Introduction to R: Final Project

## Demographic and Health Survey

The Demographic and Health Surveys are nationally representative population-based serial cross-sectional surveys sponsored by the United States Agency for International Development (USAID) and administered by ICF International. Currently data is collected from 44 countries/regions in sub-Saharan Africa as well as countries in North Africa, West Asia, Europe, South and Southeast Asia, Latin America and the Caribbean. Households in these regions are selected using a probability sample from census frames or, in cases where no census frame exists, from a complete list of villages or communities. The researchers employed a two-stage cluster sampling procedure where a cluster of households was randomly selected from the sampling frame. Subsequent to the household interviews, every eligible woman and man were interviewed. The domestic violence module of the DHS was only administered to females in a subset of the countries that participated in the DHS. Eligibility for the domestic violence module included being able to administer the survey in the absence of other people and currently/previously been married and/or living with a man that was her intimate partner.

## Sample

The present study includes countries in sub-Saharan Africa that participated in the sixth DHS survey (DHS-6) between 2010-2014 and were selected to receive the domestic violence module, including information on alcohol use. Hence, the study contains DHS-6 data from Burkina Faso, Cameroon, Comoros, the Democratic Republic of the Congo, Gabon, Cote d’Ivoire, Malawi, Mali, Mozambique, Nigeria, Sierra Leone, Tanzania, Uganda and Zimbabwe. This resulted in a final sample of 103,453 women aged 15-49 years.

## Measures

Information on basic demographics was taken from the general women’s questionnaire and includes data on current marital status, residency type, wealth and a literacy assessment. Data on partner’s alcohol use and intimate partner violence (IPV) perpetrated by the husband/partner against the participant was collected in the domestic violence module. For the items included in this study, all questions were asked of the female participant in regards to the behaviors of her current or most recent husband/partner. Items assessing the participant’s current or most recent partner’s controlling behaviors, emotional violence, physical (severe and less severe) violence and sexual violence were collected.

## Research Aims

For the purpose of this project, I plan to: 1. Characterize the sample based on social and demographic variables 2. Examine the prevalence of alcohol use and various forms of IPV 3. Evaluate the relationships between socio-demographic variables, alcohol use and IPV

dhs = read.csv(file = "/Users/Desktop/DHS\_CSV/DHSMERGE.csv", header=TRUE, as.is=TRUE, na.strings="")

To begin, I removed variables that are irrelevant to my research question.

dhsvars = names(dhs) %in% c("D113", "country", "D104", "D106", "D107", "D108", "V149", "V501", "V140", "V155", "V190", "V447A", "D103A", "D103B", "D103C", "D105A", "D105B", "D105C", "D105D", "D105E", "D105G", "D105H", "D105I", "D102", "D101A", "D101B", "D101C", "D101D", "D101E", "D101F", "D114")
dhs = dhs[dhsvars]

I then added in variable names for the remaining 31 variables in my dataset.

ColNames = c("Residence", "Education", "Literacy", "Wealth", "Age", "Marital\_Status", "Control\_Jealous", "Control\_Unfaithful", "Control\_Friends", "Control\_Family", "Control\_Whereabouts", "Control\_Trust","Control\_N", "EV\_Humiliate", "EV\_Threaten", "EV\_Insult", "Emotional\_Violence", "PV\_Push", "PV\_Slap", "PV\_Twist", "PV\_Kick", "PV\_StrangleBurn", "PV\_Weapon", "SV\_Sex", "SV\_OthSex", "Less\_Severe\_Violence", "Severe\_Violence", "Sexual\_Violence", "Alcohol", "Drunk\_Freq", "Country" )
length(ColNames)

## [1] 31

colnames(dhs) = ColNames

The majority of variables in this dataset are categorical and required some recoding. In the following code chunk I also recoding missing data as NA:

dhs[(dhs==7) | (dhs==8) | (dhs==9)] = NA
dhs$Urban = ifelse(dhs$Residence==2, 0, 1)
dhs$SomeEduc = ifelse(dhs$Education>=1, 1, 0)
dhs$Married = ifelse(dhs$Marital\_Status==1, 2, ifelse(dhs$Marital\_Status==2, 1, 0))
dhs$Literacy[(dhs$Literacy==3) | (dhs$Literacy==4)] = NA
dhs$Literacy = ifelse(dhs$Literacy==0, 0, ifelse(dhs$Literacy==1, 1, 2))
dhs$Country = ifelse(dhs$Country==0, "Burkina\_Faso", ifelse(dhs$Country==1, "Cameroon", ifelse(dhs$Country==2, "Comoros", ifelse(dhs$Country==3, "DRC", ifelse(dhs$Country==4, "Gabon", ifelse(dhs$Country==5, "Ivory\_Coast", ifelse(dhs$Country==6, "Malawi", ifelse(dhs$Country==7, "Mali", ifelse(dhs$Country==8, "Mozambique", ifelse(dhs$Country==9, "Nigeria", ifelse(dhs$Country==10, "Sierra\_Leone", ifelse(dhs$Country==11, "Tanzania", ifelse(dhs$Country==12, "Uganda", "Zimbabwe")))))))))))))

## Aim 1: Characterize the sample based on social and demographic variables

To characterize the sample, I described the average age (in years), % illiterate, % residing in a rural locale, % without any formal education, and % divorced/widowed/separated stratified by country.

Index = split(1:nrow(dhs), dhs$Country)
lapply(Index, head)

## $Burkina\_Faso
## [1] 1 2 3 4 5 6
##
## $Cameroon
## [1] 11364 11365 11366 11367 11368 11369
##
## $Comoros
## [1] 16407 16408 16409 16410 16411 16412
##
## $DRC
## [1] 19748 19749 19750 19751 19752 19753
##
## $Gabon
## [1] 26559 26560 26561 26562 26563 26564
##
## $Ivory\_Coast
## [1] 32116 32117 32118 32119 32120 32121
##
## $Malawi
## [1] 38467 38468 38469 38470 38471 38472
##
## $Sierra\_Leone
## [1] 82624 82625 82626 82627 82628 82629
##
## $Tanzania
## [1] 87809 87810 87811 87812 87813 87814
##
## $Uganda
## [1] 94856 94857 94858 94859 94860 94861
##
## $Zimbabwe
## [1] 96912 96913 96914 96915 96916 96917

#Age
T1\_Age\_Mean = tapply(dhs$Age, dhs$Country, mean)
T1\_Age\_SD = tapply(dhs$Age, dhs$Country, sd)
table1 = matrix(paste(signif(T1\_Age\_Mean, 4), " (SD=", signif(T1\_Age\_SD,2),")",sep=""), nrow =1, ncol=11)
colnames(table1) = list("Burkina Faso", "Cameroon", "Comoros", "DRC", "Gabon", "Ivory Coast", "Malawi", "Sierra Leone", "Tanzania", "Uganda", "Zimbabwe")
rownames(table1) = list("Age (Yrs)")

#Education
T1\_Educ = sapply(Index, function(x) table (dhs$SomeEduc[x]))
T1\_Educ\_P = (signif(prop.table(T1\_Educ,2),3))\*100
table1 = rbind(table1, T1\_Educ\_P[1,])
rownames(table1)[nrow(table1)] = "% Uneducated"

#Residence
T1\_Urban = sapply(Index, function(x) table (dhs$Urban[x]))
T1\_Urban\_P = (signif(prop.table(T1\_Urban,2),3))\*100
table1 = rbind(table1, T1\_Urban\_P[1,])
rownames(table1)[nrow(table1)] = "% Rural"

#Marital Status
T1\_Married = sapply(Index, function(x) table (dhs$Married[x]))
T1\_Married\_P = (signif(prop.table(T1\_Married,2),3))\*100
table1 = rbind(table1, T1\_Married\_P[1,])
rownames(table1)[nrow(table1)] = "% Divorced/Widowed/Separated"

#Literacy
T1\_Literacy = sapply(Index, function(x) table (dhs$Literacy[x]))
T1\_Literacy\_P = (signif(prop.table(T1\_Literacy,2),3))\*100
table1 = rbind(table1, T1\_Literacy\_P[1,])
rownames(table1)[nrow(table1)] = "% Illiterate"

#Adding p-values
agereg = lm(formula = Age~Country, data=dhs) #p<.001
chisq.test(dhs$SomeEduc, dhs$Country, correct=TRUE, p=rep(1/length(dhs$Literacy), length(dhs$Literacy)), rescale.p=FALLSE, simulate.p.value=FALSE, B=2000)

##
## Pearson's Chi-squared test
##
## data: dhs$SomeEduc and dhs$Country
## X-squared = 20637.8, df = 10, p-value < 2.2e-16

chisq.test(dhs$Urban, dhs$Country, correct=TRUE, p=rep(1/length(dhs$Literacy), length(dhs$Literacy)), rescale.p=FALLSE, simulate.p.value=FALSE, B=2000)

##
## Pearson's Chi-squared test
##
## data: dhs$Urban and dhs$Country
## X-squared = 4830.059, df = 10, p-value < 2.2e-16

chisq.test(dhs$Married, dhs$Country, correct=TRUE, p=rep(1/length(dhs$Literacy), length(dhs$Literacy)), rescale.p=FALLSE, simulate.p.value=FALSE, B=2000)

##
## Pearson's Chi-squared test
##
## data: dhs$Married and dhs$Country
## X-squared = 11529.55, df = 20, p-value < 2.2e-16

chisq.test(dhs$Literacy, dhs$Country, correct=TRUE, p=rep(1/length(dhs$Literacy), length(dhs$Literacy)), rescale.p=FALLSE, simulate.p.value=FALSE, B=2000)

##
## Pearson's Chi-squared test
##
## data: dhs$Literacy and dhs$Country
## X-squared = 16323.28, df = 20, p-value < 2.2e-16

pvalues = c("<.001", "<.001", "<.001", "<.001", "<.001")
dim(table1)

## [1] 5 11

table1 = cbind(table1, pvalues)
colnames(table1)[ncol(table1)] = "p-value"

## Table 1: Social and demographic characteristics by country

table1

## Burkina Faso Cameroon
## Age (Yrs) "NA (SD=NA)" "28.58 (SD=9.2)"
## % Uneducated "75.6" "18.4"
## % Rural "70.5" "51.4"
## % Divorced/Widowed/Separated "15.1" "29"
## % Illiterate "78.8" "29.7"
## Comoros DRC Gabon
## Age (Yrs) "28.21 (SD=8.8)" "NA (SD=NA)" "NA (SD=NA)"
## % Uneducated "34" "19.7" "5.16"
## % Rural "58.9" "68.6" "34.5"
## % Divorced/Widowed/Separated "31.1" "24.8" "35.9"
## % Illiterate "40.1" "43.9" "16.8"
## Ivory Coast Malawi Sierra Leone
## Age (Yrs) "NA (SD=NA)" "28.57 (SD=8.9)" "NA (SD=NA)"
## % Uneducated "60.9" "16.1" "63.2"
## % Rural "58.1" "87.4" "63.9"
## % Divorced/Widowed/Separated "26.9" "26.5" "22.3"
## % Illiterate "69.1" "33.5" "71.8"
## Tanzania Uganda Zimbabwe
## Age (Yrs) "29.53 (SD=9.2)" "NA (SD=NA)" "28.7 (SD=8.9)"
## % Uneducated "20.1" "16.8" "2.55"
## % Rural "76.7" "73.3" "65.7"
## % Divorced/Widowed/Separated "30" "29.6" "32.5"
## % Illiterate "28.5" "38.6" "6.76"
## p-value
## Age (Yrs) "<.001"
## % Uneducated "<.001"
## % Rural "<.001"
## % Divorced/Widowed/Separated "<.001"
## % Illiterate "<.001"

To visualize this data, I created a box plot for age by country and bar plots with proportions for the remaining socio-demographic variables by country. Several of the variables had more than one level. For the purpose of the bar plots, I allowed for more than 2 levels to be displayed. Also, several of the countries did not contain some of these questions in their survey which is why you will notice that some of the variables are missing data on entire countries. For the purpose of this project, I'll assume that there are no systematic differences between countries or within-country correlation that could invalidate the inferential analyses.

#Graphs
#Age
boxplot(dhs$Age ~ dhs$Country, las=3, main="Distribution of Age by Country", ylab="Age(Yrs)", cex.lab=0.6) #Not all countries asked about age



#Education
tableeduc = table(dhs$SomeEduc, dhs$Country)
barplot(prop.table(tableeduc, 2), las=3, main="% Educated by Country", ylab="Proportion", col=c("grey", "white"), legend=(c("No Education", "Some Education")), cex.names=0.6)



#Residence
tableurban = table(dhs$Urban, dhs$Country)
barplot(prop.table(tableurban, 2), las=3, main="Place of Residence by Country", ylab="Proportion", col=c("grey", "white"), legend=(c("Rural", "Urban")), cex.names=0.6)



#Marital Status
tablemarried = table(dhs$Married, dhs$Country)
barplot(prop.table(tablemarried, 2), las=3, main="Marital Status by Country", ylab="Proportion", col=c("darkgrey", "grey", "white"), legend=(c("Widowed/Divorced/Separated", "Co-Habitating", "Married")), cex.names=0.6)



#Literacy
tableliteracy = table(dhs$Literacy, dhs$Country)
barplot(prop.table(tableliteracy, 2), las=3, main="Literacy by Country", ylab="Proportion", col=c("darkgrey", "grey", "white"), legend=(c("Illiterate", "Somewhat Literate", "Literate")), cex.names=0.6)



Results from table 1 and the graphs above reveal that the average age of participants was between 28.2-29.5 years. The majority of participants had received education and were literate, with the exception of respondents in Burkina Faso, Ivory Coast, and Sierra Leone. The majority of participants lived in rural areas except in Gabon. The majority of participants were married or co-habitating with their husband/partner. There is clearly some variation in several of these sociodemographic indicators by country.

## Aim 2: Examine the prevalence of alcohol use and various forms of IPV

To explore the prevalence of intimate partner violence (IPV) and partner's alcohol use, I calculated the prevalence of each type of violence and partner's alcohol use.

#Calculating Prevalence
EV = table(dhs$Emotional\_Violence)
EV\_Prev = EV[2]/(EV[1]+EV[2])
LSV = table(dhs$Less\_Severe\_Violence)
LSV\_Prev = LSV[2]/(LSV[1]+LSV[2])
PV = table(dhs$Severe\_Violence)
PV\_Prev = PV[2]/(PV[1]+PV[2])
SV = table(dhs$Sexual\_Violence)
SV\_Prev = SV[2]/(SV[1]+SV[2])
Alc = table(dhs$Alcohol)
Alc\_Prev = Alc[2]/(Alc[1]+Alc[2])

#Creating Table 2 by IPV Type
table2 = matrix(paste(EV), nrow=1, ncol=2)
rownames(table2) = list("Emotional Violence")
table2 = rbind(table2, LSV)
rownames(table2)[nrow(table2)] = "Physical Violence (Less Severe)"
table2 = rbind(table2, PV)
rownames(table2)[nrow(table2)] = "Physical Violence (Severe)"
table2 = rbind(table2, SV)
rownames(table2)[nrow(table2)] = "Sexual Violence"
table2 = rbind(table2, Alc)
rownames(table2)[nrow(table2)] = "Partner's Alcohol Use"
Prevalence = c(EV\_Prev, LSV\_Prev, PV\_Prev, SV\_Prev, Alc\_Prev)
table2 = cbind(table2, (Prevalence\*100))
colnames(table2) = list("No", "Yes", "Prevalence (%)")

## Table 2: Prevalence of IPV and Partner's Alcohol Use by Country

table2

## No Yes Prevalence (%)
## Emotional Violence "63735" "21201" "24.9611472167279"
## Physical Violence (Less Severe) "63311" "21598" "25.4366439364496"
## Physical Violence (Severe) "77368" "7509" "8.84691966021419"
## Sexual Violence "76520" "8379" "9.86937419757594"
## Partner's Alcohol Use "57370" "27508" "32.4088692004995"

As can be seen from Table 2 above, the prevalence of partner's alcohol use as reported by the female respondent was 32.4%. The prevalence of the types of intimate partner violence are as follows:

Emotional Violence: 24.96% Physical Violence (Less Severe): 25.44% Physical Violence (Severe): 8.85% Sexual Violence: 9.87%

#Plot of IPV Prevalence by Type
prevmatrix = matrix(c(24.96, 25.44, 8.85, 9.87), 1, 4, byrow=TRUE)
colnames(prevmatrix) = c("EV", "PV", "SPV", "SV")
barplot(prevmatrix, main="Prevalence of IPV by Type", ylab = "Prevalence", ylim=c(0, 100))

 Footnote: EV=Emotional Violence, PV = Physical Violence, SPV = Severe Physical Violence, SV=Sexual Violence

## Aim 3:Evaluate the relationships between socio-demographic variables, alcohol use and IPV

To investigate Aim 3, I examined the univariate relationship between partner's alcohol use and IPV using chi-squared analyses. Subsequently, I examined this relationship controlling for potential confounders of this relationship including current marital status, literacy, education level, place of residence (urban vs. rural) and age. While not all countries administered questions for all of the aforementioned potential confounders, for the purpose of this project we are going to assume that missing data is not systematic or informative.

#Chi-Squared Analyses

#Emotional Violence
EV\_ChiSq = chisq.test(dhs$Emotional\_Violence, dhs$Alcohol)
names(EV\_ChiSq)

## [1] "statistic" "parameter" "p.value" "method" "data.name" "observed"
## [7] "expected" "residuals" "stdres"

table3 = matrix(paste(round(EV\_ChiSq$statistic, digits=3)), nrow=1, ncol=1)

#Less Severe Physical Violence
LSV\_ChiSq = chisq.test(dhs$Less\_Severe\_Violence, dhs$Alcohol)
table3 = rbind(table3, round(LSV\_ChiSq$statistic, digits=3))

#Severe Physical Violence
PV\_ChiSq = chisq.test(dhs$Severe\_Violence, dhs$Alcohol)
table3 = rbind(table3, round(PV\_ChiSq$statistic, digits=3))

#Sexual Violence
SV\_ChiSq = chisq.test(dhs$Sexual\_Violence, dhs$Alcohol)
table3 = rbind(table3, round(SV\_ChiSq$statistic, digits=3))

pval\_ChiSq = signif(c(EV\_ChiSq$p.value, LSV\_ChiSq$p.value, PV\_ChiSq$p.value, SV\_ChiSq$p.value), digits=3)
table3 = cbind(table3, c("<.001", "<.001", "<.001", "<.001"))
colnames(table3) = list("X-Squared", "p-value")
rownames(table3) = list("Emotional Violence", "Physical Violence(Less Severe)", "Physical Violence(Severe)", "Sexual Violence")

## Table 3: Univariate Relationship Between Partner's Alcohol Use and IPV

table3

## X-Squared p-value
## Emotional Violence "3947.586" "<.001"
## Physical Violence(Less Severe) "5376.13" "<.001"
## Physical Violence(Severe) "2624.279" "<.001"
## Sexual Violence "1779.968" "<.001"

## Logistic Regression Models Evaluating the Relationship Between Partner's Alcohol Use and IPV Controlling for Education, Literacy, Marital Status, Residence, Age and Wealth

The first model evaluated the relationship between partner's alcohol use and emotional violence

#Emotional Violence
EVlogistic = glm(Emotional\_Violence ~ factor(Alcohol) + factor(SomeEduc) + factor(Literacy) + factor(Married) + factor(Urban) + Wealth + Age, data=dhs, family=binomial())
summary(EVlogistic)

##
## Call:
## glm(formula = Emotional\_Violence ~ factor(Alcohol) + factor(SomeEduc) +
## factor(Literacy) + factor(Married) + factor(Urban) + Wealth +
## Age, family = binomial(), data = dhs)
##
## Deviance Residuals:
## Min 1Q Median 3Q Max
## -1.2844 -0.6999 -0.6121 -0.5479 2.0414
##
## Coefficients:
## Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.034708 0.053697 -19.269 < 2e-16 \*\*\*
## factor(Alcohol)1 0.947844 0.020247 46.814 < 2e-16 \*\*\*
## factor(SomeEduc)1 0.373718 0.030063 12.431 < 2e-16 \*\*\*
## factor(Literacy)1 -0.155612 0.041050 -3.791 0.00015 \*\*\*
## factor(Literacy)2 -0.167573 0.029320 -5.715 1.10e-08 \*\*\*
## factor(Married)1 -0.331985 0.038893 -8.536 < 2e-16 \*\*\*
## factor(Married)2 -0.510688 0.030429 -16.783 < 2e-16 \*\*\*
## factor(Urban)1 0.040683 0.024346 1.671 0.09471 .
## Wealth -0.042204 0.008601 -4.907 9.26e-07 \*\*\*
## Age -0.001433 0.001144 -1.253 0.21036
## ---
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 70142 on 62890 degrees of freedom
## Residual deviance: 66680 on 62881 degrees of freedom
## (40562 observations deleted due to missingness)
## AIC: 66700
##
## Number of Fisher Scoring iterations: 4

OR\_EV = exp(coef(EVlogistic))
OR\_EV

## (Intercept) factor(Alcohol)1 factor(SomeEduc)1 factor(Literacy)1
## 0.3553302 2.5801411 1.4531267 0.8558912
## factor(Literacy)2 factor(Married)1 factor(Married)2 factor(Urban)1
## 0.8457149 0.7174983 0.6000827 1.0415222
## Wealth Age
## 0.9586739 0.9985682

Results from this model suggest that partner's alcohol use was associated with increased odds of Emotional Violence (OR=2.58, p<0.001).

The second model evaluated the relationship between partner's alcohol use and less severe physical violence.

#Less Severe Physical Violence
LSVlogistic = glm(Less\_Severe\_Violence ~ factor(Alcohol) + factor(SomeEduc) + factor(Literacy) + factor(Married) + factor(Urban) + Wealth + Age, data=dhs, family=binomial())
summary(LSVlogistic)

##
## Call:
## glm(formula = Less\_Severe\_Violence ~ factor(Alcohol) + factor(SomeEduc) +
## factor(Literacy) + factor(Married) + factor(Urban) + Wealth +
## Age, family = binomial(), data = dhs)
##
## Deviance Residuals:
## Min 1Q Median 3Q Max
## -1.3633 -0.6721 -0.5456 -0.4736 2.2245
##
## Coefficients:
## Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.036510 0.055773 -18.584 < 2e-16 \*\*\*
## factor(Alcohol)1 1.157711 0.020777 55.722 < 2e-16 \*\*\*
## factor(SomeEduc)1 0.437501 0.031446 13.913 < 2e-16 \*\*\*
## factor(Literacy)1 -0.124886 0.042458 -2.941 0.003268 \*\*
## factor(Literacy)2 -0.100124 0.030220 -3.313 0.000922 \*\*\*
## factor(Married)1 -0.140146 0.038974 -3.596 0.000323 \*\*\*
## factor(Married)2 -0.620827 0.031096 -19.965 < 2e-16 \*\*\*
## factor(Urban)1 0.087310 0.025263 3.456 0.000548 \*\*\*
## Wealth -0.073927 0.008933 -8.275 < 2e-16 \*\*\*
## Age -0.005580 0.001198 -4.658 3.2e-06 \*\*\*
## ---
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 67796 on 62878 degrees of freedom
## Residual deviance: 62415 on 62869 degrees of freedom
## (40574 observations deleted due to missingness)
## AIC: 62435
##
## Number of Fisher Scoring iterations: 4

OR\_LSV = exp(coef(LSVlogistic))
OR\_LSV

## (Intercept) factor(Alcohol)1 factor(SomeEduc)1 factor(Literacy)1
## 0.3546903 3.1826393 1.5488318 0.8825972
## factor(Literacy)2 factor(Married)1 factor(Married)2 factor(Urban)1
## 0.9047255 0.8692314 0.5374999 1.0912349
## Wealth Age
## 0.9287397 0.9944359

Results from this model suggest that partner's alcohol use was associated with increased odds of physical violence (less severe type; OR=23.18, p<0.001).

The third model evaluated the relationship between partner's alcohol use and severe physical violence.

#Severe Physical Violence
PVlogistic = glm(Severe\_Violence ~ factor(Alcohol) + factor(SomeEduc) + factor(Literacy) + factor(Married) + factor(Urban) + Wealth + Age, data=dhs, family=binomial())
summary(PVlogistic)

##
## Call:
## glm(formula = Severe\_Violence ~ factor(Alcohol) + factor(SomeEduc) +
## factor(Literacy) + factor(Married) + factor(Urban) + Wealth +
## Age, family = binomial(), data = dhs)
##
## Deviance Residuals:
## Min 1Q Median 3Q Max
## -0.9022 -0.4354 -0.2994 -0.2381 2.7897
##
## Coefficients:
## Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.662023 0.085241 -31.229 < 2e-16 \*\*\*
## factor(Alcohol)1 1.206866 0.032390 37.261 < 2e-16 \*\*\*
## factor(SomeEduc)1 0.725579 0.049013 14.804 < 2e-16 \*\*\*
## factor(Literacy)1 -0.186040 0.062315 -2.985 0.00283 \*\*
## factor(Literacy)2 -0.179928 0.043171 -4.168 3.08e-05 \*\*\*
## factor(Married)1 -0.305929 0.051490 -5.942 2.82e-09 \*\*\*
## factor(Married)2 -0.874432 0.041612 -21.014 < 2e-16 \*\*\*
## factor(Urban)1 -0.111446 0.038478 -2.896 0.00377 \*\*
## Wealth -0.034930 0.013474 -2.592 0.00953 \*\*
## Age 0.001551 0.001842 0.842 0.39973
## ---
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 34481 on 62848 degrees of freedom
## Residual deviance: 31378 on 62839 degrees of freedom
## (40604 observations deleted due to missingness)
## AIC: 31398
##
## Number of Fisher Scoring iterations: 6

OR\_PV = exp(coef(PVlogistic))
OR\_PV

## (Intercept) factor(Alcohol)1 factor(SomeEduc)1 factor(Literacy)1
## 0.06980689 3.34299217 2.06592593 0.83024045
## factor(Literacy)2 factor(Married)1 factor(Married)2 factor(Urban)1
## 0.83533061 0.73643856 0.41709892 0.89453981
## Wealth Age
## 0.96567326 1.00155208

Results from this model suggest that partner's alcohol use was associated with increased odds of severe physical violence (OR=3.34, p<0.001).

The fourth model evaluated the relationship between partner's alcohol use and sexual violence.

#Sexual Violence
SVlogistic = glm(Sexual\_Violence ~ factor(Alcohol) + factor(SomeEduc) + factor(Literacy) + factor(Married) + factor(Urban) + Wealth + Age, data=dhs, family=binomial())
summary(SVlogistic)

##
## Call:
## glm(formula = Sexual\_Violence ~ factor(Alcohol) + factor(SomeEduc) +
## factor(Literacy) + factor(Married) + factor(Urban) + Wealth +
## Age, family = binomial(), data = dhs)
##
## Deviance Residuals:
## Min 1Q Median 3Q Max
## -0.9272 -0.4677 -0.3638 -0.3154 2.7348
##
## Coefficients:
## Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.592210 0.077765 -20.475 < 2e-16 \*\*\*
## factor(Alcohol)1 0.850696 0.029931 28.422 < 2e-16 \*\*\*
## factor(SomeEduc)1 0.338601 0.045862 7.383 1.55e-13 \*\*\*
## factor(Literacy)1 -0.011853 0.061163 -0.194 0.846332
## factor(Literacy)2 0.142700 0.043010 3.318 0.000907 \*\*\*
## factor(Married)1 -0.486963 0.053847 -9.044 < 2e-16 \*\*\*
## factor(Married)2 -0.615806 0.041339 -14.897 < 2e-16 \*\*\*
## factor(Urban)1 -0.259497 0.037207 -6.974 3.07e-12 \*\*\*
## Wealth -0.097463 0.012663 -7.697 1.40e-14 \*\*\*
## Age -0.015525 0.001734 -8.954 < 2e-16 \*\*\*
## ---
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 37803 on 62869 degrees of freedom
## Residual deviance: 36072 on 62860 degrees of freedom
## (40583 observations deleted due to missingness)
## AIC: 36092
##
## Number of Fisher Scoring iterations: 5

OR\_SV = exp(coef(SVlogistic))
OR\_SV

## (Intercept) factor(Alcohol)1 factor(SomeEduc)1 factor(Literacy)1
## 0.2034755 2.3412758 1.4029834 0.9882165
## factor(Literacy)2 factor(Married)1 factor(Married)2 factor(Urban)1
## 1.1533838 0.6144896 0.5402053 0.7714394
## Wealth Age
## 0.9071362 0.9845945

Results from this model suggest that partner's alcohol use was associated with increased odds of sexual violence (OR=2.34, p<0.001).

In summary, alcohol use by ones partner appeared to increase the odds of emotional, physical and sexual violence.