R Analysis Example Replication C6

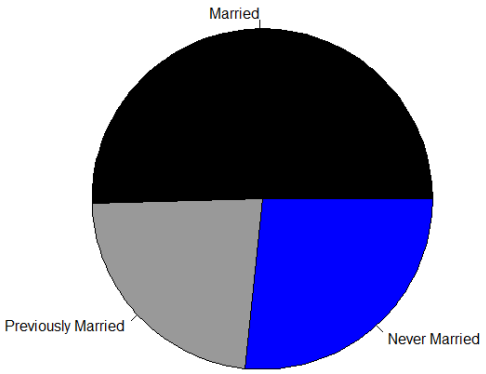
```
# Note: all data management and survey design setup code included in Chapter 5 document
# ASDA2 Chapter 6 analysis examples replication
# Example 6.1
(ex61 <- svymean(~factor(irregular), subnhanes, se=T, na.rm=T, deff=T, ci=T, keep.vars=T))
confint(ex61)
ex61p <- svyciprop(~factor(irregular), subnhanes, se=T, na.rm=T, deff=T, ci=T, keep.vars=T)
ex61p

# Example 6.2 NHANES ADULT DATA
ex62 <- svymean(~factor(racec), design=subnhanes, se=T, na.rm=T, deff=T, ci=T, keep.vars=T))
ex62
confint(ex62)

# Example 6.3 NHANES ADULT DATA
ex63 <- svymean(~factor(bp_catc), subnhanes, se=T, na.rm=T, deff=T, ci=T, keep.vars=T)
ex63
confint(ex63)



# EXAMPLE 6.4 ESS6 Russian Federation Data, Proportions of Russian 15+ by Marital Status
# GOF TOOL WITH PRE-SET PROPORTIONS (NOT AVAILABLE IN R)

rfdata <- read_sas("P:/ASDA 2/Data sets/ess6 russia/ess6_russia_20aug2016.sas7bdat")
summary(rfdata)

#create factor variables
rfdata$marcatc <- factor(rfdata$marcat, levels = 1:3, labels =c("Married", "Previous", "Never"))

rfsvy <- svydesign(strata=~stratify, id=~psu, weights=~PSPWGHT, data=rfdata, nest=T)

ex6_4 <- svymean(~factor(marcatc), design=rfsvy, na.rm=T, se=T, deff=T, ci=T, keep.vars=T)
print(ex6_4)

# Analysis Example 6.5 PIE AND BAR CHARTS

# Pie of Marital Status Russian Federation Data
ex6_5 <- svymean(~factor(marcatc), rfsvy, se=T, na.rm=T, deff=T, ci=T, keep.vars=T)
pie(ex6_5, col=c("black", "grey60", "blue", "red"), c("Married", "Previously Married", "Never Married"))

# Bar chart of marital status
barplot(ex6_5, legend=c("Married", "Previously Married", "Never Married"), col=c("black","blue", "red"))

# Analysis Example 6.6, NCS-R DATA
(ex6_6 <- svymean(~interaction (SEX, mde), ncsrsvyp1, se=T, na.rm=T, ci=T, keep.vars=T))
# obtain confidence intervals
confint(ex6_6)
# svyby analysis gives mean of mde by sex
(ex6_6a <- svyby(~mde, ~SEX, ncsrsvyp1, svymean, se=T, na.rm=T, ci=T, keep.vars=T))

#CODES FOR SEX 1=MALE 2=FEMALE
#svychisq provides a 2 by 2 chisq test (F)
svychisq(~SEX+mde, ncsrsvyp1, statistic="F")
```

```
# Analysis Example 6.7 MEAN OF MDE OVER SEX AND LINEAR COMPARISON TEST
(ex6_7 <- svyby(~mde, ~sexc, ncsrsvyp1, svymean, se=T, na.rm=T, ci=T, keep.vars=T))
svycontrast(ex6_7,list(avg=c(.5,.5), diff=c(1,-1)))

# Analysis Example 6.8 Independence of MDE and Gender
(ex6_8 <- svyby(~mde, ~SEX, ncsrsvyp1, svymean, se=T, na.rm=T, ci=T, keep.vars=T))
#CODES FOR SEX 1=MALE 2=FEMALE
#svychisq provides a 2 by 2 chisq test (F)
svychisq(~SEX+mde, ncsrsvyp1, statistic="F")


# Analysis Example 6.9 Independence of Education and Alcohol Dependence
ex6_9 <- svyby (~ald,~ed4catc, subset(ncsrsvyp2, AGE < 29 & !is.na(ED4CAT) & !is.na(ald)), svymean, na.rm=T,
ci=T)

# CODES FOR ED4CAT 1=0-11 2=12 3=13-15 4=16+ YEARS OF EDUCATION
print(ex6_9)
summary(ex6_9,statistic="ChiSq")
svychisq(~ald+ ed4catc, subset(ncsrsvyp2, AGE < 29  & !is.na(ED4CAT) & !is.na(ald)), na.rm=T, statistic = "F")

# Analysis Example 6.10 Logistic Regression MDE regressed on Gender
(ex6_10 <- svyglm (mde~sexm, design=ncsrsvyp1, family=quasibinomial))
summary(ex6_10)
#note can use exponent function with beta to obtain OR

# Figure 6.8 Bar Chart of Marital Status in Russian Federation data
fig6_8 <- svyby(~factor(marcatc), ~GNDR, rfsvy, svymean, na.rm=T)

# Bar chart of marital status by gender
print(fig6_8)
barplot(fig6_8, legend=c("Married", "Previously Married", "Never Married"), col=c("black", "blue", "red"),
xlab=c("Male","Female") )

# Analysis Example 6.11 Independence of Gender and MDE while controlling for Age, not available in R Survey
Package

# Analysis Example 6.12 Loglinear Model Examining Relationship Between MDE and Male
#null model run first
null <-svyloglin(~mde+sexm,ncsrsvyp1)
summary(null)

# saturated model, update null model with interaction of mde and sexm
saturated <- update(null, ~.+mde:sexm)
summary(saturated)

# obtain F test
svychisq(~mde+sexm, ncsrsvyp1)
# obtain ChiSQ test
svychisq(~mde+sexm, ncsrsvyp1, statistic="Chisq")
```

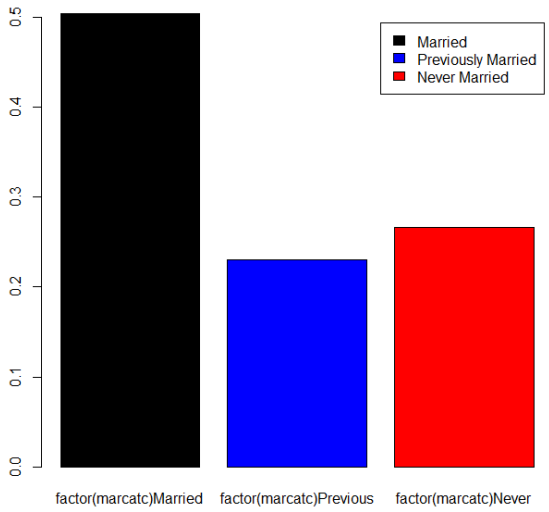
Output R Analysis Example Replication C6

```
> # Example 6.1
> (ex61 <- svymean(~factor(irregular), subnhanes, se=T, na.rm=T, deff=T, ci=T, keep.vars=T))
                        mean        SE   DEff
factor(irregular)0 0.9835810 0.0016779 0.9367
factor(irregular)1 0.0164190 0.0016779 0.9367
> confint(ex61)
                        2.5 %     97.5 %
factor(irregular)0 0.98029247 0.98686955
factor(irregular)1 0.01313045 0.01970753
> ex61p <- svyciprop(~factor(irregular), subnhanes, se=T, na.rm=T, deff=T, ci=T, keep.vars=T)
> ex61p
                           2.5% 97.5%
factor(irregular) 0.0164 0.0132  0.02

> # Example 6.2 NHANES ADULT DATA
> ex62 <- svymean(~factor(racec), design=subnhanes, se=T, na.rm=T, deff=T, ci=T, keep.vars=T)
> ex62
                                mean       SE    DEff
factor(racec)Mexican        0.079168 0.017251 22.9171
factor(racec)Other Hispanic 0.066224 0.015193 20.9551
factor(racec)White          0.659386 0.038892 37.8086
factor(racec)Black          0.117185 0.023370 29.6387
factor(racec)Other          0.078037 0.010917  9.3001
> confint(ex62)
                                 2.5 %     97.5 %
factor(racec)Mexican        0.04535738 0.11297837
factor(racec)Other Hispanic 0.03644730 0.09600144
factor(racec)White          0.58316044 0.73561243
factor(racec)Black          0.07138031 0.16298887
factor(racec)Other          0.05663959 0.09943387


> # Example 6.3 NHANES ADULT DATA
> ex63 <- svymean(~factor(bp_catc), subnhanes, se=T, na.rm=T, deff=T, ci=T, keep.vars=T)
> ex63
                                mean        SE   DEff
factor(bp_catc)Normal      0.4722233 0.0155209 5.1762
factor(bp_catc)Pre-HBP     0.4279854 0.0120353 3.1685
factor(bp_catc)Stage 1 HBP 0.0797780 0.0058154 2.4669
factor(bp_catc)Stage 2 HBP 0.0200133 0.0043847 5.2494
> confint(ex63)
                                 2.5 %     97.5 %
factor(bp_catc)Normal      0.44180279 0.50264372
factor(bp_catc)Pre-HBP     0.40439669 0.45157418
factor(bp_catc)Stage 1 HBP 0.06837999 0.09117605
factor(bp_catc)Stage 2 HBP 0.01141946 0.02860712
```

```
> # Example 6.4 ESS6 Russian Federation Data, Proportions of Russian 15+ by Marital Status
> # GOF TOOL WITH PRE-SET PROPORTIONS (NOT AVAILABLE IN R)

> rfdata <- read_sas("P:/ASDA 2/Data sets/ess6 russia/ess6_russia_20aug2016.sas7bdat")
> #create factor variables
> rfdata$marcatc <- factor(rfdata$marcat, levels = 1:3, labels =c("Married", "Previous", "Never"))

> rfsvy <- svydesign(strata=~stratify, id=~psu, weights=~PSPWGHT, data=rfdata, nest=T)

> ex6_4 <- svymean(~factor(marcatc), design=rfsvy, na.rm=T, se=T, deff=T, ci=T, keep.vars=T)
> print(ex6_4)
                            mean       SE   DEff
factor(marcatc)Married  0.503860 0.012878 442.85
factor(marcatc)Previous 0.230066 0.011536 501.57
factor(marcatc)Never    0.266074 0.013401 613.97

> # Pie of Marital Status Russian Federation Data
> ex6_5 <- svymean(~factor(marcatc), rfsvy, se=T, na.rm=T, deff=T, ci=T, keep.vars=T)
> pie(ex6_5, col=c("black", "grey60", "blue", "red"), c("Married", "Previously Married", "Never Married"))
```



```
> # Bar chart of marital status
> barplot(ex6_5, legend=c("Married", "Previously Married", "Never Married"), col=c("black","blue", "red"))
```



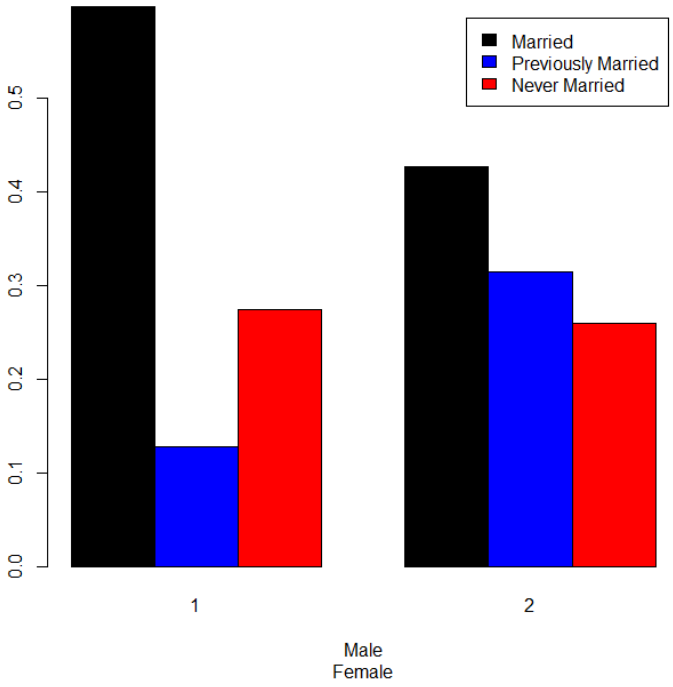
```
> # Analysis Example 6.6, NCS-R DATA
> (ex6_6 <- svymean(~interaction (SEX, mde), ncsrsvyp1, se=T, na.rm=T, ci=T, keep.vars=T))
                             mean     SE
interaction(SEX, mde)1.0 0.406644 0.0070
interaction(SEX, mde)2.0 0.401644 0.0054
interaction(SEX, mde)1.1 0.072208 0.0034
interaction(SEX, mde)2.1 0.119504 0.0030
> # obtain confidence intervals
> confint(ex6_6)
                               2.5 %     97.5 %
interaction(SEX, mde)1.0 0.39296383 0.42032513
interaction(SEX, mde)2.0 0.39113771 0.41215085
interaction(SEX, mde)1.1 0.06546993 0.07894551
interaction(SEX, mde)2.1 0.11356911 0.12543793

> # svyby analysis gives mean of mde by sex
> (ex6_6a <- svyby(~mde, ~SEX, ncsrsvyp1, svymean, se=T, na.rm=T, ci=T, keep.vars=T))
  SEX       mde          se
1   1 0.1507933 0.007747811
2   2 0.2293083 0.005647255
>
> #CODES FOR SEX 1=MALE 2=FEMALE
> #svychisq provides a 2 by 2 chisq test (F)
> svychisq(~SEX+mde, ncsrsvyp1, statistic="F")
        Pearson's X^2: Rao & Scott adjustment
data:  svychisq(~SEX + mde, ncsrsvyp1, statistic = "F")
F = 57.978, ndf = 1, ddf = 42, p-value = 1.947e-09



> # Analysis Example 6.7 MEAN OF MDE OVER SEX AND LINEAR COMPARISON TEST
> (ex6_7 <- svyby(~mde, ~sexc, ncsrsvyp1, svymean, se=T, na.rm=T, ci=T, keep.vars=T))
         sexc       mde          se
Male     Male 0.1507933 0.007747811
Female Female 0.2293083 0.005647255
> svycontrast(ex6_7,list(avg=c(.5,.5), diff=c(1,-1)))
      contrast     SE
avg   0.190051 0.0048
diff -0.078515 0.0096
Warning message:
In vcov.svyby(stat) : Only diagonal elements of vcov() available
```

```
> # Analysis Example 6.8 Independence of MDE and Gender
> (ex6_8 <- svyby(~mde, ~SEX, ncsrsvyp1, svymean, se=T, na.rm=T, ci=T, keep.vars=T))
  SEX       mde          se
1   1 0.1507933 0.007747811
2   2 0.2293083 0.005647255
> #CODES FOR SEX 1=MALE 2=FEMALE
> #svychisq provides a 2 by 2 chisq test (F)
> svychisq(~SEX+mde, ncsrsvyp1, statistic="F")
        Pearson's X^2: Rao & Scott adjustment
data:  svychisq(~SEX + mde, ncsrsvyp1, statistic = "F")
F = 57.978, ndf = 1, ddf = 42, p-value = 1.947e-09
```

```
> # Analysis Example 6.9 Independence of Education and Alcohol Dependence
> ex6_9 <- svyby (~ald,~ed4catc, subset(ncsrsvyp2, AGE < 29 & !is.na(ED4CAT) & !is.na(ald)), svymean, na.rm=T,
ci=T)
>
> # CODES FOR ED4CAT 1=0-11 2=12 3=13-15 4=16+ YEARS OF EDUCATION
> print(ex6_9)
      ed4catc        ald         se
0-11     0-11 0.09128575 0.02937999
12         12 0.04855850 0.01345971
13-15   13-15 0.04895775 0.01004206
16+       16+ 0.06903765 0.01364029
> summary(ex6_9,statistic="ChiSq")
  ed4catc       ald                se
 0-11 :1   Min.   :0.04856   Min.   :0.01004
 12   :1   1st Qu.:0.04886   1st Qu.:0.01261
 13-15:1   Median :0.05900   Median :0.01355
 16+  :1   Mean   :0.06446   Mean   :0.01663
           3rd Qu.:0.07460   3rd Qu.:0.01758
           Max.   :0.09129   Max.   :0.02938
> svychisq(~ald+ ed4catc, subset(ncsrsvyp2, AGE < 29  & !is.na(ED4CAT) & !is.na(ald)), na.rm=T, statistic = "F")
        Pearson's X^2: Rao & Scott adjustment
data:  svychisq(~ald + ed4catc, subset(ncsrsvyp2, AGE < 29 & !is.na(ED4CAT) &     !is.na(ald)), na.rm = T,
statistic = "F")
F = 1.6498, ndf = 2.7506, ddf = 112.7800, p-value = 0.1858
```

```
> # Analysis Example 6.10 Logistic Regression MDE regressed on Gender
> (ex6_10 <- svyglm (mde~sexm, design=ncsrsvyp1, family=quasibinomial))
Stratified 1 - level Cluster Sampling design (with replacement)
With (84) clusters.
svydesign(strata = ~SESTRAT, id = ~SECLUSTR, weights = ~NCSRWTSH,
    data = ncsr, nest = T)
Call:  svyglm(formula = mde ~ sexm, design = ncsrsvyp1, family = quasibinomial)
Coefficients:
(Intercept)         sexm
    -1.2122      -0.5162
Degrees of Freedom: 9281 Total (i.e. Null);  41 Residual
Null Deviance:      9072
Residual Deviance: 8979         AIC: NA
> summary(ex6_10)
Call:
svyglm(formula = mde ~ sexm, design = ncsrsvyp1, family = quasibinomial)
Survey design:
svydesign(strata = ~SESTRAT, id = ~SECLUSTR, weights = ~NCSRWTSH,
    data = ncsr, nest = T)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.21222    0.03195 -37.935  < 2e-16 ***
sexm        -0.51617    0.06820  -7.568 2.63e-09 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
(Dispersion parameter for quasibinomial family taken to be 1.000108)
Number of Fisher Scoring iterations: 4
> #note can use exponent function with beta to obtain OR
```

```
> # Figure 6.8 Bar Chart of Marital Status in Russian Federation data
> fig6_8 <- svyby(~factor(marcatc), ~GNDR, rfsvy, svymean, na.rm=T)

> # Bar chart of marital status by gender
> print(fig6_8)
  GNDR factor(marcatc)Married factor(marcatc)Previous factor(marcatc)Never se.factor(marcatc)Married
se.factor(marcatc)Previous se.factor(marcatc)Never
1    1              0.5972364              0.1281528           0.2746108               0.01707498
0.01197434              0.01631450
2    2              0.4261892              0.3148384           0.2589724               0.01632922
0.01610291              0.01722477
> barplot(fig6_8, legend=c("Married", "Previously Married", "Never Married"), col=c("black", "blue", "red"),
xlab=c("Male","Female") )
```




```
> # Analysis Example 6.11 Independence of Gender and MDE while controlling for Age, not available in R Survey
Package
```

```
> # Analysis Example 6.12 Loglinear Model Examining Relationship Between MDE and Male
> #null model run first
> null <-svyloglin(~mde+sexm,ncsrsvyp1)
> summary(null)
Loglinear model: svyloglin(~mde + sexm, ncsrsvyp1)
            coef         se            p
mde1  0.71946455 0.01573591 0.000000e+00
sexm1 0.04232084 0.01065001 7.073982e-05
>
> # saturated model, update null model with interaction of mde and sexm
> saturated <- update(null, ~.+mde:sexm)
> summary(saturated)
Loglinear model: update(null, ~. + mde:sexm)
                coef         se            p
mde1        0.7351533 0.01716099 0.000000e+00
sexm1       0.1228566 0.01151009 1.349733e-26
mde1:sexm1 -0.1290428 0.01705080 3.786148e-14
>
> # obtain F test
> svychisq(~mde+sexm, ncsrsvyp1)
        Pearson's X^2: Rao & Scott adjustment
data:  svychisq(~mde + sexm, ncsrsvyp1)
F = 57.978, ndf = 1, ddf = 42, p-value = 1.947e-09
> # obtain ChiSQ test
> svychisq(~mde+sexm, ncsrsvyp1, statistic="Chisq")
        Pearson's X^2: Rao & Scott adjustment
data:  svychisq(~mde + sexm, ncsrsvyp1, statistic = "Chisq")
X-squared = 92.15, df = 1, p-value = 2.65e-14
```