
Subject: Re: How to create composite variables
Posted by [Janet-DHS](#) on Fri, 22 Mar 2024 20:06:40 GMT
[View Forum Message](#) <> [Reply to Message](#)

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

It appears that you are using the HR file, which has one record for each household, rather than the PR file, which has one record for each person in the household. The subscripts _01, _02, etc., are the line numbers of individuals in the household, and if you switch to the PR file the line number is given with a separate variable, hvidx. I recommend using the PR file.

In the PR file for the 2022 survey, TZPR82 FL.dta, hml32 is the result of the malaria rapid test for each person. hml35 is in the data but in this survey it was not actually used. It is NA (coded with a dot) for each case. The other variable you mention, sb115, is not in the data at all.

In the PR file for the 2015-16 survey, TZPR7BFL.dta, hml32 is the result of the smear test and hml35 is the result of the rapid test. sb115 is also the result of the rapid test. I recommend that you ignore sb115. It simply duplicates hml35 but is reverse coded and does not apply to as many cases (see below). Let us know if you have other questions.

```
. correlate hml35 sb115
```

```
(obs=9,322)
```

```
      |  hml35  sb115  
-----+-----  
hml35 |  1.0000  
sb115 | -1.0000  1.0000
```

```
. tab hml35 sb115,m
```

```
result of |
```

```
  malaria |      result for malaria rdt  
  
rapid test | positive  negative  other  . | Total  
-----+-----  
negative |      0    8,206      0  1,348 |  9,554  
positive |   1,116      0      0   265 |  1,381  
  . |      0      0      3 53,942 | 53,945  
-----+-----  
Total |   1,116    8,206      3 55,555 | 64,880
```
