

# 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

Do you understand how and why each variable is constructed? If not, indicate to the author where more comments are needed:

#### Check merges:

Are any observations dropped?

If so, is there a clear justification for that?

Does the number of observations increase?

If so, is there a clear justification for that?

If any observations didn't match, is that explained in the comments?

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

How are missing values treated in these operations?

Does the sort order of the data matter?

Or, is the result uniquely determined?

#### Check winsorization and other techniques for handling outliers

Is the reason for the chosen technique/method explained?

Is there documentation explaining how parameters such as cutoff percentiles were chosen and why one or both tails were altered?

**Check creation of new variables:**

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

Is the correct function being used?

**Constructed dataset**

Is the constructed dataset tidy (that is, **each row is an observation**, AND **each variable is a column**)?

Are variable names informative?

Are variable labels informative?

Are value labels informative?

Are all labels grammatically correct and free of special characters?

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