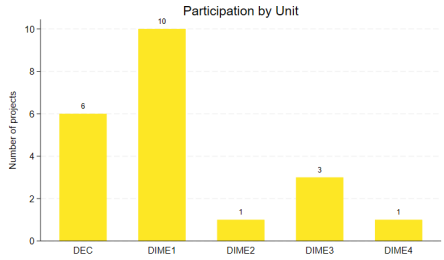
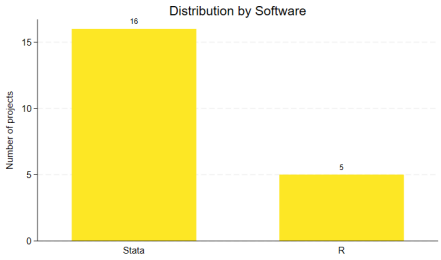
Peer Code Review Summary - FY24 Q3

DIME Analytics

A total of **21 research assistants** joined the peer code review held in the week of February 19, 2024, and reviewed code from **21 different projects**. DIME1 was the most-represented unit, followed by DEC. Most projects used Stata as the main coding language.



(a) Participant Units

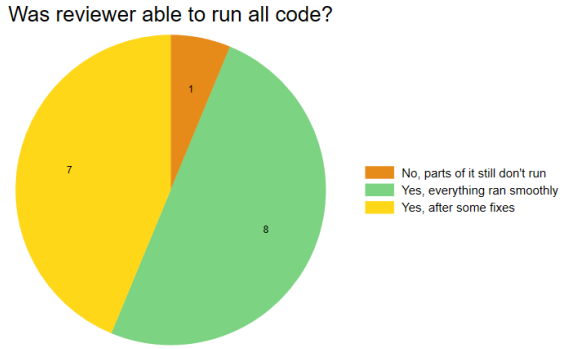


(b) Participant Software Used

Reproducibility

Out of the 21 code packages reviewed, **16 included de-identified data, and were evaluated for reproducibility.**

Encouragingly, 94% of these code packages were **reproducible**: the code file could be run by the reviewer with either minor fixes, or no changes at all. There was only one package which could not run even after attempted fixes due to missing dataset(s).

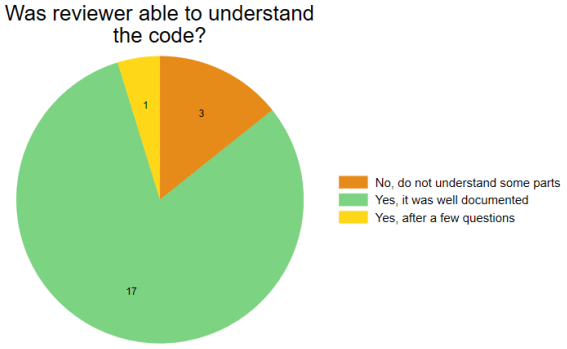


Ease of Use

81% reviewers said that the code they received was **easy to understand and well-documented**.

14% indicated that additional details in the GitHub README and more comments in the code would have been helpful.

In terms of **transferability**, 95% reviewers said they would be able to take over the project with either no communication at all with the original coder, or with just a few questions.



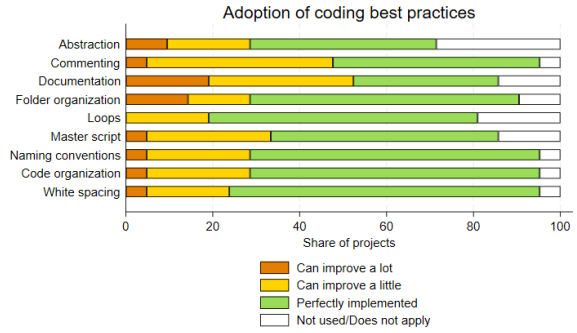
Moreover, 86% reviewers said that it would take them **2 days or less** to be able to understand the code well enough to make contributions to it. The code for 52% projects was rated **easy to maintain**. There were **10 projects** for which making adaptations would require changes in multiple places, making it hard to build on existing code.

Adoption of Coding Best Practices

The **mean number of best practices adopted was an excellent 7.9** - out of 9 in total. Further, 52% of projects **correctly implemented** each of the best practices.

The reviewers identified the **most room for improvement** in documentation, and use of helpful comments throughout the code.

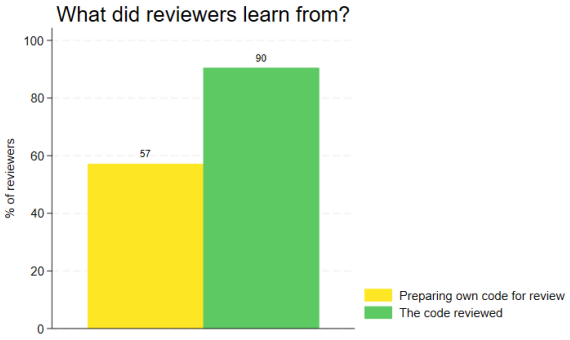
The **most widely adopted best practices**, include code organization, and use of white spaces.



Feedback and Challenges

All participants reported learning something new from the code review exercise. 7 reviewers said they learned from the code reviewed, 2 reviewers reported learning from preparing their own code for submission, and 12 reviewers reported learning from both. In addition, **participants also reported learning** the following:

- Use of GIS data and shape files
- Creating modular scripts
- Efficient code organization
- Use of new commands and functions



Finally, the **primary challenges** identified during the exercise include **constraints** in setting aside time to work on the code review, and **communication issues** arising from unavailability of some participants due to work-related deadlines. Participants acknowledged the clarity in instructions and organization. For future rounds of code review, we aim to facilitate more one-on-one interaction between participants, and ensure TTLs are aware of their RA's participation in the code review exercise.

Participant Comments

"I think the current DIME WIKI resources are well designed for this course, and the RRF course is pretty exhaustive, and serves as good training for this exercise."

"The instructions were clear and the checklists are helpful to make the review and to remind us of those resources."