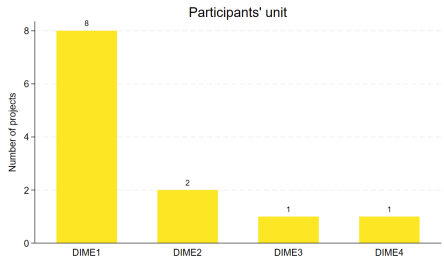
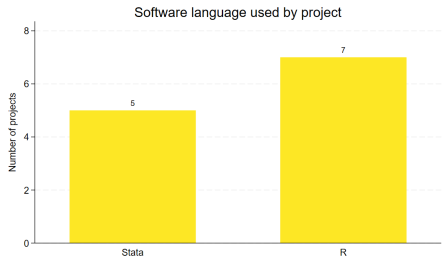
Peer Code Review Summary - FY24 Q2

DIME Analytics

A total of **13 research assistants** joined the peer code review held in the weeks of November 13th, 2023, and reviewed code from **12 different projects**. DIME1 was the most-represented unit, followed by DIME2. Most projects used R as the main coding language.



(a) Participant Units

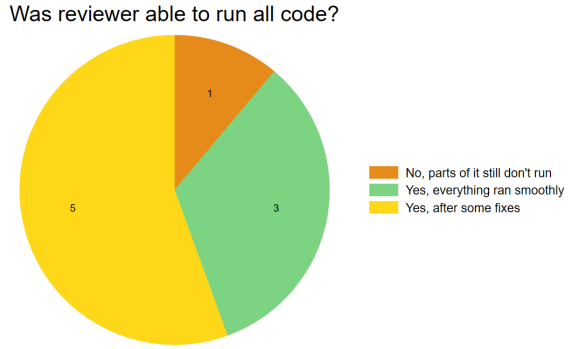


(b) Participant Software Used

Reproducibility

Out of the 12 code packages reviewed, **9 included de-identified data, and were evaluated for reproducibility.**

Encouragingly, 89% 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).

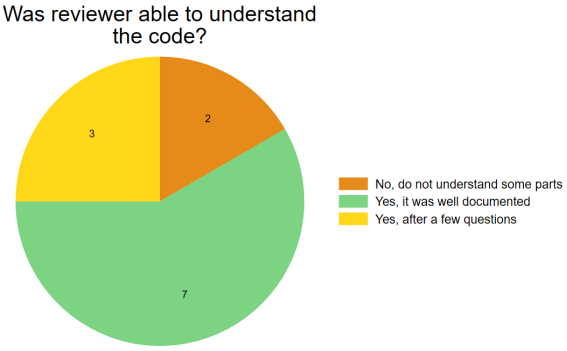


Ease of Use

83% reviewers said that the code they received was **easy to understand and well-documented**.

17% indicated that additional details in the GitHub README and a codebook for key variables would be helpful.

In terms of **transferability**, 100% 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, 83% 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 50% projects was rated **easy to maintain**. There

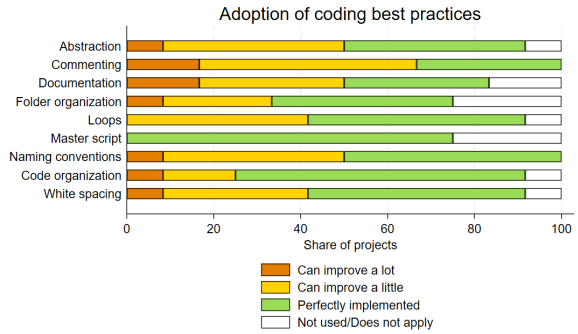
were **6 projects** for which making adaptations would require changes in multiple places, making it hard to build on existing code.

Adoption of Coding Best Practices

The **mean number of best practices adopted was a very healthy 5.91** - out of 9 in total. Further, 42% of projects **correctly implemented** each of the best practices.

The reviewers identified the **most room for improvement** in use of main scripts, folder organization, and use of helpful comments throughout the code.

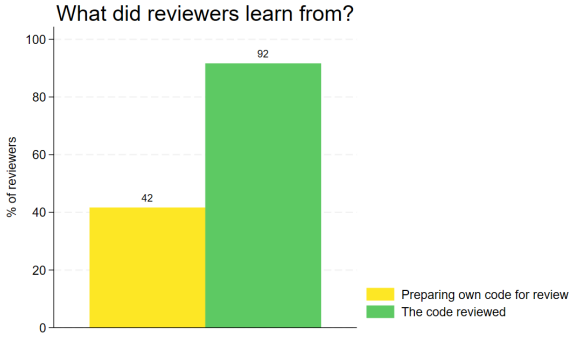
The **most widely adopted best practices**, include folder organization, and use of white spaces.



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 5 reviewers reported learning from both. In addition, **participants also reported learning** the following:

- Developing high-frequency checks (HFCs)
- Better systems of organizing code files
- New commands and functions
- Writing code that is reproducible



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 coordinating with participants in different time zones. Participants acknowledged the clarity in instructions and organization. For future rounds of code review, we aim to facilitate better communication between participants, and ensure TTLs are aware of their RA's participation in the code review exercise.

Participant Comments

"I learned how easy it was to repurpose a section of my code for reproducibility given the practices learned during the Reproducible Research Fundamentals course and DIME Wiki."

"Overall, I found the process well-organized and helpful. My reviewer couldn't be more helpful with their code and patient with mine."