DIME Analytics
Peer Code Review Guidance Note

Overview

DIME Analytics organizes the **Peer Code Review** as a real-time code and data-quality assurance process, providing a structured opportunity for participants to exchange, run, and provide feedback on each other's code. Beyond improving reproducibility, transparency, and adherence to best practices, it is also a valuable learning experience—code writers gain fresh perspectives on their work, while reviewers are exposed to different coding styles and practices.

The Peer Code Review is especially beneficial for **new Research Assistants**, as it helps them understand expectations and learn from others' code.

The review is designed for scripts that are modular enough to be understood on their own (i.e., a reviewer can follow the script without referring to other project files) and short enough to be reviewed **within one day** (depending on complexity). If participants wish to submit longer scripts, they will be paired with others submitting similarly long code to ensure a balanced review process. We recommend submitting recently completed tasks to allow for timely corrections and improvements.

The time commitment for each round is **8-10 hours**, including two dedicated sessions that participants are expected to attend.

How It Works

1. Teams sign up to participate in the code review.
    - All DIME Research Assistants working on code are expected to participate regularly.
    - Any Bank staff or consultant seeking feedback on their code is welcome to join.
2. Participants are paired for the review based on their preferred statistical software and the length of their code.
3. All participants prepare a code package to share with their assigned code review partners, following this checklist.

- Submitted code should ideally be **around 1,000 lines**, excluding the main scripts (Stata, R).
- This is a recommendation rather than a strict limit—longer submissions are allowed but may require more time to review.
- Teams should specify which aspects of their code they would like reviewed, such as: Data cleaning, Indicator construction, Analysis, Sampling and randomization, High-frequency checks, etc.
- Teams should also indicate whether they want the reviewer to assess **computational reproducibility**.
- To enable reproducibility checks, the submission package **must** include a **de-identified dataset**.

4. The **peer code review** takes place over the course of one week and includes:

- A **hybrid kickoff session** outlining the process and assigning code review partners. Attendance is **optional** for those who have participated before.
- A **hybrid group review session** with technical support from DIME Analytics.
  - Pairs work together, reviewing each other's code and addressing questions.
  - Attendance is **mandatory**—if the scheduled session is not feasible, Analytics will facilitate an alternative meeting.
- This activity is most effective when teams can collaborate **in person**, allowing for real-time discussions and clarifications.

5. Reviewers assess the code using the following resources:

- **General:** Main Feedback Checklist
- **Task-Specific Checklists:**
  - Cleaning Code Review Checklist
  - Indicator Construction Checklist
  - Analysis Code Review Checklist
  - Sampling & Randomization Checklist
  - High-Frequency Checks Checklist

6. Reviewers have one week to submit feedback using this form and upload their completed task-specific checklists.

What Teams Receive

- **Within 10 days**, TTLs and research assistants receive a standardized report indicating adherence to best practices and areas for improvement.

- Research assistants also receive detailed feedback through the task-specific checklists submitted by the reviewers.
- All participating TTLs and DIME Managers receive a summary report highlighting common strengths and weaknesses.