R Analysis Example Replication C7

```
# Note: all data management and initial survey design setup code included in Chapter 5 document
nhanesdata$bpxdi1_1 <-nhanesdata$BPXDI1
nhanesdata$bpxdi1_1 [nhanesdata$bpxdi1_1 ==0] <- NA
nhanesdata$agec <- (nhanesdata$age-46.36)
nhanesdata$agecsq <- (nhanesdata$agec*nhanesdata$agec)
nhanesdata$genderc <- factor(nhanesdata$RIAGENDR, levels = 1: 2, labels =c("M", "F"))

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

#EXAMPLE 7.5 BIVARIATE TESTING OF EACH FACTOR VARIABLE: RACE NHANES ADULT DATA
summary(ex75_race <- svyglm(bpxdi1_1 ~racec, design=subnhanes))
regTermTest(ex75_race, ~racec)
# EXAMPLE 7.5 BIVARIATE TEST OF MARITAL STATUS
summary(ex75_marital <- svyglm(bpxdi1_1 ~marcatc, design=subnhanes))
regTermTest(ex75_marital, ~marcatc)
# EXAMPLE 7.5 BIVARIATE TEST OF GENDER
summary(ex75_sex <- svyglm(bpxdi1_1 ~genderc, design=subnhanes))
regTermTest(ex75_sex, ~genderc)
# EXAMPLE 7.5 BIVARIATE TEST OF CENTERED AGE
summary(ex75_age <- svyglm(bpxdi1_1 ~agec, design=subnhanes))

#UNWEIGHTED OLS REGRESSION
(ex75_nowt <- lm(bpxdi1_1 ~ racec + genderc + agec, data=nhanesdata, age >=18 ))
summary(ex75_nowt)

#WEIGHTED LINEAR REGRESSION WITHOUT COMPLEX SAMPLE CORRECTION
(ex75_wt <- lm(bpxdi1_1 ~ racec + genderc + agec, data= nhanesdata, age >=18, weight=WTMEC2YR ))
summary(ex75_wt)

#EXAMPLE 7.5 WITH COMPLEX SAMPLE ADJUSTMENT AND WEIGHTS USING SVYGLM
summary(ex75_svyglm <- svyglm(bpxdi1_1 ~ racec + genderc + agec, design=subnhanes))
plot(ex75_svyglm)
#ex 7.5 with AgeC Squared
summary(ex75_svyglm_agesq <- svyglm(bpxdi1_1 ~ racec + genderc + agec + agecsq, design=subnhanes))
ex75_svyglm_agesq
plot(ex75_svyglm_agesq)
#note: additional plots could be done with more coding and plotting work, not shown here

#EXAMPLE 7.5 TEST OF INTERACTION OF AGE and AGESQUARED*RACE/ETHNICITY
ex75_raceint <- svyglm(bpxdi1_1 ~ genderc + agec*factor(racec) + agecsq*factor(racec), subnhanes)
summary(ex75_raceint, df.resid=Inf)
#note that Wald Test is used in regTermTest command
regTermTest(ex75_raceint, ~agec:factor(racec)+ agecsq:factor(racec))

# EXAMPLE 7.5 AGE TIMES GENDER INTERACTION TEST
ex75_sexint <- svyglm(bpxdi1_1 ~factor(genderc)*agec + factor(genderc)*agecsq + racec, subnhanes)
summary(ex75_sexint)
# Test of interactions, note that R uses a different df formula than Stata, see documentation for details
regTermTest(ex75_sexint, ~factor(genderc):agec + factor(genderc):agecsq)
```

```
#Final Model including interactions of race and age plus gender and age
ex75_final <- svyglm(bpxdi1_1 ~ agec*factor(racec) + agecsq*factor(racec) + factor(genderc)*agec +
factor(genderc)*agecsq, subnhanes)
summary(ex75_final, df.resid=Inf)
margins(ex75_final, at(~agec(-30,(5),30)))

#R Survey Diagnostics package from R. Valliant are currently available only directly from Dr. Valliant, request
by email at rvalliant@survey.umd.edu.

#Until the package is available from CRAN, we refer readers to examples in book rather than repeat here.

# Q Approach for Weighting, Pfefferman
# Step 1 linear model with weight regressed on race, gender and agec
q_wgt <- lm(WTMEC2YR ~ racec + genderc + agec, nhanesdata)
summary(q_wgt)
w_hat <- predict(q_wgt)
nhanesdata$q_wtmec2yr <- (nhanesdata$WTMEC2YR / w_hat)
names(nhanesdata)
summary(nhanesdata$q_wtmec2yr)
# design object and subset for analysis
nhanessvyq <- svydesign(strata=~SDMVSTRA, id=~SDMVPSU, weights=~q_wtmec2yr, data=nhanesdata, nest=T)
subnhanesq <- subset(nhanessvyq , age >= 18)
# Final Model with Q Weight
ex75_finalq <- svyglm(bpxdi1_1 ~ agec*factor(racec) + agecsq*factor(racec) + factor(genderc)*agec +
factor(genderc)*agecsq, subnhanesq)
summary(ex75_finalq, df.resid=Inf)
```

Output R Analysis Example Replication C7

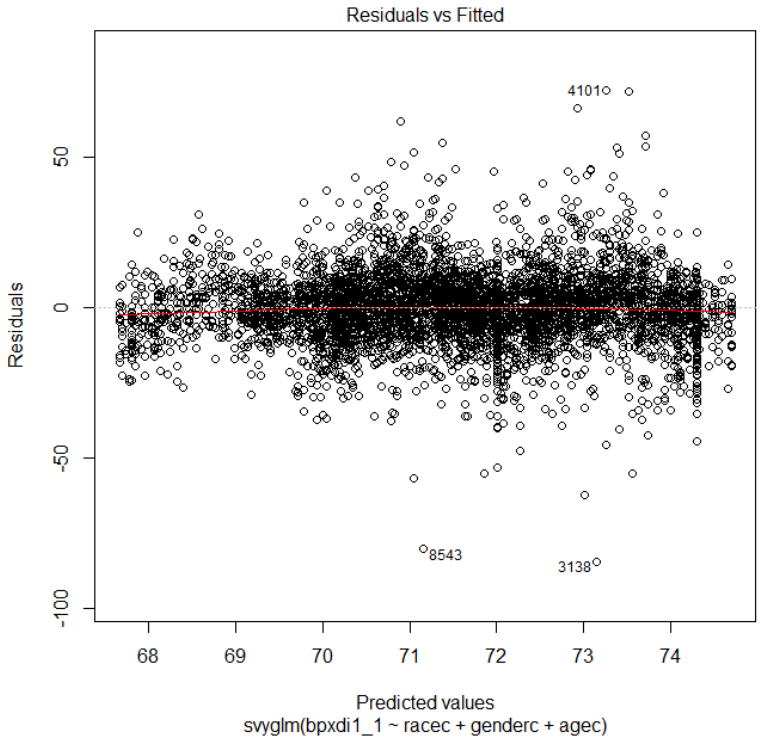
```
> #create new variable bpxdi1_1 where 0 is set to missing
> nhanesdata$bpxdi1_1 <-nhanesdata$BPXDI1
> nhanesdata$bpxdi1_1 [nhanesdata$bpxdi1_1 ==0] <- NA
> nhanesdata$agec <- (nhanesdata$age-46.36)
> nhanesdata$agecsq <- (nhanesdata$agec*nhanesdata$agec)
> nhanesdata$genderc <- factor(nhanesdata$RIAGENDR, levels = 1: 2, labels =c("M", "F"))
> nhanessvy2 <- svydesign(strata=~SDMVSTRA, id=~SDMVPSU, weights=~WTMEC2YR, data=nhanesdata, nest=T)
> subnhanes <- subset(nhanessvy2 , age >= 18)


> #EXAMPLE 7.5 BIVARIATE TESTING OF EACH FACTOR VARIABLE: RACE
> summary(ex75_race <- svyglm(bpxdi1_1 ~racec, design=subnhanes))
Call:
svyglm(formula = bpxdi1_1 ~ racec, design = subnhanes)
Survey design:
subset(nhanessvy2, age >= 18)
Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)          69.8041     0.4532 154.013  < 2e-16 ***
racecOther Hispanic  -0.1549     1.4556  -0.106  0.91688
racecWhite            2.1847     0.7427   2.942  0.01145 *
racecBlack            2.2902     0.7030   3.258  0.00623 **
racecOther            1.3056     0.7044   1.853  0.08665 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
(Dispersion parameter for gaussian family taken to be 137.0384)
Number of Fisher Scoring iterations: 2
> regTermTest(ex75_race, ~racec)
Wald test for racec
 in svyglm(formula = bpxdi1_1 ~ racec, design = subnhanes)
F =  4.771214  on  4  and  13  df: p= 0.013705


> # EXAMPLE 7.5 BIVARIATE TEST OF MARITAL STATUS
> summary(ex75_marital <- svyglm(bpxdi1_1 ~marcatc, design=subnhanes))
Call:
svyglm(formula = bpxdi1_1 ~ marcatc, design = subnhanes)
Survey design:
subset(nhanessvy2, age >= 18)
Coefficients:
                         Estimate Std. Error t value Pr(>|t|)
(Intercept)               72.1796     0.5149 140.172   <2e-16 ***
marcatcPreviously Married  -0.1451     0.6978  -0.208    0.838
marcatcNever Married       -1.1210     0.8437  -1.329    0.204
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 137.5607)
Number of Fisher Scoring iterations: 2
> regTermTest(ex75_marital, ~marcatc)
Wald test for marcatc
 in svyglm(formula = bpxdi1_1 ~ marcatc, design = subnhanes)
F =  0.9023684  on  2  and  15  df: p= 0.42653
```

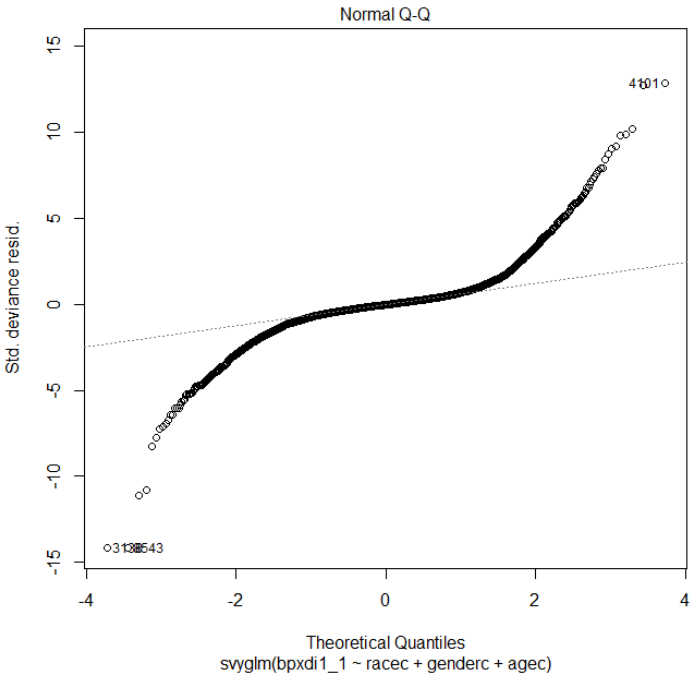
```
> # EXAMPLE 7.5 BIVARIATE TEST OF GENDER
> summary(ex75_sex <- svyglm(bpxdi1_1 ~genderc, design=subnhanes))

Call:
svyglm(formula = bpxdi1_1 ~ genderc, design = subnhanes)

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

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  72.7255     0.5901 123.245  < 2e-16 ***
gendercF     -2.2004     0.5679  -3.875  0.00134 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 136.4476)

Number of Fisher Scoring iterations: 2

> regTermTest(ex75_sex, ~genderc)
Wald test for genderc
 in svyglm(formula = bpxdi1_1 ~ genderc, design = subnhanes)
F =  15.01184  on  1  and  16  df: p= 0.0013441



> # EXAMPLE 7.5 BIVARIATE TEST OF CENTERED AGE
> summary(ex75_age <- svyglm(bpxdi1_1 ~agec, design=subnhanes))

Call:
svyglm(formula = bpxdi1_1 ~ agec, design = subnhanes)

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

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 71.60363    0.50024 143.140   <2e-16 ***
agec         0.03941    0.01889   2.087   0.0533 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 137.2234)

Number of Fisher Scoring iterations: 2
```

```
> #UNWEIGHTED OLS REGRESSION
> (ex75_nowt <- lm(bpxdi1_1 ~ racec + genderc + agec, data=nhanesdata, age >=18 ))
Call:
lm(formula = bpxdi1_1 ~ racec + genderc + agec, data = nhanesdata,
    subset = age >= 18)
Coefficients:
        (Intercept)  racecOther Hispanic           racecWhite
           70.78353               0.25519              1.19254
         racecBlack           racecOther             gendercF
            2.20541               2.01311             -2.40368
               agec
            0.04136
> summary(ex75_nowt)
Call:
lm(formula = bpxdi1_1 ~ racec + genderc + agec, data = nhanesdata,
    subset = age >= 18)
Residuals:
    Min      1Q  Median      3Q     Max
-60.964  -7.299   0.190   7.337  47.140
Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)         70.783525   0.547861 129.200  < 2e-16 ***
racecOther Hispanic  0.255192   0.737979   0.346 0.729508
racecWhite           1.192541   0.597384   1.996 0.045957 *
racecBlack           2.205414   0.615401   3.584 0.000342 ***
racecOther           2.013111   0.661517   3.043 0.002353 **
gendercF            -2.403677   0.331488  -7.251 4.75e-13 ***
agec                 0.041356   0.009034   4.578 4.81e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Residual standard error: 11.84 on 5105 degrees of freedom
  (752 observations deleted due to missingness)
Multiple R-squared:  0.01798,   Adjusted R-squared:  0.01682
F-statistic: 15.58 on 6 and 5105 DF,  p-value: < 2.2e-16
```

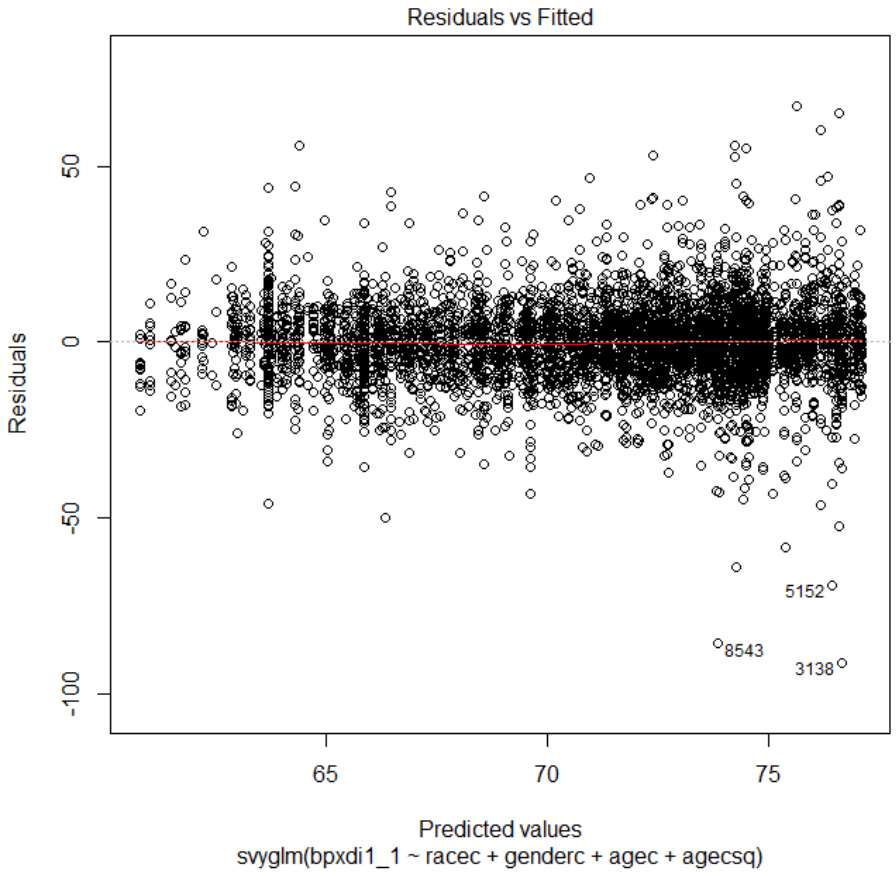
```
> #WEIGHTED LINEAR REGRESSION WITHOUT COMPLEX SAMPLE CORRECTION
> (ex75_wt <- lm(bpxdi1_1 ~ racec + genderc + agec, data= nhanesdata, age >=18, weight=WTMEC2YR ))
Call:
lm(formula = bpxdi1_1 ~ racec + genderc + agec, data = nhanesdata,
    subset = age >= 18, weights = WTMEC2YR)
Coefficients:
        (Intercept)  racecOther Hispanic           racecWhite
           71.14870             -0.14141              1.90420
         racecBlack           racecOther             gendercF
            2.30195              1.26179             -2.29114
               agec
            0.03682
> summary(ex75_wt)
Call:
lm(formula = bpxdi1_1 ~ racec + genderc + agec, data = nhanesdata,
    subset = age >= 18, weights = WTMEC2YR)
Weighted Residuals:
     Min       1Q   Median       3Q      Max
-16771.9  -1300.0    -58.2   1137.9  14319.1
Coefficients:
                     Estimate Std. Error t value Pr(>|t|)
(Intercept)         71.148697   0.591569 120.271  < 2e-16 ***
racecOther Hispanic -0.141412   0.841612  -0.168  0.86657
racecWhite           1.904199   0.607718   3.133  0.00174 **
racecBlack           2.301953   0.734488   3.134  0.00173 **
racecOther           1.261786   0.805232   1.567  0.11718
gendercF            -2.291136   0.318160  -7.201 6.84e-13 ***
agec                 0.036823   0.009289   3.964 7.47e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Residual standard error: 2315 on 5105 degrees of freedom
  (752 observations deleted due to missingness)
Multiple R-squared:  0.01742,   Adjusted R-squared:  0.01627
F-statistic: 15.09 on 6 and 5105 DF,  p-value: < 2.2e-16
```

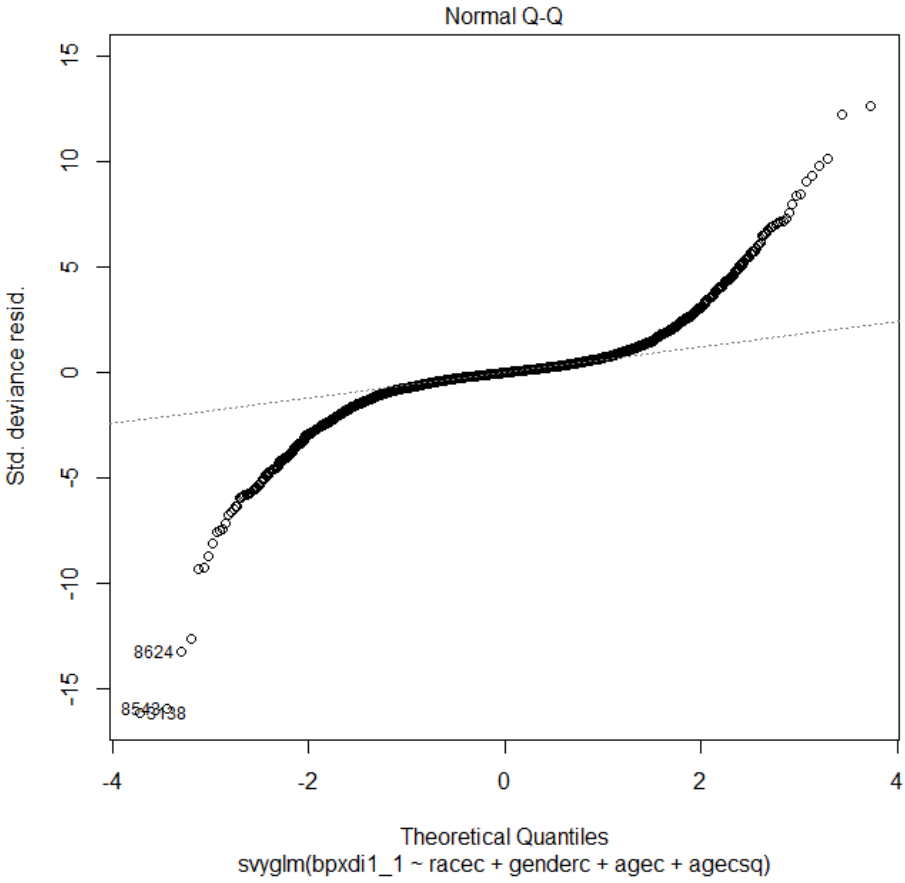
```
> #EXAMPLE 7.5 WITH COMPLEX SAMPLE ADJUSTMENT AND WEIGHTS USING SVYGLM
> summary(ex75_svyglm <- svyglm(bpxdi1_1 ~ racec + genderc + agec, design=subnhanes))

Call:
svyglm(formula = bpxdi1_1 ~ racec + genderc + agec, design = subnhanes)

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

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)         71.14870    0.51796 137.364  < 2e-16 ***
racecOther Hispanic -0.14141    1.37461  -0.103  0.91991
racecWhite           1.90420    0.80908   2.354  0.03825 *
racecBlack           2.30195    0.66462   3.464  0.00530 **
racecOther           1.26179    0.70668   1.786  0.10174
gendercF            -2.29114    0.54835  -4.178  0.00154 **
agec                 0.03682    0.02081   1.770  0.10445
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 135.3213)

Number of Fisher Scoring iterations: 2

> plot(ex75_svyglm)
```




Normal Q-Q
Std. deviance resid.
15
10
5
0
-5
-10
-15
-4
-2
0
2
4
Theoretical Quantiles
svyglm(bpxdi1_1 ~ racec + genderc + agec)

```
#ex 7.5 with AgeC Squared
summary(ex75_svyglm_agesq <- svyglm(bpxdi1_1 ~ racec + genderc + agec + agecsq, design=subnhanes))
ex75_svyglm_agesq
> #ex 7.5 with AgeC Squared
> summary(ex75_svyglm_agesq <- svyglm(bpxdi1_1 ~ racec + genderc + agec + agecsq, design=subnhanes))
Call:
svyglm(formula = bpxdi1_1 ~ racec + genderc + agec + agecsq,
    design = subnhanes)
Survey design:
subset(nhanessvy2, age >= 18)
Coefficients:
                     Estimate Std. Error t value Pr(>|t|)
(Intercept)         74.4622832  0.5652594 131.731  < 2e-16 ***
racecOther Hispanic  0.2178048  1.2171779   0.179 0.861556
racecWhite           2.0844882  0.8572122   2.432 0.035347 *
racecBlack           2.5108637  0.7336554   3.422 0.006521 **
racecOther           1.4095682  0.6873427   2.051 0.067424 .
gendercF            -2.1692000  0.4892870  -4.433 0.001267 **
agec                 0.0748534  0.0155878   4.802 0.000721 ***
agecsq              -0.0116898  0.0007179 -16.284 1.58e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
(Dispersion parameter for gaussian family taken to be 122.0121)
Number of Fisher Scoring iterations: 2

plot(ex75_svyglm_agesq)
```




Normal Q-Q
Std. deviance resid.
15
10
5
0
-5
-10
-15
8624
-4
-2
0
2
4
Theoretical Quantiles
svyglm(bpxdi1_1 ~ racec + genderc + agec + agecsq)

```
> #EXAMPLE 7.5 TEST OF INTERACTION OF AGE and AGESQUARED*RACE/ETHNICITY
> ex75_raceint <- svyglm(bpxdi1_1 ~ genderc + agec*factor(racec) + agecsq*factor(racec), subnhanes)
> summary(ex75_raceint, df.resid=Inf)

Call:
svyglm(formula = bpxdi1_1 ~ genderc + agec * factor(racec) +
    agecsq * factor(racec), subnhanes)

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

Coefficients:
                                  Estimate Std. Error t value Pr(>|t|)
(Intercept)                      74.859201   0.760761  98.400  < 2e-16 ***
gendercF                         -2.168455   0.489885  -4.426 9.58e-06 ***
agec                              0.061160   0.032933   1.857 0.063297 .
factor(racec)Other Hispanic       0.224080   0.927686   0.242 0.809131
factor(racec)White                1.398993   0.906972   1.542 0.122955
factor(racec)Black                3.341583   0.961994   3.474 0.000514 ***
factor(racec)Other                1.084784   0.899712   1.206 0.227933
agecsq                           -0.013611   0.001821  -7.476 7.69e-14 ***
agec:factor(racec)Other Hispanic  0.055855   0.047357   1.179 0.238219
agec:factor(racec)White          -0.001124   0.049801  -0.023 0.981991
agec:factor(racec)Black           0.040001   0.036043   1.110 0.267076
agec:factor(racec)Other           0.019038   0.045655   0.417 0.676686
factor(racec)Other Hispanic:agecsq  0.001209   0.003214   0.376 0.706858
factor(racec)White:agecsq         0.002960   0.001567   1.889 0.058886 .
factor(racec)Black:agecsq        -0.001948   0.001801  -1.081 0.279525
factor(racec)Other:agecsq         0.001728   0.002710   0.637 0.523843
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 121.6729)

Number of Fisher Scoring iterations: 2

> #note that Wald Test is used in regTermTest command
> regTermTest(ex75_raceint, ~agec:factor(racec)+ agecsq:factor(racec))
Wald test for agec:factor(racec) factor(racec):agecsq
 in svyglm(formula = bpxdi1_1 ~ genderc + agec * factor(racec) +
    agecsq * factor(racec), subnhanes)
F =  11.8955  on  8  and  2  df: p= 0.079828
```

```
> # EXAMPLE 7.5 AGE TIMES GENDER INTERACTION TEST
> ex75_sexint <- svyglm(bpxdi1_1 ~factor(genderc)*agec + factor(genderc)*agecsq + racec, subnhanes)
> summary(ex75_sexint)

Call:
svyglm(formula = bpxdi1_1 ~ factor(genderc) * agec + factor(genderc) *
    agecsq + racec, subnhanes)

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

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)             74.9846596  0.6457104 116.127 3.38e-14 ***
factor(genderc)F        -3.1707086  0.7572255  -4.187  0.00305 **
agec                     0.0481536  0.0163147   2.952  0.01838 *
agecsq                  -0.0135697  0.0008394 -16.165 2.15e-07 ***
racecOther Hispanic      0.2056088  1.2087808   0.170  0.86916
racecWhite               2.0990068  0.8453388   2.483  0.03793 *
racecBlack               2.5401774  0.7328015   3.466  0.00849 **
racecOther               1.4274416  0.6919988   2.063  0.07304 .
factor(genderc)F:agec    0.0476044  0.0229564   2.074  0.07182 .
factor(genderc)F:agecsq  0.0033007  0.0016223   2.035  0.07631 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 121.4838)

Number of Fisher Scoring iterations: 2

> # Test of interactions, note that R uses a different df formula than Stata, see documentation for details
> regTermTest(ex75_sexint, ~factor(genderc):agec + factor(genderc):agecsq)
Wald test for factor(genderc):agec factor(genderc):agecsq
 in svyglm(formula = bpxdi1_1 ~ factor(genderc) * agec + factor(genderc) *
    agecsq + racec, subnhanes)
F =  5.356873  on  2  and  8  df: p= 0.033398
```

```
> #Final Model including interactions of race and age plus gender and age
> ex75_final <- svyglm(bpxdi1_1 ~ agec*factor(racec) + agecsq*factor(racec) + factor(genderc)*agec +
factor(genderc)*agecsq, subnhanes)
> summary(ex75_final, df.resid=Inf)

Call:
svyglm(formula = bpxdi1_1 ~ agec * factor(racec) + agecsq * factor(racec) +
    factor(genderc) * agec + factor(genderc) * agecsq, subnhanes)

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

Coefficients:
                                  Estimate Std. Error t value Pr(>|t|)
(Intercept)                     75.3464498  0.8190674  91.991  < 2e-16 ***
agec                             0.0392266  0.0397727   0.986 0.324002
factor(racec)Other Hispanic      0.2714371  0.9210014   0.295 0.768208
factor(racec)White               1.4611713  0.9104608   1.605 0.108522
factor(racec)Black               3.4500173  0.9610532   3.590 0.000331 ***
factor(racec)Other               1.1441363  0.8948624   1.279 0.201052
agecsq                          -0.0152356  0.0018081  -8.426  < 2e-16 ***
factor(genderc)F                -3.1953718  0.7592912  -4.208 2.57e-05 ***
agec:factor(racec)Other Hispanic 0.0496377  0.0496923   0.999 0.317843
agec:factor(racec)White         -0.0044755  0.0531791  -0.084 0.932931
agec:factor(racec)Black          0.0345488  0.0387275   0.892 0.372339
agec:factor(racec)Other          0.0149059  0.0492193   0.303 0.762006
factor(racec)Other Hispanic:agecsq  0.0008365  0.0034480   0.243 0.808301
factor(racec)White:agecsq        0.0026623  0.0017149   1.552 0.120562
factor(racec)Black:agecsq       -0.0023680  0.0019817  -1.195 0.232102
factor(racec)Other:agecsq        0.0014446  0.0029821   0.484 0.628085
agec:factor(genderc)F            0.0454944  0.0234701   1.938 0.052575 .
agecsq:factor(genderc)F          0.0033864  0.0016627   2.037 0.041683 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 121.1483)

Number of Fisher Scoring iterations: 2
```

#NOTE: Predicted Margins with Continuous Variable not currently available in R. This feature may be added in near future by software developer and will be included on ASDA website if this occurs.

#R Survey Diagnostics package from R. Valliant are currently available only directly from Dr. Valliant, request by email at rvalliant@survey.umd.edu.
#Until the package is available from CRAN, we refer readers to examples in book rather than repeat here.

```
> # Q Approach for Weighting, Pfefferman
> # Step 1 linear model with weight regressed on race, gender and agec
>
> q_wgt <- lm(WTMEC2YR ~ racec + genderc + agec, nhanesdata)
> summary(q_wgt)

Call:
lm(formula = WTMEC2YR ~ racec + genderc + agec, data = nhanesdata)

Residuals:
   Min     1Q Median     3Q    Max
-72371  -9694  -1784   5723 160998

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)         24948.55     826.84  30.173  < 2e-16 ***
racecOther Hispanic -3410.17    1105.37  -3.085  0.00204 **
racecWhite          40374.19     901.34  44.794  < 2e-16 ***
racecBlack          -9093.88     903.70 -10.063  < 2e-16 ***
racecOther          -8421.65     989.51  -8.511  < 2e-16 ***
gendercF             1716.86     546.43   3.142  0.00168 **
agec                  158.50      11.35  13.964  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 26980 on 9749 degrees of freedom
Multiple R-squared:  0.413,     Adjusted R-squared:  0.4126
F-statistic:  1143 on 6 and 9749 DF,  p-value: < 2.2e-16

> w_hat <- predict(q_wgt)
>
> nhanesdata$q_wtmec2yr <- (nhanesdata$WTMEC2YR / w_hat)

> # design object and subset for analysis
> nhanessvyq <- svydesign(strata=~SDMVSTRA, id=~SDMVPSU, weights=~q_wtmec2yr, data=nhanesdata, nest=T)
> subnhanesq <- subset(nhanessvyq , age >= 18)
>
> # Final Model with Q Weight
> ex75_finalq <- svyglm(bpxdi1_1 ~ agec*factor(racec) + agecsq*factor(racec) + factor(genderc)*agec +
factor(genderc)*agecsq, subnhanesq)
> summary(ex75_finalq, df.resid=Inf)

Call:
svyglm(formula = bpxdi1_1 ~ agec * factor(racec) + agecsq * factor(racec) +
    factor(genderc) * agec + factor(genderc) * agecsq, subnhanesq)

Survey design:
subset(nhanessvyq, age >= 18)

Coefficients:
                                  Estimate Std. Error t value Pr(>|t|)
(Intercept)                      75.413392   0.772493  97.623  < 2e-16 ***
agec                              0.046734   0.040405   1.157 0.247420
factor(racec)Other Hispanic       0.247589   0.953863   0.260 0.795200
factor(racec)White                1.500999   0.894770   1.678 0.093440 .
factor(racec)Black                3.566390   0.989244   3.605 0.000312 ***
factor(racec)Other                1.237697   0.892330   1.387 0.165429
agecsq                           -0.014820   0.001685  -8.794  < 2e-16 ***
factor(genderc)F                 -3.429000   0.632731  -5.419 5.98e-08 ***
agec:factor(racec)Other Hispanic  0.048245   0.047486   1.016 0.309641
agec:factor(racec)White          -0.005508   0.051294  -0.107 0.914481
```

```
agec:factor(racec)Black             0.035959   0.037019   0.971 0.331364
agec:factor(racec)Other             0.012669   0.046463   0.273 0.785110
factor(racec)Other Hispanic:agecsq  0.001037   0.003368   0.308 0.758164
factor(racec)White:agecsq           0.002517   0.001712   1.470 0.141460
factor(racec)Black:agecsq          -0.002684   0.001961  -1.368 0.171237
factor(racec)Other:agecsq           0.001244   0.003058   0.407 0.684076
agec:factor(genderc)F               0.034390   0.026440   1.301 0.193368
agecsq:factor(genderc)F             0.002924   0.001646   1.776 0.075704 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 123.0622)

Number of Fisher Scoring iterations: 2
```