
Subject: Appending Multi-phase Nigerian DHS Surveys

Posted by [Oby](#) on Fri, 30 Aug 2024 07:41:15 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hello,

I'm working with multiple NDHS datasets from different survey years (2003, 2008, 2013, and 2018) and have encountered some challenges. I'd appreciate your guidance on the following issues:

1. Inconsistent Value Labels Across Years:

When attempting to append these datasets using `bind_rows()`, I receive warnings about conflicting value labels for certain variables.

Question: Should I convert these labelled variables to factors using a function like `as_factor()` before appending, or is there a better approach to standardize value labels across these different datasets? What is the best practice for ensuring that the labels are consistent before appending?

2. Creating a Survey Design with Different Sampling Designs:

Each of the four surveys has a different sampling design. After appending the datasets, I need to create a combined survey design object for analysis.

Questions: a.) How should I go about creating a survey design object that appropriately accounts for the different sampling designs across the four surveys?

b.) Are there other specific adjustments or considerations I need to make when combining these datasets for analysis?

Thank you for your assistance! I look forward to your advice on these issues.

Regards,
Oby
