Subject: Re: Combining PR and KR to get child-specific malaria test results
Posted by coffeehelps on Mon, 29 Aug 2016 17:46:13 GMT

Thanks very much Tom for this detective work and the detailed description!

There are several other datasets that seem to have similar problems with non-unique observations. I paste my overview below in case that's useful for your CIS DHS. (The variable "mergeid" is my creation and you can ignore it.)

Could you kindly post here or send me a message if/when the source files might be corrected? I may not be able to implement your detailed detective work for all these files and prefer to abstain from making corrections on the fly. For now I'll just drop these duplicates and will update the input data later.

Thanks much!

** List of duplicates
** "mergeid" is not an original DHS variable, please ignore

.. AO2006 - have 4 duplicates

```
+------------------------------------------------------+
| v000   v001   v002   v003   b16   bidx     mergeid |
|------------------------------------------------------|
|  AO5    58    53     4     7     1   AO5 58 53 7 |
|  AO5    58    53     6     7     1   AO5 58 53 7 |
|  AO5    96    69     3     5     1   AO5 96 69 5 |
|  AO5    96    69     7     5     1   AO5 96 69 5 |
+------------------------------------------------------+
```

.. AO2011 - have 2 duplicates

```
+------------------------------------------------------+
| v000   v001   v002   v003   b16   bidx     mergeid |
|------------------------------------------------------|
|  AO5   226    32     1     6     1   AO5 226 32 6 |
|  AO5   226    32     2     6     1   AO5 226 32 6 |
+------------------------------------------------------+
```

.. BJ2012 - have 2 duplicates

```
+------------------------------------------------------+
| v000   v001   v002   v003   b16   bidx     mergeid |
|------------------------------------------------------|
```

```
| BJ6   469   11    2    7    1   BJ6 469 11 7 |
| BJ6   469   11    3    7    1   BJ6 469 11 7 |
+------------------------------------------------------+
```

.. LS2009 - have 2 duplicates

```
+----------------------------------------------------------+
| v000   v001   v002   v003   b16   bidx      mergeid |
|----------------------------------------------------------|
| LS5   192   177    4    10    1   LS5 192 177 10 |
| LS5   192   177    6    10    1   LS5 192 177 10 |
+----------------------------------------------------------+
```

.. MW2004 - have 2 duplicates

```
+--------------------------------------------------------+
| v000   v001   v002   v003   b16   bidx      mergeid |
|--------------------------------------------------------|
| MW4    499    3    1    3    1   MW4 499 3 3 |
| MW4    499    3    1    3    2   MW4 499 3 3 |
+--------------------------------------------------------+
```