

```
Finding duplicate records I

* Both tables are sorted by county;

* If need to find duplicates in multiple vars;

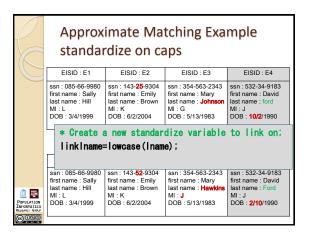
* Combine the multiple vars into one variable first, then run same code;

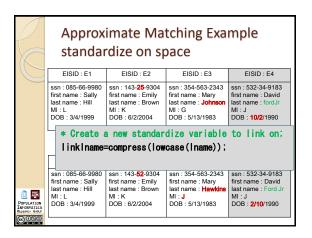
data dupenty;
merge tab1 tab2;
by county;
if ! (first. county & last. county);
```

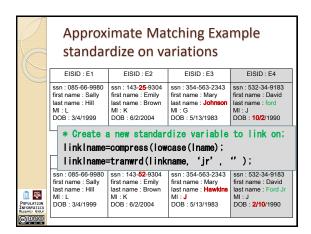
```
Finding duplicate records II

* Both tables are sorted by county;
data duporty;
merge tab1(in=aa) tab2(in=bb);
by county;
src=aa*10+bb;

proc freq;
tables src;
proc print data=duporty(obs=30);
where src~=11;
```







```
Other useful functions

• Appendix 2 (p59) of ARHQ Report

vto=translate(vfrom, ' ', "()', -. ");
vto=lowcase(compress(vto, ', 't'));
vto=tranwrd(vto, "ctr", "center");
vto=tranwrd(vto, "medical", "med");

* vto=tranwrd(vto, "med", "medical");

* medical center = ?;
vto=tranwrd(vto, "texas", "tx");
vto=tranwrd(vto, "hospital", "hosp");
```

