

---

Subject: Merging DHS Pakistan 2017-2018  
Posted by [shabina1129](#) on Thu, 14 Nov 2024 22:08:07 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Hi,

My current research is focusing on mother-in-law's educational attainment and how that can alter their daughter-in-law's autonomy within the household. I have 2 questions.

1) I am trying to merge PR and IR files together. This is what I have done so far:  
use "/Users/shabina/Desktop/Fall 2024 /PK\_2017-18\_DHS\_08072024\_1755\_200211/Household Member/household 1.dta"

```
rename hv001 v001
```

```
rename hv002 v002
```

```
rename hvidx v003
```

```
sort v001 v002 v003
```

```
save "/Users/shabina/Desktop/Fall 2024  
/PK_2017-18_DHS_08072024_1755_200211/Household Member/household 1.dta", replace
```

```
*open womens data
```

```
save " /Users/shabina/Desktop/kristinbietsch-MIL-Analysis-c4d7973/w omen1.dta "
```

```
use " /Users/shabina/Desktop/kristinbietsch-MIL-Analysis-c4d7973/w omen1.dta "
```

```
sort v001 v002 v003
```

```
keep if _merge==3
```

\*Attempted to merge the household file onto the womens only file.

/\*

Result	Number of obs
Not matched	85,801
from master	0 (_merge==1)
from using	85,801 (_merge==2)
Matched	15,068 (_merge==3)

This resulted in only 15,068 matches which the original womens only file contains, Leaving out around 30,000 of the sample. Is there another way to merge the datasets while keeping the sample I have which is 51,044?

I keep seeing that this is the only way to merge PR and IR but I wanted to know if there is another way to merge these datasets while still keeping the number of obs.

2) My second question is that if I were to try to run a regression on mother-in-law's education on standardized

autonomy there would either be no cases if I only run the regression on daughter-in-laws, or the regression would run solely on mother-in-laws since they are the only rows that have values for both the educational level obtained and standardized autonomy scale.

- Is there a way that we can construct the mother-in-law education variable to be available or on the same row as the daughter-in-law respondents so that regression analyses can be executed? If you know the codes for this. It would be so appreciated. Thanks.

---