Analysis of health and drug use
Exemplar 5

Red text is commands and blue is results

```
> summary(ex5.des)

Independent Sampling design
svydesign(id = ~1, weights = ~WEIGHT, data = ex5)
Probabilities:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.3282  0.7363  0.8910  0.8962  1.0960  1.3170
Data variables:
 [1] "AGE"     "WEIGHT"  "EMPL"    "QUAL"    "GENHELF" "Q85A"    "Q85B"    "SACC"
"SINC"    "GWEIGHT" "LIVING"

> #
> # now we can use the svymean commands to get first the weighted proportions
> # for categorical variables

> svymean(~GENHELF,ex5.des,deff=T)
                       mean        SE   DEff
GENHELFexcellent 0.1866598 0.0207625 1.0250
GENHELFvery good 0.3776246 0.0266726 1.0928
GENHELFgood      0.3435661 0.0263967 1.1153
GENHELFfair      0.0700424 0.0143276 1.1377
GENHELFpoor      0.0221072 0.0094428 1.4890

> svymean(~Q85A,ex5.des,deff=T)
                                mean SE DEff
Q85Anever used                    NA NA   NA
Q85Atried once or twice           NA NA   NA
Q85Ause daily                     NA NA   NA
Q85Ause weekly                    NA NA   NA
Q85Aused in last month            NA NA   NA
Q85Aused more than a month ago    NA NA   NA
> #
> # this fails for missing values so need to get subset of design without missing
values
> #

> svymean(~Q85A,subset(ex5.des,!is.na(Q85A)),deff=T)
                                     mean        SE   DEff
Q85Anever used                  0.5831140 0.0274052 1.1061
Q85Atried once or twice         0.2497045 0.0240431 1.1046
Q85Ause daily                   0.0340624 0.0101163 1.1135
Q85Ause weekly                  0.0240867 0.0088028 1.1801
Q85Aused in last month          0.0277697 0.0084588 0.9488
Q85Aused more than a month ago 0.0812627 0.0159420 1.2187
> #
> # and expressing this as percentages
> #
> round(print(svymean(~Q85A,subset(ex5.des,!is.na(Q85A))))*100,1)
                                mean  SE
Q85Anever used                  58.3 2.7
Q85Atried once or twice         25.0 2.4
Q85Ause daily                    3.4 1.0
Q85Ause weekly                   2.4 0.9
Q85Aused in last month           2.8 0.8
Q85Aused more than a month ago   8.1 1.6
>
```

```
> # now the means of some continuous variables
> #
> svymean(~as.numeric(GENHELF),ex5.des,deff=T)  # this treats it as a score not a
category
                         mean       SE   DEff
as.numeric(GENHELF) 2.363313 0.052741 1.1445

> svymean(~SINC,ex5.des,deff=T)
        mean      SE   DEff
SINC 18.7117  0.6946 1.3158

> svymean(~SACC,ex5.des,deff=T)
           mean        SE   DEff
SACC -0.067572  0.032883 1.0413


> summary(svyglm(as.numeric(GENHELF)~CANSCORE,ex5.des))

Call:
svyglm(formula = as.numeric(GENHELF) ~ CANSCORE, design = ex5.des)

Survey design:
update.survey.design(ex5.des, AMPSCORE = AMPSCORE)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.28376    0.05666  40.304  < 2e-16 ***
CANSCORE     0.45467    0.14982   3.035  0.00258 **
---
Signif. codes:  0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1

(Dispersion parameter for gaussian family taken to be 0.803459)

Number of Fisher Scoring iterations: 2


> #
> #  Now make a table to understand results
> #
> can.table<-svytable(~GENHELF+Q85A,ex5.des)
> can.tprint<-round(sweep(can.table,2,apply(can.table,2,sum),"/")*100) #
calculating column percents

> can.tprint<-rbind(can.tprint,table(ex5$Q85A))                    # add a row
below for base numbers
> #
> # at any stage type tyhe name of an object to see what it is
> #
> dimnames(can.tprint)[[1]][6]<-'BASE'
> can.tprint
               never    once/twice daily      weekly last month    >month ago
excellent         22         19          6          9       0           8
very good         40         35         41         30      17          42
good              29         37         53         42      54          48
fair               7          8          0         19      20           2
poor               2          0          0          0       8           0
BASE             210         90         12          8      11          27
>
>
```