
Subject: merging KR and PR - Multi-country and multi years
Posted by [ekowababio](#) on Wed, 31 May 2023 21:41:12 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi,

I am working with multi-country and multiple years data. I have seen discussions or solutions here (including <https://userforum.dhsprogram.com/index.php?t=msg&th=7086&start=0&>) regarding how to merge the PR and KR datasets. The problem is I keep getting error message when try to merge the two datasets indicating the observations are not uniquely identified in the KR data (the master). I tried

```
use "KR_Data.dta", clear
keep if b16>0 & b16<.
**(346,087 observations deleted) missing child line number in household
```

```
gen in_KR=1
gen sex_child=b4
gen age_child=hw1
gen line_number_of_child=b16
gen line_number_of_mother=v003
gen hv001=v001
gen hv002=v002
gen hv003=v003
gen hv024=v024
gen hvidx=b16
```

```
merge 1:1 hv024 hv001 hv002 sex_child age_child line_number_of_mother line_number_of_child
using "PR_Data.dta"
```

I even tried to create a unique id to see if there are duplicates in the KR using the code below

```
egen CN_BIRTHID = concat(State hv024 hv003 hv001 hv002 hvidx sex_child age_child
line_number_of_mother line_number_of_child), punct(-)
isid CN_BIRTHID
sort CN_BIRTHID
by CN_BIRTHID:generate nobs=_N
list State CN_BIRTHID if nobs>1, sepby(CN_BIRTHID)
```

```
isid CN_BIRTHID - return
variable CN_BIRTHID does not uniquely identify the observations
```

and the list returns

```
State CN_BIRTHID
```

```
346700. ML4 ML4-1-2-14-1-4-2-50-2-4
346701. ML4 ML4-1-2-14-1-4-2-50-2-4
```

346739. ML4 ML4-1-2-15-2-4-2-7-2-4
346740. ML4 ML4-1-2-15-2-4-2-7-2-4

346754. ML4 ML4-1-2-16-1-6-2-36-2-6
346755. ML4 ML4-1-2-16-1-6-2-36-2-6

346822. ML4 ML4-1-2-20-1-3-1-41-2-3
346823. ML4 ML4-1-2-20-1-3-1-41-2-3

346908. ML4 ML4-1-2-25-1-5-1-53-2-5
346909. ML4 ML4-1-2-25-1-5-1-53-2-5

347131. ML4 ML4-1-2-36-2-6-1-55-2-6
347132. ML4 ML4-1-2-36-2-6-1-55-2-6

347327. ML4 ML4-1-2-45-1-6-2-13-2-6
347328. ML4 ML4-1-2-45-1-6-2-13-2-6

348214. ML4 ML4-2-2-100-3-3-1-14-2-3
348215. ML4 ML4-2-2-100-3-3-1-14-2-3

349470. ML4 ML4-2-3-71-1-9-2-31-3-9
349471. ML4 ML4-2-3-71-1-9-2-31-3-9

350321. ML4 ML4-3-2-129-2-7-1-15-2-7
350322. ML4 ML4-3-2-129-2-7-1-15-2-7

350346. ML4 ML4-3-2-130-1-6-2-27-2-6
350347. ML4 ML4-3-2-130-1-6-2-27-2-6

350395. ML4 ML4-3-2-133-1-3-2-13-2-3
350396. ML4 ML4-3-2-133-1-3-2-13-2-3

350405. ML4 ML4-3-2-133-1-6-2-32-2-6
350406. ML4 ML4-3-2-133-1-6-2-32-2-6

350416. ML4 ML4-3-2-133-2-4-2-13-2-4
350417. ML4 ML4-3-2-133-2-4-2-13-2-4

350706. ML4 ML4-3-2-144-2-3-1-56-2-3
350707. ML4 ML4-3-2-144-2-3-1-56-2-3

350709. ML4 ML4-3-2-144-2-4-1-27-2-4
350710. ML4 ML4-3-2-144-2-4-1-27-2-4

350935. ML4 ML4-3-2-154-1-5-1-35-2-5
350936. ML4 ML4-3-2-154-1-5-1-35-2-5

350937. ML4 ML4-3-2-154-1-5-1-38-2-5
350938. ML4 ML4-3-2-154-1-5-1-38-2-5

350972. ML4 ML4-3-2-155-2-5-1-33-2-5
350973. ML4 ML4-3-2-155-2-5-1-33-2-5

351178. ML4 ML4-3-2-165-1-6-2-44-2-6
351179. ML4 ML4-3-2-165-1-6-2-44-2-6

351251. ML4 ML4-3-2-168-1-4-2-25-2-4
351252. ML4 ML4-3-2-168-1-4-2-25-2-4

351478. ML4 ML4-3-3-144-1-6-1-11-3-6
351479. ML4 ML4-3-3-144-1-6-1-11-3-6

352051. ML4 ML4-4-2-173-1-3-1-36-2-3
352052. ML4 ML4-4-2-173-1-3-1-36-2-3

352057. ML4 ML4-4-2-173-1-3-2-7-2-3
352058. ML4 ML4-4-2-173-1-3-2-7-2-3

352130. ML4 ML4-4-2-177-2-5-2-32-2-5
352131. ML4 ML4-4-2-177-2-5-2-32-2-5

352493. ML4 ML4-4-2-195-1-4-2-10-2-4
352494. ML4 ML4-4-2-195-1-4-2-10-2-4

352961. ML4 ML4-4-2-219-1-5-1-10-2-5
352962. ML4 ML4-4-2-219-1-5-1-10-2-5

352973. ML4 ML4-4-2-219-1-8-2-11-2-8
352974. ML4 ML4-4-2-219-1-8-2-11-2-8

353143. ML4 ML4-4-2-227-1-7-1-34-2-7
353144. ML4 ML4-4-2-227-1-7-1-34-2-7

353148. ML4 ML4-4-2-227-1-8-2-3-2-8
353149. ML4 ML4-4-2-227-1-8-2-3-2-8

353594. ML4 ML4-5-2-230-1-5-1-37-2-5
353595. ML4 ML4-5-2-230-1-5-1-37-2-5

354127. ML4 ML4-5-2-256-1-5-2-0-2-5
354128. ML4 ML4-5-2-256-1-5-2-0-2-5

354168. ML4 ML4-5-2-259-1-3-1-24-2-3
354169. ML4 ML4-5-2-259-1-3-1-24-2-3

354276. ML4 ML4-5-2-266-1-5-1-0-2-5
354277. ML4 ML4-5-2-266-1-5-1-0-2-5

354321. ML4 ML4-5-2-268-1-4-1-53-2-4
354322. ML4 ML4-5-2-268-1-4-1-53-2-4

354356. ML4 ML4-5-2-271-1-3-1-3-2-3
354357. ML4 ML4-5-2-271-1-3-1-3-2-3

354715. ML4 ML4-5-2-292-1-4-1-28-2-4
354716. ML4 ML4-5-2-292-1-4-1-28-2-4

355096. ML4 ML4-6-2-301-1-3-1-2-2-3
355097. ML4 ML4-6-2-301-1-3-1-2-2-3

355167. ML4 ML4-6-2-305-2-3-2-14-2-3
355168. ML4 ML4-6-2-305-2-3-2-14-2-3

355257. ML4 ML4-6-2-309-1-3-2-6-2-3
355258. ML4 ML4-6-2-309-1-3-2-6-2-3

355657. ML4 ML4-7-3-312-4-4-1-2-3-4
355658. ML4 ML4-7-3-312-4-4-1-2-3-4

355991. ML4 ML4-9-2-339-1-4-2-3-2-4
355992. ML4 ML4-9-2-339-1-4-2-3-2-4

356050. ML4 ML4-9-2-342-2-4-1-47-2-4
356051. ML4 ML4-9-2-342-2-4-1-47-2-4

356577. ML4 ML4-9-2-381-2-4-1-5-2-4
356578. ML4 ML4-9-2-381-2-4-1-5-2-4

356729. ML4 ML4-9-2-396-1-5-1-16-2-5
356730. ML4 ML4-9-2-396-1-5-1-16-2-5

398095. MW4 MW4-1-1-499-3-3-1-26-1-3
398096. MW4 MW4-1-1-499-3-3-1-26-1-3

590644. RW4 RW4-12-2-445-53-5-2-2-2-5
590645. RW4 RW4-12-2-445-53-5-2-2-2-5

591243. RW4 RW4-2-2-139-130-6-1-13-2-6
591244. RW4 RW4-2-2-139-130-6-1-13-2-6

592496. RW4 RW4-3-2-175-92-3-2-29-2-3
592497. RW4 RW4-3-2-175-92-3-2-29-2-3

595905. RW4 RW4-6-2-279-44-8-2-14-2-8
595906. RW4 RW4-6-2-279-44-8-2-14-2-8

599316. RW4 RW4-9-2-21-66-5-1-10-2-5
599317. RW4 RW4-9-2-21-66-5-1-10-2-5

any explanation for this? Thank you.

Subject: Re: merging KR and PR - Multi-country and multi years
Posted by [Janet-DHS](#) on Fri, 02 Jun 2023 14:19:09 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from DHS staff member, Tom Pullum:

Users are asked to say which survey they are using. It looks like you are having trouble with one of the older (phase 4) Mali surveys (I see "ML4" in your listing). There were several DHS surveys in Francophone Africa around DHS-4 that included sub-household identifiers. For them, you have to include the sub-household identifier or use hhid or caseid rather than just hv001 hv002 hvidx. Fortunately there were very few such surveys.

Also, if a survey does not include b16, we (DHS staff) cannot provide any guidance on PR/KR merges. We have been unable to come up with a foolproof strategy for merges without b16.

I strongly recommend that you do the merges first and then append, rather than appending first and merging second. This was discussed in earlier posts.

But mainly I suggest that you reconsider WHY you want to merge the KR and PR files. Which child-specific variables do you need that are in the PR file and not in the KR file? The KR file already includes virtually every item of information about the child that is in the PR file. It also includes virtually every item of information about the mother that is in the IR file. If you want to add household-specific variables such as sanitation and source of water, then you just merge with the HR file, using the cluster and household ID codes and forget the line number. Please let us know if there is some specific variable in the household data that you cannot get with this simpler approach.

Subject: Re: merging KR and PR - Multi-country and multi years
Posted by [ekowababio](#) on Fri, 02 Jun 2023 16:14:01 GMT
[View Forum Message](#) <> [Reply to Message](#)

Thank you, Janet.

Thanks for the insight, Tom. As you rightly suspected, I want to include household-specific variables such as cooking fuel, sanitation and source of water hence merging with the PR. I tried with the HR datasets and had similar issues hence reverting to PR thinking that would address the problem Also, I am using all DHS surveys for sub-Saharan African Countries. I will follow the recommendation of merging the two datasets first and then appending them later. Given the volume of the datasets, is there an effective way of merging the individual KR with their matching PR datasets or do I have to do the merging one by one in Stata? Thank you.
