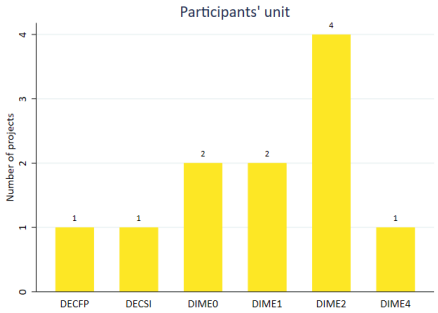
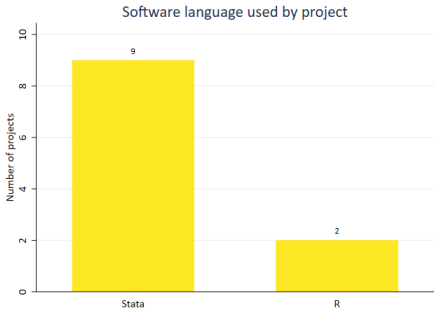
Peer Code Review Summary - FY24 Q1

DIME Analytics

A total of **11 research assistants** joined the peer code review held in the weeks of August 14th and August 21st, 2023, and reviewed code from **11 different projects**. DIME2 was the most-represented unit, followed by DIME0 and DIME1. Most projects used Stata as the main coding language.



(a) Participant Units



(b) Participant Software Used

Reproducibility

Out of the 11 code packages exchanged, **10 included de-identified data, and were evaluated for reproducibility.**

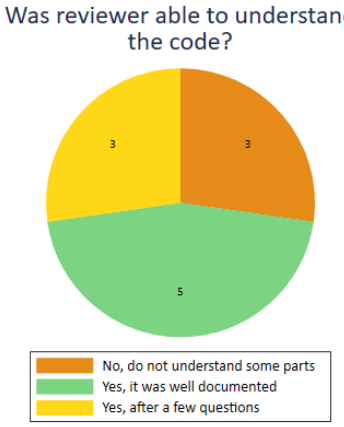
Encouragingly, 90% of these code packages were **reproducible**: the code file could be run by the reviewer with either minor fixes, or no changes at all. There was only one package which could not run even after attempted fixes due to unclear folder structure.

Ease of Use

45% reviewers said that the code they received was **easy to understand and well-documented**.

28% 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.



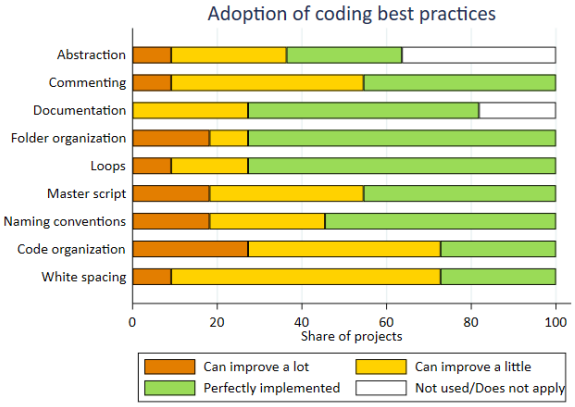
Moreover, 55% reviewers said that it would take them **2 days or less** to be able to understand the code well enough to make contributions to it. The code for 72% projects was rated **easy to maintain**. There were **only 3 projects** for which making adaptations would require changes in multiple places, making it hard to build on existing code.

Adoption of Coding Best Practices

The **mean number of best practices adopted was a very healthy 7.3** - out of 9 in total. Further, 55% of projects **correctly implemented** each of the best practices.

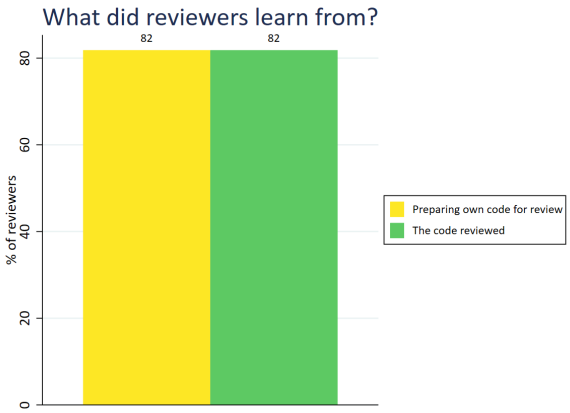
The reviewers identified the **most room for improvement** in code organization, documentation, and abstraction - that is, use of special commands to simplify repetitive tasks.

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



Feedback and Challenges

All participants reported learning something new from the code review exercise. 9 reviewers said they learned from the code reviewed, 9 reviewers reported learning from preparing their own code for submission, and 2 reviewers reported learning from both. In addition, **participants also reported learning** the following:



- New ways to present high-frequency checks (HFCs)
- Better systems of organizing code files
- DIME's coding best-practices
- Use of programs to simplify repetitive tasks

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 inability to run partner's shared code. Participants acknowledged the extended flexibility in submitting feedback based on previous rounds. 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

"The process, and its stages were well explained by the team; in addition, any questions/clarifications requested were quickly answered by DIME Analytics."

"As a new coder to the Bank I was unaware of DIME's standards and so having them now in my back pocket will help streamline and improve my work!"