DIME Analytics
Peer Code Review - Construction Checklist
v1.0

Reviewer:

Coder:

NOTE: Make sure that fill out this checklist **ONLY IF** your partner's submission includes **construction** tasks.

Indicator construction tasks

This checklist lists important factors to consider while reviewing your code review partner's **construction** scripts/code. Please fill this checklist, and submit it as an attachment when you submit this detailed form.

Construction script

I understand how and why each variable is constructed. In some cases, it was not clear - need more comments in the following lines of code:

Check merges:

If any observations are dropped, there is a clear justification in the code

If the number of observations increase/decrease, there is a clear justification in the code

Check collapses, reshapes, and group-wise calculations (e.g. `bysort` and `egen`):

The code clearly explains how missing values are treated in these operations

If the sort order of data matters, stable sorts are used (e.g. using **`set seed`**)

Check winsorization and other techniques for handling outliers

The reason for the chosen technique/method is clearly explained

There is documentation explaining how parameters such as cutoff percentiles were chosen and why one or both tails were altered **(if applicable)**

Check creation of new variables:

The code matches the variable definition in the documentation (such as a codebook of all variables)?

The correct function is used and clearly explained

Constructed dataset

The constructed dataset is tidy (that is, **each row is an observation**, AND **each variable is a column**

Variable names are informative

Variable labels are informative

Value labels are informative

All labels are grammatically correct and free of special characters

There is clear documentation (variable dictionary, variable labels, value labels, notes, comments) about variable definition? (e.g. codebook constructed using `iecodebook`)