
Subject: Re: Accessing DHS Data with Stata/IC
Posted by [user-rhs](#) on Thu, 06 Feb 2014 22:52:41 GMT
[View Forum Message](#) <> [Reply to Message](#)

Kei, Stata IC only allows for 2,048 variables max. To work around this issue, you can do one of a number of different things:

1.) Pull only the variables you need - If you already know the variable names you need, then specify it when you call the dataset in your do file, i.e.

```
use v012 v201 v025 using "c:\user-rhs\datasets\countryx.dta",clear
```

2.) If you don't feel like sifting through the FRQ file for the variables you need, or if you are not sure which ones you need, you can pull batches of the variables at a time using the wildcard (*) and remove the variables where all obs'ns have a missing value, save as separate datasets, then merge all of them together by caseid. Often, this will result in a dataset that is <2000 vbls, i.e.

*pull vbls that begin with V and S

```
use caseid v* s* using "c:\user-rhs\datasets\countryx.dta",clear
```

```
findname,all(missing(@)) /*running this command will spit out a list of vbls where all obsns have  
missing value,  
you can then drop these variables safely from your dataset */
```

```
/* copy and paste the variables that were listed as a result of the -findname- command above, use  
/// at the end of each line to tell stata that the list of vbls to remove goes beyond that line. Make  
sure there is a space b/w vbl name and ///, i.e. do v120 ///, don't do v120/// */
```

```
/*EXAMPLE*/ drop v017    v305_18  v3a00l  v453    v469i  v470xy  v759c  v762bp  
v770f  s215c_19 ///  
v040   v307_18  v3a00m  v454    v469j  v470xz  v759d  v762bq  v770g  s215m_20  
///  
v122   v305_19  v3a00n  v455    v469k  v630j  v759e  v762br  v770h  s215y_20 ///  
v128   v307_19  v3a00o  v456    v469m  v630k  v759f  v762bs  v770i  s215c_20 ///  
v156   v305_20  v3a00p  v457    v469n  v630l  v759g  v762bt  v771   s406a_2
```

```
sort caseid
```

```
*save the modified data  
save countryx_vsonly.dta
```

```
*pull vbls that begin with M  
use caseid m* b* using "c:\user-rhs\datasets\countryx.dta",clear
```

```
findname,all(missing(@))
```

```
drop b12_01 m40e_1 m40g_2 m37g_3 m2k_4 m55i_4 m49a_5 m37e_6 m55d_6  
mm15_12 ///  
b8_16 m40f_1 m40i_2 m37i_3 m2l_4 m55j_4 m49b_5 m37f_6 m55e_6  
mm5_13 ///  
b9_16 m40g_1 m40j_2 m37j_3 m2m_4 m55k_4 m49c_5 m37g_6 m55f_6  
mm10_13 ///  
b16_16 m40i_1 m40k_2 m37k_3 m2n_4 m55l_4 m49d_5 m37h_6 m55g_6  
mm11_13 ///  
b8_17 m40j_1 m40m_2 m37m_3 m3a_4 m55m_4 m49e_5 m37i_6 m55h_6  
mm12_13
```

sort caseid

```
*save modified data  
save countryx_mbonly.dta, replace
```

merge 1:1 caseid using countryx_vsonly.dta

```
*if the merge was satisfactory, you can drop the _merge vbl  
drop _merge  
save countryx_mod.dta,replace
```

Let me know whether this works for you.

HTH,
RHS