

Hello,

Thank you for reaching out. At the DHS Program we are beginning to prepare more resources for users to learn DHS data analysis in R but these are not completed yet and please keep an eye on our social media accounts for updates on this. Please also check our code share library on GitHub which provides code for constructing DHS indicators in R:
<https://github.com/DHSProgram/DHS-Indicators-R>

In the meantime I can answer your questions.

1. To view labels in R you just can just use the command `print_labels` from the `haven` library. See below an example:

```
library(haven)
print_labels(NigeriaIR$b4)
```

2. The merge code you have is correct but there is no reason to merge the IR and the KR file. The IR file is the women's file and the unit of analysis is the woman. The KR file is the file for children under five and the unit of analysis is the child. This file also contains all the information about each child's mother, that is why there is no need to merge these files. The reason why you see v001.x etc is because there can be more than one child for each mother. You can learn about the different data files here: <https://www.youtube.com/watch?v=fzLNQkkvDel&t=105s>

Let's say instead you want to merge the HR file (household file where the unit of analysis is the household) with the IR file (woman's file). Perhaps there is some information that is only in the household file that you need for your analysis of women in the IR file. Then you would do the following:

```
IRdata <- merge(IRdata,HRdata,by=c("v001", "v002"))
```

Here you don't need the v003 since there can be several woman in the same household. So this is a many to one merge.

Another example:

If you want to merge the PR file (person's file) with the MR file (men's file). This is a one to one merge. You would do the following:

```
MRdata <- merge(MRdata,PRdata,by=c("mv001", "mv002", "mv003"))
```

3. the survey design code is also correct except for using `.x` for the variables. This was a result of your incorrect merge of the IR and KR files. For instance in an IR or KR file you would do the following.

```
# creating the sampling weight variable.
```

```
IRdata$wt <- IRdata$v005/1000000
```

```
mysurvey<-svydesign(id=IRdata$v021, data=IRdata, strata=IRdata$v022, weight=IRdata$wt,  
nest=T)  
options(survey.lonely.psu="adjust")
```

Hope this helps.

Best,
Shireen Assaf
The DHS Program