
Subject: Re: merge IR and KR: duplicate line number in IR
Posted by [bun_2019fall](#) on Fri, 27 Aug 2021 15:51:56 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi Bruno,

For example, I used "BOIR01FL", and below the dup variable (dup by v000 - country v001-cluster v002-household v003-line number) that I created have values > 0.

```
use "V:\DHS raw survey/BOIR01DT/BOIR01FL", clear
duplicates tag v000 v001 v002 v003, generate(dup)
tab dup
```

And below are all the IR files that I found the duplicates (note, I cut the "IR" string from the filename, so the original filename should be, for example, BR21FL=BRIR21FL). These are all raw DHS files and I did not modified anything in the raw data.

data_name	Freq.	Percent	Cum.
BO01FL.DTA	6,669	7.13	7.13
BR21FL.DTA	54	0.06	7.19
CI51FL.DTA	7,442	7.96	15.14
CM22FL.DTA	3,329	3.56	18.70
CM31FL.DTA	4,259	4.55	23.25
CO01FL.DTA	4,721	5.05	28.30
DR21FL.DTA	6,354	6.79	35.09
HN52FL.DTA	196	0.21	35.30
HT31FL.DTA	4,438	4.74	40.05
IA42FL.DTA	34,892	37.30	77.35
LK02FL.DTA	2	0.00	77.35
ML32FL.DTA	589	0.63	77.98
ML41FL.DTA	9,300	9.94	87.92
NG21FL.DTA	4	0.00	87.92
NI22FL.DTA	4,608	4.93	92.85
NI31FL.DTA	47	0.05	92.90
PE01FL.DTA	197	0.21	93.11
PE6IFL.DTA	272	0.29	93.40
SN02FL.DTA	237	0.25	93.65
SN4AFL.DTA	4	0.00	93.66
TD31FL.DTA	5,923	6.33	99.99
TN02FL.DTA	8	0.01	100.00
UG01FL.DTA	2	0.00	100.00
-----+			
Total	93,547	100.00	

I wonder if I missed anything here, as to whether I correctly identified the unique women in each IR files, thus to link back to women who have children in the KR files? Thank you very much

again!
