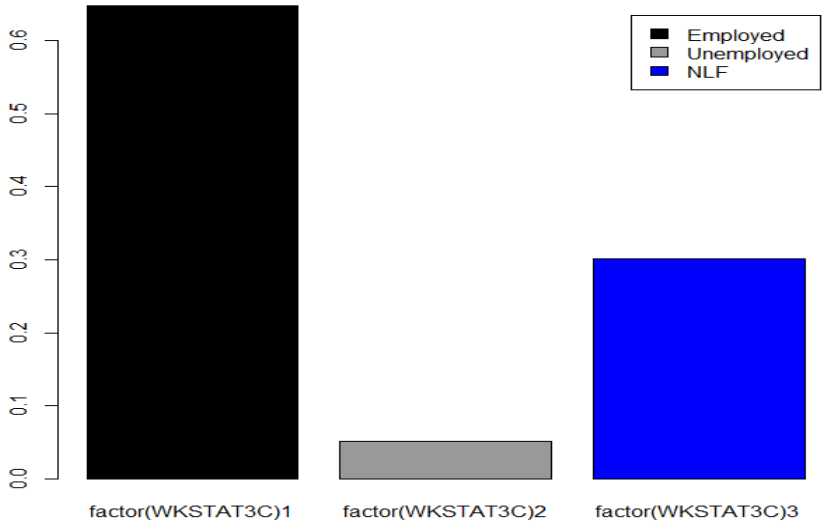
R Analysis Example Replication C9

```
# Note: all data management and survey design setup code included in Chapter 5 document
# Chapter 9 ASDA2 Analysis Examples Replication

# Figure 9.2 Bar chart of work status NCS-R data
fig92 <- svymean( ~factor(WKSTAT3C), ncsrsvyp2, na.rm=T)
barplot(fig92, legend=c("Employed", "Unemployed", "NLF") , col=c("black", "grey60", "blue"))

# Tests for Potential Predictors of Work Status
svychisq(~WKSTAT3C+SEX, ncsrsvyp2, statistic="F")
svychisq(~WKSTAT3C+ald, ncsrsvyp2, statistic="F")
svychisq(~WKSTAT3C+mde, ncsrsvyp2, statistic="F")
svychisq(~WKSTAT3C+ED4CAT, ncsrsvyp2, statistic="F")
svychisq(~WKSTAT3C+ag4cat, ncsrsvyp2, statistic="F")
svychisq(~WKSTAT3C+MAR3CAT, ncsrsvyp2, statistic="F")

# Note: No Multinomial Logit Option in R Survey package

# Ordinal Regression with Russian Federation Data
# Read Data and set design variables
# Use C9 version of data for this example
rfdata_c9 <- read_sas("P:/ASDA 2/Data sets/ess6 russia/c9_russia_1jun2017.sas7bdat")
summary(rfdata_c9)

#create factor variables
rfdata_c9$marcatc <- factor(rfdata_c9$marcat, levels = 1:3, labels =c("Married", "Previous", "Never"))
rfsvy <- svydesign(strata=~stratify, id=~psu, weights=~PSPWGHT, data=rfdata_c9, nest=T)

ex936 <- svymean(~factor(stflife2), design=rfsvy, na.rm=T, se=T, deff=T, ci=T, keep.vars=T)
print(ex936)
barplot(ex936, legend=c("0-1", "2-4", "5", "6-8", "9-10") , col=c("black", "grey60", "blue", "red", "green"))

#ordinal logistic using satisfaction with life
summary(ex936_ordinal <- svyolr (factor(stflife2) ~ factor(agecat) + male + marcatc, design=rfsvy))
exp(ex936_ordinal$coef)


# HRS data for Poisson Regression using Number of Falls
#histogram of number of falls 24
svyhist(~numfalls24 , subset (hrssvyr, NAGE >=65), main="", col="grey80", xlab ="Histogram of Number of Falls
Past 24 Months")

# Poisson model
summary(hrs)

# Create variables for example
hrs$nage_c=(hrs$NAGE - 74.5)
hrs$bmi_c=(hrs$R11BMI - 27.7)
hrs$male=(hrs$GENDER ==1)

# Create Design Object and then subset for 65 Plus
hrssvyr <- svydesign(strata=~STRATUM, id=~SECU, weights=~NWGTR , data=hrs, nest=T)
```

```
summary(hrssvyr)
hrssvy65 <- subset(hrssvyr, age65p==1)

summary(ex947_poisson <- svyglm(numfalls24 ~ male + nage_c + arthritis + diabetes + bmi_c, design=hrssvy65,
family=quasipoisson(log)))
exp(ex947_poisson$coef)


# Negative binomial (not available with survey correction, dispersion is accounted for in svyglm, per lumley)
# Zero inflated negative binomial not available in R survey package
```

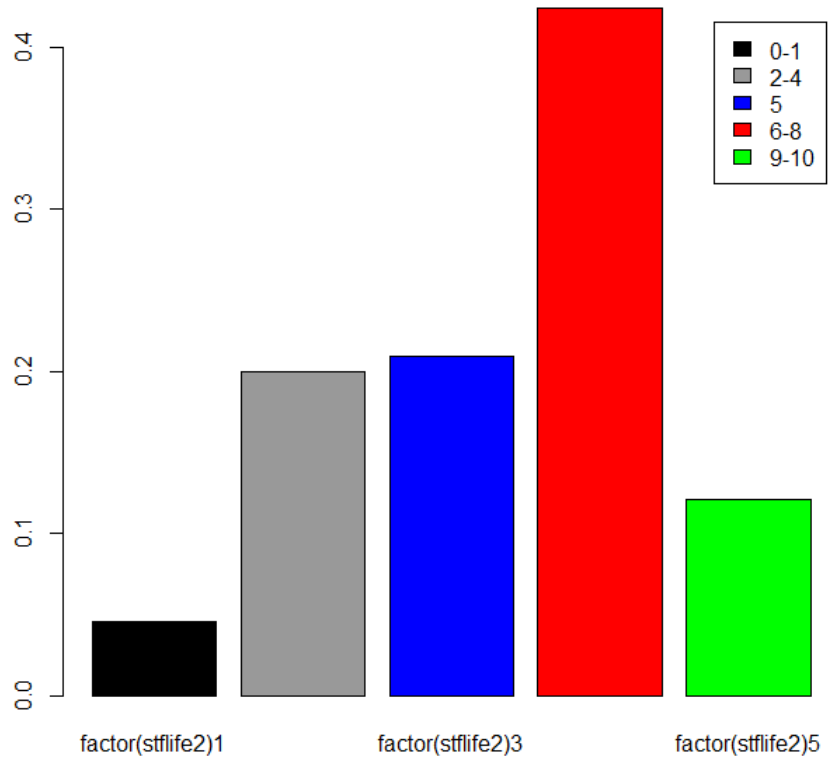
Output R Analysis Example Replication C9

```
# Chapter 9 ASDA2 Analysis Examples Replication

> # Figure 9.2 Bar chart of work status NCS-R data

> fig92 <- svymean( ~factor(WKSTAT3C), ncsrsvyp2, na.rm=T)
> barplot(fig92, legend=c("Employed", "Unemployed", "NLF") , col=c("black", "grey60", "blue"))
```



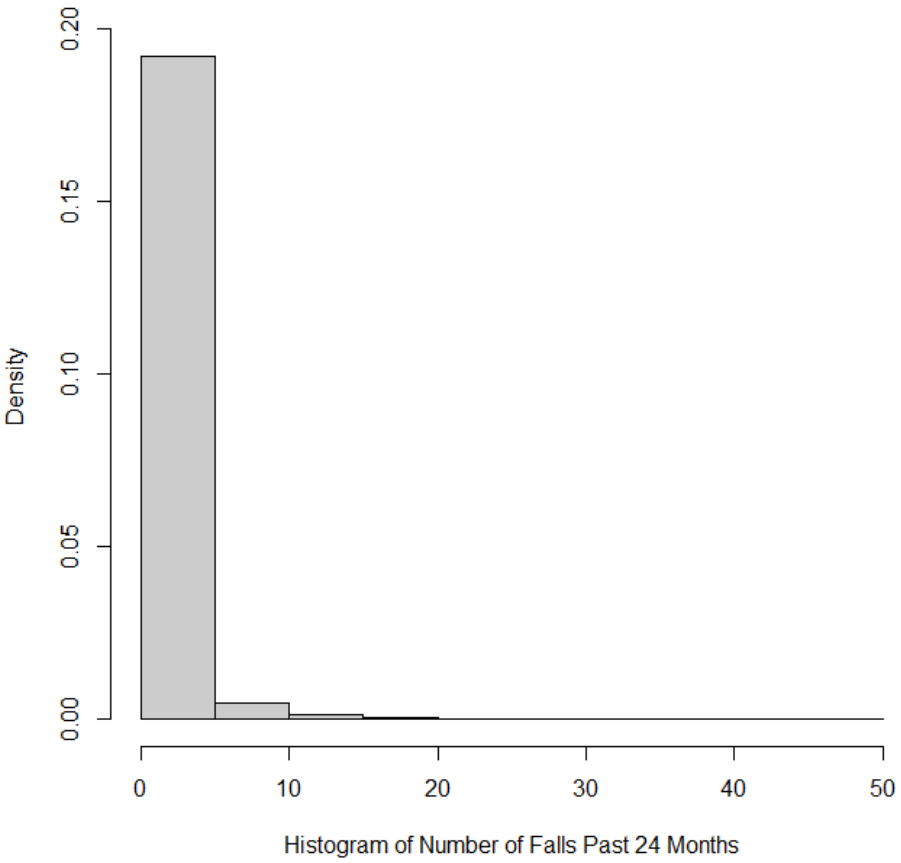
```
# Tests for Potential Predictors of Work Status
> svychisq(~WKSTAT3C+SEX, ncsrsvyp2, statistic="F")
        Pearson's X^2: Rao & Scott adjustment
data:  svychisq(~WKSTAT3C + SEX, ncsrsvyp2, statistic = "F")
F = 27.329, ndf = 1.875, ddf = 78.748, p-value = 2.171e-09

> svychisq(~WKSTAT3C+ald, ncsrsvyp2, statistic="F")
        Pearson's X^2: Rao & Scott adjustment
data:  svychisq(~WKSTAT3C + ald, ncsrsvyp2, statistic = "F")
F = 3.1249, ndf = 1.7248, ddf = 72.4410, p-value = 0.05716

> svychisq(~WKSTAT3C+mde, ncsrsvyp2, statistic="F")
        Pearson's X^2: Rao & Scott adjustment
data:  svychisq(~WKSTAT3C + mde, ncsrsvyp2, statistic = "F")
F = 4.6693, ndf = 1.7348, ddf = 72.8610, p-value = 0.01605

> svychisq(~WKSTAT3C+ED4CAT, ncsrsvyp2, statistic="F")
        Pearson's X^2: Rao & Scott adjustment
data:  svychisq(~WKSTAT3C + ED4CAT, ncsrsvyp2, statistic = "F")
F = 27.64, ndf = 5.1457, ddf = 216.1200, p-value < 2.2e-16

> svychisq(~WKSTAT3C+ag4cat, ncsrsvyp2, statistic="F")
        Pearson's X^2: Rao & Scott adjustment
data:  svychisq(~WKSTAT3C + ag4cat, ncsrsvyp2, statistic = "F")
F = 113.49, ndf = 4.9646, ddf = 208.5100, p-value < 2.2e-16

> svychisq(~WKSTAT3C+MAR3CAT, ncsrsvyp2, statistic="F")
        Pearson's X^2: Rao & Scott adjustment
data:  svychisq(~WKSTAT3C + MAR3CAT, ncsrsvyp2, statistic = "F")
F = 23.124, ndf = 3.1985, ddf = 134.3400, p-value = 1.229e-12


# Note: No Multinomial Logit Option in R Survey package
```

```
# Ordinal Regression with Russian Federation Data
> # Read Data and set design variables
> # Use C9 version of data for this example

> rfdata_c9 <- read_sas("P:/ASDA 2/Data sets/ess6 russia/c9_russia_1jun2017.sas7bdat")

#create factor variables
> rfdata_c9$marcatc <- factor(rfdata_c9$marcat, levels = 1:3, labels =c("Married", "Previous", "Never"))
> rfsvy <- svydesign(strata=~stratify, id=~psu, weights=~PSPWGHT, data=rfdata_c9, nest=T)

> ex936 <- svymean(~factor(stflife2), design=rfsvy, na.rm=T, se=T, deff=T, ci=T, keep.vars=T)
Warning message:
In svymean.survey.design2(~factor(stflife2), design = rfsvy, na.rm = T,  :
  Sample size greater than population size: are weights correctly scaled?
> print(ex936)
                       mean        SE DEff
factor(stflife2)1 0.0452316 0.0065542   NA
factor(stflife2)2 0.2001318 0.0135063   NA
factor(stflife2)3 0.2092193 0.0122451   NA
factor(stflife2)4 0.4240569 0.0174742   NA
factor(stflife2)5 0.1213603 0.0092441   NA

> barplot(ex936, legend=c("0-1", "2-4", "5", "6-8", "9-10") , col=c("black", "grey60", "blue", "red", "green"))
```



```
> summary(ex936_ordinal <- svyolr (factor(stflife2) ~ factor(agecat) + male + marcatc, design=rfsvy))
Call:
svyolr(factor(stflife2) ~ factor(agecat) + male + marcatc, design = rfsvy)
Coefficients:
                     Value Std. Error   t value
factor(agecat)2 -0.5293373 0.13614998 -3.887899
factor(agecat)3 -0.7455312 0.14332027 -5.201855
factor(agecat)4 -0.8080856 0.16561383 -4.879337
male            -0.1096234 0.09523049 -1.151138
marcatcPrevious -0.2088768 0.10542193 -1.981341
marcatcNever    -0.1371932 0.13208274 -1.038692
Intercepts:
    Value    Std. Error t value
1|2  -3.7111   0.2144   -17.3066
2|3  -1.7928   0.1668   -10.7469
3|4  -0.8348   0.1592    -5.2433
4|5   1.3841   0.1536     9.0095
(69 observations deleted due to missingness)
> exp(ex936_ordinal$coef)
factor(agecat)2 factor(agecat)3 factor(agecat)4           male marcatcPrevious
      0.5889952       0.4744822       0.4457105       0.8961715       0.8114952
   marcatcNever
```

```
> # HRS data for Poisson Regression using Number of Falls
> #histogram of number of falls 24
> svyhist(~numfalls24 , subset (hrssvyr, NAGE >=65), main="", col="grey80", xlab ="Histogram of Number of Falls
Past 24 Months")
```




```
> # Poisson model

> # Create variables for example
> hrs$nage_c=(hrs$NAGE - 74.5)
> hrs$bmi_c=(hrs$R11BMI - 27.7)
> hrs$male=(hrs$GENDER ==1)

> # Create Design Object and then subset for 65 Plus
> hrssvyr <- svydesign(strata=~STRATUM, id=~SECU, weights=~NWGTR , data=hrs, nest=T)
> hrssvy65 <- subset(hrssvyr, age65p==1)
```

```
> summary(ex947_poisson <- svyglm(numfalls24 ~ male + nage_c + arthritis + diabetes + bmi_c, design=hrssvy65,
family=quasipoisson(log)))

Call:
svyglm(formula = numfalls24 ~ male + nage_c + arthritis + diabetes +
    bmi_c, design = hrssvy65, family = quasipoisson(log))

Survey design:
subset(hrssvyr, age65p == 1)

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.635588   0.074371  -8.546 3.33e-11 ***
maleTRUE     0.257180   0.080487   3.195 0.002470 **
nage_c       0.014666   0.004425   3.314 0.001753 **
arthritis    0.736177   0.077384   9.513 1.26e-12 ***
diabetes     0.247533   0.070193   3.526 0.000938 ***
bmi_c        0.004126   0.008402   0.491 0.625615
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 8.755473)

Number of Fisher Scoring iterations: 6

Warning messages:
1: In summary.glm(g) :
  observations with zero weight not used for calculating dispersion
2: In summary.glm(glm.object) :
  observations with zero weight not used for calculating dispersion
> exp(ex947_poisson$coef)
(Intercept)    maleTRUE      nage_c   arthritis    diabetes       bmi_c
  0.5296241   1.2932785   1.0147744   2.0879372   1.2808623   1.0041347
```

> # NEGATIVE BINOMIAL (NOT AVAILABLE WITH SURVEY CORRECTION, DISPERSION IS ACCOUNTED FOR IN SVYGLM, PER LUMLEY)
> # ZERO INFLATED NEGATIVE BINOMIAL NOT AVAILABLE IN R SURVEY PACKAGE