CHAPTER 7 ANALYSIS EXAMPLES REPLICATION-R SURVEY PACKAGE 3.22

GENERAL NOTES ABOUT ANALYSIS EXAMPLES REPLICATION

These examples are intended to provide guidance on how to use the commands/procedures for analysis of complex sample survey data and assume all data management and other preliminary work is done. The relevant syntax for the procedure of interest is shown first along with the associated output for that procedure(s). In some examples, there may be more than one block of syntax and in this case all syntax is first presented followed by the output produced.

In some software packages certain procedures or options are not available but we have made every attempt to demonstrate how to match the output produced by Stata 10+ in the textbook. Check the ASDA website for updates to the various software tools we cover.

GENERAL NOTES ABOUT CHAPTER 7 ANALYSES IN R SURVEY PACKAGE 3.22 (WITH R 2.7)

The R survey package used in these examples is version 3.22 and was run under R v2.7 on a PC.

The R survey package offers a very good range of svy commands for the analyses of this chapter: svyglm with the default link is used for linear regression. Other commands used in this chapter include: the lm command with and without weights for SRS (simple random sample) linear regression, use of the factor statement for categorical variables as well as indicator variables as predictors, the regTermTest command for testing of groups of parameters including interactions in models, and the plot command with a model object for default regression diagnostics. Additional plots could be obtained with more coding and work, see the R documentation for details.

```
#Data production and set up of design objects
#remember to load package first survey package

#NHANES
nhanesdata <- read.table(file = "f:/applied_analysis_book/r/nhanes_final.txt", sep = "\t", header = T, as.is=T)

#create factor variables
nhanesdata$racec <- factor(nhanesdata$RIDRETH1, levels = 1: 5 , labels =c("Mexican", "Other Hispanic", "White",
"Black", "Other"))
nhanesdata$marcatc <- factor(nhanesdata$marcat, levels = 1: 3, labels =c("Married", "Previously Married", "Never
Married"))
nhanesdata$edcatc <- factor(nhanesdata$edcat, levels = 1: 4, labels =c("0-11", "12", "13-15","16+"))
nhanesdata$bp_catc <- factor(nhanesdata$bp_cat, levels = 1: 4, labels =c("Normal", "Pre-HBP", "Stage 1
HBP","Stage 2 HBP"))
nhanesdata$agesq <- (nhanesdata$agecent * nhanesdata$agecent )
names(nhanesdata)

nhanessvy2 <- svydesign(strata=~SDMVSTRA, id=~SDMVPSU, weights=~WTMEC2YR, data=nhanesdata, nest=T)
subnhanes <- subset(nhanessvy2 , RIDAGEYR >= 18)

#NCS-R
ncsr <- read.table(file = "f:/applied_analysis_book/r/ncsr2010.txt", sep = "\t", header = T, as.is=T)
names(ncsr)

#create factor versions with labels
ncsr$racec <- factor(ncsr$racecat, levels = 1: 4, labels =c("Other", "Hispanic", "Black", "White"))
ncsr$marcatc <- factor(ncsr$MAR3CAT, levels = 1: 3, labels =c("Married", "Previously Married", "Never Married"))
ncsr$edcatc <- factor(ncsr$ED4CAT, levels = 1: 4, labels =c("0-11", "12", "13-15","16+"))
ncsr$sexc <- factor(ncsr$SEX, levels = 1:2, labels=c("Male","Female"))
ncsr$agcatc <- factor(ncsr$ag4cat, levels = 1:4, labels=c("18-29", "30-44", "45-59", "60+"))

ncsrsvyp1 <- svydesign(strata=~SESTRAT, id=~SECLUSTR, weights=~NCSRWTSH, data=ncsr, nest=T)
ncsrsvyp2 <- svydesign(strata=~SESTRAT, id=~SECLUSTR, weights=~NCSRWTLG, data=ncsr, nest=T)
ncsrsvypop <- svydesign(strata=~SESTRAT, id=~SECLUSTR, weights=~popweight, data=ncsr, nest=T)

#HRS
#both hh and r weights are needed plus financial respondent for hh level analysis
hrs <- read.table(file = "f:/applied_analysis_book/r/hrs2010.txt", sep = "\t", header = T, as.is=T)
hrssvyhh <- svydesign(strata=~STRATUM, id=~SECU, weights=~KWGTHH , data=hrs, nest=T)
summary(hrssvyhh)
hrssvysub <-subset(hrssvyhh, KFINR==1)

hrssvyr <- svydesign(strata=~STRATUM, id=~SECU, weights=~KWGTR , data=hrs, nest=T)
summary(hrssvyr)
```

```
#EXAMPLE 7.5 BIVARIATE TESTING OF EACH FACTOR VARIABLE: RACE NHANES ADULT DATA
> ex75_race
Stratified 1 - level Cluster Sampling design (with replacement)
With (30) clusters.
subset(svynhanes, RIDAGEYR >= 18)

Call:  svyglm(bpxdi1_1 ~ racec, design = subnhanes)

Coefficients:
        (Intercept)  racecOther Hispanic           racecWhite           racecBlack           racecOther
             68.300                1.592                2.428                3.728                1.785

Degrees of Freedom: 4580 Total (i.e. Null);  11 Residual
  (982 observations deleted due to missingness)
Null Deviance:      132.5
Residual Deviance: 131.9        AIC: 37690

> summary(ex75_race <- svyglm(bpxdi1_1 ~racec, design=subnhanes))

Call:
svyglm(bpxdi1_1 ~ racec, design = subnhanes)

Survey design:
subset(svynhanes, RIDAGEYR >= 18)

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)          68.2996     0.4125 165.587  < 2e-16 ***
racecOther Hispanic   1.5924     1.1088   1.436 0.178802
racecWhite            2.4276     0.5543   4.380 0.001100 **
racecBlack            3.7278     0.7533   4.949 0.000437 ***
racecOther            1.7847     1.0298   1.733 0.110991
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 131.9065)

Number of Fisher Scoring iterations: 2


> regTermTest(ex75_race, ~racec, df==4)
Wald test for racec
 in svyglm(bpxdi1_1 ~ racec, design = subnhanes)
Chisq =  31.14746  on  4  df: p= 2.8565e-06
```

```
# EXAMPLE 7.5 BIVARIATE TEST OF MARITAL STATUS
> (ex75_marital <- svyglm(bpxdi1_1 ~marcatc, design=subnhanes))
Stratified 1 - level Cluster Sampling design (with replacement)
With (30) clusters.
subset(svynhanes, RIDAGEYR >= 18)

Call:  svyglm(bpxdi1_1 ~ marcatc, design = subnhanes)

Coefficients:
              (Intercept)  marcatcPreviously Married       marcatcNever Married
                 71.39171                   -0.07331                   -4.38617

Degrees of Freedom: 4577 Total (i.e. Null);  13 Residual
  (985 observations deleted due to missingness)
Null Deviance:      132.3
Residual Deviance: 129.9        AIC: 37590


> summary(ex75_marital)

Call:
svyglm(bpxdi1_1 ~ marcatc, design = subnhanes)

Survey design:
subset(svynhanes, RIDAGEYR >= 18)

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               71.39171    0.46754 152.696  < 2e-16 ***
marcatcPreviously Married -0.07331    0.68114  -0.108    0.916
marcatcNever Married      -4.38617    0.57305  -7.654 3.62e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 129.9686)

Number of Fisher Scoring iterations: 2


> regTermTest(ex75_marital, ~marcatc, df==2)
Wald test for marcatc
 in svyglm(bpxdi1_1 ~ marcatc, design = subnhanes)
Chisq =  80.31409  on  2  df: p= < 2.22e-16
```

```
# EXAMPLE 7.5 BIVARIATE TEST OF GENDER
> (ex75_sex <- svyglm(bpxdi1_1 ~RIAGENDR, design=subnhanes))
Stratified 1 - level Cluster Sampling design (with replacement)
With (30) clusters.
subset(svynhanes, RIDAGEYR >= 18)

Call:  svyglm(bpxdi1_1 ~ RIAGENDR, design = subnhanes)

Coefficients:
(Intercept)     RIAGENDR
     74.914       -2.844

Degrees of Freedom: 4580 Total (i.e. Null);  14 Residual
  (982 observations deleted due to missingness)
Null Deviance:      132.5
Residual Deviance: 130.7        AIC: 37640



> summary(ex75_sex)

Call:
svyglm(bpxdi1_1 ~ RIAGENDR, design = subnhanes)

Survey design:
subset(svynhanes, RIDAGEYR >= 18)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  74.9136     0.7271 103.036  < 2e-16 ***
RIAGENDR     -2.8442     0.3786  -7.512 2.83e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 130.7400)

Number of Fisher Scoring iterations: 2



> regTermTest(ex75_sex, ~RIAGENDR)
Wald test for RIAGENDR
 in svyglm(bpxdi1_1 ~ RIAGENDR, design = subnhanes)
Chisq =  56.42996  on  1  df: p= 5.8236e-14
```

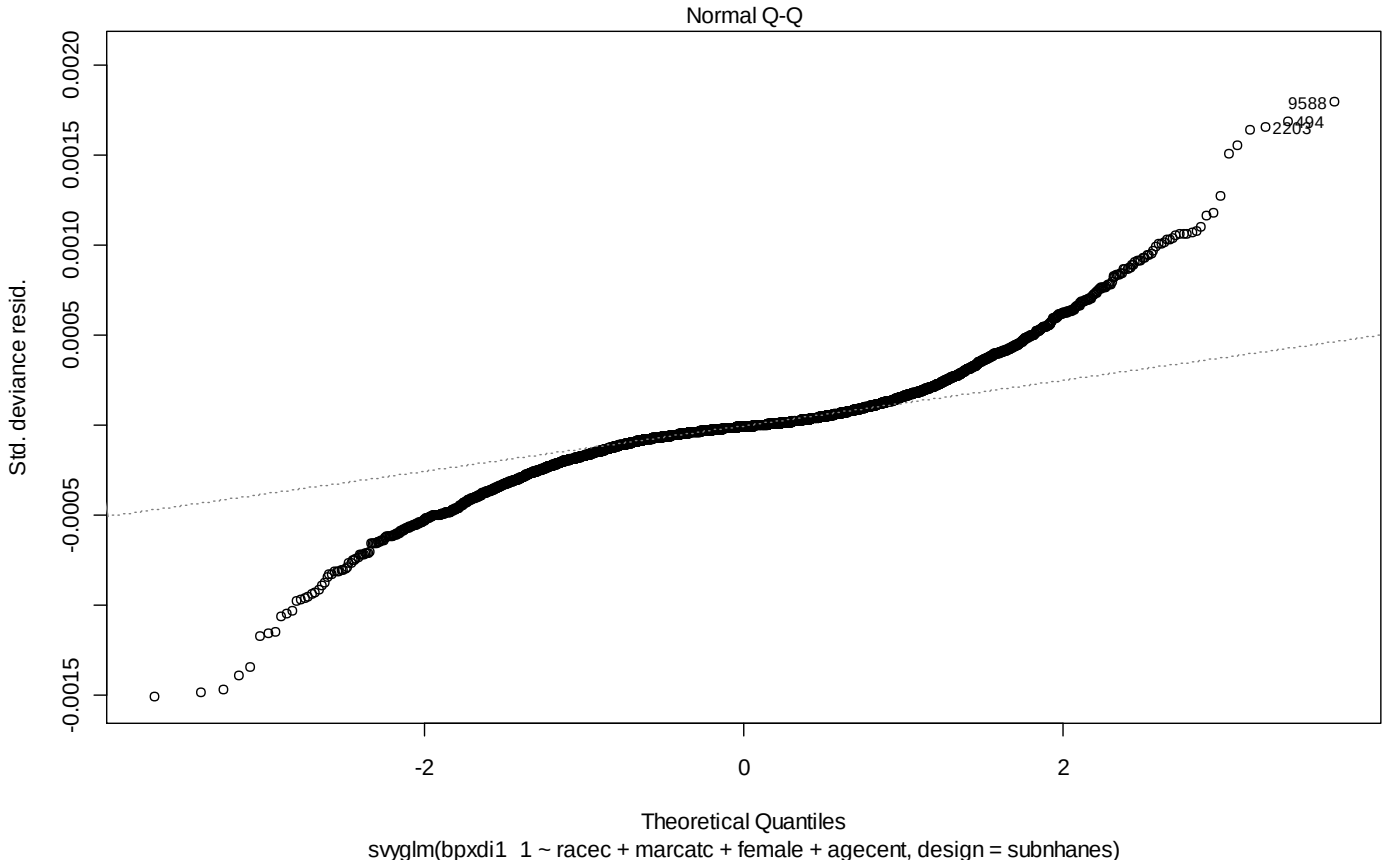
```
# EXAMPLE 7.5 BIVARIATE TEST OF CENTERED AGE

> (ex75_age <- svyglm(bpxdi1_1 ~agecent, design=subnhanes))
Stratified 1 - level Cluster Sampling design (with replacement)
With (30) clusters.
subset(svynhanes, RIDAGEYR >= 18)

Call:  svyglm(bpxdi1_1 ~ agecent, design = subnhanes)

Coefficients:
(Intercept)      agecent
   70.61552      0.05727

Degrees of Freedom: 4580 Total (i.e. Null);  14 Residual
  (982 observations deleted due to missingness)
Null Deviance:      132.5
Residual Deviance: 131.6        AIC: 37670


> summary(ex75_age)

Call:
svyglm(bpxdi1_1 ~ agecent, design = subnhanes)

Survey design:
subset(svynhanes, RIDAGEYR >= 18)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 70.61552    0.34968 201.942   <2e-16 ***
agecent      0.05727    0.02065   2.774   0.0149 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 131.6469)

Number of Fisher Scoring iterations: 2
```

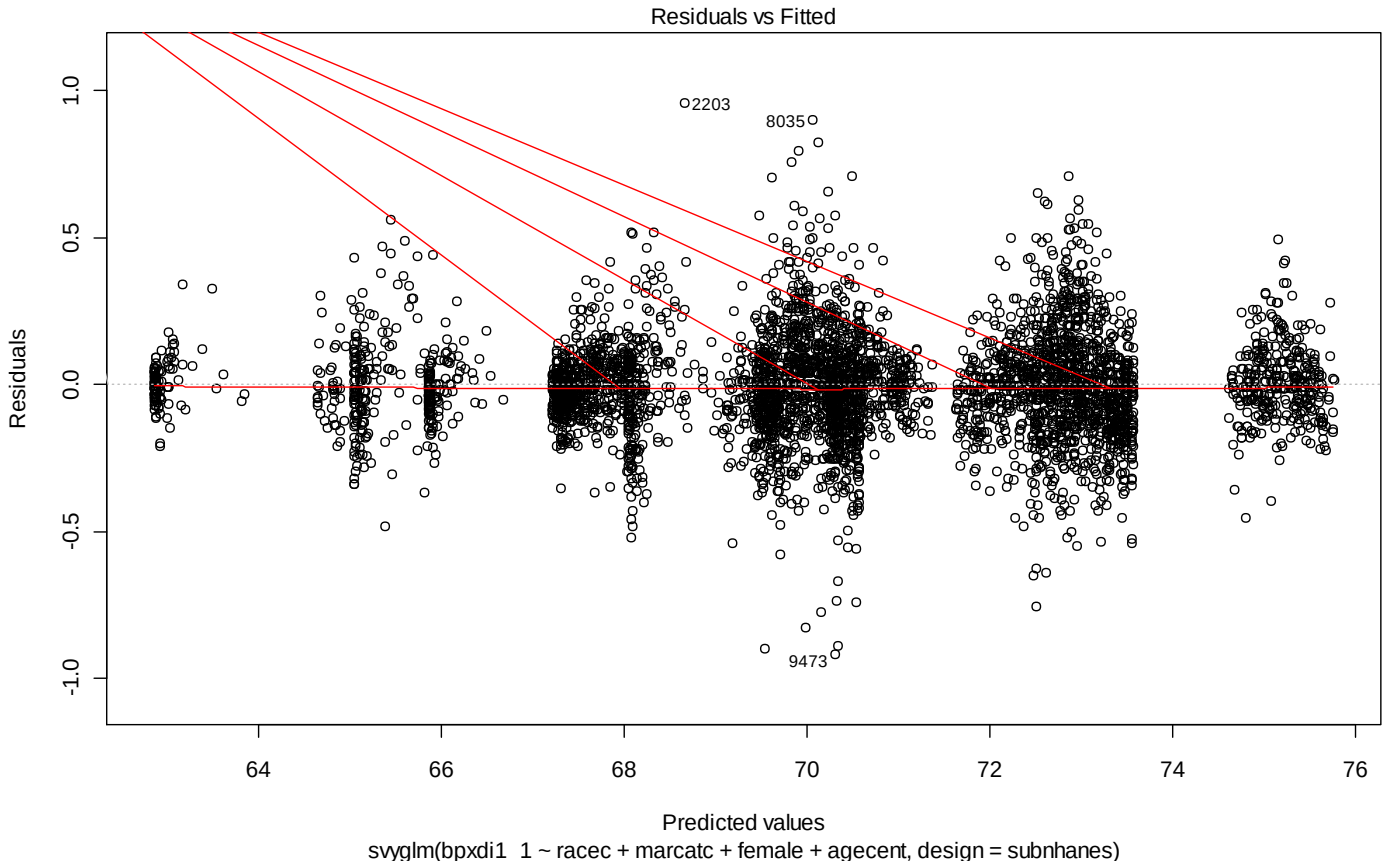
```
#EXAMPLE 7.5  UNWEIGHTED OLS REGRESSION
> (ex75_nowt <- lm(bpxdi1_1 ~ racec + marcatc + female + agecent, data= nhanesdata, RIDAGEYR >=18 ))

Call:
lm(formula = bpxdi1_1 ~ racec + marcatc + female + agecent, data = nhanesdata,     subset = RIDAGEYR >= 18)

Coefficients:
              (Intercept)        racecOther Hispanic                 racecWhite                 racecBlack
                 69.67211                    1.89823                    1.67193                    4.50813
               racecOther  marcatcPreviously Married       marcatcNever Married                     female
                  2.31195                    0.32691                   -4.21636                   -3.40181
                  agecent
                  0.03898

> summary(ex75_nowt)

Call:
lm(formula = bpxdi1_1 ~ racec + marcatc + female + agecent, data = nhanesdata,
    subset = RIDAGEYR >= 18)

Residuals:
     Min       1Q   Median       3Q      Max
-64.8883  -8.0284   0.2348   7.7130  54.3511

Coefficients:
                        Estimate Std. Error t value Pr(>|t|)
(Intercept)             69.67211    0.46435 150.043  < 2e-16 ***
racecOther Hispanic      1.89823    1.12538   1.687 0.091720 .
racecWhite               1.67193    0.49147   3.402 0.000675 ***
racecBlack               4.50813    0.56347   8.001 1.56e-15 ***
racecOther               2.31195    1.00454   2.302 0.021408 *
marcatcPreviously Married  0.32691    0.52221   0.626 0.531343
marcatcNever Married     -4.21636    0.51006  -8.266  < 2e-16 ***
female                   -3.40181    0.37459  -9.081  < 2e-16 ***
agecent                   0.03898    0.01146   3.402 0.000675 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 12.49 on 4569 degrees of freedom
  (985 observations deleted due to missingness)
Multiple R-squared: 0.05989,    Adjusted R-squared: 0.05824
F-statistic: 36.38 on 8 and 4569 DF,  p-value: < 2.2e-16
```

```
#EXAMPLE 7.5 WEIGHTED LINEAR REGRESSION WITHOUT COMPLEX SAMPLE CORRECTION (SRS ASSUMPTION)


> (ex75_wt <- lm(bpxdi1_1 ~ racec + marcatc + female + agecent, data= nhanesdata, RIDAGEYR >=18, weight=WTMEC2YR
))

Call:
lm(formula = bpxdi1_1 ~ racec + marcatc + female + agecent, data = nhanesdata,     subset = RIDAGEYR >= 18,
weights = WTMEC2YR)

Coefficients:
              (Intercept)        racecOther Hispanic                 racecWhite                 racecBlack
                 70.67812                    1.78651                    2.19191                    4.40863
               racecOther  marcatcPreviously Married       marcatcNever Married                     female
                  1.95845                    0.01725                   -4.35623                   -2.99734
                  agecent
                  0.01703

> summary(ex75_wt)

Call:
lm(formula = bpxdi1_1 ~ racec + marcatc + female + agecent, data = nhanesdata,
    subset = RIDAGEYR >= 18, weights = WTMEC2YR)

Residuals:
     Min       1Q   Median       3Q      Max
-13529.1  -1457.4   -177.3   1112.6  14142.0

Coefficients:
                         Estimate Std. Error t value Pr(>|t|)
(Intercept)              70.67812    0.66677 106.001  < 2e-16 ***
racecOther Hispanic       1.78651    1.16011   1.540  0.12364
racecWhite                2.19191    0.67357   3.254  0.00115 **
racecBlack                4.40863    0.84061   5.245 1.64e-07 ***
racecOther                1.95845    1.00650   1.946  0.05174 .
marcatcPreviously Married 0.01725    0.50332   0.034  0.97266
marcatcNever Married     -4.35623    0.52403  -8.313  < 2e-16 ***
female                   -2.99734    0.36059  -8.312  < 2e-16 ***
agecent                   0.01703    0.01200   1.420  0.15576
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2462 on 4569 degrees of freedom
  (985 observations deleted due to missingness)
Multiple R-squared: 0.03903,    Adjusted R-squared: 0.03735
F-statistic:  23.2 on 8 and 4569 DF,  p-value: < 2.2e-16
```

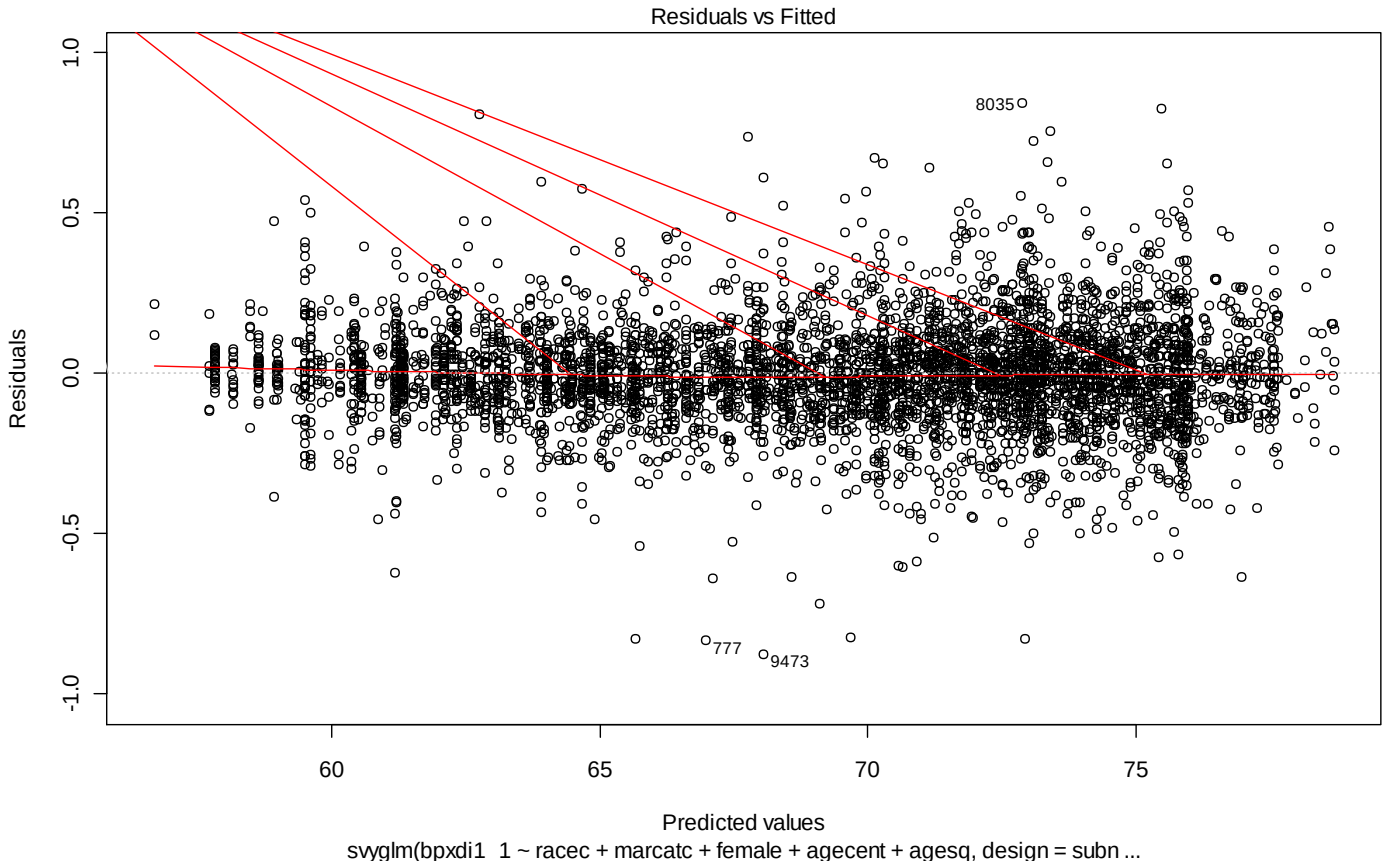
```
#EXAMPLE 7.5 WITH COMPLEX SAMPLE ADJUSTMENT AND WEIGHTS USING SVYGLM

> (ex75_svyglm <- svyglm(bpxdi1_1 ~ racec + marcatc + female + agecent, design=subnhanes))

Stratified 1 - level Cluster Sampling design (with replacement)
With (30) clusters.
subset(svynhanes, RIDAGEYR >= 18)

Call:  svyglm(bpxdi1_1 ~ racec + marcatc + female + agecent, design = subnhanes)

Coefficients:
              (Intercept)        racecOther Hispanic                 racecWhite                   racecBlack
                 70.67812                    1.78651                    2.19191                      4.40863
               racecOther  marcatcPreviously Married       marcatcNever Married                       female
                  1.95845                    0.01725                   -4.35623                     -2.99734
                  agecent
                  0.01703

Degrees of Freedom: 4577 Total (i.e. Null);  7 Residual
  (985 observations deleted due to missingness)
Null Deviance:      132.3
Residual Deviance: 127.2        AIC: 37510
> summary(ex75_svyglm)

Call:
svyglm(bpxdi1_1 ~ racec + marcatc + female + agecent, design = subnhanes)

Survey design:
subset(svynhanes, RIDAGEYR >= 18)

Coefficients:
                         Estimate Std. Error t value Pr(>|t|)
(Intercept)              70.67812    0.50076 141.141 2.36e-13 ***
racecOther Hispanic       1.78651    1.14219   1.564 0.161770
racecWhite                2.19191    0.60482   3.624 0.008464 **
racecBlack                4.40863    0.76116   5.792 0.000669 ***
racecOther                1.95845    0.98808   1.982 0.087913 .
marcatcPreviously Married 0.01725    0.71777   0.024 0.981496
marcatcNever Married     -4.35623    0.56499  -7.710 0.000115 ***
female                   -2.99734    0.33112  -9.052 4.11e-05 ***
agecent                   0.01703    0.02187   0.779 0.461500
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 127.2097)

Number of Fisher Scoring iterations: 2
```

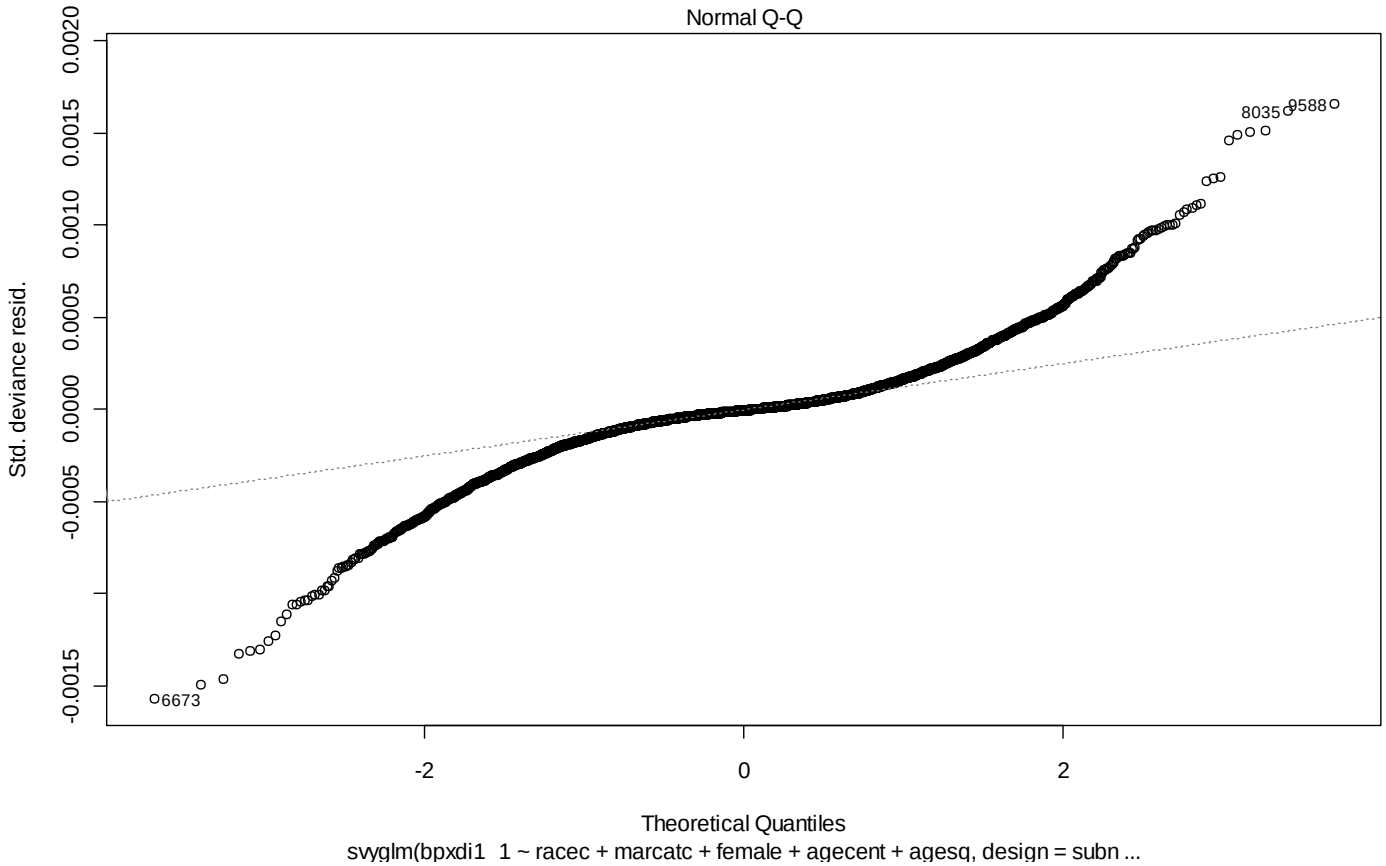
```
#ADD SELECTED PLOTS FROM DEFAULT OF PLOTS PROVIDED BY THE PLOT COMMAND

> plot(ex75_svyglm)
```




Residuals vs Fitted
2203
8035
9473
Residuals
1.0
0.5
0.0
-0.5
-1.0
64
66
68
70
72
74
76
Predicted values
svyglm(bpxdi1 1 ~ racec + marcatc + female + agecent, design = subnhanes)

```
#EXAMPLE 7.5 WITH AGE CENTERED SQUARED ADDED TO MODEL

> summary(ex75_svyglm_agesq <- svyglm(bpxdi1_1 ~ racec + marcatc + female + agecent + agesq , design=subnhanes))

Call:
svyglm(bpxdi1_1 ~ racec + marcatc + female + agecent + agesq,
    design = subnhanes)

Survey design:
subset(nhanessvy2, RIDAGEYR >= 18)

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)               73.8590162  0.4548829 162.369 3.68e-12 ***
racecOther Hispanic        1.1891589  1.0866940   1.094 0.315801
racecWhite                 1.7805528  0.6306574   2.823 0.030222 *
racecBlack                 3.4651170  0.7792454   4.447 0.004344 **
racecOther                 1.1885852  0.9341707   1.272 0.250334
marcatcPreviously Married  1.0404757  0.6217367   1.673 0.145255
marcatcNever Married      -0.3432436  0.5818098  -0.590 0.576745
female                    -2.7211812  0.3375608  -8.061 0.000195 ***
agecent                    0.1252717  0.0148188   8.454 0.000150 ***
agesq                     -0.0124771  0.0007638 -16.336 3.35e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 114.6982)

Number of Fisher Scoring iterations: 2

> ex75_svyglm_agesq
Stratified 1 - level Cluster Sampling design (with replacement)
With (30) clusters.
subset(nhanessvy2, RIDAGEYR >= 18)

Call:  svyglm(bpxdi1_1 ~ racec + marcatc + female + agecent + agesq,      design = subnhanes)

Coefficients:
              (Intercept)        racecOther Hispanic                 racecWhite                  racecBlack
                 73.85902                    1.18916                    1.78055                     3.46512
               racecOther  marcatcPreviously Married       marcatcNever Married                      female
                  1.18859                    1.04048                   -0.34324                    -2.72118
                  agecent                      agesq
                  0.12527                   -0.01248

Degrees of Freedom: 4577 Total (i.e. Null);  6 Residual
  (985 observations deleted due to missingness)
Null Deviance:      132.3
Residual Deviance: 114.7        AIC: 37040
```


Residuals vs Fitted
Residuals
Predicted values
svyglm(bpxdi1 1 ~ racec + marcatc + female + agecent + agesq, design = subn ...
8035
777
9473
1.0
0.5
0.0
-0.5
-1.0
60
65
70
75


Normal Q-Q
Std. deviance resid.
Theoretical Quantiles
svyglm(bpxdi1 1 ~ racec + marcatc + female + agecent + agesq, design = subn ...
8035
9588
6673
0.0020
0.0015
0.0010
0.0005
0.0000
-0.0005
-0.0015
-2
0
2

```
#EXAMPLE 7.5 TEST OF INTERACTION OF AGE*RACE/ETHNICITY

> ex75_raceint <- svyglm(bpxdi1_1 ~ prevmar + nevmar + female + othhis + white + black + other + agecent + agesq
+ othhis*agecent + white*agecent + black*agecent + other*agecent + othhis*agesq + white*agesq + black*agesq +
other*agesq  , subnhanes)

> summary(ex75_raceint, df.resid=Inf)

Call:
svyglm(bpxdi1_1 ~ prevmar + nevmar + female + othhis + white +
    black + other + agecent + agesq + othhis * agecent + white *
    agecent + black * agecent + other * agecent + othhis * agesq +
    white * agesq + black * agesq + other * agesq, subnhanes)

Survey design:
subset(nhanessvy2, RIDAGEYR >= 18)

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    74.220028   0.465564 159.419  < 2e-16 ***
prevmar         0.990076   0.624523   1.585 0.112891
nevmar         -0.335654   0.585910  -0.573 0.566728
female         -2.720991   0.342034  -7.955 1.79e-15 ***
othhis          0.608453   1.251902   0.486 0.626951
white           1.423877   0.566570   2.513 0.011966 *
black           3.022178   0.917176   3.295 0.000984 ***
other           0.706689   1.179878   0.599 0.549206
agecent         0.133699   0.030683   4.357 1.32e-05 ***
agesq          -0.013551   0.001130 -11.993  < 2e-16 ***
othhis:agecent  0.067328   0.077694   0.867 0.386170
white:agecent  -0.013260   0.039618  -0.335 0.737864
black:agecent   0.041140   0.036590   1.124 0.260862
other:agecent  -0.091053   0.053263  -1.709 0.087359 .
othhis:agesq    0.004039   0.003469   1.164 0.244297
white:agesq     0.001113   0.001150   0.968 0.333092
black:agesq     0.001976   0.001686   1.172 0.241101
other:agesq     0.000203   0.002907   0.070 0.944337
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 114.4961)

Number of Fisher Scoring iterations: 2

#note that Wald Test is used in regTermTest command

> regTermTest(ex75_raceint, ~othhis:agecent + white:agecent + black:agecent + other:agecent + othhis:agesq +
white:agesq + black:agesq + other:agesq , df==8)

Wald test for othhis:agecent agecent:white agecent:black agecent:other othhis:agesq white:agesq black:agesq
other:agesq
 in svyglm(bpxdi1_1 ~ prevmar + nevmar + female + othhis + white +
    black + other + agecent + agesq + othhis * agecent + white *
    agecent + black * agecent + other * agecent + othhis * agesq +
    white * agesq + black * agesq + other * agesq, subnhanes)
Chisq =  14.75220  on  8  df: p= 0.064147
```

```
# EXAMPLE OF AGE TIMES GENDER INTERACTION TEST
> ex75_sexint <- svyglm(bpxdi1_1 ~ prevmar + nevmar + female + othhis + white + black + other + agecent + agesq
+ female*agecent + female*agesq, subnhanes)


> summary(ex75_sexint)

Call:
svyglm(bpxdi1_1 ~ prevmar + nevmar + female + othhis + white +
    black + other + agecent + agesq + female * agecent + female *
    agesq, subnhanes)

Survey design:
subset(nhanessvy2, RIDAGEYR >= 18)

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    74.138327   0.567257 130.696 2.06e-08 ***
prevmar         0.907239   0.652628   1.390 0.236848
nevmar         -0.346201   0.584881  -0.592 0.585742
female         -3.237223   0.713458  -4.537 0.010518 *
othhis          1.200924   1.096066   1.096 0.334766
white           1.796412   0.631708   2.844 0.046691 *
black           3.492023   0.777452   4.492 0.010892 *
other           1.207868   0.932990   1.295 0.265129
agecent         0.117836   0.019524   6.036 0.003799 **
agesq          -0.013467   0.001287 -10.466 0.000471 ***
female:agecent  0.014012   0.027755   0.505 0.640215
female:agesq    0.001782   0.001654   1.077 0.341919
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 114.5881)

Number of Fisher Scoring iterations: 2

> regTermTest(ex75_sexint, ~female:agecent + female:agesq, df==2)

Wald test for female:agecent female:agesq
 in svyglm(bpxdi1_1 ~ prevmar + nevmar + female + othhis + white +
    black + other + agecent + agesq + female * agecent + female *
    agesq, subnhanes)
Chisq =  3.711827  on  2  df: p= 0.15631
```

```
#EXAMPLE 7.5 FINAL MODEL WITHOUT INTERACTIONS
> ex75_svyglm_agesq

Stratified 1 - level Cluster Sampling design (with replacement)
With (30) clusters.
subset(nhanessvy2, RIDAGEYR >= 18)

Call:  svyglm(bpxdi1_1 ~ racec + marcatc + female + agecent + agesq,      design = subnhanes)

Coefficients:
              (Intercept)        racecOther Hispanic                 racecWhite                  racecBlack
                 73.85902                    1.18916                    1.78055                    3.46512
               racecOther  marcatcPreviously Married       marcatcNever Married                     female
                  1.18859                    1.04048                   -0.34324                   -2.72118
                  agecent                      agesq
                  0.12527                   -0.01248

Degrees of Freedom: 4577 Total (i.e. Null);  6 Residual
  (985 observations deleted due to missingness)
Null Deviance:      132.3
Residual Deviance: 114.7        AIC: 37040


> summary(ex75_svyglm_agesq <- svyglm(bpxdi1_1 ~ racec + marcatc + female + agecent + agesq , design=subnhanes))

Call:
svyglm(bpxdi1_1 ~ racec + marcatc + female + agecent + agesq,
    design = subnhanes)

Survey design:
subset(nhanessvy2, RIDAGEYR >= 18)

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)               73.8590162  0.4548829 162.369 3.68e-12 ***
racecOther Hispanic        1.1891589  1.0866940   1.094 0.315801
racecWhite                 1.7805528  0.6306574   2.823 0.030222 *
racecBlack                 3.4651170  0.7792454   4.447 0.004344 **
racecOther                 1.1885852  0.9341707   1.272 0.250334
marcatcPreviously Married  1.0404757  0.6217367   1.673 0.145255
marcatcNever Married      -0.3432436  0.5818098  -0.590 0.576745
female                    -2.7211812  0.3375608  -8.061 0.000195 ***
agecent                    0.1252717  0.0148188   8.454 0.000150 ***
agesq                     -0.0124771  0.0007638 -16.336 3.35e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 114.6982)

Number of Fisher Scoring iterations: 2
```

```
plot(ex75_svyglm_agesq)
```