

```
clear

log using d:\temp\hps.log, replace

*****
* Merge the individual data with weights dta at Week 11
*****

use "D:\jason\workshop\Household Pulse Survey\Phase 1\week 11-(7.9-7.14)\pulse2020_puf_11.dta", clear
sort scram
save, replace

use "D:\jason\workshop\Household Pulse Survey\Phase 1\week 11-(7.9-7.14)\pulse2020_repwgt_puf_11.dta", clear
sort scram
save, replace

use "D:\jason\workshop\Household Pulse Survey\Phase 1\week 11-(7.9-7.14)\pulse2020_puf_11.dta", clear
merge 1:1 scram using "D:\jason\workshop\Household Pulse Survey\Phase 1\week 11-(7.9-7.14)\pulse2020_repwgt_puf_11.dta"
drop _merge

sort scram
save "D:\jason\workshop\Household Pulse Survey\week11.dta", replace

*****
* Merge the individual data with weights dta at Week 12
*****

use "D:\jason\workshop\Household Pulse Survey\Phase 1\week 12-(7.16-7.21)\pulse2020_puf_12.dta", clear
sort scram
save, replace

use "D:\jason\workshop\Household Pulse Survey\Phase 1\week 12-(7.16-7.21)\pulse2020_repwgt_puf_12.dta", clear
sort scram
save, replace

use "D:\jason\workshop\Household Pulse Survey\Phase 1\week 12-(7.16-7.21)\pulse2020_puf_12.dta", clear
merge 1:1 scram using "D:\jason\workshop\Household Pulse Survey\Phase 1\week 12-(7.16-7.21)\pulse2020_repwgt_puf_12.dta"
drop _merge

sort scram
save "D:\jason\workshop\Household Pulse Survey\week12.dta", replace

*****
* Analyzing single-week HPS data
*****

use "D:\jason\workshop\Household Pulse Survey\week11.dta", clear

*****
*clean up the data
*****

tab1 TBIRTH_YEAR, mis
gen agegroup = TBIRTH_YEAR
recode agegroup (1930/1940.1 = 7) (1940.2/1950.1 = 6) (1950.2/1960.1 = 5) (1960.2/1970.1 = 4) (1970.2/1980.1 = 3) (1980.

tab2 TBIRTH_YEAR agegroup, mis

tab1 income, mis

gen newincome = 1
replace newincome = 0 if inrange(income,1,4)
replace newincome = . if inlist(income,-99, .m)
tab2 income newincome, mis

*****
* The use of SVY and Subpop commands to get the correct estimated parameters
*****

gen valid1 = 1
replace valid1 = 0 if newincome ==. | agegroup ==.
tab1 valid1, mis

svyset [iw=pweight], jkrweight(PWEIGHT1- PWEIGHT80, multiplier(.05)) ///
vce(jackknife) mse
```

```
svy, subpop(valid1): tab newincome agegroup  
svy, subpop(valid1): logit newincome agegroup
```

```
*****  
* If the subpop option is not used, the estimated standard errors of the parameters are not correct  
*****
```

```
svy: tab newincome agegroup  
svy: logit newincome agegroup
```

```
*****  
* Analyzing Multiple-week HPS data  
*****
```

```
*****  
* Pooling the data from Weeks 11 and 12 together  
*****
```

```
use "D:\jason\workshop\Household Pulse Survey\week11.dta" , clear  
append using "D:\jason\workshop\Household Pulse Survey\week12.dta"
```

```
*****  
* check the data  
*****
```

```
sort scram week  
by scram: gen N = _N  
by scram: gen n = _n
```

```
tab1 N if n ==1, mis
```

```
list scram week n N ms wrkloss in 1/6000 if N ==2, sepby(scram)
```

```
*****  
*clean up the data  
*****
```

```
tab1 TBIRTH_YEAR, mis  
gen agegroup = TBIRTH_YEAR  
recode agegroup (1930/1940.1 = 7) (1940.2/1950.1 = 6) (1950.2/1960.1 = 5) (1960.2/1970.1 = 4) (1970.2/1980.1 = 3) (1980.
```

```
tab2 TBIRTH_YEAR agegroup, mis
```

```
tab1 income, mis
```

```
gen newincome = 1  
replace newincome = 0 if inrange(income,1,4)  
replace newincome =. if inlist(income,-99, .m)  
tab2 income newincome, mis
```

```
*****  
* Adjustment to the personal weights  
*****
```

```
gen pweight_11_12 = pweight/2  
label variable pweight_11_12 "new personal weights for weeks 11 and 12"
```

```
*****  
* The use of SVY and Subpop commands to get the correct estimated parameters  
*****
```

```
gen valid1 = 1  
replace valid1 =0 if newincome ==. | agegroup ==.  
tab1 valid1, mis
```

```
svyset [iw=pweight_11_12], jkrweight(PWEIGHT1- PWEIGHT80, multiplier(.05)) ///  
vce(jackknife) mse
```

```
svy, subpop(valid1): tab newincome agegroup  
svy, subpop(valid1): logit newincome agegroup
```

```
*****
```

```
* If the subpop option is not used, the estimated standard errors of the parameters are not correct
*****
svy: tab newincome agegroup
svy: logit newincome agegroup

*****
* If the new weight variable is not used, tthe estimated standard errors of the parameters are not correct
*****
svyset [iw=pweight], jkrweight(PWEIGHT1- PWEIGHT80, multiplier(.05)) ///
        vce(jackknife) mse

svy, subpop(valid1): tab newincome agegroup
svy, subpop(valid1): logit newincome agegroup

log close
```