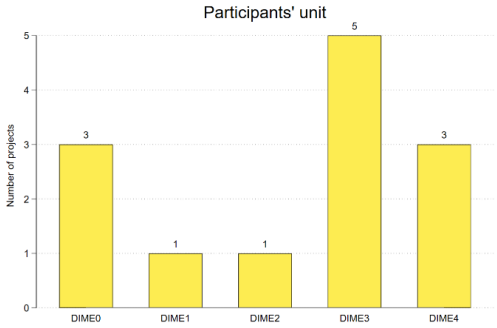
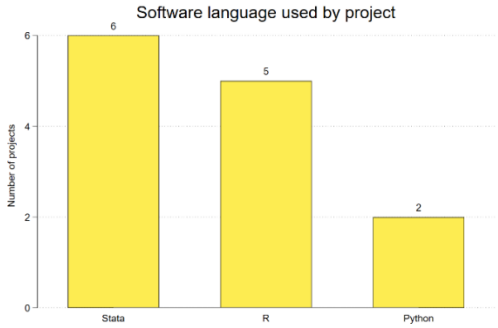
Peer Code Review Debrief - FY22 Q4

DIME Analytics

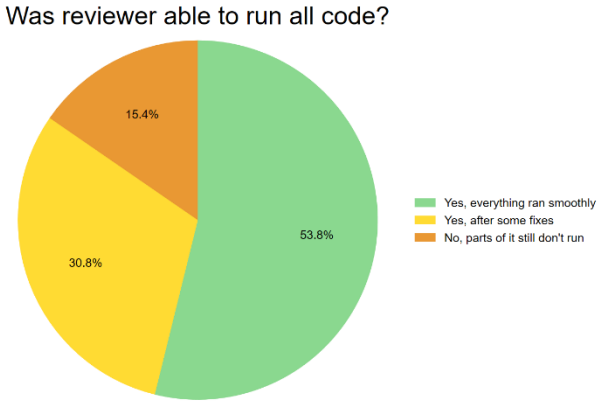
A total of **13 research assistants** joined the peer code review the week of July 25th, 2022, and reviewed code from **12 different projects**. Most of the participants are part of DIME3 and used Stata as the main coding language.





Reproducibility

Of the 12 code files, 8 were reproducible as submitted: the code file could be run by the reviewer without edits and produced identical outputs. 2 of the 12 files could not be run by reviewers even after attempted fixes. The reasons for this included incorrect file paths, variables that changed on every run, and use of incorrect commands.

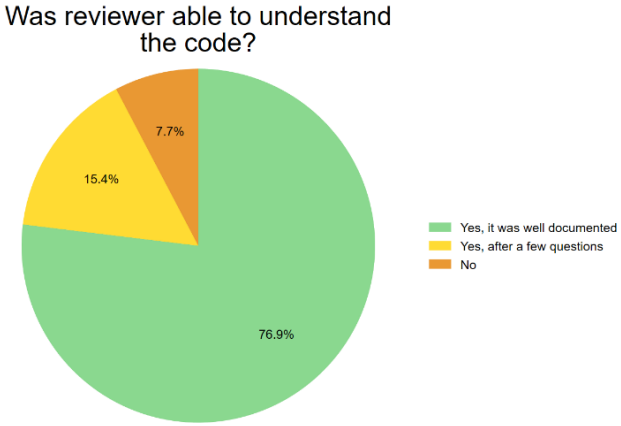


Ease of Use

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

2 reviewers indicated that more comments in the code and additional details in the GitHub README would be helpful.

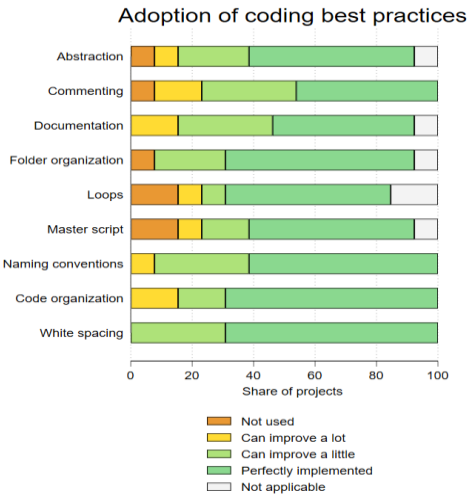
4 reviewers said the material provided would be enough to take over the project without any additional communication with the original coder.



9 reviewers said that it would take them no more than a day to be able to understand the code well enough to make contributions to it. For most projects, reviewers said the code is easy to maintain. There was only one project for which making adaptations would require changes in multiple places, making it hard to build on it.

Adoption of best practices

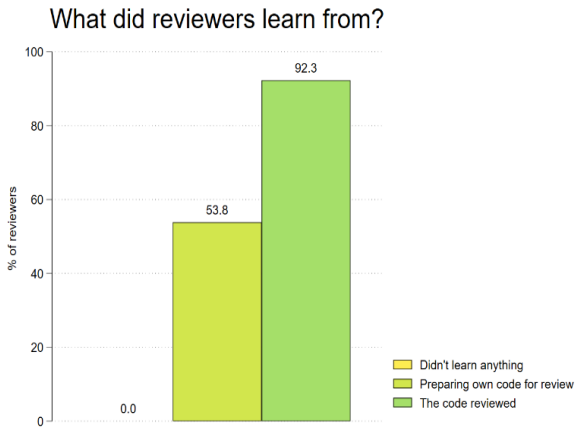
All reviewers pointed to at least one coding practice that could be improved, but all projects adopted at least 1 recommended best practice. There is most room for improvement in **use of comments, documentation, and code organization**. The most widely adopted best practices include clear naming conventions and use of white space for readability.



Feedback and Challenges

All participants reported learning something new from the code review exercise. 12 reviewers learned from the code reviewed, and **7 reviewers** reported learning from preparing their own code for submission. Participants reported learning how to:

- Organize code,
- Write reproducible code,
- Use of new command options, and
- How to write better master scripts and README files.



The **primary challenges** identified during the code review are **time constraints** and **coordination issues between the pairs**. Participants also asked for checklists to review tasks that are currently not available in the templates – for instance, ensuring replicability of code, and suggested spreading out the code review process to two weeks instead of one. Based on this feedback, we aim to plan the next round over two weeks, including a detailed debrief session with RAs to discuss commonly faced issues – and potential fixes.