem -- things break due to upstream changes every time I look away.

Packaging for distribution especially on Mac.

Provide advocacy to funding agencies to pay for development of new package and maintaining existing ones!

Building science packages that run on Windows -- important for students!

Writing *useful* documentation is hard.

Documentation

Testing

23

List some of the software or code challenges that you have encountered.

Reducing redundancy in testing

Code optimization

Going from coding for myself to wanting to code for others has been a steep learning curve (I'm a scientist by training, not a software engineer). How do I structure my code? How do I deal with dependencies? What's git? What's a python environment?

Work-site specific Security Assessments requiring significant vulnerability review and confirmation of policy compliance.

code duplication

23

Describe your biggest challenge(s) with making your research reproducible?

Rapidly evolving software stack

dependencies on unreleased third party software (e.g. github repos but not in official repos like pypi, cran etc)

Large data objects, intermediate data objects required to execute projects

Documenting the workflow

Sharing big dataset

The time it takes to be systematic given the state of today's tooling for reproducibility. Tools like Whole Tale and Binder will help.

Getting others to do the same and share code

Making data FAIR.

Data validation of special input data structures.

Describe your biggest challenge(s) with making your research reproducible?

Multiple codes used at different stages of research

Releasing to pypi

Sharing proprietary data.

Questions

? 8 questions

6 upvotes