
Subject: Issue with importing Individual Recode datasets for India for the years 2015-2016 on R

Posted by [Navya](#) on Mon, 02 May 2022 15:55:39 GMT

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

Hello,

I am currently doing a study on the effect of the Protection of Women against Domestic Violence Act (2005) on the domestic violence rate and GDP in India. For this, I need to use Individual Recode (IR) datasets for the years 2005-2006 and 2015-2016 as the base datasets and merge them with Couples Recode (CR) and Household Members Recode (PR) datasets. I have managed to do that for the year 2005-2006, however, for the years 2015-2016, I am not able to import the Individual Recode file (filename is IAIR74DT.dta) on R. When I use RDHS (the API built for R to access DHS datasets) I get the error "Error in unzip(tf, list = TRUE): zip file '/var/folders/kk/510ybpq53gb4br9_wtqyvy8r0000gn/T//RtmpRzo8A q/file436a330c3b1b' cannot be opened" (image is included in this message) and when I try to manually download the dataset I get a system timeout error. I was wondering how I should proceed with this issue.

Thank you for your help!

File Attachments

1) [Screenshot 2022-05-02 at 17.52.31.png](#), downloaded 1055 times

Subject: Re: Issue with importing Individual Recode datasets for India for the years 2015-2016 on R

Posted by [Janet-DHS](#) on Tue, 03 May 2022 14:09:30 GMT

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

Following is response from DHS Research & Data Analysis Director, Tom Pullum:

Are you sure that you need to do this merge? I have checked the CR files for these two surveys and see that they include the domestic violence (d*) variables. The CR files always include almost all of the information about the household and almost all of the information in the IR and MR files. Are you sure that the variables you need from the IR and PR files are not already in the CR file? If so, what variables are missing? Please let me know and perhaps I can help.

I hope you are aware that any merges with the India files must include region (v024 or mv024 or hv024) in the merge command.

As for difficulties unzipping the ASCII version of the IR file, unzipping the file shouldn't have anything to do with it being a DHS file. Your software may not realize that a .DAT file is equivalent to a .TXT file.

Subject: Re: Issue with importing Individual Recode datasets for India for the years 2015-2016 on R

Posted by [Navya](#) on Tue, 03 May 2022 21:34:11 GMT

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

Hi,

Thank you for your quick reply. I am sure I need to do the merge since although both CR and IR contain the domestic violence variables, I need the data for currently unmarried women from the IR file, who make up around 30% of the total women's sample in the 2005-2006 survey. Therefore, using IR as the base file rather than CR is necessary to ensure I don't lose important data.

I merged the datasets for 2005-2006 by v001, v002, v003 and v024 (which took into account the primary sampling unit, household number, respondent's/ individual's line number and state).

The issue only arises for the BR, HR and IR files for the years 2015-2016 (I only need the IR file), and gives the same error when I try to import the dataset as a .dta file rather than a .dat file. I was wondering if there's any way I can import the IR file to R using another method?

Thank you for your help!

Subject: Re: Issue with importing Individual Recode datasets for India for the years 2015-2016 on R

Posted by [Janet-DHS](#) on Thu, 05 May 2022 18:59:37 GMT

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

Following is a response from DHS Senior Sampling Specialist, Mahmoud Elkasabi:

We can only suggest that you try to download the Stata .dta files from the website and then import them into R. The files are large, and this may take time, but we don't know of another method.

Subject: Re: Issue with importing Individual Recode datasets for India for the years 2015-2016 on R

Posted by [Nidigot](#) on Fri, 09 Sep 2022 17:15:33 GMT

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

This would be the ugliest work around, but download the file from DHS and then store it in RDHS' cache:

```
data <- read_dta('/Users/Nidi/Desktop/filename.DTA')
saveRDS(data, "~/Library/Caches/rdhs/datasets/filename.rds")
```

and then call the file from there.

If you necessarily would like to have it in the "downloads list" then store it (depending on where it

should have been, but here I choose "1":

```
downloads[[1]] <- "~/Library/Caches/rdhs/datasets/filename.rds"
```

like I said: it is ugly, but it does work for your purpose
