
Subject: Re: Extracting data from HH
Posted by [Janet-DHS](#) on Mon, 13 Jun 2022 15:06:01 GMT
[View Forum Message](#) <> [Reply to Message](#)

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

You are correct that in this survey, husband's education (v701) is missing for most (about 80%) of the women who are currently married. There is no good reason for that. Your strategy is to match women with their husbands using the relation to head code (hv101) in the HR file. You will indeed capture the great majority of women and husbands from the head / spouse pairs but not all of them. The following Stata code, using the PR file (rather than the AR file) will produce a file of women who can be merged with the IR file. Note that in the India surveys, state (hv024 or v024) is part of the unique identifier.

```
use "...IAPR7AFL.DTA" , clear
```

```
keep hv001 hv002 hv024 hvidx hv101 hv104 hv108
```

```
keep if hv101==1 | hv101==2
```

```
gen hv101_1=0
```

```
gen hv101_2=0
```

```
gen head=1      if hv101==1
```

```
gen head_sex=hv104 if head==1
```

```
gen head_ed=hv108  if head==1
```

```
gen head_line=hvidx if head==1
```

```
gen spouse=1    if hv101==2
```

```
gen spouse_sex=hv104 if spouse==1
```

```
gen spouse_ed=hv108  if spouse==1
```

```
gen spouse_line=hvidx if spouse==1
```

```
collapse (min) head* spouse*, by(hv001 hv002 hv024)
```

```
drop if spouse==.
```

```
rename hv001 v001
```

```
rename hv002 v002
```

```
rename hv024 v024
```

* combinations in which the woman is spouse and the man is head

```
gen v003=spouse_line if spouse_sex==2
```

```
gen woman_ed=spouse_ed if spouse_sex==2
```

```
gen husband_ed=head_ed if spouse_sex==2
```

* combinations in which the woman is head and the man is spouse

```
replace v003=head_line if head_sex==2
```

```
replace woman_ed=head_ed if head_sex==2  
replace husband_ed=spouse_ed if head_sex==2
```

```
drop head* spouse*
```

```
* prepare the file for a merge with the IR file  
order v024 v001 v002 v003 woman_ed husband_ed
```

```
list if _n<=50, table clean  
sort v024 v001 v002 v003
```

```
* save and then merge with the sorted IR file
```
