

DIME Analytics Code Review Package Checklist

Overview

The submission package for peer code review should include all components in the checklist below. Code review participants can provide either a link to the complete package (**OneDrive** or **GitHub**), or a zipped folder - along with information on programming language(s) involved, and approximate run time. Questions should be directed to [DIME Analytics](#).

Package Requirements

1. Code:

- All scripts to be reviewed.

- A **main script** that allows the reviewer to run all code after changing only the top-level directory.

2. Data:

- All de-identified data necessary to run the code.

3. README:

- A [README file](#) containing:

 - Software and version used.

 - Approximate run time of the code.

 - Purpose of each folder.

 - High-level objective of the code (e.g., regression analysis, data visualization).

 - Line(s) in the **main script** that should be changed to run the code on a different machine.

 - Purpose of each script (if multiple scripts are used).

 - Outputs produced (if part of the review).

 - Mapping of outputs to scripts responsible for generating them.

 - Any specific requests or guidance for the reviewer.

README file path and name:

Component-Specific Requirements

Before submitting your package, ensure the following requirements are met.

1. Main Script Requirements

The main script should be structured to ensure smooth execution. See templates for [Stata](#) and [R](#) main scripts.

The main script runs from start to end after changing only directory paths in one location. This should be clearly specified in the README.

The main script installs any required packages or provides instructions to do so (e.g., SSC for Stata, CRAN for R). Refer to this session on [Building Reproducible Environments](#) for more information.

For Stata, the main do-file sets critical configurations such as version, matsize, and varabbrev, either directly or via a wrapper command (e.g., ieboilstart from ietoolkit).

For R, all necessary packages are loaded within the main script.

Approximate run time of the main script:

2. Data Requirements

The data can be shared for the review.

The data is de-identified.

All necessary datasets are included in the package.

3. Output Requirements

[Refer to this session on Creating Reproducible Outputs for details.](#)

All outputs generated by the code are included in the package.

Outputs are saved and stored in the appropriate folders (e.g., final outputs in **Outputs/Final**).