

## Results from R on what happens with data problems

```
> shs.des_nc <- svydesign(id=~PSU,
weights=~IND_WT,strata=~STRATUM,data=shs_nc)
Error in svydesign(id = ~PSU, weights = ~IND_WT, strata = ~STRATUM,
data = shs_nc) :
  Clusters not nested in strata you may want nest=T

#
# if you are smart at R you can use this code to work out which
# PSUs are causing the problem
# First get unique combinations of PSU and stratum
# and then work out for which there are >1 stratum

> check1<-
shs_nc[!duplicated(paste(shs_nc$PSU,shs_nc$STRATUM)),c(1,12)]
> check2<-table(check1[,1])
> check2[check2>1]

089019D4 266001I4 266019D3
      2      2      2

> #
> # since there are only a couple you can force nesting
> # which will split these PSUs
> #
> shs.des_nc <- svydesign(id=~PSU,
weights=~IND_WT,strata=~STRATUM,data=shs_nc,nest=T)
> svymean(~INTUSE,shs.des_nc,deff=T)
Error in svyCprod(x * pweights/psum, design$strata,
design$cluster[[1]], :
  Stratum 100H has only one sampling unit.

> #
> # this fails from a lonely PSU, need to try other options
> options(survey.lonely.psu='remove')
> svymean(~INTUSE,shs.des_nc,deff=T)
      mean      SE  DEff
INTUSE 0.341363 0.003386 1.4628
There were 50 or more warnings (use warnings() to see the first 50)
>
> options(survey.lonely.psu='adjust')
> svymean(~INTUSE,shs.des_nc,deff=T)
      mean      SE  DEff
INTUSE 0.3413627 0.0034051 1.4793
There were 50 or more warnings (use warnings() to see the first 50)
> warnings()
Warning messages:
1: Stratum 100H has only one sampling unit. in: svyCprod(x *
pweights/psum, design$strata, design$cluster[[1]], ...
2: Stratum 200K has only one sampling unit. in: svyCprod(x *
pweights/psum, design$strata, design$cluster[[1]], ...
3: Stratum 280Z has only one sampling unit. in: svyCprod(x *
pweights/psum, design$strata, design$cluster[[1]], ...
4: Stratum 110G has only one sampling unit. in: svyCprod(x *
pweights/psum, design$strata, design$cluster[[1]], ...
5: Stratum 110F has only one sampling unit. in: svyCprod(x *
pweights/psum, design$strata, design$cluster[[1]], ...
6: Stratum 120X has only one sampling unit. in: svyCprod(x *
pweights/psum, design$strata, design$cluster[[1]], ...
```

etc.