

DIME Analytics

Code Review Package Checklist

v1.0

Required Content and Practices

The code review package must be shared as a .zip file, GitHub link or shared folder containing the items listed below. All file paths in the code **must** use only forward slashes, i.e. C:/users/...

1. Package components

All scripts to be reviewed.

All data necessary to run the code to be reviewed - in deidentified form.

A README. file that contains all the information about the project and the tasks being performed in the code that may be necessary for the reviewer to understand the code and are not included in the scripts.

README. file path and name:

A **master script** that runs all the necessary code and allows the reviewer to adjust file paths only once.

2. Master script requirements

Master script clearly indicates which line needs to be changed to run in a different computer

The master script runs from start to end if the reviewer changes only directory path(s) in one location of the master script

The master script installs any commands required (e.g. from SSC in Stata, CRAN in R, etc.)

In Stata, the master do file sets critical configurations such as version, matsize, and varabbrev (either sets directly or through a wrapper command such as ieboilstart from ietoolkit)

In R, all necessary packages to run the code are loaded in the master script

Approximate time it takes to run all code from the the master script:

2. Data

The package includes:

Fully de-identified data. **Any exception where sensitive data is strictly needed to conduct the review will require additional ethics and security measures, and will be handled on a case-by-case basis.**

All data necessary to run the code to be reviewed is included

3. Outputs

All outputs created by the code to be review are included

All raw outputs to be reviewed are created in the same folder structure as those originally submitted.

It is easy to identify which code creates which output

4. Recommended Practices

Scripts for data cleaning, variable creation, and analysis are separate; and analysis scripts do not include any data processing, unless necessary for the creation of a table or graphic

Variable construction scripts include detailed comments about each variable created

Analysis scripts do not depend on having the results of other scripts in memory, except for the Master script

All scripts are well-commented and formatted, such that one can easily identify functional chunks of code and evaluate whether they correctly implement the econometric or statistical process described

Tables are output as .csv or .tex files or other raw text files when possible