Appendix 4.5: Burden of cervical cancer risk factors in West Africa R script

# Benin

# Select needed variable from women dataset in Benin  
BJDHS <- subset(BJDHSa, select = c(v001, v002, v021, v005, v025, v463z, v437, v438, v525, v212, v201, v312, v337, v836, v763a, v012, v013, v502, v106, v714, v190, v130, v717, v157, v158, v159, v171b, v171a, v481, v467b, v467c, v467d, v467f, v3a08d, s1402, s1407, s1411, s1413, s1415, s1417, s1419, s1421, s1423, s1424, s1426))  
  
# Sample weight/design for women  
BJDHS$BJwt <-BJDHS$v005/1000000  
designBJ <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~BJwt, data = BJDHS)  
  
# Demographics  
BJDHS$BJwt <- BJDHS$v005/1000000  
freq(BJDHS$v013, BJDHS$BJwt, plot = FALSE) #Age  
freq(BJDHS$v025, BJDHS$BJwt, plot = FALSE) #Residential status  
freq(BJDHS$v502, BJDHS$BJwt, plot = FALSE) #Marital status  
freq(BJDHS$v106, BJDHS$BJwt, plot = FALSE) #Education  
freq(BJDHS$v714, BJDHS$BJwt, plot = FALSE) #Employment   
freq(BJDHS$v190, BJDHS$BJwt, plot = FALSE) #Wealth index  
variable.names(BJDHS)  
freq(BJDHS$v013, plot = FALSE)  
designBJ <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~BJwt, data = BJDHS) # Mean of Age and Confidence Interval  
BJttAge <- svyttest(v012~0, designBJ)  
BJttAge  
  
#Proportion of Women With Risk Factors for Cervical Cancer  
#1. Use cigarettes and tobacco   
freq(BJDHS$v463z, BJDHS$BJwt, plot = FALSE)   
sum(is.na(BJDHS$v463z))  
#2. BMI (Missing include women that were not measured in the subsample, flagged cases e.g refused-9995, missing for other reasons- 9996)  
favstats(BJDHS$v438, BJDHS$BJwt, plot = FALSE)   
favstats(BJDHSa$v438, BJDHS$BJwt, plot = FALSE)   
is.na(BJDHS$v438)  
sum(is.na(BJDHS$v438))  
is.na(BJDHS$v438) <- which(BJDHS$v438 == 9996)  
is.na(BJDHS$v438) <- which(BJDHS$v438 == 9995)  
is.na(BJDHS$v438) <- which(BJDHS$v438 == 9994)  
favstats(BJDHS$v438, BJDHS$BJwt, plot = FALSE)  
BJDHS$BJHtM <- (BJDHS$v438/1000) #Height (Height is in cm, converted into meters. Also into 1 implied decimal)  
favstats(BJDHS$BJHtM, plot = FALSE)  
favstats(BJDHS$v437, BJDHS$BJwt, plot = FALSE)   
favstats(BJDHSa$v437, BJDHS$BJwt, plot = FALSE)  
is.na(BJDHS$v437)  
sum(is.na(BJDHS$v437))  
sum(is.na(BJDHSa$v437)) # original data  
is.na(BJDHS$v437) <- which(BJDHS$v437 == 9996)  
is.na(BJDHS$v437) <- which(BJDHS$v437 == 9995)  
is.na(BJDHS$v437) <- which(BJDHS$v437 == 9994)  
favstats(BJDHS$v437, BJDHS$BJwt, plot = FALSE)   
BJDHS$BJwtkg <- (BJDHS$v437/10)   
favstats(BJDHS$BJwtkg, plot = FALSE)  
BJDHS$BMI <- (BJDHS$BJwtkg) /(BJDHS$BJHtM \* BJDHS$BJHtM) # Calculated BMI  
favstats(BJDHS$BMI, BJDHS$BJwt, plot = FALSE)  
head(BJDHS$BMI)  
BJDHS$BMICAT <- NA  
BJDHS$BMICAT[BJDHS$BMI < 18.50000] <-1  
BJDHS$BMICAT[BJDHS$BMI >= 18.50000 & BJDHS$BMI <= 24.9999] <-2  
BJDHS$BMICAT[BJDHS$BMI >= 25.00000 & BJDHS$BMI <= 29.9999] <-3  
BJDHS$BMICAT[BJDHS$BMI >= 30.00000] <-4  
BJDHS$BMICAT <- factor(BJDHS$BMICAT, levels = c(1,2,3,4), labels = c("Underweight", "Normal", "Overweight", "Obese"))  
BJDHS$BJwt <- BJDHS$v005/1000000  
freq(BJDHS$BMICAT, BJDHS$BJwt, plot = FALSE)  
BJBiCAT <- NA # Binary Categories of BMI   
BJBiCAT[BJDHS$BMI < 25.0000] <- "low risk bmi"  
BJBiCAT[BJDHS$BMI >= 25.0000] <- "high risk bmi"  
freq(BJBiCAT, BJDHS$BJwt, plot = FALSE)   
#3. Age at First Sexual Intercourse   
freq(BJDHS$v525, BJDHS$BJwt, plot = FALSE)   
sum(is.na(BJDHS$v525))  
BJAFSI <- subset(BJDHS, v525 >=8 & v525 <=36)  
freq(BJAFSI$v525, plot = FALSE)  
BJAFSI$BJAFSIwt <- BJAFSI$v005/1000000  
designBJAFSI <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~BJAFSIwt, data = BJAFSI)  
svymean(~v525, design = designBJAFSI, na.rm = FALSE)  
confint(svymean(~v525, designBJAFSI))  
BJAFSICAT <- NA # Binary Categories of Age at First Sexual Intercourse Combine, had sex or not   
BJAFSICAT[BJDHS$v525 <=17.9999] <- "less or equal 17 years"  
BJAFSICAT[BJDHS$v525 >= 18.0000] <- "greater than 17 years"  
freq(BJAFSICAT, BJDHS$BJwt, plot = FALSE)  
#4. Age at First Birth  
freq(BJDHS$v212, BJDHS$BJwt, plot = FALSE) # NA not in the universe  
is.na(BJDHS$v212)  
sum(is.na(BJDHS$v212))  
favstats(BJDHS$v212, BJDHS$BJwt, plot = FALSE)  
BJAFB <- subset(BJDHS, v212 >=11 & v212 <=41)  
freq(BJAFB$v212, plot = FALSE)  
BJAFB$BJAFBwt <- BJAFB$v005/1000000  
designBJAFB <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~BJAFBwt, data = BJAFB)  
svymean(~v212, design = designBJAFB, na.rm = FALSE)  
confint(svymean(~v212, designBJAFB))  
BJAFBCAT <- NA # Binary Categories of Age at First Birth  
BJAFBCAT[BJDHS$v212 <=19.9999] <- "less or equal 19 years"  
BJAFBCAT[BJDHS$v212 >= 20.0000] <- "greater than 19 years"  
freq(BJAFBCAT, BJDHS$BJwt, plot = FALSE)  
#5. Total Children Ever Born (Parity)  
freq(BJDHS$v201, BJDHS$BJwt, plot = FALSE)   
sum(is.na(BJDHS$v201))  
BJP <- subset(BJDHS, v201 >=1 & v201 <=16)  
freq(BJP$v201, plot = FALSE)  
BJP$BJPwt <- BJP$v005/1000000  
designBJP <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~BJPwt, data = BJP)  
svymean(~v201, design = designBJP, na.rm = FALSE)  
confint(svymean(~v201, designBJP))  
BJPCAT <- NA # Binary Categories of Total Children Ever Born (Parity)  
BJPCAT[BJDHS$v201<=2.9999] <- "less or equal 2 births"  
BJPCAT[BJDHS$v201 >= 3.0000] <- "greater than 2 births"  
freq(BJPCAT, BJDHS$BJwt, plot = FALSE)  
#6.Total Lifetime Number of Sex Partners   
freq(BJDHS$v836, BJDHS$BJwt, plot = FALSE) # NA not in the universe  
is.na(BJDHS$v836)  
sum(is.na(BJDHS$v836))  
favstats(BJDHS$v836, BJDHS$BJwt, plot = FALSE)  
sum(is.na(BJDHS$v836))  
is.na(BJDHS$v836) <- which(BJDHS$v836 == 98)  
favstats(BJDHS$v836, BJDHS$BJwt, plot = FALSE)  
BJLSP <- subset(BJDHS, v836 >=1 & v836 <=95)  
freq(BJLSP$v836, plot = FALSE)  
BJLSP$BJLSPwt <- BJLSP$v005/1000000  
designBJLSP <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~BJLSPwt, data = BJLSP)  
svymean(~v836, design = designBJLSP, na.rm = FALSE)  
confint(svymean(~v836, designBJLSP))  
BJLSPCAT <- NA # Binary Categories of Total Lifetime Number of Sex Partners  
BJLSPCAT[BJDHS$v836 <=5.9999] <- "less or equal 5 partners"  
BJLSPCAT[BJDHS$v836 >= 6.0000] <- "greater than 5 partners"  
freq(BJLSPCAT, BJDHS$BJwt, plot = FALSE)  
#7. Contraceptive Usage  
freq(BJDHS$v312, BJDHS$BJwt, plot = FALSE)  
is.na(BJDHS$v312)  
sum(is.na(BJDHS$v312))  
nlevels(BJDHS$v312)  
table(BJDHS$v312)  
BJDHSContwt <- freq(BJDHS$v312, BJDHS$BJwt, plot = FALSE)  
BJDHSContwt  
BJCont\_recoded <- c("not using" = 0, "pill" =1, "iud"=1, "injections"=1, "diaphragm"= 2, "male condom"=2, "female sterilization"=2,  
 "male sterilization"=2, "periodic abstinence"=2, "withdrawal"=2, "other traditional"=2, "implants/norplant"=1,  
 "prolonged abstinence"= 2, "lactational amenorrhea (lam)"=2, "female condom"=2, "foam or jelly"= 2, "emergency contraception"= 1,  
 "other modern method"= 2, "standard days method (sdm)"=2, "specific method 1"= 2, "specific method 2"= 2)  
BJDHS$BJREC\_CONT <- BJCont\_recoded[BJDHS$v312]  
freq(BJDHS$BJREC\_CONT, BJDHS$BJwt, plot = FALSE)  
BJCont\_recoded <- c("not using" = 0, "pill" =1, "iud"=1, "injections"=1, "diaphragm"= 0, "male condom"=0, "female sterilization"=0,  
 "male sterilization"=0, "periodic abstinence"=0, "withdrawal"=0, "other traditional"=0, "implants/norplant"=1,  
 "prolonged abstinence"= 0, "lactational amenorrhea (lam)"=0, "female condom"=0, "foam or jelly"= 0, "emergency contraception"= 1,  
 "other modern method"= 0, "standard days method (sdm)"=0, "specific method 1"= 0, "specific method 2"= 0)  
BJDHS$BJREC\_CONT <- BJCont\_recoded[BJDHS$v312]  
freq(BJDHS$BJREC\_CONT, BJDHS$BJwt, plot = FALSE)  
# Months of use of current family planning method, For women who are currently using a contraceptive method  
freq(BJDHS$v337, BJDHS$BJwt, plot = FALSE)   
is.na(BJDHS$v337)  
sum(is.na(BJDHS$v337))  
favstats(BJDHS$v337, BJDHS$BJwt, plot = FALSE)  
summary(BJDHS$v337, BJDHS$BJwt, plot = FALSE)  
BJMFP <- NA  
BJMFP[BJDHS$v337 <60.0000]<- "Less than 5 years"  
BJMFP[BJDHS$v337 >=60.0000]<- "More than or equal 5 years"  
freq(BJMFP, BJDHS$BJwt, plot = FALSE)  
# Combine number hormonal and non hormonal to number of years  
BJDHS$BJcom <- paste(BJDHS$v312, BJDHS$v337)  
BJDHS$BJcom  
BJDHS$BJcom <- paste(BJDHS$v312, "-", BJDHS$v337)  
BJDHS$BJcom  
BJDHS$BJcom <- paste(BJDHS$v312, BJMFP, sep = "-" )  
BJDHS$BJcom  
freq(BJDHS$BJcom, plot = FALSE)  
BJcom\_recoded <- c("emergency contraception-Less than 5 years" = 0, "emergency contraception-More than or equal 5 years"= 1, "female sterilization-Less than 5 years" = 0, "female sterilization-More than or equal 5 years" = 0,   
 "implants/norplant-Less than 5 years" = 0, "implants/norplant-More than or equal 5 years" = 1, "injections-Less than 5 years" =0, "injections-More than or equal 5 years" =1,  
 "iud-Less than 5 years" = 0, "iud-More than or equal 5 years"= 1, "lactational amenorrhea (lam)-Less than 5 years"= 0, "male condom-Less than 5 years" =0, "male condom-More than or equal 5 years" = 0,  
 "not using-NA" =0, "other modern method-Less than 5 years" = 0, "other traditional-Less than 5 years"= 0, "other traditional-More than or equal 5 years"=0, "periodic abstinence-Less than 5 years"=0,  
 "periodic abstinence-More than or equal 5 years"=0, "pill-Less than 5 years" =0, "pill-More than or equal 5 years"=0, "standard days method (sdm)-Less than 5 years"=0, "standard days method (sdm)-More than or equal 5 years"=0,  
 "withdrawal-Less than 5 years"=0, "withdrawal-More than or equal 5 years"= 0)  
BJDHS$BJyears\_FP <- BJcom\_recoded[BJDHS$BJcom]  
freq(BJDHS$BJyears\_FP, BJDHS$BJwt, plot = FALSE)  
#8. Sexually Transmitted Infection in last 12 months ("don't know" added to "no")  
freq(BJDHS$v763a, BJDHS$BJwt, plot = FALSE)  
sum(is.na(BJDHS$v763a))  
BJSTI\_recoded <- c("no"= 0, "yes" = 1, "don't know"= 0)  
BJDHS$BJREC\_STI <- BJSTI\_recoded[BJDHS$v763a]  
freq(BJDHS$BJREC\_STI, BJDHS$BJwt, plot = FALSE)  
  
  
# Number of Risk Factors   
# Smoking  
freq(BJDHS$v463z, BJDHS$BJwt, plot = FALSE)  
BJsmoke\_recoded <- c("no" = 1, "yes, smokes nothing" = 0)  
BJDHS$BJREC\_SMOKE <- BJsmoke\_recoded[BJDHS$v463z]  
freq(BJDHS$BJREC\_SMOKE, BJDHS$BJwt, plot = FALSE)  
#BMI  
freq(BJBiCAT, BJDHS$BJwt, plot = FALSE)  
BJbmi\_recoded <- c("high risk bmi" = 1, "low risk bmi" = 0)  
BJDHS$BJREC\_BMI <- BJbmi\_recoded[BJBiCAT]  
freq(BJDHS$BJREC\_BMI, BJDHS$BJwt, na.rm= TRUE, plot = FALSE)  
#AFSI  
freq(BJAFSICAT, BJDHS$BJwt, plot = FALSE)  
BJAFSI\_recoded <- c("less or equal 17 years" = 1, "greater than 17 years" = 0)  
BJDHS$BJREC\_AFSI <- BJAFSI\_recoded[BJAFSICAT]  
freq(BJDHS$BJREC\_AFSI, BJDHS$BJwt, na.rm= TRUE, plot = FALSE)  
#AFB  
freq(BJAFBCAT, BJDHS$BJwt, na.rm= TRUE, plot = FALSE)  
BJAFB\_recoded <- c("less or equal 19 years" = 1, "greater than 19 years" = 0)  
BJDHS$BJREC\_AFB <- BJAFB\_recoded[BJAFBCAT]  
freq(BJDHS$BJREC\_AFB, BJDHS$BJwt, na.rm= TRUE, plot = FALSE)  
#Parity  
freq(BJPCAT, BJDHS$BJwt, plot = FALSE)  
BJParity\_recoded <- c("greater than 2 births" = 1, "less or equal 2 births" = 0)  
BJDHS$BJREC\_Parity <- BJParity\_recoded[BJPCAT]  
freq(BJDHS$BJREC\_Parity, BJDHS$BJwt, plot = FALSE)  
#FP  
freq(BJDHS$BJcom, plot = FALSE)  
BJcom\_recoded <- c("emergency contraception-Less than 5 years" = 0, "emergency contraception-More than or equal 5 years"= 1, "female sterilization-Less than 5 years" = 0, "female sterilization-More than or equal 5 years" = 0,   
 "implants/norplant-Less than 5 years" = 0, "implants/norplant-More than or equal 5 years" = 1, "injections-Less than 5 years" =0, "injections-More than or equal 5 years" =1,  
 "iud-Less than 5 years" = 0, "iud-More than or equal 5 years"= 1, "lactational amenorrhea (lam)-Less than 5 years"= 0, "male condom-Less than 5 years" =0, "male condom-More than or equal 5 years" = 0,  
 "not using-NA" =0, "other modern method-Less than 5 years" = 0, "other traditional-Less than 5 years"= 0, "other traditional-More than or equal 5 years"=0, "periodic abstinence-Less than 5 years"=0,  
 "periodic abstinence-More than or equal 5 years"=0, "pill-Less than 5 years" =0, "pill-More than or equal 5 years"=0, "standard days method (sdm)-Less than 5 years"=0, "standard days method (sdm)-More than or equal 5 years"=0,  
 "withdrawal-Less than 5 years"=0, "withdrawal-More than or equal 5 years"= 0)  
BJDHS$BJyears\_FP <- BJcom\_recoded[BJDHS$BJcom]  
freq(BJDHS$BJyears\_FP, BJDHS$BJwt, plot = FALSE)  
# LSP women  
freq(BJLSPCAT, BJDHS$BJwt, plot = FALSE)  
BJLSP\_recoded <- c("greater than 5 partners" = 1, "less or equal 5 partners" = 0)  
BJDHS$BJREC\_LSP <- BJLSP\_recoded[BJLSPCAT]  
freq(BJDHS$BJREC\_LSP, BJDHS$BJwt, plot = FALSE)  
# STI  
freq(BJDHS$v763a, BJDHS$BJwt, plot = FALSE)  
BJSTI\_recoded <- c("no"= 0, "yes" = 1, "don't know"= 0)  
BJDHS$BJREC\_STI <- BJSTI\_recoded[BJDHS$v763a]  
freq(BJDHS$BJREC\_STI, BJDHS$BJwt, plot = FALSE)  
  
  
#Number of Risk Factors  
BJRiskcom <- data.frame(BJDHS$BJREC\_SMOKE, BJDHS$BJREC\_BMI, BJDHS$BJREC\_AFSI, BJDHS$BJREC\_AFB, BJDHS$BJREC\_Parity, BJDHS$BJyears\_FP, BJDHS$BJREC\_LSP, BJDHS$BJREC\_STI)  
BJRiskcom  
BJtotal <- cbind(BJRiskcom, total = rowSums(BJRiskcom, na.rm = TRUE))  
BJtotal <- cbind(BJRiskcom, total = rowSums(BJRiskcom, na.rm = FALSE))  
tail(BJtotal)  
head(BJtotal)  
head(BJRiskcom)  
  
# Regrouped number of Risk Factors  
BJTotalnew <- NA   
BJTotalnew[BJtotal$total <=4.9999] <- "less or equal 4"  
BJTotalnew[BJtotal$total >= 5.0000] <- "greater or equal 5" # >=6 number of risk factorsvar  
freq(BJTotalnew, BJDHS$BJwt, plot = FALSE)  
  
# Modelling  
# Missing Data Removed  
BJDHS$BJREC\_summary= as.numeric(BJDHS$BJREC\_SMOKE + BJDHS$BJREC\_BMI + BJDHS$BJREC\_AFSI + BJDHS$BJREC\_AFB + BJDHS$BJREC\_Parity + BJDHS$BJyears\_FP + BJDHS$BJREC\_LSP + BJDHS$BJREC\_STI)  
crosstab(BJDHS$BJREC\_summary, BJDHS$v013, BJDHS$BJwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(BJDHS$v013, BJDHS$BJREC\_summary, BJDHS$BJwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(BJDHS$BJREC\_summary, BJDHS$v106, BJDHS$BJwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(BJDHS$v106, BJDHS$BJREC\_summary, BJDHS$BJwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
summary(BJDHS$BJREC\_summary)  
favstats(BJDHS$BJREC\_summary)  
designBJ <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~BJwt, data = BJDHS)  
is.numeric(BJDHS$BJREC\_summary)  
designBJ <-update(designBJ, BJREC\_summary=BJDHS$BJREC\_SMOKE + BJDHS$BJREC\_BMI + BJDHS$BJREC\_AFSI + BJDHS$BJREC\_AFB + BJDHS$BJREC\_Parity + BJDHS$BJyears\_FP + BJDHS$BJREC\_LSP + BJDHS$BJREC\_STI)  
  
  
# Unadjusted RR  
# Age  
BJmodelAge <- svyglm(BJREC\_summary~v013, design= designBJ, family = poisson(link = "log"), na.action = na.omit)  
summary(BJmodelAge)  
exp(coef(BJmodelAge)) # calculate the RRs  
round(exp(coef(BJmodelAge)), 2) # round to make it clearer  
cbind(RR = exp(coef(BJmodelAge)), exp(confint(BJmodelAge))) # RR to CI  
round(cbind(RR = exp(coef(BJmodelAge)), exp(confint(BJmodelAge))), 2) # round to 2  
# Residential Status  
BJmodelRS <- svyglm(BJREC\_summary~v025, design= designBJ, family = poisson(link = "log"), na.action = na.omit)  
summary(BJmodelRS)  
exp(coef(BJmodelRS)) # calculate the RRs  
round(exp(coef(BJmodelRS)), 2) # round to make it clearer  
cbind(RR = exp(coef(BJmodelRS)), exp(confint(BJmodelRS))) # RR to CI  
round(cbind(RR = exp(coef(BJmodelRS)), exp(confint(BJmodelRS))), 2) # round to 2  
# Marital Status  
BJmodelMS <- svyglm(BJREC\_summary~v502, design= designBJ, family = poisson(link = "log"), na.action = na.omit)  
summary(BJmodelMS)  
exp(coef(BJmodelMS)) # calculate the RRs  
round(exp(coef(BJmodelMS)), 2) # round to make it clearer  
cbind(RR = exp(coef(BJmodelMS)), exp(confint(BJmodelMS))) # RR to CI  
round(cbind(RR = exp(coef(BJmodelMS)), exp(confint(BJmodelMS))), 2) # round to 2  
# Education  
BJmodelEdu <- svyglm(BJREC\_summary~v106, design= designBJ, family = poisson(link = "log"), na.action = na.omit)  
summary(BJmodelEdu)  
exp(coef(BJmodelEdu)) # calculate the RRs  
round(exp(coef(BJmodelEdu)), 2) # round to make it clearer  
cbind(RR = exp(coef(BJmodelEdu)), exp(confint(BJmodelEdu))) # RR to CI  
round(cbind(RR = exp(coef(BJmodelEdu)), exp(confint(BJmodelEdu))), 2) # round to 2  
# Employment  
BJmodelEmp <- svyglm(BJREC\_summary~v714, design= designBJ, family = poisson(link = "log"), na.action = na.omit)  
summary(BJmodelEmp)  
exp(coef(BJmodelEmp)) # calculate the RRs  
round(exp(coef(BJmodelEmp)), 2) # round to make it clearer  
cbind(RR = exp(coef(BJmodelEmp)), exp(confint(BJmodelEmp))) # RR to CI  
round(cbind(RR = exp(coef(BJmodelEmp)), exp(confint(BJmodelEmp))), 2) # round to 2  
# Wealth Index  
BJmodelWI <- svyglm(BJREC\_summary~v190, design= designBJ, family = poisson(link = "log"), na.action = na.omit)  
summary(BJmodelWI)  
exp(coef(BJmodelWI)) # calculate the RRs  
round(exp(coef(BJmodelWI)), 2) # round to make it clearer  
cbind(RR = exp(coef(BJmodelWI)), exp(confint(BJmodelWI))) # RR to CI  
round(cbind(RR = exp(coef(BJmodelWI)), exp(confint(BJmodelWI))), 2) # round to 2  
  
# Adjusted RR  
# Poisson Regression Modelling  
#BJmodel <- svyglm(BJREC\_SMOKE+BJREC\_BMI+BJREC\_AFSI+BJREC\_AFB+BJREC\_Parity+BJyears\_FP+BJREC\_LSP+BJREC\_STI+BJh\_LSP2~ v013 + v025 + v190 + v714 + v106 + v502, design= designBJ, family = poisson(link = "log"), na.action = na.omit)  
BJmodel <- svyglm(BJREC\_summary~v013 + v025 + v190 + v714 + v106 + v502, design= designBJ, family = poisson(link = "log"), na.action = "na.omit") # Observations removed if it contains missing values  
summary(BJmodel)   
# calculate the RRs   
exp(coef(BJmodel))  
# round to make it clearer  
round(exp(coef(BJmodel)), 2)  
# RR to CI  
cbind(RR = exp(coef(BJmodel)), exp(confint(BJmodel)))  
# round to 2  
round(cbind(RR = exp(coef(BJmodel)), exp(confint(BJmodel))), 2) # Implication is estimate of the number of cervical cancer risk factors a woman had in Benin in 2018 given   
# her age, residential status, marital status, education, employment and wealth index.  
  
  
# Missing Data Included  
# Convert cases with NA into 0 and run a different model  
# library("imputeTS")  
#BMI  
freq(BJDHS$BJREC\_BMI, BJDHS$BJwt, na.rm= TRUE, plot = FALSE)  
BJDHS$BJREC\_BMI2 <- na.replace(BJDHS$BJREC\_BMI, 0)  
freq(BJDHS$BJREC\_BMI2, BJDHS$BJwt, na.rm= TRUE, plot = FALSE)  
#AFB  
freq(BJDHS$BJREC\_AFB, BJDHS$BJwt, na.rm= TRUE, plot = FALSE)  
BJDHS$BJREC\_AFB2 <- na.replace(BJDHS$BJREC\_AFB, 0)  
freq(BJDHS$BJREC\_AFB2, BJDHS$BJwt, na.rm= TRUE, plot = FALSE)  
# LSP women  
freq(BJDHS$BJREC\_LSP, BJDHS$BJwt, plot = FALSE)  
BJDHS$BJREC\_LSP2 <- na.replace(BJDHS$BJREC\_LSP, 0)  
freq(BJDHS$BJREC\_LSP2, BJDHS$BJwt, na.rm= TRUE, plot = FALSE)  
# NB- Type of Contraceptive use was linked with month of usage. month of use contains NA i.e. not in the universe.   
# Variables with missing values- BMI (149), LSP woman (56) and LSP men (193)  
BJDHS$BJREC\_summary2= as.numeric(BJDHS$BJREC\_SMOKE + BJDHS$BJREC\_BMI2 + BJDHS$BJREC\_AFSI + BJDHS$BJREC\_AFB2 + BJDHS$BJREC\_Parity + BJDHS$BJyears\_FP + BJDHS$BJREC\_LSP2 + BJDHS$BJREC\_STI)  
summary(BJDHS$BJREC\_summary2)  
favstats(BJDHS$BJREC\_summary2)  
designBJ <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~BJwt, data = BJDHS)  
is.numeric(BJDHS$BJREC\_summary2)  
designBJ <-update(designBJ, BJREC\_summary2=BJDHS$BJREC\_SMOKE + BJDHS$BJREC\_BMI2 + BJDHS$BJREC\_AFSI + BJDHS$BJREC\_AFB2 + BJDHS$BJREC\_Parity + BJDHS$BJyears\_FP + BJDHS$BJREC\_LSP2 + BJDHS$BJREC\_STI)  
  
# Unadjusted RR  
# Age  
BJmodelAge2 <- svyglm(BJREC\_summary2~v013, design= designBJ, family = poisson(link = "log"), na.action = "na.pass")  
summary(BJmodelAge2)  
exp(coef(BJmodelAge2)) # calculate the RRs  
round(exp(coef(BJmodelAge2)), 2) # round to make it clearer  
cbind(RR = exp(coef(BJmodelAge2)), exp(confint(BJmodelAge2))) # RR to CI  
round(cbind(RR = exp(coef(BJmodelAge2)), exp(confint(BJmodelAge2))), 2) # round to 2  
# Residential Status  
BJmodelRS2 <- svyglm(BJREC\_summary2~v025, design= designBJ, family = poisson(link = "log"), na.action = "na.pass")  
summary(BJmodelRS2)  
exp(coef(BJmodelRS2)) # calculate the RRs  
round(exp(coef(BJmodelRS2)), 2) # round to make it clearer  
cbind(RR = exp(coef(BJmodelRS2)), exp(confint(BJmodelRS2))) # RR to CI  
round(cbind(RR = exp(coef(BJmodelRS2)), exp(confint(BJmodelRS2))), 2) # round to 2  
# Marital Status  
BJmodelMS2 <- svyglm(BJREC\_summary2~v502, design= designBJ, family = poisson(link = "log"), na.action = "na.pass")  
summary(BJmodelMS2)  
exp(coef(BJmodelMS2)) # calculate the RRs  
round(exp(coef(BJmodelMS2)), 2) # round to make it clearer  
cbind(RR = exp(coef(BJmodelMS2)), exp(confint(BJmodelMS2))) # RR to CI  
round(cbind(RR = exp(coef(BJmodelMS2)), exp(confint(BJmodelMS2))), 2) # round to 2  
# Education  
BJmodelEdu2 <- svyglm(BJREC\_summary2~v106, design= designBJ, family = poisson(link = "log"), na.action = "na.pass")  
summary(BJmodelEdu2)  
exp(coef(BJmodelEdu2)) # calculate the RRs  
round(exp(coef(BJmodelEdu2)), 2) # round to make it clearer  
cbind(RR = exp(coef(BJmodelEdu2)), exp(confint(BJmodelEdu2))) # RR to CI  
round(cbind(RR = exp(coef(BJmodelEdu2)), exp(confint(BJmodelEdu2))), 2) # round to 2  
# Employment  
BJmodelEmp2 <- svyglm(BJREC\_summary2~v714, design= designBJ, family = poisson(link = "log"), na.action = "na.pass")  
summary(BJmodelEmp2)  
exp(coef(BJmodelEmp2)) # calculate the RRs  
round(exp(coef(BJmodelEmp2)), 2) # round to make it clearer  
cbind(RR = exp(coef(BJmodelEmp2)), exp(confint(BJmodelEmp2))) # RR to CI  
round(cbind(RR = exp(coef(BJmodelEmp2)), exp(confint(BJmodelEmp2))), 2) # round to 2  
# Wealth Index  
BJmodelWI2 <- svyglm(BJREC\_summary2~v190, design= designBJ, family = poisson(link = "log"), na.action = "na.pass")  
summary(BJmodelWI2)  
exp(coef(BJmodelWI2)) # calculate the RRs  
round(exp(coef(BJmodelWI2)), 2) # round to make it clearer  
cbind(RR = exp(coef(BJmodelWI2)), exp(confint(BJmodelWI2))) # RR to CI  
round(cbind(RR = exp(coef(BJmodelWI2)), exp(confint(BJmodelWI2))), 2) # round to 2  
  
# Adjusted RR  
# Keep NA  
BJmodel2 <- svyglm(BJREC\_summary2~v013 + v025 + v190 + v714 + v106 + v502, design= designBJ, family = poisson(link = "log"), na.action = "na.pass") # keep all data, including NAs  
summary(BJmodel2)  
# calculate the RRs   
exp(coef(BJmodel2))  
# round to make it clearer  
round(exp(coef(BJmodel2)), 2)  
# RR to CI  
cbind(RR = exp(coef(BJmodel2)), exp(confint(BJmodel2)))  
# round to 2  
round(cbind(RR = exp(coef(BJmodel2)), exp(confint(BJmodel2))), 2)  
  
  
# Better Model  
# Comparing two models, AIC or BIC  
AIC(BJmodel, BJmodel2)  
BIC(BJmodel, BJmodel2, maximal = BJmodel2)  
plot\_summs(BJmodel, scale = TRUE, exp = TRUE)  
plot\_summs(BJmodel2, scale = TRUE, exp = TRUE)  
plot\_summs(BJmodel, BJmodel2, ci\_level = 0.95, model.names = c("Model 1"= "Model Excluded Missing", "Model 2"= "Model INcluded Missing"),   
 colors = 'CUD Bright', exp = TRUE, legend.title = "Graph Comparing Models" )  
# Graph of Effect Size BJModel (Missing Data Excluded Model)  
BJeff.pres <- allEffects(BJmodel, scale = T)  
plot(BJeff.pres, ask=F)  
BJeff.pres2 <- allEffects(BJmodel2, scale = T)  
plot(BJeff.pres2, ask=F)

# Cameroon

# Select needed variable from women dataset in Cameroon  
CMDHS <- subset(CMDHSa, select = c(v001, v002, v021, v005, v025, v463z, v437, v438, v525, v212, v201,  
 v312, v337, v836, v763a, v012, v013, v502, v106, v714, v190, v130,   
 v717, v157, v158, v159, v171b, v171a, v481, v467b, v467c, v467d, v467f, v3a08d,  
 s1202, s1207, s1211, s1213, s1215, s1217, s1219, s1221, s1223, s1224, s1226))  
  
# Sample weight/design for women  
CMAGE49$CMwt <-CMAGE49$v005/1000000  
designCM <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~CMwt, data = CMAGE49)  
  
# Demographics  
CMAGE49 <- subset(CMDHS, v012 >= 15 & v012 <= 49)  
CMAGE49$CMwt <- CMAGE49$v005/1000000  
freq(CMAGE49$v013, CMAGE49$CMwt, plot = FALSE) #Age  
freq(CMAGE49$v025, CMAGE49$CMwt, plot = FALSE) #Residential status  
freq(CMAGE49$v502, CMAGE49$CMwt, plot = FALSE) #Marital status  
freq(CMAGE49$v106, CMAGE49$CMwt, plot = FALSE) #Education  
freq(CMAGE49$v714, CMAGE49$CMwt, plot = FALSE) #Employment   
freq(CMAGE49$v190, CMAGE49$CMwt, plot = FALSE) #Wealth index  
designCM <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~CMwt, data = CMAGE49) # Mean of Age and Confidence Interval  
CMttAge <- svyttest(v012~0, designCM)  
CMttAge  
  
#Proportion of Women With Risk Factors for Cervical Cancer  
#1. Use cigarettes and tobacco   
freq(CMAGE49$v463z, CMAGE49$CMwt, plot = FALSE)   
sum(is.na(CMAGE49$v463z))  
#2. BMI   
favstats(CMAGE49$v438, CMAGE49$CMwt, plot = FALSE)   
favstats(CMDHS$v438, CMAGE49$CMwt, plot = FALSE)   
is.na(CMAGE49$v438)  
sum(is.na(CMAGE49$v438))  
is.na(CMAGE49$v438) <- which(CMAGE49$v438 == 9996)  
is.na(CMAGE49$v438) <- which(CMAGE49$v438 == 9995)  
is.na(CMAGE49$v438) <- which(CMAGE49$v438 == 9994)  
favstats(CMAGE49$v438, CMAGE49$CMwt, plot = FALSE)  
CMAGE49$CMHtM <- (CMAGE49$v438/1000) #Height (Height is in cm, converted into meters. Also into 1 implied decimal)  
favstats(CMAGE49$CMHtM)  
favstats(CMAGE49$v437, CMAGE49$CMwt, plot = FALSE)   
is.na(CMAGE49$v437)  
sum(is.na(CMAGE49$v437))  
is.na(CMAGE49$v437) <- which(CMAGE49$v437 == 9996)  
is.na(CMAGE49$v437) <- which(CMAGE49$v437 == 9995)  
is.na(CMAGE49$v437) <- which(CMAGE49$v437 == 9994)  
favstats(CMAGE49$v437, CMAGE49$CMwt, plot = FALSE)  
CMAGE49$CMwtkg <- (CMAGE49$v437/10) #Weight  
favstats(CMAGE49$CMwtkg)  
CMAGE49$BMI <- (CMAGE49$CMwtkg) /(CMAGE49$CMHtM \* CMAGE49$CMHtM) # Calculated BMI  
favstats(CMAGE49$BMI, CMAGE49$CMwt, plot = FALSE)  
CMAGE49$BMICAT <- NA  
CMAGE49$BMICAT[CMAGE49$BMI < 18.50000] <-1  
CMAGE49$BMICAT[CMAGE49$BMI >= 18.50000 & CMAGE49$BMI <= 24.9999] <-2  
CMAGE49$BMICAT[CMAGE49$BMI >= 25.00000 & CMAGE49$BMI <= 29.9999] <-3  
CMAGE49$BMICAT[CMAGE49$BMI >= 30.00000] <-4  
CMAGE49$BMICAT <- factor(CMAGE49$BMICAT, levels = c(1,2,3,4), labels = c("Underweight", "Normal", "Overweight", "Obese"))  
CMAGE49$BJwt <- CMAGE49$v005/1000000  
freq(CMAGE49$BMICAT, CMAGE49$CMwt, plot = FALSE)  
CMBiCAT <- NA # Binary Categories of BMI   
CMBiCAT[CMAGE49$BMI < 25.0000] <- "low risk bmi"  
CMBiCAT[CMAGE49$BMI >= 25.0000] <- "high risk bmi"  
freq(CMBiCAT, CMAGE49$CMwt, plot = FALSE)  
#3. Age at First Sexual Intercourse   
freq(CMAGE49$v525, CMAGE49$CMwt, plot = FALSE)   
sum(is.na(CMAGE49$v525))  
CMAFSI <- subset(CMAGE49, v525 >=8 & v525 <=39, na.rm=TRUE)  
freq(CMAFSI$v525, plot = FALSE)  
CMAFSI$CMAFSIwt <- CMAFSI$v005/1000000  
designCMAFSI <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~CMAFSIwt, data = CMAFSI)  
svymean(~v525, design = designCMAFSI, na.rm = FALSE)  
confint(svymean(~v525, designCMAFSI))  
CMAFSICAT <- NA # Binary Categories of Age at First Sexual Intercourse Combine, had sex or not   
CMAFSICAT[CMAGE49$v525 <=17.9999] <- "less or equal 17 years"  
CMAFSICAT[CMAGE49$v525 >= 18.0000] <- "greater than 17 years"  
freq(CMAFSICAT, CMAGE49$CMwt, plot = FALSE)  
#4. Age at First Birth   
freq(CMAGE49$v212,CMAGE49$CMwt, plot = FALSE) # NA not in the universe   
sum(is.na(CMAGE49$v212))  
favstats(CMAGE49$v212, CMAGE49$CMwt, plot = FALSE)   
CMAFB <- subset(CMAGE49, v212 >=10 & v212 <=41, na.rm=TRUE)  
freq(CMAFB$v212, plot = FALSE)  
CMAFB$CMAFBwt <- CMAFB$v005/1000000  
designCMAFB <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~CMAFBwt, data = CMAFB)  
svymean(~v212, design = designCMAFB, na.rm = FALSE)  
confint(svymean(~v212, designCMAFB))  
CMAFBCAT <- NA # Binary Categories of Age at First Birth  
CMAFBCAT[CMAGE49$v212 <=19.9999] <- "less or equal 19 years"  
CMAFBCAT[CMAGE49$v212 >= 20.0000] <- "greater than 19 years"  
freq(CMAFBCAT, CMAGE49$CMwt, plot = FALSE)  
#5. Total Children Ever Born (Parity)  
freq(CMAGE49$v201, CMAGE49$CMwt, plot = FALSE)   
sum(is.na(CMAGE49$v201))  
CMP <- subset(CMAGE49, v201 >=1 & v201 <=16)  
freq(CMP$v201, plot = FALSE)  
CMP$CMPwt <- CMP$v005/1000000  
designCMP <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~CMPwt, data = CMP)  
svymean(~v201, design = designCMP, na.rm = FALSE)  
confint(svymean(~v201, designCMP))  
CMPCAT <- NA # Binary Categories of Total Children Ever Born (Parity)  
CMPCAT[CMAGE49$v201<=2.9999] <- "less or equal 2 births"  
CMPCAT[CMAGE49$v201 >= 3.0000] <- "greater than 2 births"  
freq(CMPCAT,CMAGE49$CMwt, plot = FALSE)  
#6. Total Lifetime Number of Sex Partners #Missing Data Excluded(280)   
freq(CMAGE49$v836, CMAGE49$CMwt, plot = FALSE)  
sum(is.na(CMAGE49$v836))  
favstats(CMAGE49$v836, CMAGE49$CMwt, plot = FALSE)  
is.na(CMAGE49$v836) <- which(CMAGE49$v836 == 98)  
favstats(CMAGE49$v836, CMAGE49$CMwt, plot = FALSE)  
CMLSP <- subset(CMAGE49, v836 >=1 & v836 <= 95)   
freq(CMLSP$v836, plot = FALSE)  
CMLSP$CMLSPwt <- CMLSP$v005/1000000  
designCMLSP <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~CMLSPwt, data = CMLSP)  
svymean(~v836, design = designCMLSP, na.rm = FALSE)  
confint(svymean(~v836, designCMLSP))  
CMLSPCAT <- NA # Binary Categories of Total Lifetime Number of Sex Partners  
CMLSPCAT[CMAGE49$v836 <=5.9999] <- "less or equal 5 partners"  
CMLSPCAT[CMAGE49$v836 >= 6.0000] <- "greater than 5 partners"  
freq(CMLSPCAT, CMAGE49$CMwt, plot = FALSE)  
#7. Contraceptive Usage  
freq(CMAGE49$v312, CMAGE49$CMwt, plot = FALSE)  
sum(is.na(CMAGE49$v312))  
nlevels(CMAGE49$v312)  
table(CMAGE49$v312)  
CMDHSContwt <- freq(CMAGE49$v312, CMAGE49$CMwt, plot = FALSE)  
CMDHSContwt  
CMCont\_recoded <- c("not using" = 0, "pill" =1, "iud"=1, "injections"=1, "diaphragm"= 2, "male condom"=2, "female sterilization"=2,  
 "male sterilization"=2, "periodic abstinence"=2, "withdrawal"=2, "other traditional"=2, "implants/norplant"=1,  
 "prolonged abstinence"= 2, "lactational amenorrhea (lam)"=2, "female condom"=2, "foam or jelly"= 2, "emergency contraception"= 1,  
 "other modern method"= 2, "standard days method (sdm)"=2, "specific method 1"= 2, "specific method 2"= 2)  
CMAGE49$CMREC\_CONT <- CMCont\_recoded[CMAGE49$v312]  
freq(CMAGE49$CMREC\_CONT, CMAGE49$CMwt, plot = FALSE)  
CMCont\_recoded2<- c("not using" = 0, "pill" =1, "iud"=1, "injections"=1, "diaphragm"= 0, "male condom"=0, "female sterilization"=0,  
 "male sterilization"=0, "periodic abstinence"=0, "withdrawal"=0, "other traditional"=0, "implants/norplant"=1,  
 "prolonged abstinence"= 0, "lactational amenorrhea (lam)"=0, "female condom"=0, "foam or jelly"= 0, "emergency contraception"= 1,  
 "other modern method"= 0, "standard days method (sdm)"=0, "specific method 1"= 0, "specific method 2"= NA)  
CMAGE49$CMREC\_CONT2 <- CMCont\_recoded2[CMAGE49$v312]  
freq(CMAGE49$CMREC\_CONT2, CMAGE49$CMwt, plot = FALSE)  
# Months of use of current family planning method, For women who are currently using a contraceptive method  
freq(CMAGE49$v337, CMAGE49$CMwt, plot = FALSE)  
sum(is.na(CMAGE49$v337))  
favstats(CMAGE49$v337, CMAGE49$CMwt, plot = FALSE)  
CMMFP <- NA  
CMMFP[CMAGE49$v337 <60.0000]<- "Less than 5 years"  
CMMFP[CMAGE49$v337 >=60.0000]<- "More than or equal 5 years"  
freq(CMMFP, CMAGE49$CMwt, plot = FALSE)  
# Combine number hormonal and non hormonal to number of years  
CMAGE49$CMcom <- paste(CMAGE49$v312, CMAGE49$v337)  
  
CMAGE49$CMcom  
CMAGE49$CMcom <- paste(CMAGE49$v312, "-", CMAGE49$v337)  
CMAGE49$CMcom  
CMAGE49$CMcom <- paste(CMAGE49$v312, CMMFP, sep = "-" )  
CMAGE49$CMcom  
freq(CMAGE49$CMcom, plot = FALSE)  
CMcom\_recoded <- c("emergency contraception-Less than 5 years" = 0, "emergency contraception-More than or equal 5 years"= 1, "female condom-Less than 5 years"=0, "female condom-More than or equal 5 years"=0, "female sterilization-Less than 5 years" = 0, "female sterilization-More than or equal 5 years" = 0,   
 "implants/norplant-Less than 5 years" = 0, "implants/norplant-More than or equal 5 years" = 1, "injections-Less than 5 years" =0, "injections-More than or equal 5 years" =1,  
 "iud-Less than 5 years" = 0, "iud-More than or equal 5 years"= 1, "lactational amenorrhea (lam)-Less than 5 years"= 0, "male condom-Less than 5 years" =0, "male condom-More than or equal 5 years" = 0, "male sterilization-More than or equal 5 years" = 0,  
 "not using-NA" =0, "other modern method-Less than 5 years" = 0, "periodic abstinence-Less than 5 years"=0, "periodic abstinence-More than or equal 5 years"=0, "pill-Less than 5 years" =0, "pill-More than or equal 5 years"=0, "standard days method (sdm)-Less than 5 years"=0, "standard days method (sdm)-More than or equal 5 years"=0,  
 "withdrawal-Less than 5 years"=0, "withdrawal-More than or equal 5 years"= 0)  
  
CMAGE49$CMyears\_FP <- CMcom\_recoded[CMAGE49$CMcom]  
freq(CMAGE49$CMyears\_FP, CMAGE49$CMwt, plot = FALSE)  
#8. Sexually Transmitted Infection in last 12 months ("don't know" added to "no")  
freq(CMAGE49$v763a, CMAGE49$CMwt, plot = FALSE)  
sum(is.na(CMAGE49$v763a))  
CMSTI\_recoded <- c("no"= 0, "yes" = 1, "don't know"= 0)  
CMAGE49$CMREC\_STI <- CMSTI\_recoded[CMAGE49$v763a]  
freq(CMAGE49$CMREC\_STI, CMAGE49$CMwt, plot = FALSE)  
  
  
#Number of Risk Factors  
# Smoking  
freq(CMAGE49$v463z, CMAGE49$CMwt, plot = FALSE)  
CMsmoke\_recoded <- c("no" = 1, "yes, smokes nothing" = 0)  
CMAGE49$CMREC\_SMOKE <- CMsmoke\_recoded[CMAGE49$v463z]  
freq(CMAGE49$CMREC\_SMOKE, CMAGE49$CMwt, plot = FALSE)  
#BMI  
freq(CMBiCAT, CMAGE49$CMwt, plot = FALSE)  
CMbmi\_recoded <- c("high risk bmi" = 1, "low risk bmi" = 0)  
CMAGE49$CMREC\_BMI <- CMbmi\_recoded[CMBiCAT]  
freq(CMAGE49$CMREC\_BMI, CMAGE49$CMwt, plot = FALSE)  
#AFSI  
freq(CMAFSICAT, CMAGE49$CMwt, plot = FALSE)  
CMAFSI\_recoded <- c("less or equal 17 years" = 1, "greater than 17 years" = 0)  
CMAGE49$CMREC\_AFSI <- CMAFSI\_recoded[CMAFSICAT]  
freq(CMAGE49$CMREC\_AFSI, CMAGE49$CMwt, plot = FALSE)  
#AFB  
freq(CMAFBCAT, CMAGE49$CMwt, plot = FALSE)  
CMAFB\_recoded <- c("less or equal 19 years" = 1, "greater than 19 years" = 0)  
CMAGE49$CMREC\_AFB <- CMAFB\_recoded[CMAFBCAT]  
freq(CMAGE49$CMREC\_AFB, CMAGE49$CMwt, plot = FALSE)  
#Parity  
freq(CMPCAT, CMAGE49$CMwt, plot = FALSE)  
CMParity\_recoded <- c("greater than 2 births" = 1, "less or equal 2 births" = 0)  
CMAGE49$CMREC\_Parity <- CMParity\_recoded[CMPCAT]  
freq(CMAGE49$CMREC\_Parity, CMAGE49$CMwt, plot = FALSE)  
#FP  
freq(CMAGE49$CMcom, plot = FALSE)  
CMcom\_recoded <- c("emergency contraception-Less than 5 years" = 0, "emergency contraception-More than or equal 5 years"= 1, "female condom-Less than 5 years"=0, "female condom-More than or equal 5 years"=0, "female sterilization-Less than 5 years" = 0, "female sterilization-More than or equal 5 years" = 0,   
 "implants/norplant-Less than 5 years" = 0, "implants/norplant-More than or equal 5 years" = 1, "injections-Less than 5 years" =0, "injections-More than or equal 5 years" =1,  
 "iud-Less than 5 years" = 0, "iud-More than or equal 5 years"= 1, "lactational amenorrhea (lam)-Less than 5 years"= 0, "male condom-Less than 5 years" =0, "male condom-More than or equal 5 years" = 0, "male sterilization-More than or equal 5 years" = 0,  
 "not using-NA" =0, "other modern method-Less than 5 years" = 0, "periodic abstinence-Less than 5 years"=0, "periodic abstinence-More than or equal 5 years"=0, "pill-Less than 5 years" =0, "pill-More than or equal 5 years"=0, "standard days method (sdm)-Less than 5 years"=0, "standard days method (sdm)-More than or equal 5 years"=0,  
 "withdrawal-Less than 5 years"=0, "withdrawal-More than or equal 5 years"= 0)  
CMAGE49$CMyears\_FP <- CMcom\_recoded[CMAGE49$CMcom]  
freq(CMAGE49$CMyears\_FP, CMAGE49$CMwt, plot = FALSE)  
# LSP woman  
freq(CMLSPCAT, CMAGE49$CMwt, plot = FALSE)  
CMLSP\_recoded <- c("greater than 5 partners" =1, "less or equal 5 partners" =0)  
CMAGE49$CMREC\_LSP <- CMLSP\_recoded[CMLSPCAT]  
freq(CMAGE49$CMREC\_LSP, CMAGE49$CMwt, plot = FALSE)  
# STI  
freq(CMAGE49$v763a, CMAGE49$CMwt, plot = FALSE)  
CMSTI\_recoded <- c("no"= 0, "yes" = 1, "don't know"= 0)  
CMAGE49$CMREC\_STI <- CMSTI\_recoded[CMAGE49$v763a]  
freq(CMAGE49$CMREC\_STI, CMAGE49$CMwt, plot = FALSE)  
  
  
#Number of Risk Factors, \* problems with Height, weight and contraceptives  
CMRiskcom <- cbind.data.frame(CMAGE49$CMREC\_SMOKE, CMAGE49$CMREC\_BMI, CMAGE49$CMREC\_AFSI, CMAGE49$CMREC\_AFB, CMAGE49$CMREC\_Parity, CMAGE49$CMyears\_FP, CMAGE49$CMREC\_LSP, CMAGE49$CMREC\_STI)  
CMRiskcom  
CMtotal <- cbind(CMRiskcom, total = rowSums(CMRiskcom, na.rm = TRUE))  
CMtotal <- cbind(CMRiskcom, total = rowSums(CMRiskcom, na.rm = FALSE))  
tail(CMtotal)  
head(CMtotal)  
  
# Regrouped number of Risk Factors  
CMTotalnew <- NA   
CMTotalnew[CMtotal$total <=4.9999] <- "less or equal 4"  
CMTotalnew[CMtotal$total >= 5.0000] <- "greater or equal 5" # >=6 number of risk factors  
freq(CMTotalnew, CMAGE49$CMwt, plot = FALSE)  
  
# Modelling  
# Missing Data Removed  
CMAGE49$CMREC\_summary= as.numeric(CMAGE49$CMREC\_SMOKE + CMAGE49$CMREC\_BMI + CMAGE49$CMREC\_AFSI + CMAGE49$CMREC\_AFB + CMAGE49$CMREC\_Parity + CMAGE49$CMyears\_FP + CMAGE49$CMREC\_LSP + CMAGE49$CMREC\_STI)  
crosstab(CMAGE49$CMREC\_summary, CMAGE49$v013, CMAGE49$CMwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(CMAGE49$v013, CMAGE49$CMREC\_summary, CMAGE49$CMwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(CMAGE49$CMREC\_summary, CMAGE49$v106, CMAGE49$CMwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(CMAGE49$v106, CMAGE49$CMREC\_summary, CMAGE49$CMwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
summary(CMAGE49$CMREC\_summary)  
favstats(CMAGE49$CMREC\_summary)  
designCM <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~CMwt, data = CMAGE49)  
is.numeric(CMAGE49$CMREC\_summary)  
designCM <-update(designCM, CMREC\_summary=CMAGE49$CMREC\_SMOKE + CMAGE49$CMREC\_BMI + CMAGE49$CMREC\_AFSI + CMAGE49$CMREC\_AFB + CMAGE49$CMREC\_Parity + CMAGE49$CMyears\_FP + CMAGE49$CMREC\_LSP + CMAGE49$CMREC\_STI)  
  
# Unadjusted RR  
# Age  
CMmodelAge <- svyglm(CMREC\_summary~v013, design= designCM, family = poisson(link = "log"), na.action = na.omit)  
summary(CMmodelAge)  
exp(coef(CMmodelAge)) # calculate the RRs  
round(exp(coef(CMmodelAge)), 2) # round to make it clearer  
cbind(RR = exp(coef(CMmodelAge)), exp(confint(CMmodelAge))) # RR to CI  
round(cbind(RR = exp(coef(CMmodelAge)), exp(confint(CMmodelAge))), 2) # round to 2  
# Residential Status  
CMmodelRS <- svyglm(CMREC\_summary~v025, design= designCM, family = poisson(link = "log"), na.action = na.omit)  
summary(CMmodelRS)  
exp(coef(CMmodelRS)) # calculate the RRs  
round(exp(coef(CMmodelRS)), 2) # round to make it clearer  
cbind(RR = exp(coef(CMmodelRS)), exp(confint(CMmodelRS))) # RR to CI  
round(cbind(RR = exp(coef(CMmodelRS)), exp(confint(CMmodelRS))), 2) # round to 2  
# Marital Status  
CMmodelMS <- svyglm(CMREC\_summary~v502, design= designCM, family = poisson(link = "log"), na.action = na.omit)  
summary(CMmodelMS)  
exp(coef(CMmodelMS)) # calculate the RRs  
round(exp(coef(CMmodelMS)), 2) # round to make it clearer  
cbind(RR = exp(coef(CMmodelMS)), exp(confint(CMmodelMS))) # RR to CI  
round(cbind(RR = exp(coef(CMmodelMS)), exp(confint(CMmodelMS))), 2) # round to 2  
# Education  
CMmodelEdu <- svyglm(CMREC\_summary~v106, design= designCM, family = poisson(link = "log"), na.action = na.omit)  
summary(CMmodelEdu)  
exp(coef(CMmodelEdu)) # calculate the RRs  
round(exp(coef(CMmodelEdu)), 2) # round to make it clearer  
cbind(RR = exp(coef(CMmodelEdu)), exp(confint(CMmodelEdu))) # RR to CI  
round(cbind(RR = exp(coef(CMmodelEdu)), exp(confint(CMmodelEdu))), 2) # round to 2  
# Employment  
CMmodelEmp <- svyglm(CMREC\_summary~v714, design= designCM, family = poisson(link = "log"), na.action = na.omit)  
summary(CMmodelEmp)  
exp(coef(CMmodelEmp)) # calculate the RRs  
round(exp(coef(CMmodelEmp)), 2) # round to make it clearer  
cbind(RR = exp(coef(CMmodelEmp)), exp(confint(CMmodelEmp))) # RR to CI  
round(cbind(RR = exp(coef(CMmodelEmp)), exp(confint(CMmodelEmp))), 2) # round to 2  
# Wealth Index  
CMmodelWI <- svyglm(CMREC\_summary~v190, design= designCM, family = poisson(link = "log"), na.action = na.omit)  
summary(CMmodelWI)  
exp(coef(CMmodelWI)) # calculate the RRs  
round(exp(coef(CMmodelWI)), 2) # round to make it clearer  
cbind(RR = exp(coef(CMmodelWI)), exp(confint(CMmodelWI))) # RR to CI  
round(cbind(RR = exp(coef(CMmodelWI)), exp(confint(CMmodelWI))), 2) # round to 2  
# Adjusted RR  
# Poisson Regression Modelling  
CMmodel <- svyglm(CMREC\_summary~v013 + v025 + v190 + v714 + v106 + v502, design= designCM, family = poisson(link = "log"), na.action = "na.omit") # Observations removed if it contains missing values  
summary(CMmodel)   
# calculate the RRs   
exp(coef(CMmodel))  
# round to make it clearer  
round(exp(coef(CMmodel)), 2)  
# RR to CI  
cbind(RR = exp(coef(CMmodel)), exp(confint(CMmodel)))  
# round to 2  
round(cbind(RR = exp(coef(CMmodel)), exp(confint(CMmodel))), 2)   
  
# Missing Data Included  
# Convert cases with NA into 0 and run a different model  
# library("imputeTS")  
#BMI  
freq(CMAGE49$CMREC\_BMI, CMAGE49$CMwt, na.rm= TRUE, plot = FALSE)  
CMAGE49$CMREC\_BMI2 <- na.replace(CMAGE49$CMREC\_BMI, 0)  
freq(CMAGE49$CMREC\_BMI2, CMAGE49$CMwt, na.rm= TRUE, plot = FALSE)  
#AFB  
freq(CMAGE49$CMREC\_AFB, CMAGE49$CMwt, na.rm= TRUE, plot = FALSE)  
CMAGE49$CMREC\_AFB2 <- na.replace(CMAGE49$CMREC\_AFB, 0)  
freq(CMAGE49$CMREC\_AFB2, CMAGE49$CMwt, na.rm= TRUE, plot = FALSE)  
# LSP women  
freq(CMAGE49$CMREC\_LSP, CMAGE49$CMwt, plot = FALSE)  
CMAGE49$CMREC\_LSP2 <- na.replace(CMAGE49$CMREC\_LSP, 0)  
freq(CMAGE49$CMREC\_LSP2, CMAGE49$CMwt, na.rm= TRUE, plot = FALSE)  
  
CMAGE49$CMREC\_summary2= as.numeric(CMAGE49$CMREC\_SMOKE + CMAGE49$CMREC\_BMI2 + CMAGE49$CMREC\_AFSI + CMAGE49$CMREC\_AFB2 + CMAGE49$CMREC\_Parity + CMAGE49$CMyears\_FP + CMAGE49$CMREC\_LSP2 + CMAGE49$CMREC\_STI)  
summary(CMAGE49$CMREC\_summary2)  
favstats(CMAGE49$CMREC\_summary2)  
designCM <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~CMwt, data = CMAGE49)  
is.numeric(CMAGE49$CMREC\_summary2)  
designCM <-update(designCM, BJREC\_summary2=CMAGE49$CMREC\_SMOKE + CMAGE49$CMREC\_BMI2 + CMAGE49$CMREC\_AFSI + CMAGE49$CMREC\_AFB2 + CMAGE49$CMREC\_Parity + CMAGE49$CMyears\_FP + CMAGE49$CMREC\_LSP2 + CMAGE49$CMREC\_STI)  
  
# Unadjusted RR  
# Age  
CMmodelAge2 <- svyglm(CMREC\_summary2~v013, design= designCM, family = poisson(link = "log"), na.action = "na.pass")  
summary(CMmodelAge2)  
exp(coef(CMmodelAge2)) # calculate the RRs  
round(exp(coef(CMmodelAge2)), 2) # round to make it clearer  
cbind(RR = exp(coef(CMmodelAge2)), exp(confint(CMmodelAge2))) # RR to CI  
round(cbind(RR = exp(coef(CMmodelAge2)), exp(confint(CMmodelAge2))), 2) # round to 2  
# Residential Status  
CMmodelRS2 <- svyglm(CMREC\_summary2~v025, design= designCM, family = poisson(link = "log"), na.action = "na.pass")  
summary(CMmodelRS2)  
exp(coef(CMmodelRS2)) # calculate the RRs  
round(exp(coef(CMmodelRS2)), 2) # round to make it clearer  
cbind(RR = exp(coef(CMmodelRS2)), exp(confint(CMmodelRS2))) # RR to CI  
round(cbind(RR = exp(coef(CMmodelRS2)), exp(confint(CMmodelRS2))), 2) # round to 2  
# Marital Status  
CMmodelMS2 <- svyglm(CMREC\_summary2~v502, design= designCM, family = poisson(link = "log"), na.action = "na.pass")  
summary(CMmodelMS2)  
exp(coef(CMmodelMS2)) # calculate the RRs  
round(exp(coef(CMmodelMS2)), 2) # round to make it clearer  
cbind(RR = exp(coef(CMmodelMS2)), exp(confint(CMmodelMS2))) # RR to CI  
round(cbind(RR = exp(coef(CMmodelMS2)), exp(confint(CMmodelMS2))), 2) # round to 2  
# Education  
CMmodelEdu2 <- svyglm(CMREC\_summary2~v106, design= designCM, family = poisson(link = "log"), na.action = "na.pass")  
summary(CMmodelEdu2)  
exp(coef(CMmodelEdu2)) # calculate the RRs  
round(exp(coef(CMmodelEdu2)), 2) # round to make it clearer  
cbind(RR = exp(coef(CMmodelEdu2)), exp(confint(CMmodelEdu2))) # RR to CI  
round(cbind(RR = exp(coef(CMmodelEdu2)), exp(confint(CMmodelEdu2))), 2) # round to 2  
# Employment  
CMmodelEmp2 <- svyglm(CMREC\_summary2~v714, design= designCM, family = poisson(link = "log"), na.action = "na.pass")  
summary(CMmodelEmp2)  
exp(coef(CMmodelEmp2)) # calculate the RRs  
round(exp(coef(CMmodelEmp2)), 2) # round to make it clearer  
cbind(RR = exp(coef(CMmodelEmp2)), exp(confint(CMmodelEmp2))) # RR to CI  
round(cbind(RR = exp(coef(CMmodelEmp2)), exp(confint(CMmodelEmp2))), 2) # round to 2  
# Wealth Index  
CMmodelWI2 <- svyglm(CMREC\_summary2~v190, design= designCM, family = poisson(link = "log"), na.action = "na.pass")  
summary(CMmodelWI2)  
exp(coef(CMmodelWI2)) # calculate the RRs  
round(exp(coef(CMmodelWI2)), 2) # round to make it clearer  
cbind(RR = exp(coef(CMmodelWI2)), exp(confint(CMmodelWI2))) # RR to CI  
round(cbind(RR = exp(coef(CMmodelWI2)), exp(confint(CMmodelWI2))), 2) # round to 2  
  
# Adjusted RR  
# Keep NA  
CMmodel2 <- svyglm(CMREC\_summary2~v013 + v025 + v190 + v714 + v106 + v502, design= designCM, family = poisson(link = "log"), na.action = "na.pass") # keep all data, including NAs  
summary(CMmodel2)  
# calculate the RRs   
exp(coef(CMmodel2))  
# round to make it clearer  
round(exp(coef(CMmodel2)), 2)  
# RR to CI  
cbind(RR = exp(coef(CMmodel2)), exp(confint(CMmodel2)))  
# round to 2  
round(cbind(RR = exp(coef(CMmodel2)), exp(confint(CMmodel2))), 2)  
  
# Better Model  
# Comparing two models, AIC or BIC  
AIC(CMmodel, CMmodel2)  
BIC(CMmodel, CMmodel2, maximal = CMmodel2)  
plot\_summs(CMmodel, scale = TRUE, exp = TRUE)  
plot\_summs(CMmodel2, scale = TRUE, exp = TRUE)  
plot\_summs(CMmodel, CMmodel2, ci\_level = 0.95, model.names = c("Model 1"= "Model Excluded Missing", "Model 2"= "Model INcluded Missing"),   
 colors = 'CUD Bright', exp = TRUE, legend.title = "Graph Comparing Models" )  
# Graph of Effect Size BJModel (Missing Data Excluded Model)  
CMeff.pres <- allEffects(CMmodel, scale = T)  
plot(CMeff.pres, ask=F)  
CMeff.pres2 <- allEffects(CMmodel2, scale = T)  
plot(CMeff.pres2, ask=F)

# The Gambia

# Select needed variable from women dataset in The Gambia  
GMDHS <- subset(GMDHSa, select = c(v001, v002, v021, v005, v025, v463z, v437, v438, v525, v212, v201,  
 v312, v337, v836, v763a, v012, v013, v502, v106, v714, v190, v130,   
 v717, v157, v158, v159, v171b, v481, v467b, v467c, v467d, v467f))  
  
# Sample weight/design for women  
GMDHS$GMwt <-GMDHS$v005/1000000  
designGM <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~GMwt, data = GMDHS)  
  
  
# Demographics  
GMDHS$GMwt <- GMDHS$v005/1000000  
freq(GMDHS$v013, GMDHS$GMwt, plot = FALSE) #Age  
freq(GMDHS$v025, GMDHS$GMwt, plot = FALSE) #Residential status  
freq(GMDHS$v502, GMDHS$GMwt, plot = FALSE) #Marital status  
freq(GMDHS$v106, GMDHS$GMwt, plot = FALSE) #Education  
freq(GMDHS$v714, GMDHS$GMwt, plot = FALSE) #Employment   
freq(GMDHS$v190, GMDHS$GMwt, plot = FALSE) #Wealth index  
designGM <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~GMwt, data = GMDHS) # Mean of Age and Confidence Interval  
GMttAge <- svyttest(v012~0, designGM)  
GMttAge  
  
#Proportion of Women With Risk Factors for Cervical Cancer  
# Use cigarettes and tobacco  
freq(GMDHS$v463z, GMDHS$GMwt, plot = FALSE)   
sum(is.na(GMDHS$v463z))  
# BMI   
favstats(GMDHS$v438, GMDHS$GMwt, plot = FALSE)  
sum(is.na(GMDHS$v438))  
is.na(GMDHS$v438)  
is.na(GMDHS$v438) <- which(GMDHS$v438 == 9996)  
is.na(GMDHS$v438) <- which(GMDHS$v438 == 9995)  
is.na(GMDHS$v438) <- which(GMDHS$v438 == 9994)  
favstats(GMDHS$v438, GMDHS$GMwt, plot = FALSE)  
GMDHS$GMHtM <- (GMDHS$v438/1000) #Height (Height is in cm, converted into meters. Also into 1 implied decimal)  
favstats(GMDHS$GMHtM)  
favstats(GMDHS$v437, GMDHS$GMwt, plot = FALSE)  
is.na(GMDHS$v437)  
sum(is.na(GMDHS$v437))  
is.na(GMDHS$v437) <- which(GMDHS$v437 == 9996)  
is.na(GMDHS$v437) <- which(GMDHS$v437 == 9995)  
is.na(GMDHS$v437) <- which(GMDHS$v437 == 9994)  
favstats(GMDHS$v437, GMDHS$GMwt, plot = FALSE)  
GMDHS$GMwtkg <- (GMDHS$v437/10) #Weight  
favstats(GMDHS$GMwtkg)  
GMDHS$BMI <- (GMDHS$GMwtkg) /(GMDHS$GMHtM \* GMDHS$GMHtM) # Calculated BMI  
favstats(GMDHS$BMI, GMDHS$GMwt, plot = FALSE)  
GMDHS$BMICAT <- NA  
GMDHS$BMICAT[GMDHS$BMI < 18.50000] <-1  
GMDHS$BMICAT[GMDHS$BMI >= 18.50000 & GMDHS$BMI <= 24.9999] <-2  
GMDHS$BMICAT[GMDHS$BMI >= 25.00000 & GMDHS$BMI <= 29.9999] <-3  
GMDHS$BMICAT[GMDHS$BMI >= 30.00000] <-4  
GMDHS$BMICAT <- factor(GMDHS$BMICAT, levels = c(1,2,3,4), labels = c("Underweight", "Normal", "Overweight", "Obese"))  
GMDHS$GMwt <- GMDHS$v005/1000000  
freq(GMDHS$BMICAT, GMDHS$GMwt, plot = FALSE)  
GMBiCAT <- NA # Binary Categories of BMI   
GMBiCAT[GMDHS$BMI < 25.0000] <- "low risk bmi"  
GMBiCAT[GMDHS$BMI >= 25.0000] <- "high risk bmi"  
freq(GMBiCAT, GMDHS$GMwt, plot = FALSE)  
# Age at First Sexual Intercourse   
freq(GMDHS$v525, GMDHS$GMwt, plot = FALSE)   
sum(is.na(GMDHS$v525))  
GMAFSI <- subset(GMDHS, v525 >=8 & v525 <=42)  
freq(GMAFSI$v525, plot = FALSE)  
GMAFSI$GMAFSIwt <- GMAFSI$v005/1000000  
designGMAFSI <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~GMAFSIwt, data = GMAFSI)  
svymean(~v525, design = designGMAFSI, na.rm = FALSE)  
confint(svymean(~v525, designGMAFSI))  
GMAFSICAT <- NA # Binary Categories of Age at First Sexual Intercourse   
GMAFSICAT[GMDHS$v525 <=17.9999] <- "less or equal 17 years"  
GMAFSICAT[GMDHS$v525 >= 18.0000] <- "greater than 17 years"  
freq(GMAFSICAT, GMDHS$GMwt, plot = FALSE)  
#Age at First Birth   
freq(GMDHS$v212, GMDHS$GMwt, plot = FALSE)   
sum(is.na(GMDHS$v212))  
favstats(GMDHS$v212, GMDHS$GMwt, plot = FALSE)  
GMAFB <- subset(GMDHS, v212 >=12 & v525 <=48)  
freq(GMAFB$v212, plot = FALSE)  
GMAFB$GMAFBwt <- GMAFB$v005/1000000  
designGMAFB <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~GMAFBwt, data = GMAFB)  
svymean(~v212, design = designGMAFB, na.rm = FALSE)  
confint(svymean(~v212, designGMAFB))  
GMAFBCAT <- NA # Binary Categories of Age at First Birth  
GMAFBCAT[GMDHS$v212 <=19.9999] <- "less or equal 19 years"  
GMAFBCAT[GMDHS$v212 >= 20.0000] <- "greater than 19 years"  
freq(GMAFBCAT, GMDHS$GMwt, plot = FALSE)  
#Total Children Ever Born (Parity)  
freq(GMDHS$v201, GMDHS$GMwt, plot = FALSE)   
sum(is.na(GMDHS$v201))  
GMP <- subset(GMDHS, v201 >=1 & v201 <=14)  
freq(GMP$v201, plot = FALSE)  
GMP$GMPwt <- GMP$v005/1000000  
designGMP <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~GMPwt, data = GMP)  
svymean(~v201, design = designGMP, na.rm = FALSE)  
confint(svymean(~v201, designGMP))  
GMPCAT <- NA # Binary Categories of Total Children Ever Born (Parity)  
GMPCAT[GMDHS$v201<=2.9999] <- "less or equal 2 births"  
GMPCAT[GMDHS$v201 >= 3.0000] <- "greater than 2 births"  
freq(GMPCAT, GMDHS$GMwt, plot = FALSE)  
# Total Lifetime Number of Sex Partners   
freq(GMDHS$v836, GMDHS$GMwt, plot = FALSE)   
sum(is.na(GMDHS$v836))  
favstats(GMDHS$v836, GMDHS$GMwt, plot = FALSE)  
is.na(GMDHS$v836) <- which(GMDHS$v836 == 98)  
favstats(GMDHS$v836, GMDHS$GMwt, plot = FALSE)  
GMLSP <- subset(GMDHS, v836 >=1 & v836 <=95)  
freq(GMLSP$v836, plot = FALSE)  
GMLSP$GMLSPwt <- GMLSP$v005/1000000  
designGMLSP <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~GMLSPwt, data = GMLSP)  
svymean(~v836, design = designGMLSP, na.rm = FALSE)  
confint(svymean(~v836, designGMLSP))  
GMLSPCAT <- NA # Binary Categories of Total Lifetime Number of Sex Partners  
GMLSPCAT[GMDHS$v836 <=5.9999] <- "less or equal 5 partners"  
GMLSPCAT[GMDHS$v836 >= 6.0000] <- "greater than 5 partners"  
freq(GMLSPCAT, GMDHS$GMwt, plot = FALSE)  
#7. Contraceptive Usage  
freq(GMDHS$v312, GMDHS$GMwt, plot = FALSE)  
is.na(GMDHS$v312)  
sum(is.na(GMDHS$v312))  
nlevels(GMDHS$v312)  
table(GMDHS$v312)  
GMDHSContwt <- freq(GMDHS$v312, GMDHS$GMwt, plot = FALSE)  
GMDHSContwt  
GMCont\_recoded <- c("not using" = 0, "pill" =1, "iud"=1, "injections"=1, "diaphragm"= 2, "male condom"=2, "female sterilization"=2,  
 "male sterilization"=2, "periodic abstinence"=2, "withdrawal"=2, "other traditional"=2, "implants/norplant"=1,  
 "prolonged abstinence"= 2, "lactational amenorrhea (lam)"=2, "female condom"=2, "foam or jelly"= 2, "emergency contraception"= 1,  
 "other modern method"= 2, "standard days method (sdm)"=2, "specific method 1"= 2, "specific method 2"= 2)  
GMDHS$GMREC\_CONT <- GMCont\_recoded[GMDHS$v312]  
freq(GMDHS$GMREC\_CONT, GMDHS$GMwt, plot = FALSE)  
GMCont\_recoded <- c("not using" = 0, "pill" =1, "iud"=1, "injections"=1, "diaphragm"= 0, "male condom"=0, "female sterilization"=0,  
 "male sterilization"=0, "periodic abstinence"=0, "withdrawal"=0, "other traditional"=0, "implants/norplant"=1,  
 "prolonged abstinence"= 0, "lactational amenorrhea (lam)"=0, "female condom"=0, "foam or jelly"= 0, "emergency contraception"= 1,  
 "other modern method"= 0, "standard days method (sdm)"=0, "specific method 1"= 0, "specific method 2"= 0)  
GMDHS$GMREC\_CONT <- GMCont\_recoded[GMDHS$v312]  
freq(GMDHS$GMREC\_CONT, GMDHS$GMwt, plot = FALSE)  
# Months of use of current family planning method, For women who are currently using a contraceptive method  
freq(GMDHS$v337, GMDHS$GMwt, plot = FALSE)   
is.na(GMDHS$v337)  
sum(is.na(GMDHS$v337))  
favstats(GMDHS$v337, GMDHS$GMwt, plot = FALSE)  
summary(GMDHS$v337, GMDHS$GMwt, plot = FALSE)  
GMMFP <- NA  
GMMFP[GMDHS$v337 <60.0000]<- "Less than 5 years"  
GMMFP[GMDHS$v337 >=60.0000]<- "More than or equal 5 years"  
freq(GMMFP, GMDHS$GMwt, plot = FALSE)  
# Combine number hormonal and non hormonal to number of years  
GMDHS$GMcom <- paste(GMDHS$v312, GMDHS$v337)  
GMDHS$GMcom  
GMDHS$GMcom <- paste(GMDHS$v312, "-", GMDHS$v337)  
GMDHS$GMcom  
GMDHS$GMcom <- paste(GMDHS$v312, GMMFP, sep = "-" )  
GMDHS$GMcom  
freq(GMDHS$GMcom, plot = FALSE)  
GMcom\_recoded <- c("female condom-Less than 5 years" = 0, "female sterilization-Less than 5 years"= 0, "female sterilization-More than or equal 5 years" = 0, "implants/norplant-Less than 5 years" = 0,   
 "implants/norplant-More than or equal 5 years" = 1, "injections-Less than 5 years" =0, "injections-More than or equal 5 years" =1,  
 "iud-Less than 5 years" = 0, "iud-More than or equal 5 years"= 1, "lactational amenorrhea (lam)-Less than 5 years"= 0, "male condom-Less than 5 years" =0, "male condom-More than or equal 5 years" = 0,  
 "male sterilization-More than or equal 5 years" = 0, "not using-NA" =0, "other traditional-Less than 5 years"= 0, "other traditional-More than or equal 5 years"=0, "periodic abstinence-Less than 5 years"=0,  
 "periodic abstinence-More than or equal 5 years"=0, "pill-Less than 5 years" =0, "pill-More than or equal 5 years"=1, "standard days method (sdm)-Less than 5 years"=0,   
 "withdrawal-Less than 5 years"=0)  
GMDHS$GMyears\_FP <- GMcom\_recoded[GMDHS$GMcom]  
freq(GMDHS$GMyears\_FP, GMDHS$GMwt, plot = FALSE)  
#8.Sexually Transmitted Infection in last 12 months ("don't know" added to "no")  
freq(GMDHS$v763a, GMDHS$GMwt, plot = FALSE)  
sum(is.na(GMDHS$v763a))  
GMSTI\_recoded <- c("no"= 0, "yes" = 1, "don't know"= 0)  
GMDHS$GMREC\_STI <- GMSTI\_recoded[GMDHS$v763a]  
freq(GMDHS$GMREC\_STI, GMDHS$GMwt, plot = FALSE)  
  
  
#Number of Risk Factors  
# Smoking  
freq(GMDHS$v463z, GMDHS$GMwt, plot = FALSE)  
GMsmoke\_recoded <- c("no" = 1, "yes, smokes nothing" = 0)  
GMDHS$GMREC\_SMOKE <- GMsmoke\_recoded[GMDHS$v463z]  
freq(GMDHS$GMREC\_SMOKE, GMDHS$GMwt, plot = FALSE)  
#BMI  
freq(GMBiCAT, GMDHS$GMwt, plot = FALSE)  
GMbmi\_recoded <- c("high risk bmi" = 1, "low risk bmi" = 0)  
GMDHS$GMREC\_BMI <- GMbmi\_recoded[GMBiCAT]  
freq(GMDHS$GMREC\_BMI, GMDHS$GMwt, plot = FALSE)  
#AFSI  
freq(GMAFSICAT, GMDHS$GMwt, plot = FALSE)  
GMAFSI\_recoded <- c("less or equal 17 years" = 1, "greater than 17 years" = 0)  
GMDHS$GMREC\_AFSI <- GMAFSI\_recoded[GMAFSICAT]  
freq(GMDHS$GMREC\_AFSI, GMDHS$GMwt, plot = FALSE)  
#AFB  
freq(GMAFBCAT, GMDHS$GMwt, plot = FALSE)  
GMAFB\_recoded <- c("less or equal 19 years" = 1, "greater than 19 years" = 0)  
GMDHS$GMREC\_AFB <- GMAFB\_recoded[GMAFBCAT]  
freq(GMDHS$GMREC\_AFB, GMDHS$GMwt, plot = FALSE)  
#Parity  
freq(GMPCAT, GMDHS$GMwt, plot = FALSE)  
GMParity\_recoded <- c("greater than 2 births" = 1, "less or equal 2 births" = 0)  
GMDHS$GMREC\_Parity <- GMParity\_recoded[GMPCAT]  
freq(GMDHS$GMREC\_Parity, GMDHS$GMwt, plot = FALSE)  
#FP  
freq(GMDHS$GMcom, plot = FALSE)  
GMcom\_recoded <- c("female condom-Less than 5 years" = 0, "female sterilization-Less than 5 years"= 0, "female sterilization-More than or equal 5 years" = 0, "implants/norplant-Less than 5 years" = 0,   
 "implants/norplant-More than or equal 5 years" = 1, "injections-Less than 5 years" =0, "injections-More than or equal 5 years" =1,  
 "iud-Less than 5 years" = 0, "iud-More than or equal 5 years"= 1, "lactational amenorrhea (lam)-Less than 5 years"= 0, "male condom-Less than 5 years" =0, "male condom-More than or equal 5 years" = 0,  
 "male sterilization-More than or equal 5 years" = 0, "not using-NA" =0, "other traditional-Less than 5 years"= 0, "other traditional-More than or equal 5 years"=0, "periodic abstinence-Less than 5 years"=0,  
 "periodic abstinence-More than or equal 5 years"=0, "pill-Less than 5 years" =0, "pill-More than or equal 5 years"=1, "standard days method (sdm)-Less than 5 years"=0,   
 "withdrawal-Less than 5 years"=0)  
GMDHS$GMyears\_FP <- GMcom\_recoded[GMDHS$GMcom]  
freq(GMDHS$GMyears\_FP, GMDHS$GMwt, plot = FALSE)  
# LSP  
freq(GMLSPCAT, GMDHS$GMwt, plot = FALSE)  
GMLSP\_recoded <- c("greater than 5 partners" = 1, "less or equal 5 partners" = 0)  
GMDHS$GMREC\_LSP <- GMLSP\_recoded[GMLSPCAT]  
freq(GMDHS$GMREC\_LSP, GMDHS$GMwt, plot = FALSE)  
# STI  
freq(GMDHS$v763a, GMDHS$GMwt, plot = FALSE)  
GMSTI\_recoded <- c("no"= 0, "yes" = 1, "don't know"= 0)  
GMDHS$GMREC\_STI <- GMSTI\_recoded[GMDHS$v763a]  
freq(GMDHS$GMREC\_STI, GMDHS$GMwt, plot = FALSE)  
  
  
#Number of Risk Factors,   
GMRiskcom <- data.frame(GMDHS$GMREC\_SMOKE, GMDHS$GMREC\_BMI, GMDHS$GMREC\_AFSI, GMDHS$GMREC\_AFB, GMDHS$GMREC\_Parity, GMDHS$GMyears\_FP, GMDHS$GMREC\_LSP, GMDHS$GMREC\_STI)  
GMRiskcom  
GMtotal <- cbind(GMRiskcom, total = rowSums(GMRiskcom, na.rm = TRUE))  
GMtotal <- cbind(GMRiskcom, total = rowSums(GMRiskcom, na.rm = FALSE))  
tail(GMtotal)  
head(GMtotal)  
# Regrouped number of Risk Factors  
GMTotalnew <- NA   
GMTotalnew[GMtotal$total <=4.9999] <- "less or equal 4"  
GMTotalnew[GMtotal$total >= 5.0000] <- "greater or equal 5"   
freq(GMTotalnew, GMDHS$GMwt, plot = FALSE)  
  
# Modelling  
# Missing Data Removed  
GMDHS$GMREC\_summary= as.numeric(GMDHS$GMREC\_SMOKE + GMDHS$GMREC\_BMI + GMDHS$GMREC\_AFSI + GMDHS$GMREC\_AFB + GMDHS$GMREC\_Parity + GMDHS$GMyears\_FP + GMDHS$GMREC\_LSP + GMDHS$GMREC\_STI)  
crosstab(GMDHS$GMREC\_summary, GMDHS$v013, GMDHS$GMwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(GMDHS$v013, GMDHS$GMREC\_summary, GMDHS$GMwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(GMDHS$GMREC\_summary, GMDHS$v106, GMDHS$GMwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(GMDHS$v106, GMDHS$GMREC\_summary, GMDHS$GMwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
summary(GMDHS$GMREC\_summary)  
favstats(GMDHS$GMREC\_summary)  
designGM <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~GMwt, data = GMDHS)  
is.numeric(GMDHS$GMREC\_summary)  
designGM <-update(designGM, GMREC\_summary=GMDHS$GMREC\_SMOKE + GMDHS$GMREC\_BMI + GMDHS$GMREC\_AFSI + GMDHS$GMREC\_AFB + GMDHS$GMREC\_Parity + GMDHS$GMyears\_FP + GMDHS$GMREC\_LSP + GMDHS$GMREC\_STI)  
  
# Unadjusted RR  
# Age  
GMmodelAge <- svyglm(GMREC\_summary~v013, design= designGM, family = poisson(link = "log"), na.action = na.omit)  
summary(GMmodelAge)  
exp(coef(GMmodelAge)) # calculate the RRs  
round(exp(coef(GMmodelAge)), 2) # round to make it clearer  
cbind(RR = exp(coef(GMmodelAge)), exp(confint(GMmodelAge))) # RR to CI  
round(cbind(RR = exp(coef(GMmodelAge)), exp(confint(GMmodelAge))), 2) # round to 2  
# Residential Status  
GMmodelRS <- svyglm(GMREC\_summary~v025, design= designGM, family = poisson(link = "log"), na.action = na.omit)  
summary(GMmodelRS)  
exp(coef(GMmodelRS)) # calculate the RRs  
round(exp(coef(GMmodelRS)), 2) # round to make it clearer  
cbind(RR = exp(coef(GMmodelRS)), exp(confint(GMmodelRS))) # RR to CI  
round(cbind(RR = exp(coef(GMmodelRS)), exp(confint(GMmodelRS))), 2) # round to 2  
# Marital Status  
GMmodelMS <- svyglm(GMREC\_summary~v502, design= designGM, family = poisson(link = "log"), na.action = na.omit)  
summary(GMmodelMS)  
exp(coef(GMmodelMS)) # calculate the RRs  
round(exp(coef(GMmodelMS)), 2) # round to make it clearer  
cbind(RR = exp(coef(GMmodelMS)), exp(confint(GMmodelMS))) # RR to CI  
round(cbind(RR = exp(coef(GMmodelMS)), exp(confint(GMmodelMS))), 2) # round to 2  
# Education  
GMmodelEdu <- svyglm(GMREC\_summary~v106, design= designGM, family = poisson(link = "log"), na.action = na.omit)  
summary(GMmodelEdu)  
exp(coef(GMmodelEdu)) # calculate the RRs  
round(exp(coef(GMmodelEdu)), 2) # round to make it clearer  
cbind(RR = exp(coef(GMmodelEdu)), exp(confint(GMmodelEdu))) # RR to CI  
round(cbind(RR = exp(coef(GMmodelEdu)), exp(confint(GMmodelEdu))), 2) # round to 2  
# Employment  
GMmodelEmp <- svyglm(GMREC\_summary~v714, design= designGM, family = poisson(link = "log"), na.action = na.omit)  
summary(GMmodelEmp)  
exp(coef(GMmodelEmp)) # calculate the RRs  
round(exp(coef(GMmodelEmp)), 2) # round to make it clearer  
cbind(RR = exp(coef(GMmodelEmp)), exp(confint(GMmodelEmp))) # RR to CI  
round(cbind(RR = exp(coef(GMmodelEmp)), exp(confint(GMmodelEmp))), 2) # round to 2  
# Wealth Index  
GMmodelWI <- svyglm(GMREC\_summary~v190, design= designGM, family = poisson(link = "log"), na.action = na.omit)  
summary(GMmodelWI)  
exp(coef(GMmodelWI)) # calculate the RRs  
round(exp(coef(GMmodelWI)), 2) # round to make it clearer  
cbind(RR = exp(coef(GMmodelWI)), exp(confint(GMmodelWI))) # RR to CI  
round(cbind(RR = exp(coef(GMmodelWI)), exp(confint(GMmodelWI))), 2) # round to 2  
# Adjusted RR  
# Poisson Regression Modelling  
#BJmodel <- svyglm(BJREC\_SMOKE+BJREC\_BMI+BJREC\_AFSI+BJREC\_AFB+BJREC\_Parity+BJyears\_FP+BJREC\_LSP+BJREC\_STI+BJh\_LSP2~ v013 + v025 + v190 + v714 + v106 + v502, design= designBJ, family = poisson(link = "log"), na.action = na.omit)  
GMmodel <- svyglm(GMREC\_summary~v013 + v025 + v190 + v714 + v106 + v502, design= designGM, family = poisson(link = "log"), na.action = "na.omit") # Observations removed if it contains missing values  
summary(GMmodel)   
# calculate the RRs   
exp(coef(GMmodel))  
# round to make it clearer  
round(exp(coef(GMmodel)), 2)  
# RR to CI  
cbind(RR = exp(coef(GMmodel)), exp(confint(GMmodel)))  
# round to 2  
round(cbind(RR = exp(coef(GMmodel)), exp(confint(GMmodel))), 2) # Implication is estimate of the number of cervical cancer risk factors a woman had in Benin in 2018 given   
# her age, residential status, marital status, education, employment and wealth index.  
  
# Missing Data Included  
# Convert cases with NA into 0 and run a different model  
# library("imputeTS")  
#BMI  
freq(GMDHS$GMREC\_BMI, GMDHS$GMwt, na.rm= TRUE, plot = FALSE)  
GMDHS$GMREC\_BMI2 <- na.replace(GMDHS$GMREC\_BMI, 0)  
freq(GMDHS$GMREC\_BMI2, GMDHS$GMwt, na.rm= TRUE, plot = FALSE)  
#AFB  
freq(GMDHS$GMREC\_AFB, GMDHS$GMwt, na.rm= TRUE, plot = FALSE)  
GMDHS$GMREC\_AFB2 <- na.replace(GMDHS$GMREC\_AFB, 0)  
freq(GMDHS$GMREC\_AFB2, GMDHS$GMwt, na.rm= TRUE, plot = FALSE)  
# LSP women  
freq(GMDHS$GMREC\_LSP, GMDHS$GMwt, plot = FALSE)  
GMDHS$GMREC\_LSP2 <- na.replace(GMDHS$GMREC\_LSP, 0)  
freq(GMDHS$GMREC\_LSP2, GMDHS$GMwt, na.rm= TRUE, plot = FALSE)  
# NB- Type of Contraceptive use was linked with month of usage. month of use contains NA i.e. not in the universe.   
# Variables with missing values- BMI (149), LSP woman (56) and LSP men (193)  
GMDHS$GMREC\_summary2= as.numeric(GMDHS$GMREC\_SMOKE + GMDHS$GMREC\_BMI2 + GMDHS$GMREC\_AFSI + GMDHS$GMREC\_AFB2 + GMDHS$GMREC\_Parity + GMDHS$GMyears\_FP + GMDHS$GMREC\_LSP2 + GMDHS$GMREC\_STI)  
summary(GMDHS$GMREC\_summary2)  
favstats(GMDHS$GMREC\_summary2)  
designGM <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~GMwt, data = GMDHS)  
is.numeric(GMDHS$GMREC\_summary2)  
designGM <-update(designGM, GMREC\_summary2=GMDHS$GMREC\_SMOKE + GMDHS$GMREC\_BMI2 + GMDHS$GMREC\_AFSI + GMDHS$GMREC\_AFB2 + GMDHS$GMREC\_Parity + GMDHS$GMyears\_FP + GMDHS$GMREC\_LSP2 + GMDHS$GMREC\_STI)  
  
# Unadjusted RR  
# Age  
GMmodelAge2 <- svyglm(GMREC\_summary2~v013, design= designGM, family = poisson(link = "log"), na.action = "na.pass")  
summary(GMmodelAge2)  
exp(coef(GMmodelAge2)) # calculate the RRs  
round(exp(coef(GMmodelAge2)), 2) # round to make it clearer  
cbind(RR = exp(coef(GMmodelAge2)), exp(confint(GMmodelAge2))) # RR to CI  
round(cbind(RR = exp(coef(GMmodelAge2)), exp(confint(GMmodelAge2))), 2) # round to 2  
# Residential Status  
GMmodelRS2 <- svyglm(GMREC\_summary2~v025, design= designGM, family = poisson(link = "log"), na.action = "na.pass")  
summary(GMmodelRS2)  
exp(coef(GMmodelRS2)) # calculate the RRs  
round(exp(coef(GMmodelRS2)), 2) # round to make it clearer  
cbind(RR = exp(coef(GMmodelRS2)), exp(confint(GMmodelRS2))) # RR to CI  
round(cbind(RR = exp(coef(GMmodelRS2)), exp(confint(GMmodelRS2))), 2) # round to 2  
# Marital Status  
GMmodelMS2 <- svyglm(GMREC\_summary2~v502, design= designGM, family = poisson(link = "log"), na.action = "na.pass")  
summary(GMmodelMS2)  
exp(coef(GMmodelMS2)) # calculate the RRs  
round(exp(coef(GMmodelMS2)), 2) # round to make it clearer  
cbind(RR = exp(coef(GMmodelMS2)), exp(confint(GMmodelMS2))) # RR to CI  
round(cbind(RR = exp(coef(GMmodelMS2)), exp(confint(GMmodelMS2))), 2) # round to 2  
# Education  
GMmodelEdu2 <- svyglm(GMREC\_summary2~v106, design= designGM, family = poisson(link = "log"), na.action = "na.pass")  
summary(GMmodelEdu2)  
exp(coef(GMmodelEdu2)) # calculate the RRs  
round(exp(coef(GMmodelEdu2)), 2) # round to make it clearer  
cbind(RR = exp(coef(GMmodelEdu2)), exp(confint(GMmodelEdu2))) # RR to CI  
round(cbind(RR = exp(coef(GMmodelEdu2)), exp(confint(GMmodelEdu2))), 2) # round to 2  
# Employment  
GMmodelEmp2 <- svyglm(GMREC\_summary2~v714, design= designGM, family = poisson(link = "log"), na.action = "na.pass")  
summary(GMmodelEmp2)  
exp(coef(GMmodelEmp2)) # calculate the RRs  
round(exp(coef(GMmodelEmp2)), 2) # round to make it clearer  
cbind(RR = exp(coef(GMmodelEmp2)), exp(confint(GMmodelEmp2))) # RR to CI  
round(cbind(RR = exp(coef(GMmodelEmp2)), exp(confint(GMmodelEmp2))), 2) # round to 2  
# Wealth Index  
GMmodelWI2 <- svyglm(GMREC\_summary2~v190, design= designGM, family = poisson(link = "log"), na.action = "na.pass")  
summary(GMmodelWI2)  
exp(coef(GMmodelWI2)) # calculate the RRs  
round(exp(coef(GMmodelWI2)), 2) # round to make it clearer  
cbind(RR = exp(coef(GMmodelWI2)), exp(confint(GMmodelWI2))) # RR to CI  
round(cbind(RR = exp(coef(GMmodelWI2)), exp(confint(GMmodelWI2))), 2) # round to 2  
  
# Adjusted RR  
# Keep NA  
GMmodel2 <- svyglm(GMREC\_summary2~v013 + v025 + v190 + v714 + v106 + v502, design= designGM, family = poisson(link = "log"), na.action = "na.pass") # keep all data, including NAs  
summary(GMmodel2)  
# calculate the RRs   
exp(coef(GMmodel2))  
# round to make it clearer  
round(exp(coef(GMmodel2)), 2)  
# RR to CI  
cbind(RR = exp(coef(GMmodel2)), exp(confint(GMmodel2)))  
# round to 2  
round(cbind(RR = exp(coef(GMmodel2)), exp(confint(GMmodel2))), 2)  
  
# Better Model  
# Comparing two models, AIC or BIC  
AIC(GMmodel, GMmodel2)  
BIC(GMmodel, GMmodel2, maximal = GMmodel2)  
plot\_summs(GMmodel, scale = TRUE, exp = TRUE)  
plot\_summs(GMmodel2, scale = TRUE, exp = TRUE)  
plot\_summs(GMmodel, GMmodel2, ci\_level = 0.95, model.names = c("Model 1"= "Model Excluded Missing", "Model 2"= "Model INcluded Missing"),   
 colors = 'CUD Bright', exp = TRUE, legend.title = "Graph Comparing Models" )  
# Graph of Effect Size BJModel (Missing Data Excluded Model)  
GMeff.pres <- allEffects(GMmodel, scale = T)  
plot(GMeff.pres, ask=F)  
GMeff.pres2 <- allEffects(GMmodel2, scale = T)  
plot(GMeff.pres2, ask=F)

# Guinea

# Select needed variable from women dataset in Guinea  
GNDHS <- subset(GNDHSa, select = c(v001, v002, v021, v005, v025, v463z, v437, v438, v525, v212, v201,  
 v312, v337, v836, v763a, v012, v013, v502, v106, v714, v190, v130,   
 v717, v157, v158, v159, v171b, v481, v467b, v467c, v467d, v467f))  
  
# Sample weight/design for women  
GNDHS$GNwt <-GNDHS$v005/1000000  
designGN <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~GNwt, data = GNDHS)  
  
  
# Demographics  
GNDHS$GNwt <- GNDHS$v005/1000000  
freq(GNDHS$v013, GNDHS$GNwt, plot = FALSE) #Age  
freq(GNDHS$v025, GNDHS$GNwt, plot = FALSE) #Residential status  
freq(GNDHS$v502, GNDHS$GNwt, plot = FALSE) #Marital status  
freq(GNDHS$v106, GNDHS$GNwt, plot = FALSE) #Education  
freq(GNDHS$v714, GNDHS$GNwt, plot = FALSE) #Employment   
freq(GNDHS$v190, GNDHS$GNwt, plot = FALSE) #Wealth index  
designGN <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~GNwt, data = GNDHS) # Mean of Age and Confidence Interval  
GNttAge <- svyttest(v012~0, designGN)  
GNttAge  
  
#Proportion of Women With Risk Factors for Cervical Cancer  
#1. Use cigarettes and tobacco   
freq(GNDHS$v463z, GNDHS$GNwt, plot = FALSE)   
sum(is.na(GNDHS$v463z))  
#2. BMI   
favstats(GNDHS$v438, GNDHS$GNwt, plot = FALSE)  
sum(is.na(GNDHS$v438))  
is.na(GNDHS$v438)  
is.na(GNDHS$v438) <- which(GNDHS$v438 == 9996)  
is.na(GNDHS$v438) <- which(GNDHS$v438 == 9995)  
is.na(GNDHS$v438) <- which(GNDHS$v438 == 9994)  
favstats(GNDHS$v438, GNDHS$GNwt, plot = FALSE)  
GNDHS$GNHtM <- (GNDHS$v438/1000) #Height (Height is in cm, converted into meters. Also into 1 implied decimal)  
favstats(GNDHS$GNHtM)  
favstats(GNDHS$v437, GNDHS$GNwt, plot = FALSE)  
is.na(GNDHS$v437)  
sum(is.na(GNDHS$v437))  
is.na(GNDHS$v437) <- which(GNDHS$v437 == 9996)  
is.na(GNDHS$v437) <- which(GNDHS$v437 == 9995)  
is.na(GNDHS$v437) <- which(GNDHS$v437 == 9994)  
favstats(GNDHS$v437, GNDHS$GNwt, plot = FALSE)  
GNDHS$GNwtkg <- (GNDHS$v437/10) #Weight  
favstats(GNDHS$GNwtkg)  
GNDHS$BMI <- (GNDHS$GNwtkg) /(GNDHS$GNHtM \* GNDHS$GNHtM) # Calculated BMI  
favstats(GNDHS$BMI, GNDHS$GNwt, plot = FALSE)  
GNDHS$BMICAT <- NA  
GNDHS$BMICAT[GNDHS$BMI < 18.50000] <-1  
GNDHS$BMICAT[GNDHS$BMI >= 18.50000 & GNDHS$BMI <= 24.9999] <-2  
GNDHS$BMICAT[GNDHS$BMI >= 25.00000 & GNDHS$BMI <= 29.9999] <-3  
GNDHS$BMICAT[GNDHS$BMI >= 30.00000] <-4  
GNDHS$BMICAT <- factor(GNDHS$BMICAT, levels = c(1,2,3,4), labels = c("Underweight", "Normal", "Overweight", "Obese"))  
GNDHS$GNwt <- GNDHS$v005/1000000  
freq(GNDHS$BMICAT, GNDHS$GNwt, plot = FALSE)  
GNBiCAT <- NA # Binary Categories of BMI   
GNBiCAT[GNDHS$BMI < 25.0000] <- "low risk bmi"  
GNBiCAT[GNDHS$BMI >= 25.0000] <- "high risk bmi"  
freq(GNBiCAT, GNDHS$GNwt, plot = FALSE)  
#3. Age at First Sexual Intercourse   
freq(GNDHS$v525, GNDHS$GNwt, plot = FALSE)  
sum(is.na(GNDHS$v525))  
GNAFSI <- subset(GNDHS, v525 >=8 & v525 <=37)  
freq(GNAFSI$v525, plot = FALSE)  
GNAFSI$GNAFSIwt <- GNAFSI$v005/1000000  
designGNAFSI <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~GNAFSIwt, data = GNAFSI)  
svymean(~v525, design = designGNAFSI, na.rm = FALSE)  
confint(svymean(~v525, designGNAFSI))  
GNAFSICAT <- NA # Binary Categories of Age at First Sexual Intercourse   
GNAFSICAT[GNDHS$v525 <=17.9999] <- "less or equal 17 years"  
GNAFSICAT[GNDHS$v525 >= 18.0000] <- "greater than 17 years"  
freq(GNAFSICAT, GNDHS$GNwt, plot = FALSE)  
#4. Age at First Birth   
freq(GNDHS$v212, GNDHS$GNwt, plot = FALSE)   
sum(is.na(GNDHS$v212))  
favstats(GNDHS$v212, GNDHS$GNwt, plot = FALSE)  
GNAFB <- subset(GNDHS, v212 >=11 & v525 <=44)  
freq(GNAFB$v212, plot = FALSE)  
GNAFB$GNAFBwt <- GNAFB$v005/1000000  
designGNAFB <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~GNAFBwt, data = GNAFB)  
svymean(~v212, design = designGNAFB, na.rm = FALSE)  
confint(svymean(~v212, designGNAFB))  
GNAFBCAT <- NA # Binary Categories of Age at First Birth  
GNAFBCAT[GNDHS$v212 <=19.9999] <- "less or equal 19 years"  
GNAFBCAT[GNDHS$v212 >= 20.0000] <- "greater than 19 years"  
freq(GNAFBCAT, GNDHS$GNwt, plot = FALSE)  
#5. Total Children Ever Born (Parity)   
freq(GNDHS$v201, GNDHS$GNwt, plot = FALSE)   
sum(is.na(GNDHS$v201))  
GNP <- subset(GNDHS, v201 >=1 & v201 <=14)  
freq(GNP$v201, plot = FALSE)  
GNP$GNPwt <- GNP$v005/1000000  
designGNP <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~GNPwt, data = GNP)  
svymean(~v201, design = designGNP, na.rm = FALSE)  
confint(svymean(~v201, designGNP))  
GNPCAT <- NA # Binary Categories of Total Children Ever Born (Parity)  
GNPCAT[GNDHS$v201<=2.9999] <- "less or equal 2 births"  
GNPCAT[GNDHS$v201 >= 3.0000] <- "greater than 2 births"  
freq(GNPCAT, GNDHS$GNwt, plot = FALSE)  
#6. Total Lifetime Number of Sex Partners   
freq(GNDHS$v836, GNDHS$GNwt, plot = FALSE)   
sum(is.na(GNDHS$v836))  
favstats(GNDHS$v836, GNDHS$GNwt, plot = FALSE)  
is.na(GNDHS$v836) <- which(GNDHS$v836 == 98)  
favstats(GNDHS$v836, GNDHS$GNwt, plot = FALSE)  
GNLSP <- subset(GNDHS, v836 >=1 & v836 <=95)  
freq(GNLSP$v836, plot = FALSE)  
GNLSP$GNLSPwt <- GNLSP$v005/1000000  
designGNLSP <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~GNLSPwt, data = GNLSP)  
svymean(~v836, design = designGNLSP, na.rm = FALSE)  
confint(svymean(~v836, designGNLSP))  
GNLSPCAT <- NA # Binary Categories of Total Lifetime Number of Sex Partners  
GNLSPCAT[GNDHS$v836 <=5.9999] <- "less or equal 5 partners"  
GNLSPCAT[GNDHS$v836 >= 6.0000] <- "greater than 5 partners"  
freq(GNLSPCAT, GNDHS$GNwt, plot = FALSE)  
#7. Contraceptive Usage  
freq(GNDHS$v312, GNDHS$GNwt, plot = FALSE)  
is.na(GNDHS$v312)  
sum(is.na(GNDHS$v312))  
nlevels(GNDHS$v312)  
table(GNDHS$v312)  
GNDHSContwt <- freq(GNDHS$v312, GNDHS$GNwt, plot = FALSE)  
GNDHSContwt  
GNCont\_recoded <- c("not using" = 0, "pill" =1, "iud"=1, "injections"=1, "diaphragm"= 2, "male condom"=2, "female sterilization"=2,  
 "male sterilization"=2, "periodic abstinence"=2, "withdrawal"=2, "other traditional"=2, "implants/norplant"=1,  
 "prolonged abstinence"= 2, "lactational amenorrhea (lam)"=2, "female condom"=2, "foam or jelly"= 2, "emergency contraception"= 1,  
 "other modern method"= 2, "standard days method (sdm)"=2, "specific method 1"= 2, "specific method 2"= 2)  
GNDHS$GNREC\_CONT <- GNCont\_recoded[GNDHS$v312]  
freq(GNDHS$GNREC\_CONT, GNDHS$GNwt, plot = FALSE)  
GNCont\_recoded <- c("not using" = 0, "pill" =1, "iud"=1, "injections"=1, "diaphragm"= 0, "male condom"=0, "female sterilization"=0,  
 "male sterilization"=0, "periodic abstinence"=0, "withdrawal"=0, "other traditional"=0, "implants/norplant"=1,  
 "prolonged abstinence"= 0, "lactational amenorrhea (lam)"=0, "female condom"=0, "foam or jelly"= 0, "emergency contraception"= 1,  
 "other modern method"= 0, "standard days method (sdm)"=0, "specific method 1"= 0, "specific method 2"= NA)  
GNDHS$GNREC\_CONT <- GNCont\_recoded[GNDHS$v312]  
freq(GNDHS$GNREC\_CONT, GNDHS$GNwt, plot = FALSE)  
# Months of use of current family planning method, For women who are currently using a contraceptive method  
freq(GNDHS$v337, GNDHS$GNwt, plot = FALSE)   
is.na(GNDHS$v337)  
sum(is.na(GNDHS$v337))  
favstats(GNDHS$v337, GNDHS$GNwt, plot = FALSE)  
summary(GNDHS$v337, GNDHS$GNwt, plot = FALSE)  
GNMFP <- NA  
GNMFP[GNDHS$v337 <60.0000]<- "Less than 5 years"  
GNMFP[GNDHS$v337 >=60.0000]<- "More than or equal 5 years"  
freq(GNMFP, GNDHS$GNwt, plot = FALSE)  
# Combine number hormonal and non hormonal to number of years  
GNDHS$GNcom <- paste(GNDHS$v312, GNDHS$v337)  
GNDHS$GNcom  
GNDHS$GNcom <- paste(GNDHS$v312, "-", GNDHS$v337)  
GNDHS$GNcom  
GNDHS$GNcom <- paste(GNDHS$v312, GNMFP, sep = "-" )  
GMDHS$GMcom  
freq(GNDHS$GNcom, plot = FALSE)  
GNcom\_recoded <- c("emergency contraception-Less than 5 years" = 0, "female sterilization-Less than 5 years" = 0, "female sterilization-More than or equal 5 years" = 0,   
 "implants/norplant-Less than 5 years" = 0, "implants/norplant-More than or equal 5 years" = 1, "injections-Less than 5 years" =0, "injections-More than or equal 5 years" =1,  
 "iud-Less than 5 years" = 0, "iud-More than or equal 5 years"= 1, "lactational amenorrhea (lam)-Less than 5 years"= 0, "lactational amenorrhea (lam)-More than or equal 5 years"=0,  
 "male condom-Less than 5 years" =0, "male condom-More than or equal 5 years" = 0, "male sterilization-Less than 5 years" =0, "male sterilization-More than or equal 5 years" =0, "not using-NA" =0,   
 "other modern method-Less than 5 years"=0, "other traditional-Less than 5 years"= 0, "other traditional-More than or equal 5 years"=0, "periodic abstinence-Less than 5 years"=0,  
 "periodic abstinence-More than or equal 5 years"=0, "pill-Less than 5 years" =0, "pill-More than or equal 5 years"=1, "standard days method (sdm)-Less than 5 years"=0,   
 "standard days method (sdm)-More than or equal 5 years"=0,"withdrawal-Less than 5 years"=0)  
GNDHS$GNyears\_FP <- GNcom\_recoded[GNDHS$GNcom]  
freq(GNDHS$GNyears\_FP, GNDHS$GNwt, plot = FALSE)  
#8. Sexually Transmitted Infection in last 12 months ("don't know" added to "no")  
freq(GNDHS$v763a, GNDHS$GNwt, plot = FALSE)  
sum(is.na(GNDHS$v763a))  
GNSTI\_recoded <- c("no"= 0, "yes" = 1, "don't know"= 0)  
GNDHS$GNREC\_STI <- GNSTI\_recoded[GNDHS$v763a]  
freq(GNDHS$GNREC\_STI, GNDHS$GNwt, plot = FALSE)  
  
  
#Number of Risk Factors  
# Smoking  
freq(GNDHS$v463z, GNDHS$GNwt, plot = FALSE)  
GNsmoke\_recoded <- c("no" = 1, "yes, smokes nothing" = 0)  
GNDHS$GNREC\_SMOKE <- GNsmoke\_recoded[GNDHS$v463z]  
freq(GNDHS$GNREC\_SMOKE, GNDHS$GNwt, plot = FALSE)  
#BMI  
freq(GNBiCAT, GNDHS$GNwt, plot = FALSE)  
GNbmi\_recoded <- c("high risk bmi" = 1, "low risk bmi" = 0)  
GNDHS$GNREC\_BMI <- GNbmi\_recoded[GNBiCAT]  
freq(GNDHS$GNREC\_BMI, GNDHS$GNwt, plot = FALSE)  
#AFSI  
freq(GNAFSICAT, GNDHS$GNwt, plot = FALSE)  
GNAFSI\_recoded <- c("less or equal 17 years" = 1, "greater than 17 years" = 0)  
GNDHS$GNREC\_AFSI <- GNAFSI\_recoded[GNAFSICAT]  
freq(GNDHS$GNREC\_AFSI, GNDHS$GNwt, plot = FALSE)  
#AFB  
freq(GNAFBCAT, GNDHS$GNwt, plot = FALSE)  
GNAFB\_recoded <- c("less or equal 19 years" = 1, "greater than 19 years" = 0)  
GNDHS$GNREC\_AFB <- GNAFB\_recoded[GNAFBCAT]  
freq(GNDHS$GNREC\_AFB, GNDHS$GNwt, plot = FALSE)  
#Parity  
freq(GNPCAT, GNDHS$GNwt, plot = FALSE)  
GNParity\_recoded <- c("greater than 2 births" = 1, "less or equal 2 births" = 0)  
GNDHS$GNREC\_Parity <- GNParity\_recoded[GNPCAT]  
freq(GNDHS$GNREC\_Parity, GNDHS$GNwt, plot = FALSE)  
#FP  
freq(GNDHS$GNcom, plot = FALSE)  
GNcom\_recoded <- c("emergency contraception-Less than 5 years" = 0, "female sterilization-Less than 5 years" = 0, "female sterilization-More than or equal 5 years" = 0,   
 "implants/norplant-Less than 5 years" = 0, "implants/norplant-More than or equal 5 years" = 1, "injections-Less than 5 years" =0, "injections-More than or equal 5 years" =1,  
 "iud-Less than 5 years" = 0, "iud-More than or equal 5 years"= 1, "lactational amenorrhea (lam)-Less than 5 years"= 0, "lactational amenorrhea (lam)-More than or equal 5 years"=0,  
 "male condom-Less than 5 years" =0, "male condom-More than or equal 5 years" = 0, "male sterilization-Less than 5 years" =0, "male sterilization-More than or equal 5 years" =0, "not using-NA" =0,   
 "other modern method-Less than 5 years"=0, "other traditional-Less than 5 years"= 0, "other traditional-More than or equal 5 years"=0, "periodic abstinence-Less than 5 years"=0,  
 "periodic abstinence-More than or equal 5 years"=0, "pill-Less than 5 years" =0, "pill-More than or equal 5 years"=1, "standard days method (sdm)-Less than 5 years"=0,   
 "standard days method (sdm)-More than or equal 5 years"=0,"withdrawal-Less than 5 years"=0)  
GNDHS$GNyears\_FP <- GNcom\_recoded[GNDHS$GNcom]  
freq(GNDHS$GNyears\_FP, GNDHS$GNwt, plot = FALSE)  
# LSP  
freq(GNLSPCAT, GNDHS$GNwt, plot = FALSE)  
GNLSP\_recoded <- c("greater than 5 partners" = 1, "less or equal 5 partners" = 0)  
GNDHS$GNREC\_LSP <- GNLSP\_recoded[GNLSPCAT]  
freq(GNDHS$GNREC\_LSP, GNDHS$GNwt, plot = FALSE)  
# STI  
freq(GNDHS$v763a, GNDHS$GNwt, plot = FALSE)  
GNSTI\_recoded <- c("no"= 0, "yes" = 1, "don't know"= 0)  
GNDHS$GNREC\_STI <- GNSTI\_recoded[GNDHS$v763a]  
freq(GNDHS$GNREC\_STI, GNDHS$GNwt, plot = FALSE)  
  
#Number of Risk Factors, \* problems with Height, weight and contraceptives  
GNRiskcom <- data.frame(GNDHS$GNREC\_SMOKE, GNDHS$GNREC\_BMI, GNDHS$GNREC\_AFSI, GNDHS$GNREC\_AFB, GNDHS$GNREC\_Parity, GNDHS$GNyears\_FP, GNDHS$GNREC\_LSP, GNDHS$GNREC\_STI)  
GNRiskcom  
GNtotal <- cbind(GNRiskcom, total = rowSums(GNRiskcom, na.rm = TRUE))  
GNtotal <- cbind(GNRiskcom, total = rowSums(GNRiskcom, na.rm = FALSE))  
tail(GNtotal)  
head(GNtotal)  
  
# Regrouped number of Risk Factor  
GNTotalnew <- NA   
GNTotalnew[GNtotal$total <=4.9999] <- "less or equal 4"  
GNTotalnew[GNtotal$total >= 5.0000] <- "greater or equal 5"   
freq(GNTotalnew, GNDHS$GNwt, plot = FALSE)  
  
# Modelling  
# Missing Data Removed  
GNDHS$GNREC\_summary= as.numeric(GNDHS$GNREC\_SMOKE + GNDHS$GNREC\_BMI + GNDHS$GNREC\_AFSI + GNDHS$GNREC\_AFB + GNDHS$GNREC\_Parity + GNDHS$GNyears\_FP + GNDHS$GNREC\_LSP + GNDHS$GNREC\_STI)  
crosstab(GNDHS$GNREC\_summary, GNDHS$v013, GNDHS$GNwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(GNDHS$v013, GNDHS$GNREC\_summary, GNDHS$GNwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(GNDHS$GNREC\_summary, GNDHS$v106, GNDHS$GNwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(GNDHS$v106, GNDHS$GNREC\_summary, GNDHS$GNwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
summary(GNDHS$GNREC\_summary)  
favstats(GNDHS$GNREC\_summary)  
designGN <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~GNwt, data = GNDHS)  
is.numeric(GNDHS$GNREC\_summary)  
designGN <-update(designGN, GNREC\_summary=GNDHS$GNREC\_SMOKE + GNDHS$GNREC\_BMI + GNDHS$GNREC\_AFSI + GNDHS$GNREC\_AFB + GNDHS$GNREC\_Parity + GNDHS$GNyears\_FP + GNDHS$GNREC\_LSP + GNDHS$GNREC\_STI)  
  
  
# Unadjusted RR  
# Age  
GNmodelAge <- svyglm(GNREC\_summary~v013, design= designGN, family = poisson(link = "log"), na.action = na.omit)  
summary(GNmodelAge)  
exp(coef(GNmodelAge)) # calculate the RRs  
round(exp(coef(GNmodelAge)), 2) # round to make it clearer  
cbind(RR = exp(coef(GNmodelAge)), exp(confint(GNmodelAge))) # RR to CI  
round(cbind(RR = exp(coef(GNmodelAge)), exp(confint(GNmodelAge))), 2) # round to 2  
# Residential Status  
GNmodelRS <- svyglm(GNREC\_summary~v025, design= designGN, family = poisson(link = "log"), na.action = na.omit)  
summary(GNmodelRS)  
exp(coef(GNmodelRS)) # calculate the RRs  
round(exp(coef(GNmodelRS)), 2) # round to make it clearer  
cbind(RR = exp(coef(GNmodelRS)), exp(confint(GNmodelRS))) # RR to CI  
round(cbind(RR = exp(coef(GNmodelRS)), exp(confint(GNmodelRS))), 2) # round to 2  
# Marital Status  
GNmodelMS <- svyglm(GNREC\_summary~v502, design= designGN, family = poisson(link = "log"), na.action = na.omit)  
summary(GNmodelMS)  
exp(coef(GNmodelMS)) # calculate the RRs  
round(exp(coef(GNmodelMS)), 2) # round to make it clearer  
cbind(RR = exp(coef(GNmodelMS)), exp(confint(GNmodelMS))) # RR to CI  
round(cbind(RR = exp(coef(GNmodelMS)), exp(confint(GNmodelMS))), 2) # round to 2  
# Education  
GNmodelEdu <- svyglm(GNREC\_summary~v106, design= designGN, family = poisson(link = "log"), na.action = na.omit)  
summary(GNmodelEdu)  
exp(coef(GNmodelEdu)) # calculate the RRs  
round(exp(coef(GNmodelEdu)), 2) # round to make it clearer  
cbind(RR = exp(coef(GNmodelEdu)), exp(confint(GNmodelEdu))) # RR to CI  
round(cbind(RR = exp(coef(GNmodelEdu)), exp(confint(GNmodelEdu))), 2) # round to 2  
# Employment  
GNmodelEmp <- svyglm(GNREC\_summary~v714, design= designGN, family = poisson(link = "log"), na.action = na.omit)  
summary(GNmodelEmp)  
exp(coef(GNmodelEmp)) # calculate the RRs  
round(exp(coef(GNmodelEmp)), 2) # round to make it clearer  
cbind(RR = exp(coef(GNmodelEmp)), exp(confint(GNmodelEmp))) # RR to CI  
round(cbind(RR = exp(coef(GNmodelEmp)), exp(confint(GNmodelEmp))), 2) # round to 2  
# Wealth Index  
GNmodelWI <- svyglm(GNREC\_summary~v190, design= designGN, family = poisson(link = "log"), na.action = na.omit)  
summary(GNmodelWI)  
exp(coef(GNmodelWI)) # calculate the RRs  
round(exp(coef(GNmodelWI)), 2) # round to make it clearer  
cbind(RR = exp(coef(GNmodelWI)), exp(confint(GNmodelWI))) # RR to CI  
round(cbind(RR = exp(coef(GNmodelWI)), exp(confint(GNmodelWI))), 2) # round to 2  
# Adjusted RR  
# Poisson Regression Modelling  
GNmodel <- svyglm(GNREC\_summary~v013 + v025 + v190 + v714 + v106 + v502, design= designGN, family = poisson(link = "log"), na.action = "na.omit") # Observations removed if it contains missing values  
summary(GNmodel)   
# calculate the RRs   
exp(coef(GNmodel))  
# round to make it clearer  
round(exp(coef(GNmodel)), 2)  
# RR to CI  
cbind(RR = exp(coef(GNmodel)), exp(confint(GNmodel)))  
# round to 2  
round(cbind(RR = exp(coef(GNmodel)), exp(confint(GNmodel))), 2)   
  
# Missing Data Included  
# Convert cases with NA into 0 and run a different model  
# library("imputeTS")  
#BMI  
freq(GNDHS$GNREC\_BMI, GNDHS$GNwt, na.rm= TRUE, plot = FALSE)  
GNDHS$GNREC\_BMI2 <- na.replace(GNDHS$GNREC\_BMI, 0)  
freq(GNDHS$GNREC\_BMI2, GNDHS$GNwt, na.rm= TRUE, plot = FALSE)  
#AFB  
freq(GNDHS$GNREC\_AFB, GNDHS$GNwt, na.rm= TRUE, plot = FALSE)  
GNDHS$GNREC\_AFB2 <- na.replace(GNDHS$GNREC\_AFB, 0)  
freq(GNDHS$GNREC\_AFB2, GNDHS$GNwt, na.rm= TRUE, plot = FALSE)  
# LSP women  
freq(GNDHS$GNREC\_LSP, GNDHS$GNwt, plot = FALSE)  
GNDHS$GNREC\_LSP2 <- na.replace(GNDHS$GNREC\_LSP, 0)  
freq(GNDHS$GNREC\_LSP2, GNDHS$GNwt, na.rm= TRUE, plot = FALSE)  
# NB- Type of Contraceptive use was linked with month of usage. month of use contains NA i.e. not in the universe.   
# Variables with missing values- BMI (149), LSP woman (56) and LSP men (193)  
GNDHS$GNREC\_summary2= as.numeric(GNDHS$GNREC\_SMOKE + GNDHS$GNREC\_BMI2 + GNDHS$GNREC\_AFSI + GNDHS$GNREC\_AFB2 + GNDHS$GNREC\_Parity + GNDHS$GNyears\_FP + GNDHS$GNREC\_LSP2 + GNDHS$GNREC\_STI)  
summary(GNDHS$GNREC\_summary2)  
favstats(GNDHS$GNREC\_summary2)  
designGN <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~GNwt, data = GNDHS)  
is.numeric(GNDHS$GNREC\_summary2)  
designGN <-update(designGN, GNREC\_summary2=GNDHS$GNREC\_SMOKE + GNDHS$GNREC\_BMI2 + GNDHS$GNREC\_AFSI + GNDHS$GNREC\_AFB2 + GNDHS$GNREC\_Parity + GNDHS$GNyears\_FP + GNDHS$GNREC\_LSP2 + GNDHS$GNREC\_STI)  
  
  
# Unadjusted RR  
# Age  
GNmodelAge2 <- svyglm(GNREC\_summary2~v013, design= designGN, family = poisson(link = "log"), na.action = "na.pass")  
summary(GNmodelAge2)  
exp(coef(GNmodelAge2)) # calculate the RRs  
round(exp(coef(GNmodelAge2)), 2) # round to make it clearer  
cbind(RR = exp(coef(GNmodelAge2)), exp(confint(GNmodelAge2))) # RR to CI  
round(cbind(RR = exp(coef(GNmodelAge2)), exp(confint(GNmodelAge2))), 2) # round to 2  
# Residential Status  
GNmodelRS2 <- svyglm(GNREC\_summary2~v025, design= designGN, family = poisson(link = "log"), na.action = "na.pass")  
summary(GNmodelRS2)  
exp(coef(GNmodelRS2)) # calculate the RRs  
round(exp(coef(GNmodelRS2)), 2) # round to make it clearer  
cbind(RR = exp(coef(GNmodelRS2)), exp(confint(GNmodelRS2))) # RR to CI  
round(cbind(RR = exp(coef(GNmodelRS2)), exp(confint(GNmodelRS2))), 2) # round to 2  
# Marital Status  
GNmodelMS2 <- svyglm(GNREC\_summary2~v502, design= designGN, family = poisson(link = "log"), na.action = "na.pass")  
summary(GNmodelMS2)  
exp(coef(GNmodelMS2)) # calculate the RRs  
round(exp(coef(GNmodelMS2)), 2) # round to make it clearer  
cbind(RR = exp(coef(GNmodelMS2)), exp(confint(GNmodelMS2))) # RR to CI  
round(cbind(RR = exp(coef(GNmodelMS2)), exp(confint(GNmodelMS2))), 2) # round to 2  
# Education  
GNmodelEdu2 <- svyglm(GNREC\_summary2~v106, design= designGN, family = poisson(link = "log"), na.action = "na.pass")  
summary(GNmodelEdu2)  
exp(coef(GNmodelEdu2)) # calculate the RRs  
round(exp(coef(GNmodelEdu2)), 2) # round to make it clearer  
cbind(RR = exp(coef(GNmodelEdu2)), exp(confint(GNmodelEdu2))) # RR to CI  
round(cbind(RR = exp(coef(GNmodelEdu2)), exp(confint(GNmodelEdu2))), 2) # round to 2  
# Employment  
GNmodelEmp2 <- svyglm(GNREC\_summary2~v714, design= designGN, family = poisson(link = "log"), na.action = "na.pass")  
summary(GNmodelEmp2)  
exp(coef(GNmodelEmp2)) # calculate the RRs  
round(exp(coef(GNmodelEmp2)), 2) # round to make it clearer  
cbind(RR = exp(coef(GNmodelEmp2)), exp(confint(GNmodelEmp2))) # RR to CI  
round(cbind(RR = exp(coef(GNmodelEmp2)), exp(confint(GNmodelEmp2))), 2) # round to 2  
# Wealth Index  
GNmodelWI2 <- svyglm(GNREC\_summary2~v190, design= designGN, family = poisson(link = "log"), na.action = "na.pass")  
summary(GNmodelWI2)  
exp(coef(GNmodelWI2)) # calculate the RRs  
round(exp(coef(GNmodelWI2)), 2) # round to make it clearer  
cbind(RR = exp(coef(GNmodelWI2)), exp(confint(GNmodelWI2))) # RR to CI  
round(cbind(RR = exp(coef(GNmodelWI2)), exp(confint(GNmodelWI2))), 2) # round to 2  
  
# Adjusted RR  
# Keep NA  
GNmodel2 <- svyglm(GNREC\_summary2~v013 + v025 + v190 + v714 + v106 + v502, design= designGN, family = poisson(link = "log"), na.action = "na.pass") # keep all data, including NAs  
summary(GNmodel2)  
# calculate the RRs   
exp(coef(GNmodel2))  
# round to make it clearer  
round(exp(coef(GNmodel2)), 2)  
# RR to CI  
cbind(RR = exp(coef(GNmodel2)), exp(confint(GNmodel2)))  
# round to 2  
round(cbind(RR = exp(coef(GNmodel2)), exp(confint(GNmodel2))), 2)  
  
# Better Model  
# Comparing two models, AIC or BIC  
AIC(GNmodel, GNmodel2)  
BIC(GNmodel, GNmodel2, maximal = GNmodel2)  
plot\_summs(GNmodel, scale = TRUE, exp = TRUE)  
plot\_summs(GNmodel2, scale = TRUE, exp = TRUE)  
plot\_summs(GNmodel, GNmodel2, ci\_level = 0.95, model.names = c("Model 1"= "Model Excludes Missing", "Model 2"= "Model Includes Missing"),   
 colors = 'CUD Bright', exp = TRUE, legend.title = "Graph Comparing Models" )  
# Graph of Effect Size BJModel (Missing Data Excluded Model)  
GNeff.pres <- allEffects(GNmodel, scale = T)  
plot(GNeff.pres, ask=F)  
GNeff.pres2 <- allEffects(GNmodel2, scale = T)  
plot(GNeff.pres2, ask=F)

# Liberia

# Select needed variable from women dataset in Liberia  
LBDHS <- subset(LBDHSa, select = c(v001, v002, v021, v005, v025, v463z, v437, v438, v525, v212, v201,  
 v312, v337, v836, v763a, v012, v013, v502, v106, v714, v190, v130,   
 v717, v157, v158, v159, v171b, v481, v467b, v467c, v467d, v467f))  
  
# Sample weight/design for women  
LBDHS$LBwt <-LBDHS$v005/1000000  
designLB <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~LBwt, data = LBDHS)  
  
  
# Demographics  
LBDHS$LBwt <- LBDHS$v005/1000000  
freq(LBDHS$v013, LBDHS$LBwt, plot = FALSE) #Age  
freq(LBDHS$v025, LBDHS$LBwt, plot = FALSE) #Residential status  
freq(LBDHS$v502, LBDHS$LBwt, plot = FALSE) #Marital status  
freq(LBDHS$v106, LBDHS$LBwt, plot = FALSE) #Education  
freq(LBDHS$v714, LBDHS$LBwt, plot = FALSE) #Employment   
freq(LBDHS$v190, LBDHS$LBwt, plot = FALSE) #Wealth index  
designLB <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~LBwt, data = LBDHS) # Mean of Age and Confidence Interval  
LBttAge <- svyttest(v012~0, designLB)  
LBttAge  
  
#Proportion of Women With Risk Factors for Cervical Cancer  
#1. Use cigarettes and tobacco   
freq(LBDHS$v463z, LBDHS$LBwt, plot = FALSE)   
sum(is.na(LBDHS$v463z))  
#2. BMI   
favstats(LBDHS$v438, LBDHS$LBwt, plot = FALSE)  
sum(is.na(LBDHS$v438))  
is.na(LBDHS$v438)  
is.na(LBDHS$v438) <- which(LBDHS$v438 == 9996)  
is.na(LBDHS$v438) <- which(LBDHS$v438 == 9995)  
is.na(LBDHS$v438) <- which(LBDHS$v438 == 9994)  
favstats(LBDHS$v438, LBDHS$LBwt, plot = FALSE)  
LBDHS$LBHtM <- (LBDHS$v438/1000) #Height (Height is in cm, converted into meters. Also into 1 implied decimal)  
favstats(LBDHS$LBHtM, plot = FALSE)  
favstats(LBDHS$v437, LBDHS$LBwt, plot = FALSE)  
is.na(LBDHS$v437)  
sum(is.na(LBDHS$v437))  
is.na(LBDHS$v437) <- which(LBDHS$v437 == 9996)  
is.na(LBDHS$v437) <- which(LBDHS$v437 == 9995)  
is.na(LBDHS$v437) <- which(LBDHS$v437 == 9994)  
favstats(LBDHS$v437, LBDHS$LBwt, plot = FALSE)  
LBDHS$LBwtkg <- (LBDHS$v437/10) #Weight  
favstats(LBDHS$LBwtkg, plot = FALSE)  
LBDHS$BMI <- (LBDHS$LBwtkg) /(LBDHS$LBHtM \* LBDHS$LBHtM)   
favstats(LBDHS$BMI, LBDHS$LBwt, plot = FALSE)  
LBDHS$BMICAT <- NA  
LBDHS$BMICAT[LBDHS$BMI < 18.50000] <-1  
LBDHS$BMICAT[LBDHS$BMI >= 18.50000 & LBDHS$BMI <= 24.9999] <-2  
LBDHS$BMICAT[LBDHS$BMI >= 25.00000 & LBDHS$BMI <= 29.9999] <-3  
LBDHS$BMICAT[LBDHS$BMI >= 30.00000] <-4  
LBDHS$BMICAT <- factor(LBDHS$BMICAT, levels = c(1,2,3,4), labels = c("Underweight", "Normal", "Overweight", "Obese"))  
LBDHS$LBwt <- LBDHS$v005/1000000  
freq(LBDHS$BMICAT, LBDHS$LBwt, plot = FALSE)  
LBBiCAT <- NA # Binary Categories of BMI   
LBBiCAT[LBDHS$BMI < 25.0000] <- "low risk bmi"  
LBBiCAT[LBDHS$BMI >= 25.0000] <- "high risk bmi"  
freq(LBBiCAT, LBDHS$LBwt, plot = FALSE)  
#3. Age at First Sexual Intercourse   
freq(LBDHS$v525, LBDHS$LBwt, plot = FALSE)  
sum(is.na(LBDHS$v525))  
LBAFSI <- subset(LBDHS, v525 >=8 & v525 <=27)  
freq(LBAFSI$v525, plot = FALSE)  
LBAFSI$LBAFSIwt <- LBAFSI$v005/1000000  
designLBAFSI <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~LBAFSIwt, data = LBAFSI)  
svymean(~v525, design = designLBAFSI, na.rm = FALSE)  
confint(svymean(~v525, designLBAFSI))  
LBAFSICAT <- NA # Binary Categories of Age at First Sexual Intercourse Combine, had sex or not   
LBAFSICAT[LBDHS$v525 <=17.9999] <- "less or equal 17 years"  
LBAFSICAT[LBDHS$v525 >= 18.0000] <- "greater than 17 years"  
freq(LBAFSICAT, LBDHS$LBwt, plot = FALSE)  
#4. Age at First Birth   
freq(LBDHS$v212, LBDHS$LBwt, plot = FALSE)   
sum(is.na(LBDHS$v212))  
favstats(LBDHS$v212, LBDHS$LBwt, plot = FALSE)  
LBAFB <- subset(LBDHS, v212 >=10 & v212 <=43)  
freq(LBAFB$v212, plot = FALSE)  
LBAFB$LBAFBwt <- LBAFB$v005/1000000  
designLBAFB <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~LBAFBwt, data = LBAFB)  
svymean(~v212, design = designLBAFB, na.rm = FALSE)  
confint(svymean(~v212, designLBAFB))  
LBAFBCAT <- NA # Binary Categories of Age at First Birth  
LBAFBCAT[LBDHS$v212 <=19.9999] <- "less or equal 19 years"  
LBAFBCAT[LBDHS$v212 >= 20.0000] <- "greater than 19 years"  
freq(LBAFBCAT, LBDHS$LBwt, plot = FALSE)  
#5. Total Children Ever Born (Parity)   
freq(LBDHS$v201,LBDHS$LBwt, plot = FALSE)   
sum(is.na(LBDHS$v201))  
LBP <- subset(LBDHS, v201 >=1 & v201 <=16)  
freq(LBP$v201, plot = FALSE)  
LBP$LBPwt <- LBP$v005/1000000  
designLBP <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~LBPwt, data = LBP)  
svymean(~v201, design = designLBP, na.rm = FALSE)  
confint(svymean(~v201, designLBP))  
LBPCAT <- NA # Binary Categories of Total Children Ever Born (Parity)  
LBPCAT[LBDHS$v201<=2.9999] <- "less or equal 2 births"  
LBPCAT[LBDHS$v201 >= 3.0000] <- "greater than 2 births"  
freq(LBPCAT, LBDHS$LBwt, plot = FALSE)  
#6.Total Lifetime Number of Sex Partners   
freq(LBDHS$v836, LBDHS$LBwt, plot = FALSE)   
sum(is.na(LBDHS$v836))  
favstats(LBDHS$v836, LBDHS$LBwt, plot = FALSE)  
is.na(LBDHS$v836) <- which(LBDHS$v836 == 98)  
favstats(LBDHS$v836, LBDHS$LBwt, plot = FALSE)  
LBLSP <- subset(LBDHS, v836 >=1 & v836 <=95)  
freq(LBLSP$v836, plot = FALSE)  
LBLSP$LBLSPwt <- LBLSP$v005/1000000  
designLBLSP <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~LBLSPwt, data = LBLSP)  
svymean(~v836, design = designLBLSP, na.rm = FALSE)  
confint(svymean(~v836, designLBLSP))  
LBLSPCAT <- NA # Binary Categories of Total Lifetime Number of Sex Partners  
LBLSPCAT[LBDHS$v836 <=5.9999] <- "less or equal 5 partners"  
LBLSPCAT[LBDHS$v836 >= 6.0000] <- "greater than 5 partners"  
freq(LBLSPCAT, LBDHS$LBwt, plot = FALSE)  
#7. Contraceptive Usage  
freq(LBDHS$v312, LBDHS$LBwt, plot = FALSE)  
is.na(LBDHS$v312)  
sum(is.na(LBDHS$v312))  
nlevels(LBDHS$v312)  
table(LBDHS$v312)  
LBDHSContwt <- freq(LBDHS$v312, LBDHS$LBwt, plot = FALSE)  
LBDHSContwt  
LBCont\_recoded <- c("not using" = 0, "pill" =1, "iud"=1, "injections"=1, "diaphragm"= 2, "male condom"=2, "female sterilization"=2,  
 "male sterilization"=2, "periodic abstinence"=2, "withdrawal"=2, "other traditional"=2, "implants/norplant"=1,  
 "prolonged abstinence"= 2, "lactational amenorrhea (lam)"=2, "female condom"=2, "foam or jelly"= 2, "emergency contraception"= 1,  
 "other modern method"= 2, "standard days method (sdm)"=2, "specific method 1"= 2, "specific method 2"= 2)  
LBDHS$LBREC\_CONT <- LBCont\_recoded[LBDHS$v312]  
freq(LBDHS$LBREC\_CONT, LBDHS$LBwt, plot = FALSE)  
LBCont\_recoded <- c("not using" = 0, "pill" =1, "iud"=1, "injections"=1, "diaphragm"= 0, "male condom"=0, "female sterilization"=0,  
 "male sterilization"=0, "periodic abstinence"=0, "withdrawal"=0, "other traditional"=0, "implants/norplant"=1,  
 "prolonged abstinence"= 0, "lactational amenorrhea (lam)"=0, "female condom"=0, "foam or jelly"= 0, "emergency contraception"= 1,  
 "other modern method"= 0, "standard days method (sdm)"=0, "specific method 1"= 0, "specific method 2"= NA)  
LBDHS$LBREC\_CONT <- LBCont\_recoded[LBDHS$v312]  
freq(LBDHS$LBREC\_CONT, LBDHS$LBwt, plot = FALSE)  
# Months of use of current family planning method, For women who are currently using a contraceptive method  
freq(LBDHS$v337, LBDHS$LBwt, plot = FALSE)   
is.na(LBDHS$v337)  
sum(is.na(LBDHS$v337))  
favstats(LBDHS$v337, LBDHS$LBwt, plot = FALSE)  
summary(LBDHS$v337, LBDHS$LBwt, plot = FALSE)  
LBMFP <- NA  
LBMFP[LBDHS$v337 <60.0000]<- "Less than 5 years"  
LBMFP[LBDHS$v337 >=60.0000]<- "More than or equal 5 years"  
freq(LBMFP, LBDHS$LBwt, plot = FALSE)  
# Combine number hormonal and non hormonal to number of years  
LBDHS$LBcom <- paste(LBDHS$v312, LBDHS$v337)  
LBDHS$LBcom  
LBDHS$LBcom <- paste(LBDHS$v312, "-", LBDHS$v337)  
LBDHS$LBcom  
LBDHS$LBcom <- paste(LBDHS$v312, LBMFP, sep = "-" )  
LBDHS$LBcom  
freq(LBDHS$LBcom, plot = FALSE)  
LBcom\_recoded <- c("emergency contraception-Less than 5 years" = 0, "female sterilization-Less than 5 years" = 0, "female sterilization-More than or equal 5 years" = 0,   
 "implants/norplant-Less than 5 years" = 0, "implants/norplant-More than or equal 5 years" = 1, "injections-Less than 5 years" =0, "injections-More than or equal 5 years" =1,  
 "iud-Less than 5 years" = 0, "iud-More than or equal 5 years"= 1, "lactational amenorrhea (lam)-Less than 5 years"= 0, "male condom-Less than 5 years" =0, "male condom-More than or equal 5 years" = 0, "not using-NA" =0,   
 "other modern method-Less than 5 years"=0, "periodic abstinence-Less than 5 years"=0, "periodic abstinence-More than or equal 5 years"=0, "pill-Less than 5 years" =0, "pill-More than or equal 5 years"=1, "standard days method (sdm)-Less than 5 years"=0,   
 "standard days method (sdm)-More than or equal 5 years"=0,"withdrawal-Less than 5 years"=0, "withdrawal-More than or equal 5 years"=0)  
  
LBDHS$LByears\_FP <- LBcom\_recoded[LBDHS$LBcom]  
freq(LBDHS$LByears\_FP, LBDHS$LBwt, plot = FALSE)  
#8. Sexually Transmitted Infection in last 12 months ("don't know" added to "no")  
freq(LBDHS$v763a, LBDHS$LBwt, plot = FALSE)  
sum(is.na(LBDHS$v763a))  
LBSTI\_recoded <- c("no"= 0, "yes" = 1, "don't know"= 0)  
LBDHS$LBREC\_STI <- LBSTI\_recoded[LBDHS$v763a]  
freq(LBDHS$LBREC\_STI, LBDHS$LBwt, plot = FALSE)  
  
  
#Number of Risk Factors  
# Smoking  
freq(LBDHS$v463z, LBDHS$LBwt, plot = FALSE)  
LBsmoke\_recoded <- c("no" = 1, "yes, smokes nothing" = 0)  
LBDHS$LBREC\_SMOKE <- LBsmoke\_recoded[LBDHS$v463z]  
freq(LBDHS$LBREC\_SMOKE, LBDHS$LBwt, plot = FALSE)  
#BMI  
freq(LBBiCAT, LBDHS$LBwt, plot = FALSE)  
LBbmi\_recoded <- c("high risk bmi" = 1, "low risk bmi" = 0)  
LBDHS$LBREC\_BMI <- LBbmi\_recoded[LBBiCAT]  
freq(LBDHS$LBREC\_BMI, LBDHS$LBwt, plot = FALSE)  
#AFSI  
freq(LBAFSICAT, LBDHS$LBwt, plot = FALSE)  
LBAFSI\_recoded <- c("less or equal 17 years" = 1, "greater than 17 years" = 0)  
LBDHS$LBREC\_AFSI <- LBAFSI\_recoded[LBAFSICAT]  
freq(LBDHS$LBREC\_AFSI, LBDHS$LBwt, plot = FALSE)  
#AFB  
freq(LBAFBCAT, LBDHS$LBwt, plot = FALSE)  
LBAFB\_recoded <- c("less or equal 19 years" = 1, "greater than 19 years" = 0)  
LBDHS$LBREC\_AFB <- LBAFB\_recoded[LBAFBCAT]  
freq(LBDHS$LBREC\_AFB, LBDHS$LBwt, plot = FALSE)  
#Parity  
freq(LBPCAT, LBDHS$LBwt, plot = FALSE)  
LBParity\_recoded <- c("greater than 2 births" = 1, "less or equal 2 births" = 0)  
LBDHS$LBREC\_Parity <- LBParity\_recoded[LBPCAT]  
freq(LBDHS$LBREC\_Parity, LBDHS$LBwt, plot = FALSE)  
#FP  
freq(LBDHS$LBcom, plot = FALSE)  
LBcom\_recoded <- c("emergency contraception-Less than 5 years" = 0, "female sterilization-Less than 5 years" = 0, "female sterilization-More than or equal 5 years" = 0,   
 "implants/norplant-Less than 5 years" = 0, "implants/norplant-More than or equal 5 years" = 1, "injections-Less than 5 years" =0, "injections-More than or equal 5 years" =1,  
 "iud-Less than 5 years" = 0, "iud-More than or equal 5 years"= 1, "lactational amenorrhea (lam)-Less than 5 years"= 0, "male condom-Less than 5 years" =0, "male condom-More than or equal 5 years" = 0, "not using-NA" =0,   
 "other modern method-Less than 5 years"=0, "periodic abstinence-Less than 5 years"=0, "periodic abstinence-More than or equal 5 years"=0, "pill-Less than 5 years" =0, "pill-More than or equal 5 years"=1, "standard days method (sdm)-Less than 5 years"=0,   
 "standard days method (sdm)-More than or equal 5 years"=0,"withdrawal-Less than 5 years"=0, "withdrawal-More than or equal 5 years"=0)  
LBDHS$LByears\_FP <- LBcom\_recoded[LBDHS$LBcom]  
freq(LBDHS$LByears\_FP, LBDHS$LBwt, plot = FALSE)  
# LSP  
freq(LBLSPCAT, LBDHS$LBwt, plot = FALSE)  
LBLSP\_recoded <- c("greater than 5 partners" = 1, "less or equal 5 partners" = 0)  
LBDHS$LBREC\_LSP <- LBLSP\_recoded[LBLSPCAT]  
freq(LBDHS$LBREC\_LSP, LBDHS$LBwt, plot = FALSE)  
# STI  
freq(LBDHS$v763a, LBDHS$LBwt, plot = FALSE)  
LBSTI\_recoded <- c("no"= 0, "yes" = 1, "don't know"= 0)  
LBDHS$LBREC\_STI <- LBSTI\_recoded[LBDHS$v763a]  
freq(LBDHS$LBREC\_STI, LBDHS$LBwt, plot = FALSE)  
  
  
  
#Number of Risk Factors,   
LBRiskcom <- cbind.data.frame(LBDHS$LBREC\_SMOKE, LBDHS$LBREC\_BMI, LBDHS$LBREC\_AFSI, LBDHS$LBREC\_AFB, LBDHS$LBREC\_Parity, LBDHS$LByears\_FP, LBDHS$LBREC\_LSP, LBDHS$LBREC\_STI)  
LBRiskcom  
LBtotal <- cbind(LBRiskcom, total = rowSums(LBRiskcom, na.rm = TRUE))  
LBtotal <- cbind(LBRiskcom, total = rowSums(LBRiskcom, na.rm = FALSE))  
tail(LBtotal)  
head(LBtotal)  
  
# Regrouped number of Risk Factors  
LBTotalnew <- NA   
LBTotalnew[LBtotal$total <=4.9999] <- "less or equal 4"  
LBTotalnew[LBtotal$total >= 5.0000] <- "greater or equal 5"   
freq(LBTotalnew, LBDHS$LBwt, plot = FALSE)  
  
# Modelling  
# Missing Data Removed  
LBDHS$LBREC\_summary= as.numeric(LBDHS$LBREC\_SMOKE + LBDHS$LBREC\_BMI + LBDHS$LBREC\_AFSI + LBDHS$LBREC\_AFB + LBDHS$LBREC\_Parity + LBDHS$LByears\_FP + LBDHS$LBREC\_LSP + LBDHS$LBREC\_STI)  
crosstab(LBDHS$LBREC\_summary, LBDHS$v013, LBDHS$LBwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(LBDHS$v013, LBDHS$LBREC\_summary, LBDHS$LBwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(LBDHS$LBREC\_summary, LBDHS$v106, LBDHS$LBwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(LBDHS$v106, LBDHS$LBREC\_summary, LBDHS$LBwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
summary(LBDHS$LBREC\_summary)  
favstats(LBDHS$LBREC\_summary)  
designLB <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~LBwt, data = LBDHS)  
is.numeric(LBDHS$LBREC\_summary)  
designLB <-update(designLB, LBREC\_summary=LBDHS$LBREC\_SMOKE + LBDHS$LBREC\_BMI + LBDHS$LBREC\_AFSI + LBDHS$LBREC\_AFB + LBDHS$LBREC\_Parity + LBDHS$LByears\_FP + LBDHS$LBREC\_LSP + LBDHS$LBREC\_STI)  
  
  
# Unadjusted RR  
# Age  
LBmodelAge <- svyglm(LBREC\_summary~v013, design= designLB, family = poisson(link = "log"), na.action = na.omit)  
summary(LBmodelAge)  
exp(coef(LBmodelAge)) # calculate the RRs  
round(exp(coef(LBmodelAge)), 2) # round to make it clearer  
cbind(RR = exp(coef(LBmodelAge)), exp(confint(LBmodelAge))) # RR to CI  
round(cbind(RR = exp(coef(LBmodelAge)), exp(confint(LBmodelAge))), 2) # round to 2  
# Residential Status  
LBmodelRS <- svyglm(LBREC\_summary~v025, design= designLB, family = poisson(link = "log"), na.action = na.omit)  
summary(LBmodelRS)  
exp(coef(LBmodelRS)) # calculate the RRs  
round(exp(coef(LBmodelRS)), 2) # round to make it clearer  
cbind(RR = exp(coef(LBmodelRS)), exp(confint(LBmodelRS))) # RR to CI  
round(cbind(RR = exp(coef(LBmodelRS)), exp(confint(LBmodelRS))), 2) # round to 2  
# Marital Status  
LBmodelMS <- svyglm(LBREC\_summary~v502, design= designLB, family = poisson(link = "log"), na.action = na.omit)  
summary(LBmodelMS)  
exp(coef(LBmodelMS)) # calculate the RRs  
round(exp(coef(LBmodelMS)), 2) # round to make it clearer  
cbind(RR = exp(coef(LBmodelMS)), exp(confint(LBmodelMS))) # RR to CI  
round(cbind(RR = exp(coef(LBmodelMS)), exp(confint(LBmodelMS))), 2) # round to 2  
# Education  
LBmodelEdu <- svyglm(LBREC\_summary~v106, design= designLB, family = poisson(link = "log"), na.action = na.omit)  
summary(LBmodelEdu)  
exp(coef(LBmodelEdu)) # calculate the RRs  
round(exp(coef(LBmodelEdu)), 2) # round to make it clearer  
cbind(RR = exp(coef(LBmodelEdu)), exp(confint(LBmodelEdu))) # RR to CI  
round(cbind(RR = exp(coef(LBmodelEdu)), exp(confint(LBmodelEdu))), 2) # round to 2  
# Employment  
LBmodelEmp <- svyglm(LBREC\_summary~v714, design= designLB, family = poisson(link = "log"), na.action = na.omit)  
summary(LBmodelEmp)  
exp(coef(LBmodelEmp)) # calculate the RRs  
round(exp(coef(LBmodelEmp)), 2) # round to make it clearer  
cbind(RR = exp(coef(LBmodelEmp)), exp(confint(LBmodelEmp))) # RR to CI  
round(cbind(RR = exp(coef(LBmodelEmp)), exp(confint(LBmodelEmp))), 2) # round to 2  
# Wealth Index  
LBmodelWI <- svyglm(LBREC\_summary~v190, design= designLB, family = poisson(link = "log"), na.action = na.omit)  
summary(LBmodelWI)  
exp(coef(LBmodelWI)) # calculate the RRs  
round(exp(coef(LBmodelWI)), 2) # round to make it clearer  
cbind(RR = exp(coef(LBmodelWI)), exp(confint(LBmodelWI))) # RR to CI  
round(cbind(RR = exp(coef(LBmodelWI)), exp(confint(LBmodelWI))), 2) # round to 2  
# Adjusted RR  
# Poisson Regression Modelling  
LBmodel <- svyglm(LBREC\_summary~v013 + v025 + v190 + v714 + v106 + v502, design= designLB, family = poisson(link = "log"), na.action = "na.omit") # Observations removed if it contains missing values  
summary(LBmodel)   
# calculate the RRs   
exp(coef(LBmodel))  
# round to make it clearer  
round(exp(coef(LBmodel)), 2)  
# RR to CI  
cbind(RR = exp(coef(LBmodel)), exp(confint(LBmodel)))  
# round to 2  
round(cbind(RR = exp(coef(LBmodel)), exp(confint(LBmodel))), 2)   
  
# Missing Data Included  
# Convert cases with NA into 0 and run a different model  
# library("imputeTS")  
#BMI  
freq(LBDHS$LBREC\_BMI, LBDHS$LBwt, na.rm= TRUE, plot = FALSE)  
LBDHS$LBREC\_BMI2 <- na.replace(LBDHS$LBREC\_BMI, 0) #gtools package  
freq(LBDHS$LBREC\_BMI2, LBDHS$LBwt, na.rm= TRUE, plot = FALSE)  
#AFB  
freq(LBDHS$LBREC\_AFB, LBDHS$LBwt, na.rm= TRUE, plot = FALSE)  
LBDHS$LBREC\_AFB2 <- na.replace(LBDHS$LBREC\_AFB, 0)  
freq(LBDHS$LBREC\_AFB2, LBDHS$LBwt, na.rm= TRUE, plot = FALSE)  
# LSP women  
freq(LBDHS$LBREC\_LSP, LBDHS$LBwt, plot = FALSE)  
LBDHS$LBREC\_LSP2 <- na.replace(LBDHS$LBREC\_LSP, 0)  
freq(LBDHS$LBREC\_LSP2, LBDHS$LBwt, na.rm= TRUE, plot = FALSE)  
# NB- Type of Contraceptive use was linked with month of usage. month of use contains NA i.e. not in the universe.   
# Variables with missing values- BMI (149), LSP woman (56) and LSP men (193)  
LBDHS$LBREC\_summary2= as.numeric(LBDHS$LBREC\_SMOKE + LBDHS$LBREC\_BMI2 + LBDHS$LBREC\_AFSI + LBDHS$LBREC\_AFB2 + LBDHS$LBREC\_Parity + LBDHS$LByears\_FP + LBDHS$LBREC\_LSP2 + LBDHS$LBREC\_STI)  
summary(LBDHS$LBREC\_summary2)  
favstats(LBDHS$LBREC\_summary2)  
designLB <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~LBwt, data = LBDHS)  
is.numeric(LBDHS$LBREC\_summary2)  
designLB <-update(designLB, LBREC\_summary2=LBDHS$LBREC\_SMOKE + LBDHS$LBREC\_BMI2 + LBDHS$LBREC\_AFSI + LBDHS$LBREC\_AFB2 + LBDHS$LBREC\_Parity + LBDHS$LByears\_FP + LBDHS$LBREC\_LSP2 + LBDHS$LBREC\_STI)  
  
  
# Unadjusted RR  
# Age  
LBmodelAge2 <- svyglm(LBREC\_summary2~v013, design= designLB, family = poisson(link = "log"), na.action = "na.pass")  
summary(LBmodelAge2)  
exp(coef(LBmodelAge2)) # calculate the RRs  
round(exp(coef(LBmodelAge2)), 2) # round to make it clearer  
cbind(RR = exp(coef(LBmodelAge2)), exp(confint(LBmodelAge2))) # RR to CI  
round(cbind(RR = exp(coef(LBmodelAge2)), exp(confint(LBmodelAge2))), 2) # round to 2  
# Residential Status  
LBmodelRS2 <- svyglm(LBREC\_summary2~v025, design= designLB, family = poisson(link = "log"), na.action = "na.pass")  
summary(LBmodelRS2)  
exp(coef(LBmodelRS2)) # calculate the RRs  
round(exp(coef(LBmodelRS2)), 2) # round to make it clearer  
cbind(RR = exp(coef(LBmodelRS2)), exp(confint(LBmodelRS2))) # RR to CI  
round(cbind(RR = exp(coef(LBmodelRS2)), exp(confint(LBmodelRS2))), 2) # round to 2  
# Marital Status  
LBmodelMS2 <- svyglm(LBREC\_summary2~v502, design= designLB, family = poisson(link = "log"), na.action = "na.pass")  
summary(LBmodelMS2)  
exp(coef(LBmodelMS2)) # calculate the RRs  
round(exp(coef(LBmodelMS2)), 2) # round to make it clearer  
cbind(RR = exp(coef(LBmodelMS2)), exp(confint(LBmodelMS2))) # RR to CI  
round(cbind(RR = exp(coef(LBmodelMS2)), exp(confint(LBmodelMS2))), 2) # round to 2  
# Education  
LBmodelEdu2 <- svyglm(LBREC\_summary2~v106, design= designLB, family = poisson(link = "log"), na.action = "na.pass")  
summary(LBmodelEdu2)  
exp(coef(LBmodelEdu2)) # calculate the RRs  
round(exp(coef(LBmodelEdu2)), 2) # round to make it clearer  
cbind(RR = exp(coef(LBmodelEdu2)), exp(confint(LBmodelEdu2))) # RR to CI  
round(cbind(RR = exp(coef(LBmodelEdu2)), exp(confint(LBmodelEdu2))), 2) # round to 2  
# Employment  
LBmodelEmp2 <- svyglm(LBREC\_summary2~v714, design= designLB, family = poisson(link = "log"), na.action = "na.pass")  
summary(LBmodelEmp2)  
exp(coef(LBmodelEmp2)) # calculate the RRs  
round(exp(coef(LBmodelEmp2)), 2) # round to make it clearer  
cbind(RR = exp(coef(LBmodelEmp2)), exp(confint(LBmodelEmp2))) # RR to CI  
round(cbind(RR = exp(coef(LBmodelEmp2)), exp(confint(LBmodelEmp2))), 2) # round to 2  
# Wealth Index  
LBmodelWI2 <- svyglm(LBREC\_summary2~v190, design= designLB, family = poisson(link = "log"), na.action = "na.pass")  
summary(LBmodelWI2)  
exp(coef(LBmodelWI2)) # calculate the RRs  
round(exp(coef(LBmodelWI2)), 2) # round to make it clearer  
cbind(RR = exp(coef(LBmodelWI2)), exp(confint(LBmodelWI2))) # RR to CI  
round(cbind(RR = exp(coef(LBmodelWI2)), exp(confint(LBmodelWI2))), 2) # round to 2  
  
# Adjusted RR  
# Keep NA  
LBmodel2 <- svyglm(LBREC\_summary2~v013 + v025 + v190 + v714 + v106 + v502, design= designLB, family = poisson(link = "log"), na.action = "na.pass") # keep all data, including NAs  
summary(LBmodel2)  
# calculate the RRs   
exp(coef(LBmodel2))  
# round to make it clearer  
round(exp(coef(LBmodel2)), 2)  
# RR to CI  
cbind(RR = exp(coef(LBmodel2)), exp(confint(LBmodel2)))  
# round to 2  
round(cbind(RR = exp(coef(LBmodel2)), exp(confint(LBmodel2))), 2)  
  
# Better Model  
# Comparing two models, AIC or BIC  
AIC(LBmodel, LBmodel2)  
BIC(LBmodel, LBmodel2, maximal = LBmodel2)  
plot\_summs(LBmodel, scale = TRUE, exp = TRUE)  
plot\_summs(LBmodel2, scale = TRUE, exp = TRUE)  
plot\_summs(LBmodel, LBmodel2, ci\_level = 0.95, model.names = c("Model 1"= "Model Excludes Missing", "Model 2"= "Model Includes Missing"),   
 colors = 'CUD Bright', exp = TRUE, legend.title = "Graph Comparing Models" )  
# Graph of Effect Size BJModel (Missing Data Excluded Model)  
LBeff.pres <- allEffects(LBmodel, scale = T)  
plot(LBeff.pres, ask=F)  
LBeff.pres2 <- allEffects(LBmodel2, scale = T)  
plot(LBeff.pres2, ask=F)

# Mali

# Select needed variable from women dataset in Mali  
MLDHS <- subset(MLDHSa, select = c(v001, v002, v021, v005, v025, v463z, v437, v438, v525, v212, v201,  
 v312, v337, v836, v763a, v012, v013, v502, v106, v714, v190, v130,   
 v717, v157, v158, v159, v171b, v481, v467b, v467c, v467d, v467f))  
  
# Sample weight/design for women  
MLDHS$MLwt <- MLDHS$v005/1000000  
designML <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~MLwt, data = MLDHS)  
  
  
# Demographics  
MLDHS$MLwt <- MLDHS$v005/1000000  
freq(MLDHS$v013, MLDHS$MLwt, plot = FALSE) #Age  
freq(MLDHS$v025, MLDHS$MLwt, plot = FALSE) #Residential status  
freq(MLDHS$v502, MLDHS$MLwt, plot = FALSE) #Marital status  
freq(MLDHS$v106, MLDHS$MLwt, plot = FALSE) #Education  
freq(MLDHS$v714, MLDHS$MLwt, plot = FALSE) #Employment   
freq(MLDHS$v190, MLDHS$MLwt, plot = FALSE) #Wealth index  
designML <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~MLwt, data = MLDHS) # Mean of Age and Confidence Interval  
MLttAge <- svyttest(v012~0, designML)  
MLttAge  
  
  
# Proportion of Women With Risk Factors for Cervical Cancer  
#1. Use cigarettes and tobacco   
freq(MLDHS$v463z, MLDHS$MLwt, plot = FALSE)   
sum(is.na(MLDHS$v463z))  
#2. BMI   
favstats(MLDHS$v438, MLDHS$MLwt, plot = FALSE)  
sum(is.na(MLDHS$v438))  
is.na(MLDHS$v438)  
is.na(MLDHS$v438) <- which(MLDHS$v438 == 9996)  
is.na(MLDHS$v438) <- which(MLDHS$v438 == 9995)  
is.na(MLDHS$v438) <- which(MLDHS$v438 == 9994)  
favstats(MLDHS$v438, MLDHS$MLwt, plot = FALSE)  
MLDHS$MLHtM <- (MLDHS$v438/1000) #Height (Height is in cm, converted into meters. Also into 1 implied decimal)  
favstats(MLDHS$MLHtM, plot = FALSE)  
favstats(MLDHS$v437, MLDHS$MLwt, plot = FALSE)  
is.na(MLDHS$v437)  
sum(is.na(MLDHS$v437))  
is.na(MLDHS$v437) <- which(MLDHS$v437 == 9996)  
is.na(MLDHS$v437) <- which(MLDHS$v437 == 9995)  
is.na(MLDHS$v437) <- which(MLDHS$v437 == 9994)  
favstats(MLDHS$v437, MLDHS$MLwt, plot = FALSE)  
MLDHS$MLwtkg <- (MLDHS$v437/10) #Weight  
favstats(MLDHS$MLwtkg, plot = FALSE)  
MLDHS$BMI <- (MLDHS$MLwtkg) /(MLDHS$MLHtM \* MLDHS$MLHtM) # Calculated BMI  
favstats(MLDHS$BMI, MLDHS$MLwt, plot = FALSE)  
MLDHS$BMICAT <- NA  
MLDHS$BMICAT[MLDHS$BMI < 18.50000] <-1  
MLDHS$BMICAT[MLDHS$BMI >= 18.50000 & MLDHS$BMI <= 24.9999] <-2  
MLDHS$BMICAT[MLDHS$BMI >= 25.00000 & MLDHS$BMI <= 29.9999] <-3  
MLDHS$BMICAT[MLDHS$BMI >= 30.00000] <-4  
MLDHS$BMICAT <- factor(MLDHS$BMICAT, levels = c(1,2,3,4), labels = c("Underweight", "Normal", "Overweight", "Obese"))  
MLDHS$MLwt <- MLDHS$v005/1000000  
freq(MLDHS$BMICAT, MLDHS$MLwt, plot = FALSE)  
MLBiCAT <- NA # Binary Categories of BMI   
MLBiCAT[MLDHS$BMI < 25.0000] <- "low risk bmi"  
MLBiCAT[MLDHS$BMI >= 25.0000] <- "high risk bmi"  
freq(MLBiCAT, MLDHS$MLwt, plot = FALSE)  
#3. Age at First Sexual Intercourse  
freq(MLDHS$v525, MLDHS$MLwt, plot = FALSE)   
sum(is.na(MLDHS$v525))  
MLAFSI <- subset(MLDHS, v525 >=8 & v525 <=36)  
freq(MLAFSI$v525, plot = FALSE)  
MLAFSI$MLAFSIwt <- MLAFSI$v005/1000000  
designMLAFSI <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~MLAFSIwt, data = MLAFSI)  
svymean(~v525, design = designMLAFSI, na.rm = FALSE)  
confint(svymean(~v525, designMLAFSI))  
MLAFSICAT <- NA # Binary Categories of Age at First Sexual Intercourse Combine, had sex or not   
MLAFSICAT[MLDHS$v525 <=17.9999] <- "less or equal 17 years"  
MLAFSICAT[MLDHS$v525 >= 18.0000] <- "greater than 17 years"  
freq(MLAFSICAT, MLDHS$MLwt, plot = FALSE)  
#4. Age at First Birth   
freq(MLDHS$v212, MLDHS$MLwt, plot = FALSE)   
sum(is.na(MLDHS$v212))  
favstats(MLDHS$v212, MLDHS$MLwt, plot = FALSE)  
MLAFB <- subset(MLDHS, v212 >=12 & v525 <=44)  
freq(MLAFB$v212, plot = FALSE)  
MLAFB$MLAFBwt <- MLAFB$v005/1000000  
designMLAFB <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~MLAFBwt, data = MLAFB)  
svymean(~v212, design = designMLAFB, na.rm = FALSE)  
confint(svymean(~v212, designMLAFB))  
MLAFBCAT <- NA # Binary Categories of Age at First Birth  
MLAFBCAT[MLDHS$v212 <=19.9999] <- "less or equal 19 years"  
MLAFBCAT[MLDHS$v212 >= 20.0000] <- "greater than 19 years"  
freq(MLAFBCAT, MLDHS$MLwt, plot = FALSE)  
#5. Total Children Ever Born (Parity)   
freq(MLDHS$v201, MLDHS$MLwt, plot = FALSE)   
sum(is.na(MLDHS$v201))  
MLP <- subset(MLDHS, v201 >=1 & v201 <=16)  
freq(MLP$v201, plot = FALSE)  
MLP$MLPwt <- MLP$v005/1000000  
designMLP <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~MLPwt, data = MLP)  
svymean(~v201, design = designMLP, na.rm = FALSE)  
confint(svymean(~v201, designMLP))  
MLPCAT <- NA # Binary Categories of Total Children Ever Born (Parity)  
MLPCAT[MLDHS$v201<=2.9999] <- "less or equal 2 births"  
MLPCAT[MLDHS$v201 >= 3.0000] <- "greater than 2 births"  
freq(MLPCAT, MLDHS$MLwt, plot = FALSE)  
#6. Total Lifetime Number of Sex Partners #Missing Data Excluded(30)  
freq(MLDHS$v836, MLDHS$MLwt, plot = FALSE)   
sum(is.na(MLDHS$v836))  
favstats(MLDHS$v836, MLDHS$MLwt, plot = FALSE)  
is.na(MLDHS$v836) <- which(MLDHS$v836 == 98)  
favstats(MLDHS$v836, MLDHS$MLwt, plot = FALSE)  
MLLSP <- subset(MLDHS, v836 >=1 & v836 <=95)  
freq(MLLSP$v836, plot = FALSE)  
MLLSP$MLLSPwt <- MLLSP$v005/1000000  
designMLLSP <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~MLLSPwt, data = MLLSP)  
svymean(~v836, design = designMLLSP, na.rm = FALSE)  
confint(svymean(~v836, designMLLSP))  
MLLSPCAT <- NA # Binary Categories of Total Lifetime Number of Sex Partners  
MLLSPCAT[MLDHS$v836 <=5.9999] <- "less or equal 5 partners"  
MLLSPCAT[MLDHS$v836 >= 6.0000] <- "greater than 5 partners"  
freq(MLLSPCAT, MLDHS$MLwt, plot = FALSE)  
#7. Contraceptive Usage  
freq(MLDHS$v312, MLDHS$MLwt, plot = FALSE)  
is.na(MLDHS$v312)  
sum(is.na(MLDHS$v312))  
nlevels(MLDHS$v312)  
table(MLDHS$v312)  
MLDHSContwt <- freq(MLDHS$v312, MLDHS$MLwt, plot = FALSE)  
MLDHSContwt  
MLCont\_recoded <- c("not using" = 0, "pill" =1, "iud"=1, "injections"=1, "diaphragm"= 2, "male condom"=2, "female sterilization"=2,  
 "male sterilization"=2, "periodic abstinence"=2, "withdrawal"=2, "other traditional"=2, "implants/norplant"=1,  
 "prolonged abstinence"= 2, "lactational amenorrhea (lam)"=2, "female condom"=2, "foam or jelly"= 2, "emergency contraception"= 1,  
 "other modern method"= 2, "standard days method (sdm)"=2, "specific method 1"= 2, "specific method 2"= 2)  
MLDHS$MLREC\_CONT <- MLCont\_recoded[MLDHS$v312]  
freq(MLDHS$MLREC\_CONT, MLDHS$MLwt, plot = FALSE)  
MLCont\_recoded <- c("not using" = 0, "pill" =1, "iud"=1, "injections"=1, "diaphragm"= 0, "male condom"=0, "female sterilization"=0,  
 "male sterilization"=0, "periodic abstinence"=0, "withdrawal"=0, "other traditional"=0, "implants/norplant"=1,  
 "prolonged abstinence"= 0, "lactational amenorrhea (lam)"=0, "female condom"=0, "foam or jelly"= 0, "emergency contraception"= 1,  
 "other modern method"= 0, "standard days method (sdm)"=0, "specific method 1"= 0, "specific method 2"= NA)  
MLDHS$MLREC\_CONT <- MLCont\_recoded[MLDHS$v312]  
freq(MLDHS$MLREC\_CONT, MLDHS$MLwt, plot = FALSE)  
# Months of use of current family planning method, For women who are currently using a contraceptive method  
freq(MLDHS$v337, MLDHS$MLwt, plot = FALSE)  
is.na(MLDHS$v337)  
sum(is.na(MLDHS$v337))  
favstats(MLDHS$v337, MLDHS$MLwt, plot = FALSE)  
summary(MLDHS$v337, MLDHS$MLwt, plot = FALSE)  
MLMFP <- NA  
MLMFP[MLDHS$v337 <60.0000]<- "Less than 5 years"  
MLMFP[MLDHS$v337 >=60.0000]<- "More than or equal 5 years"  
freq(MLMFP, MLDHS$MLwt, plot = FALSE)  
# Combine number hormonal and non hormonal to number of years  
MLDHS$MLcom <- paste(MLDHS$v312, MLDHS$v337)  
MLDHS$MLcom  
MLDHS$MLcom <- paste(MLDHS$v312, "-", MLDHS$v337)  
MLDHS$MLcom  
MLDHS$MLcom <- paste(MLDHS$v312, MLMFP, sep = "-" )  
MLDHS$MLcom  
freq(MLDHS$MLcom, plot = FALSE)  
MLcom\_recoded <- c("female condom-Less than 5 years"=0, "female sterilization-Less than 5 years"=0, "female sterilization-More than or equal 5 years"=0,  
 "implants/norplant-Less than 5 years"= 0, "implants/norplant-More than or equal 5 years"=1, "injections-Less than 5 years"=0,  
 "injections-More than or equal 5 years"=1, "iud-Less than 5 years"=0, "iud-More than or equal 5 years"=1,  
 "lactational amenorrhea (lam)-Less than 5 years"=0, "male condom-Less than 5 years"=0, "male condom-More than or equal 5 years"=0,  
 "not using-NA"=0, "other traditional-Less than 5 years"=0, "periodic abstinence-Less than 5 years"=0,   
 "periodic abstinence-More than or equal 5 years"=0, "pill-Less than 5 years"=0, "pill-More than or equal 5 years"=1,  
 "standard days method (sdm)-Less than 5 years"=0, "standard days method (sdm)-More than or equal 5 years"=0,  
 "withdrawal-Less than 5 years"=0, "withdrawal-More than or equal 5 years"=0)  
MLDHS$MLyears\_FP <- MLcom\_recoded[MLDHS$MLcom]  
freq(MLDHS$MLyears\_FP, MLDHS$MLwt, plot = FALSE)  
#8. Sexually Transmitted Infection in last 12 months ("don't know" added to "no")  
freq(MLDHS$v763a, MLDHS$MLwt, plot = FALSE)  
sum(is.na(MLDHS$v763a))  
MLSTI\_recoded <- c("no"= 0, "yes" = 1, "don't know"= 0)  
MLDHS$MLREC\_STI <- MLSTI\_recoded[MLDHS$v763a]  
freq(MLDHS$MLREC\_STI, MLDHS$MLwt, plot = FALSE)  
  
  
  
#Number of Risk Factors  
# Smoking  
freq(MLDHS$v463z, MLDHS$MLwt, plot = FALSE)  
MLsmoke\_recoded <- c("no" = 1, "yes, smokes nothing" = 0)  
MLDHS$MLREC\_SMOKE <- MLsmoke\_recoded[MLDHS$v463z]  
freq(MLDHS$MLREC\_SMOKE, MLDHS$MLwt, plot = FALSE)  
#BMI  
freq(MLBiCAT, MLDHS$MLwt, plot = FALSE)  
MLbmi\_recoded <- c("high risk bmi" = 1, "low risk bmi" = 0)  
MLDHS$MLREC\_BMI <- MLbmi\_recoded[MLBiCAT]  
freq(MLDHS$MLREC\_BMI, MLDHS$MLwt, plot = FALSE)  
#AFSI  
freq(MLAFSICAT, MLDHS$MLwt, plot = FALSE)  
MLAFSI\_recoded <- c("less or equal 17 years" = 1, "greater than 17 years" = 0)  
MLDHS$MLREC\_AFSI <- MLAFSI\_recoded[MLAFSICAT]  
freq(MLDHS$MLREC\_AFSI, MLDHS$MLwt, plot = FALSE)  
#AFB  
freq(MLAFBCAT, MLDHS$MLwt, plot = FALSE)  
MLAFB\_recoded <- c("less or equal 19 years" = 1, "greater than 19 years" = 0)  
MLDHS$MLREC\_AFB <- MLAFB\_recoded[MLAFBCAT]  
freq(MLDHS$MLREC\_AFB, MLDHS$MLwt, plot = FALSE)  
#Parity  
freq(MLPCAT, MLDHS$MLwt, plot = FALSE)  
MLParity\_recoded <- c("greater than 2 births" = 1, "less or equal 2 births" = 0)  
MLDHS$MLREC\_Parity <- MLParity\_recoded[MLPCAT]  
freq(MLDHS$MLREC\_Parity, MLDHS$MLwt, plot = FALSE)  
#FP  
freq(MLDHS$MLcom, plot = FALSE)  
MLcom\_recoded <- c("female condom-Less than 5 years"=0, "female sterilization-Less than 5 years"=0, "female sterilization-More than or equal 5 years"=0,  
 "implants/norplant-Less than 5 years"= 0, "implants/norplant-More than or equal 5 years"=1, "injections-Less than 5 years"=0,  
 "injections-More than or equal 5 years"=1, "iud-Less than 5 years"=0, "iud-More than or equal 5 years"=1,  
 "lactational amenorrhea (lam)-Less than 5 years"=0, "male condom-Less than 5 years"=0, "male condom-More than or equal 5 years"=0,  
 "not using-NA"=0, "other traditional-Less than 5 years"=0, "periodic abstinence-Less than 5 years"=0,   
 "periodic abstinence-More than or equal 5 years"=0, "pill-Less than 5 years"=0, "pill-More than or equal 5 years"=1,  
 "standard days method (sdm)-Less than 5 years"=0, "standard days method (sdm)-More than or equal 5 years"=0,  
 "withdrawal-Less than 5 years"=0, "withdrawal-More than or equal 5 years"=0)  
MLDHS$MLyears\_FP <- MLcom\_recoded[MLDHS$MLcom]  
freq(MLDHS$MLyears\_FP, MLDHS$MLwt, plot = FALSE)  
# LSP  
freq(MLLSPCAT, MLDHS$MLwt, plot = FALSE)  
MLLSP\_recoded <- c("greater than 5 partners" = 1, "less or equal 5 partners" = 0)  
MLDHS$MLREC\_LSP <- MLLSP\_recoded[MLLSPCAT]  
freq(MLDHS$MLREC\_LSP, MLDHS$MLwt, plot = FALSE)  
# STI  
freq(MLDHS$v763a, MLDHS$MLwt, plot = FALSE )  
MLSTI\_recoded <- c("no"= 0, "yes" = 1, "don't know"= 0)  
MLDHS$MLREC\_STI <- MLSTI\_recoded[MLDHS$v763a]  
freq(MLDHS$MLREC\_STI, MLDHS$MLwt, plot = FALSE)  
  
  
#Number of Risk Factors  
MLRiskcom <- data.frame(MLDHS$MLREC\_SMOKE, MLDHS$MLREC\_BMI, MLDHS$MLREC\_AFSI, MLDHS$MLREC\_AFB, MLDHS$MLREC\_Parity, MLDHS$MLyears\_FP, MLDHS$MLREC\_LSP, MLDHS$MLREC\_STI)  
MLRiskcom  
MLtotal <- cbind(MLRiskcom, total = rowSums(MLRiskcom, na.rm = TRUE))  
MLtotal <- cbind(MLRiskcom, total = rowSums(MLRiskcom, na.rm = FALSE))  
tail(MLtotal)  
head(MLtotal)  
  
# Regrouped number of Risk Factors  
MLTotalnew <- NA   
MLTotalnew[MLtotal$total <=4.9999] <- "less or equal 4"  
MLTotalnew[MLtotal$total >= 5.0000] <- "greater or equal 5"   
freq(MLTotalnew, MLDHS$MLwt, plot = FALSE)  
  
  
# Modelling  
# Missing Data Removed  
MLDHS$MLREC\_summary= as.numeric(MLDHS$MLREC\_SMOKE + MLDHS$MLREC\_BMI + MLDHS$MLREC\_AFSI + MLDHS$MLREC\_AFB + MLDHS$MLREC\_Parity + MLDHS$MLyears\_FP + MLDHS$MLREC\_LSP + MLDHS$MLREC\_STI)  
crosstab(MLDHS$MLREC\_summary, MLDHS$v013, MLDHS$MLwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(MLDHS$v013, MLDHS$MLREC\_summary, MLDHS$MLwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(MLDHS$MLREC\_summary, MLDHS$v106, MLDHS$MLwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(MLDHS$v106, MLDHS$MLREC\_summary, MLDHS$MLwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(MLTotalnew, MLDHS$v013, MLDHS$MLwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
summary(MLDHS$MLREC\_summary)  
favstats(MLDHS$MLREC\_summary)  
designML <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~MLwt, data = MLDHS)  
is.numeric(MLDHS$MLREC\_summary)  
designML <-update(designML, MLREC\_summary=MLDHS$MLREC\_SMOKE + MLDHS$MLREC\_BMI + MLDHS$MLREC\_AFSI + MLDHS$MLREC\_AFB + MLDHS$MLREC\_Parity + MLDHS$MLyears\_FP + MLDHS$MLREC\_LSP + MLDHS$MLREC\_STI)  
  
  
# Unadjusted RR  
# Age  
MLmodelAge <- svyglm(MLREC\_summary~v013, design= designML, family = poisson(link = "log"), na.action = na.omit)  
summary(MLmodelAge)  
exp(coef(MLmodelAge)) # calculate the RRs  
round(exp(coef(MLmodelAge)), 2) # round to make it clearer  
cbind(RR = exp(coef(MLmodelAge)), exp(confint(MLmodelAge))) # RR to CI  
round(cbind(RR = exp(coef(MLmodelAge)), exp(confint(MLmodelAge))), 2) # round to 2  
# Residential Status  
MLmodelRS <- svyglm(MLREC\_summary~v025, design= designML, family = poisson(link = "log"), na.action = na.omit)  
summary(MLmodelRS)  
exp(coef(MLmodelRS)) # calculate the RRs  
round(exp(coef(MLmodelRS)), 2) # round to make it clearer  
cbind(RR = exp(coef(MLmodelRS)), exp(confint(MLmodelRS))) # RR to CI  
round(cbind(RR = exp(coef(MLmodelRS)), exp(confint(MLmodelRS))), 2) # round to 2  
# Marital Status  
MLmodelMS <- svyglm(MLREC\_summary~v502, design= designML, family = poisson(link = "log"), na.action = na.omit)  
summary(MLmodelMS)  
exp(coef(MLmodelMS)) # calculate the RRs  
round(exp(coef(MLmodelMS)), 2) # round to make it clearer  
cbind(RR = exp(coef(MLmodelMS)), exp(confint(MLmodelMS))) # RR to CI  
round(cbind(RR = exp(coef(MLmodelMS)), exp(confint(MLmodelMS))), 2) # round to 2  
# Education  
MLmodelEdu <- svyglm(MLREC\_summary~v106, design= designML, family = poisson(link = "log"), na.action = na.omit)  
summary(MLmodelEdu)  
exp(coef(MLmodelEdu)) # calculate the RRs  
round(exp(coef(MLmodelEdu)), 2) # round to make it clearer  
cbind(RR = exp(coef(MLmodelEdu)), exp(confint(MLmodelEdu))) # RR to CI  
round(cbind(RR = exp(coef(MLmodelEdu)), exp(confint(MLmodelEdu))), 2) # round to 2  
# Employment  
MLmodelEmp <- svyglm(MLREC\_summary~v714, design= designML, family = poisson(link = "log"), na.action = na.omit)  
summary(MLmodelEmp)  
exp(coef(MLmodelEmp)) # calculate the RRs  
round(exp(coef(MLmodelEmp)), 2) # round to make it clearer  
cbind(RR = exp(coef(MLmodelEmp)), exp(confint(MLmodelEmp))) # RR to CI  
round(cbind(RR = exp(coef(MLmodelEmp)), exp(confint(MLmodelEmp))), 2) # round to 2  
# Wealth Index  
MLmodelWI <- svyglm(MLREC\_summary~v190, design= designML, family = poisson(link = "log"), na.action = na.omit)  
summary(MLmodelWI)  
exp(coef(MLmodelWI)) # calculate the RRs  
round(exp(coef(MLmodelWI)), 2) # round to make it clearer  
cbind(RR = exp(coef(MLmodelWI)), exp(confint(MLmodelWI))) # RR to CI  
round(cbind(RR = exp(coef(MLmodelWI)), exp(confint(MLmodelWI))), 2) # round to 2  
# Adjusted RR  
# Poisson Regression Modelling  
MLmodel <- svyglm(MLREC\_summary~v013 + v025 + v190 + v714 + v106 + v502, design= designML, family = poisson(link = "log"), na.action = "na.omit") # Observations removed if it contains missing values  
summary(MLmodel)   
# calculate the RRs   
exp(coef(MLmodel))  
# round to make it clearer  
round(exp(coef(MLmodel)), 2)  
# RR to CI  
cbind(RR = exp(coef(MLmodel)), exp(confint(MLmodel)))  
# round to 2  
round(cbind(RR = exp(coef(MLmodel)), exp(confint(MLmodel))), 2)   
  
# Missing Data Included  
# Convert cases with NA into 0 and run a different model  
# library("imputeTS")  
#BMI  
freq(MLDHS$MLREC\_BMI, MLDHS$MLwt, na.rm= TRUE, plot = FALSE)  
MLDHS$MLREC\_BMI2 <- na.replace(MLDHS$MLREC\_BMI, 0) #gtools package  
freq(MLDHS$MLREC\_BMI2, MLDHS$MLwt, na.rm= TRUE, plot = FALSE)  
#AFB  
freq(MLDHS$MLREC\_AFB, MLDHS$MLwt, na.rm= TRUE, plot = FALSE)  
MLDHS$MLREC\_AFB2 <- na.replace(MLDHS$MLREC\_AFB, 0)  
freq(MLDHS$MLREC\_AFB2, MLDHS$MLwt, na.rm= TRUE, plot = FALSE)  
# LSP women  
freq(MLDHS$MLREC\_LSP, MLDHS$MLwt, plot = FALSE)  
MLDHS$MLREC\_LSP2 <- na.replace(MLDHS$MLREC\_LSP, 0)  
freq(MLDHS$MLREC\_LSP2, MLDHS$MLwt, na.rm= TRUE, plot = FALSE)  
# NB- Type of Contraceptive use was linked with month of usage. month of use contains NA i.e. not in the universe.   
# Variables with missing values- BMI (149), LSP woman (56) and LSP men (193)  
MLDHS$MLREC\_summary2= as.numeric(MLDHS$MLREC\_SMOKE + MLDHS$MLREC\_BMI2 + MLDHS$MLREC\_AFSI + MLDHS$MLREC\_AFB2 + MLDHS$MLREC\_Parity + MLDHS$MLyears\_FP + MLDHS$MLREC\_LSP2 + MLDHS$MLREC\_STI)  
summary(MLDHS$MLREC\_summary2)  
favstats(MLDHS$MLREC\_summary2)  
designML <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~MLwt, data = MLDHS)  
is.numeric(MLDHS$MLREC\_summary2)  
designML <-update(designML, MLREC\_summary2=MLDHS$MLREC\_SMOKE + MLDHS$MLREC\_BMI2 + MLDHS$MLREC\_AFSI + MLDHS$MLREC\_AFB2 + MLDHS$MLREC\_Parity + MLDHS$MLyears\_FP + MLDHS$MLREC\_LSP2 + MLDHS$MLREC\_STI)  
  
  
# Unadjusted RR  
# Age  
MLmodelAge2 <- svyglm(MLREC\_summary2~v013, design= designML, family = poisson(link = "log"), na.action = "na.pass")  
summary(MLmodelAge2)  
exp(coef(MLmodelAge2)) # calculate the RRs  
round(exp(coef(MLmodelAge2)), 2) # round to make it clearer  
cbind(RR = exp(coef(MLmodelAge2)), exp(confint(MLmodelAge2))) # RR to CI  
round(cbind(RR = exp(coef(MLmodelAge2)), exp(confint(MLmodelAge2))), 2) # round to 2  
# Residential Status  
MLmodelRS2 <- svyglm(MLREC\_summary2~v025, design= designML, family = poisson(link = "log"), na.action = "na.pass")  
summary(MLmodelRS2)  
exp(coef(MLmodelRS2)) # calculate the RRs  
round(exp(coef(MLmodelRS2)), 2) # round to make it clearer  
cbind(RR = exp(coef(MLmodelRS2)), exp(confint(MLmodelRS2))) # RR to CI  
round(cbind(RR = exp(coef(MLmodelRS2)), exp(confint(MLmodelRS2))), 2) # round to 2  
# Marital Status  
MLmodelMS2 <- svyglm(MLREC\_summary2~v502, design= designML, family = poisson(link = "log"), na.action = "na.pass")  
summary(MLmodelMS2)  
exp(coef(MLmodelMS2)) # calculate the RRs  
round(exp(coef(MLmodelMS2)), 2) # round to make it clearer  
cbind(RR = exp(coef(MLmodelMS2)), exp(confint(MLmodelMS2))) # RR to CI  
round(cbind(RR = exp(coef(MLmodelMS2)), exp(confint(MLmodelMS2))), 2) # round to 2  
# Education  
MLmodelEdu2 <- svyglm(MLREC\_summary2~v106, design= designML, family = poisson(link = "log"), na.action = "na.pass")  
summary(MLmodelEdu2)  
exp(coef(MLmodelEdu2)) # calculate the RRs  
round(exp(coef(MLmodelEdu2)), 2) # round to make it clearer  
cbind(RR = exp(coef(MLmodelEdu2)), exp(confint(MLmodelEdu2))) # RR to CI  
round(cbind(RR = exp(coef(MLmodelEdu2)), exp(confint(MLmodelEdu2))), 2) # round to 2  
# Employment  
MLmodelEmp2 <- svyglm(MLREC\_summary2~v714, design= designML, family = poisson(link = "log"), na.action = "na.pass")  
summary(MLmodelEmp2)  
exp(coef(MLmodelEmp2)) # calculate the RRs  
round(exp(coef(MLmodelEmp2)), 2) # round to make it clearer  
cbind(RR = exp(coef(MLmodelEmp2)), exp(confint(MLmodelEmp2))) # RR to CI  
round(cbind(RR = exp(coef(MLmodelEmp2)), exp(confint(MLmodelEmp2))), 2) # round to 2  
# Wealth Index  
MLmodelWI2 <- svyglm(MLREC\_summary2~v190, design= designML, family = poisson(link = "log"), na.action = "na.pass")  
summary(MLmodelWI2)  
exp(coef(MLmodelWI2)) # calculate the RRs  
round(exp(coef(MLmodelWI2)), 2) # round to make it clearer  
cbind(RR = exp(coef(MLmodelWI2)), exp(confint(MLmodelWI2))) # RR to CI  
round(cbind(RR = exp(coef(MLmodelWI2)), exp(confint(MLmodelWI2))), 2) # round to 2  
  
# Adjusted RR  
# Keep NA  
MLmodel2 <- svyglm(MLREC\_summary2~v013 + v025 + v190 + v714 + v106 + v502, design= designML, family = poisson(link = "log"), na.action = "na.pass") # keep all data, including NAs  
summary(MLmodel2)  
# calculate the RRs   
exp(coef(MLmodel2))  
# round to make it clearer  
round(exp(coef(MLmodel2)), 2)  
# RR to CI  
cbind(RR = exp(coef(MLmodel2)), exp(confint(MLmodel2)))  
# round to 2  
round(cbind(RR = exp(coef(MLmodel2)), exp(confint(MLmodel2))), 2)  
  
# Better Model  
# Comparing two models, AIC or BIC  
AIC(MLmodel, MLmodel2)  
BIC(MLmodel, MLmodel2, maximal = MLmodel2)  
plot\_summs(MLmodel, scale = TRUE, exp = TRUE)  
plot\_summs(MLmodel2, scale = TRUE, exp = TRUE)  
plot\_summs(MLmodel, MLmodel2, ci\_level = 0.95, model.names = c("Model 1"= "Model Excludes Missing", "Model 2"= "Model Includes Missing"),   
 colors = 'CUD Bright', exp = TRUE, legend.title = "Graph Comparing Models" )  
# Graph of Effect Size BJModel (Missing Data Excluded Model)  
MLeff.pres <- allEffects(MLmodel, scale = T)  
plot(MLeff.pres, ask=F)  
MLeff.pres2 <- allEffects(MLmodel2, scale = T)  
plot(MLeff.pres2, ask=F)

# Nigeria

# Select needed variable from women dataset in Nigeria  
NGDHS <- subset(NGDHSa, select = c(v001, v002, v021, v005, v025, v463z, v437, v438, v525, v212, v201,  
 v312, v337, v836, v763a, v012, v013, v502, v106, v714, v190, v130,   
 v717, v157, v158, v159, v171b, v481, v467b, v467c, v467d, v467f))  
  
# Sample weight/design for women  
NGDHS$NGwt <- NGDHS$v005/1000000  
designNG <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~NGwt, data = NGDHS)  
  
  
# Demographics  
NGDHS$NGwt <- NGDHS$v005/1000000  
freq(NGDHS$v013, NGDHS$NGwt, plot = FALSE) #Age  
freq(NGDHS$v025, NGDHS$NGwt, plot = FALSE) #Residential status  
freq(NGDHS$v502, NGDHS$NGwt, plot = FALSE) #Marital status  
freq(NGDHS$v106, NGDHS$NGwt, plot = FALSE) #Education  
freq(NGDHS$v714, NGDHS$NGwt, plot = FALSE) #Employment   
freq(NGDHS$v190, NGDHS$NGwt, plot = FALSE) #Wealth index  
designNG <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~NGwt, data = NGDHS) # Mean of Age and Confidence Interval  
NGttAge <- svyttest(v012~0, designNG)  
NGttAge  
  
# Proportion of Women With Risk Factors for Cervical Cancer  
#1. Use cigarettes and tobacco   
freq(NGDHS$v463z, NGDHS$NGwt, plot = FALSE)   
sum(is.na(NGDHS$v463z))  
#2. BMI   
favstats(NGDHS$v438, NGDHS$NGwt, plot = FALSE)  
sum(is.na(NGDHS$v438))  
is.na(NGDHS$v438)  
is.na(NGDHS$v438) <- which(NGDHS$v438 == 9996)  
is.na(NGDHS$v438) <- which(NGDHS$v438 == 9995)  
is.na(NGDHS$v438) <- which(NGDHS$v438 == 9994)  
favstats(NGDHS$v438, NGDHS$NGwt, plot = FALSE)  
NGDHS$NGHtM <- (NGDHS$v438/1000) #Height (Height is in cm, converted into meters. Also into 1 implied decimal)  
favstats(NGDHS$NGHtM, plot = FALSE)  
favstats(NGDHS$v437, NGDHS$NGwt, plot = FALSE)  
sum(is.na(NGDHS$v437))  
is.na(NGDHS$v437)  
is.na(NGDHS$v437) <- which(NGDHS$v437 == 9996)  
is.na(NGDHS$v437) <- which(NGDHS$v437 == 9995)  
is.na(NGDHS$v437) <- which(NGDHS$v437 == 9994)  
favstats(NGDHS$v437, NGDHS$NGwt, plot = FALSE)  
NGDHS$NGwtkg <- (NGDHS$v437/10) #Weight  
favstats(NGDHS$NGwtkg, plot = FALSE)  
NGDHS$BMI <- (NGDHS$NGwtkg) /(NGDHS$NGHtM \* NGDHS$NGHtM) # Calculated BMI  
favstats(NGDHS$BMI, NGDHS$NGwt, plot = FALSE)  
NGDHS$BMICAT <- NA  
NGDHS$BMICAT[NGDHS$BMI < 18.50000] <-1  
NGDHS$BMICAT[NGDHS$BMI >= 18.50000 & NGDHS$BMI <= 24.9999] <-2  
NGDHS$BMICAT[NGDHS$BMI >= 25.00000 & NGDHS$BMI <= 29.9999] <-3  
NGDHS$BMICAT[NGDHS$BMI >= 30.00000] <-4  
NGDHS$BMICAT <- factor(NGDHS$BMICAT, levels = c(1,2,3,4), labels = c("Underweight", "Normal", "Overweight", "Obese"))  
NGDHS$NGwt <- NGDHS$v005/1000000  
freq(NGDHS$BMICAT, NGDHS$NGwt, plot = FALSE)  
NGBiCAT <- NA # Binary Categories of BMI   
NGBiCAT[NGDHS$BMI < 25.0000] <- "low risk bmi"  
NGBiCAT[NGDHS$BMI >= 25.0000] <- "high risk bmi"  
freq(NGBiCAT, NGDHS$NGwt, plot = FALSE)  
#3. Age at First Sexual Intercourse   
freq(NGDHS$v525, NGDHS$NGwt, plot = FALSE)   
sum(is.na(NGDHS$v525))  
NGAFSI <- subset(NGDHS, v525 >=8 & v525 <=45)  
freq(NGAFSI$v525, plot = FALSE)  
NGAFSI$NGAFSIwt <- NGAFSI$v005/1000000  
designNGAFSI <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~NGAFSIwt, data = NGAFSI)  
svymean(~v525, design = designNGAFSI, na.rm = FALSE)  
confint(svymean(~v525, designNGAFSI))  
NGAFSICAT <- NA # Binary Categories of Age at First Sexual Intercourse Combine, had sex or not   
NGAFSICAT[NGDHS$v525 <=17.9999] <- "less or equal 17 years"  
NGAFSICAT[NGDHS$v525 >= 18.0000] <- "greater than 17 years"  
freq(NGAFSICAT, NGDHS$NGwt, plot = FALSE)  
#4. Age at First Birth   
freq(NGDHS$v212, NGDHS$NGwt, plot = FALSE)   
sum(is.na(NGDHS$v212))  
favstats(NGDHS$v212, NGDHS$NGwt, plot = FALSE)  
NGAFB <- subset(NGDHS, v212 >=11 & v212 <=48)  
freq(NGAFB$v212, plot = FALSE)  
NGAFB$NGAFBwt <- NGAFB$v005/1000000  
designNGAFB <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~NGAFBwt, data = NGAFB)  
svymean(~v212, design = designNGAFB, na.rm = FALSE)  
confint(svymean(~v212, designNGAFB))  
NGAFBCAT <- NA # Binary Categories of Age at First Birth  
NGAFBCAT[NGDHS$v212 <=19.9999] <- "less or equal 19 years"  
NGAFBCAT[NGDHS$v212 >= 20.0000] <- "greater than 19 years"  
freq(NGAFBCAT, NGDHS$NGwt, plot = FALSE)  
#5. Total Children Ever Born (Parity)  
freq(NGDHS$v201, NGDHS$NGwt, plot = FALSE)   
sum(is.na(NGDHS$v201))  
NGP <- subset(NGDHS, v201 >=1 & v201 <=17)  
freq(NGP$v201, plot = FALSE)  
NGP$NGPwt <- NGP$v005/1000000  
designNGP <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~NGPwt, data = NGP)  
svymean(~v201, design = designNGP, na.rm = FALSE)  
confint(svymean(~v201, designNGP))  
NGPCAT <- NA # Binary Categories of Total Children Ever Born (Parity)  
NGPCAT[NGDHS$v201<=2.9999] <- "less or equal 2 births"  
NGPCAT[NGDHS$v201 >= 3.0000] <- "greater than 2 births"  
freq(NGPCAT, NGDHS$NGwt, plot = FALSE)  
#6.Total Lifetime Number of Sex Partners   
freq(NGDHS$v836, NGDHS$NGwt, plot = FALSE)   
sum(is.na(NGDHS$v836))  
favstats(NGDHS$v836, NGDHS$NGwt, plot = FALSE)  
is.na(NGDHS$v836) <- which(NGDHS$v836 == 98)  
favstats(NGDHS$v836, NGDHS$NGwt, plot = FALSE)  
NGLSP <- subset(NGDHS, v836 >=1 & v836 <=95)  
freq(NGLSP$v836, plot = FALSE)  
NGLSP$NGLSPwt <- NGLSP$v005/1000000  
designNGLSP <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~NGLSPwt, data = NGLSP)  
svymean(~v836, design = designNGLSP, na.rm = FALSE)  
confint(svymean(~v836, designNGLSP))  
NGLSPCAT <- NA # Binary Categories of Total Lifetime Number of Sex Partners  
NGLSPCAT[NGDHS$v836 <=5.9999] <- "less or equal 5 partners"  
NGLSPCAT[NGDHS$v836 >= 6.0000] <- "greater than 5 partners"  
freq(NGLSPCAT, NGDHS$NGwt, plot = FALSE)  
#7. Contraceptive Usage  
freq(NGDHS$v312, NGDHS$NGwt, plot = FALSE)  
is.na(NGDHS$v312)  
sum(is.na(NGDHS$v312))  
nlevels(NGDHS$v312)  
table(NGDHS$v312)  
NGDHSContwt <- freq(NGDHS$v312, NGDHS$NGwt, plot = FALSE)  
NGDHSContwt  
NGCont\_recoded <- c("not using" = 0, "pill" =1, "iud"=1, "injections"=1, "diaphragm"= 2, "male condom"=2, "female sterilization"=2,  
 "male sterilization"=2, "periodic abstinence"=2, "withdrawal"=2, "other traditional"=2, "implants/norplant"=1,  
 "prolonged abstinence"= 2, "lactational amenorrhea (lam)"=2, "female condom"=2, "foam or jelly"= 2, "emergency contraception"= 1,  
 "other modern method"= 2, "standard days method (sdm)"=2, "specific method 1"= 2, "specific method 2"= 2)  
NGDHS$NGREC\_CONT <- NGCont\_recoded[NGDHS$v312]  
freq(NGDHS$NGREC\_CONT, NGDHS$NGwt, plot = FALSE)  
NGCont\_recoded <- c("not using" = 0, "pill" =1, "iud"=1, "injections"=1, "diaphragm"= 0, "male condom"=0, "female sterilization"=0,  
 "male sterilization"=0, "periodic abstinence"=0, "withdrawal"=0, "other traditional"=0, "implants/norplant"=1,  
 "prolonged abstinence"= 0, "lactational amenorrhea (lam)"=0, "female condom"=0, "foam or jelly"= 0, "emergency contraception"= 1,  
 "other modern method"= 0, "standard days method (sdm)"=0, "specific method 1"= 0, "specific method 2"= NA)  
NGDHS$NGREC\_CONT <- NGCont\_recoded[NGDHS$v312]  
freq(NGDHS$NGREC\_CONT, NGDHS$NGwt, plot = FALSE)  
# Months of use of current family planning method, For women who are currently using a contraceptive method  
freq(NGDHS$v337, NGDHS$NGwt, plot = FALSE)  
is.na(NGDHS$v337)  
sum(is.na(NGDHS$v337))  
favstats(NGDHS$v337, NGDHS$NGwt, plot = FALSE)  
summary(NGDHS$v337, NGDHS$NGwt, plot = FALSE)  
NGMFP <- NA  
NGMFP[NGDHS$v337 <60.0000]<- "Less than 5 years"  
NGMFP[NGDHS$v337 >=60.0000]<- "More than or equal 5 years"  
freq(NGMFP, NGDHS$NGwt, plot = FALSE)  
# Combine number hormonal and non hormonal to number of years  
NGDHS$NGcom <- paste(NGDHS$v312, NGDHS$v337)  
NGDHS$NGcom  
NGDHS$NGcom <- paste(NGDHS$v312, "-", NGDHS$v337)  
NGDHS$NGcom  
NGDHS$NGcom <- paste(NGDHS$v312, NGMFP, sep = "-" )  
NGDHS$NGcom  
freq(NGDHS$NGcom, plot = FALSE)  
NGcom\_recoded <- c( "emergency contraception-Less than 5 years"=0, "emergency contraception-More than or equal 5 years"=1, "female condom-Less than 5 years"=0, "female sterilization-Less than 5 years"=0,  
 "female sterilization-More than or equal 5 years"=0,"implants/norplant-Less than 5 years"=0, "implants/norplant-More than or equal 5 years"=1, "injections-Less than 5 years"=0,   
 "injections-More than or equal 5 years"=1, "iud-Less than 5 years"=0, "iud-More than or equal 5 years"=1, "lactational amenorrhea (lam)-Less than 5 years"= 0, "male condom-Less than 5 years"=0,   
 "male condom-More than or equal 5 years"=0, "not using-NA"=0, "other modern method-Less than 5 years"=0, "other modern method-More than or equal 5 years"=0,"other traditional-Less than 5 years"=0,   
 "other traditional-More than or equal 5 years"=0, "periodic abstinence-Less than 5 years"=0, "periodic abstinence-More than or equal 5 years"=0, "pill-Less than 5 years"=0, "pill-More than or equal 5 years"=1,   
 "standard days method (sdm)-Less than 5 years"=0, "standard days method (sdm)-More than or equal 5 years"=0,"withdrawal-Less than 5 years"=0, "withdrawal-More than or equal 5 years"=0)  
  
NGDHS$NGyears\_FP <- NGcom\_recoded[NGDHS$NGcom]  
freq(NGDHS$NGyears\_FP, NGDHS$NGwt, plot = FALSE)  
#8. Sexually Transmitted Infection in last 12 months ("don't know" added to "no")  
freq(NGDHS$v763a, NGDHS$NGwt, plot = FALSE)  
sum(is.na(NGDHS$v763a))  
NGSTI\_recoded <- c("no"= 0, "yes" = 1, "don't know"= 0)  
NGDHS$NGREC\_STI <- NGSTI\_recoded[NGDHS$v763a]  
freq(NGDHS$NGREC\_STI, NGDHS$NGwt, plot = FALSE)  
  
  
#Number of Risk Factors  
# Smoking  
freq(NGDHS$v463z, NGDHS$NGwt, plot = FALSE)  
NGsmoke\_recoded <- c("no" = 1, "yes, smokes nothing" = 0)  
NGDHS$NGREC\_SMOKE <- NGsmoke\_recoded[NGDHS$v463z]  
freq(NGDHS$NGREC\_SMOKE, NGDHS$NGwt, plot = FALSE)  
#BMI  
freq(NGBiCAT, NGDHS$NGwt, plot = FALSE)  
NGbmi\_recoded <- c("high risk bmi" = 1, "low risk bmi" = 0)  
NGDHS$NGREC\_BMI <- NGbmi\_recoded[NGBiCAT]  
freq(NGDHS$NGREC\_BMI, NGDHS$NGwt, plot = FALSE)  
#AFSI  
freq(NGAFSICAT, NGDHS$NGwt, plot = FALSE)  
NGAFSI\_recoded <- c("less or equal 17 years" = 1, "greater than 17 years" = 0)  
NGDHS$NGREC\_AFSI <- NGAFSI\_recoded[NGAFSICAT]  
freq(NGDHS$NGREC\_AFSI, NGDHS$NGwt, plot = FALSE)  
#AFB  
freq(NGAFBCAT, NGDHS$NGwt, plot = FALSE)  
NGAFB\_recoded <- c("less or equal 19 years" = 1, "greater than 19 years" = 0)  
NGDHS$NGREC\_AFB <- NGAFB\_recoded[NGAFBCAT]  
freq(NGDHS$NGREC\_AFB, NGDHS$NGwt, plot = FALSE)  
#Parity  
freq(NGPCAT, NGDHS$NGwt, plot = FALSE)  
NGParity\_recoded <- c("greater than 2 births" = 1, "less or equal 2 births" = 0)  
NGDHS$NGREC\_Parity <- NGParity\_recoded[NGPCAT]  
freq(NGDHS$NGREC\_Parity, NGDHS$NGwt, plot = FALSE)  
#FP  
freq(NGDHS$NGcom, plot = FALSE)  
NGcom\_recoded <- c( "emergency contraception-Less than 5 years"=0, "emergency contraception-More than or equal 5 years"=1, "female condom-Less than 5 years"=0, "female sterilization-Less than 5 years"=0,  
 "female sterilization-More than or equal 5 years"=0,"implants/norplant-Less than 5 years"=0, "implants/norplant-More than or equal 5 years"=1, "injections-Less than 5 years"=0,   
 "injections-More than or equal 5 years"=1, "iud-Less than 5 years"=0, "iud-More than or equal 5 years"=1, "lactational amenorrhea (lam)-Less than 5 years"= 0, "male condom-Less than 5 years"=0,   
 "male condom-More than or equal 5 years"=0, "not using-NA"=0, "other modern method-Less than 5 years"=0, "other modern method-More than or equal 5 years"=0,"other traditional-Less than 5 years"=0,   
 "other traditional-More than or equal 5 years"=0, "periodic abstinence-Less than 5 years"=0, "periodic abstinence-More than or equal 5 years"=0, "pill-Less than 5 years"=0, "pill-More than or equal 5 years"=1,   
 "standard days method (sdm)-Less than 5 years"=0, "standard days method (sdm)-More than or equal 5 years"=0,"withdrawal-Less than 5 years"=0, "withdrawal-More than or equal 5 years"=0)  
NGDHS$NGyears\_FP <- NGcom\_recoded[NGDHS$NGcom]  
freq(NGDHS$NGyears\_FP, NGDHS$NGwt, plot = FALSE)  
# LSP  
freq(NGLSPCAT, NGDHS$NGwt, plot = FALSE)  
NGLSP\_recoded <- c("greater than 5 partners" = 1, "less or equal 5 partners" = 0)  
NGDHS$NGREC\_LSP <- NGLSP\_recoded[NGLSPCAT]  
freq(NGDHS$NGREC\_LSP, NGDHS$NGwt, plot = FALSE)  
# STI  
freq(NGDHS$v763a, NGDHS$NGwt, plot = FALSE)  
NGSTI\_recoded <- c("no"= 0, "yes" = 1, "don't know"= 0)  
NGDHS$NGREC\_STI <- NGSTI\_recoded[NGDHS$v763a]  
freq(NGDHS$NGREC\_STI, NGDHS$NGwt, plot = FALSE)  
  
#Number of Risk Factors  
NGRiskcom <- data.frame(NGDHS$NGREC\_SMOKE, NGDHS$NGREC\_BMI, NGDHS$NGREC\_AFSI, NGDHS$NGREC\_AFB, NGDHS$NGREC\_Parity, NGDHS$NGyears\_FP, NGDHS$NGREC\_LSP, NGDHS$NGREC\_STI)  
NGRiskcom  
NGtotal <- cbind(NGRiskcom, total = rowSums(NGRiskcom, na.rm = TRUE))  
NGtotal <- cbind(NGRiskcom, total = rowSums(NGRiskcom, na.rm = FALSE))  
tail(NGtotal)  
head(NGtotal)  
  
# Regrouped number of Risk Factors  
NGTotalnew <- NA   
NGTotalnew[NGtotal$total <=4.9999] <- "less or equal 4"  
NGTotalnew[NGtotal$total >= 5.0000] <- "greater or equal 5"   
freq(NGTotalnew, NGDHS$NGwt, plot = FALSE)  
  
  
# Modelling  
# Missing Data Removed  
NGDHS$NGREC\_summary= as.numeric(NGDHS$NGREC\_SMOKE + NGDHS$NGREC\_BMI + NGDHS$NGREC\_AFSI + NGDHS$NGREC\_AFB + NGDHS$NGREC\_Parity + NGDHS$NGyears\_FP + NGDHS$NGREC\_LSP + NGDHS$NGREC\_STI)  
crosstab(NGDHS$NGREC\_summary, NGDHS$v013, NGDHS$NGwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(NGDHS$v013, NGDHS$NGREC\_summary, NGDHS$NGwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(NGDHS$NGREC\_summary, NGDHS$v106, NGDHS$NGwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(NGDHS$v106, NGDHS$NGREC\_summary, NGDHS$NGwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(NGTotalnew, NGDHS$v013, NGDHS$NGwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
summary(NGDHS$NGREC\_summary)  
favstats(NGDHS$NGREC\_summary)  
designNG <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~NGwt, data = NGDHS)  
is.numeric(NGDHS$MLREC\_summary)  
designNG <-update(designNG, NGREC\_summary=NGDHS$NGREC\_SMOKE + NGDHS$NGREC\_BMI + NGDHS$NGREC\_AFSI + NGDHS$NGREC\_AFB + NGDHS$NGREC\_Parity + NGDHS$NGyears\_FP + NGDHS$NGREC\_LSP + NGDHS$NGREC\_STI)  
  
  
# Unadjusted RR  
# Age  
NGmodelAge <- svyglm(NGREC\_summary~v013, design= designNG, family = poisson(link = "log"), na.action = na.omit)  
summary(NGmodelAge)  
exp(coef(NGmodelAge)) # calculate the RRs  
round(exp(coef(NGmodelAge)), 2) # round to make it clearer  
cbind(RR = exp(coef(NGmodelAge)), exp(confint(NGmodelAge))) # RR to CI  
round(cbind(RR = exp(coef(NGmodelAge)), exp(confint(NGmodelAge))), 2) # round to 2  
# Residential Status  
NGmodelRS <- svyglm(NGREC\_summary~v025, design= designNG, family = poisson(link = "log"), na.action = na.omit)  
summary(NGmodelRS)  
exp(coef(NGmodelRS)) # calculate the RRs  
round(exp(coef(NGmodelRS)), 2) # round to make it clearer  
cbind(RR = exp(coef(NGmodelRS)), exp(confint(NGmodelRS))) # RR to CI  
round(cbind(RR = exp(coef(NGmodelRS)), exp(confint(NGmodelRS))), 2) # round to 2  
# Marital Status  
NGmodelMS <- svyglm(NGREC\_summary~v502, design= designNG, family = poisson(link = "log"), na.action = na.omit)  
summary(NGmodelMS)  
exp(coef(NGmodelMS)) # calculate the RRs  
round(exp(coef(NGmodelMS)), 2) # round to make it clearer  
cbind(RR = exp(coef(NGmodelMS)), exp(confint(NGmodelMS))) # RR to CI  
round(cbind(RR = exp(coef(NGmodelMS)), exp(confint(NGmodelMS))), 2) # round to 2  
# Education  
NGmodelEdu <- svyglm(NGREC\_summary~v106, design= designNG, family = poisson(link = "log"), na.action = na.omit)  
summary(NGmodelEdu)  
exp(coef(NGmodelEdu)) # calculate the RRs  
round(exp(coef(NGmodelEdu)), 2) # round to make it clearer  
cbind(RR = exp(coef(NGmodelEdu)), exp(confint(NGmodelEdu))) # RR to CI  
round(cbind(RR = exp(coef(NGmodelEdu)), exp(confint(NGmodelEdu))), 2) # round to 2  
# Employment  
NGmodelEmp <- svyglm(NGREC\_summary~v714, design= designNG, family = poisson(link = "log"), na.action = na.omit)  
summary(NGmodelEmp)  
exp(coef(NGmodelEmp)) # calculate the RRs  
round(exp(coef(NGmodelEmp)), 2) # round to make it clearer  
cbind(RR = exp(coef(NGmodelEmp)), exp(confint(NGmodelEmp))) # RR to CI  
round(cbind(RR = exp(coef(NGmodelEmp)), exp(confint(NGmodelEmp))), 2) # round to 2  
# Wealth Index  
NGmodelWI <- svyglm(NGREC\_summary~v190, design= designNG, family = poisson(link = "log"), na.action = na.omit)  
summary(NGmodelWI)  
exp(coef(NGmodelWI)) # calculate the RRs  
round(exp(coef(NGmodelWI)), 2) # round to make it clearer  
cbind(RR = exp(coef(NGmodelWI)), exp(confint(NGmodelWI))) # RR to CI  
round(cbind(RR = exp(coef(NGmodelWI)), exp(confint(NGmodelWI))), 2) # round to 2  
# Adjusted RR  
# Poisson Regression Modelling  
NGmodel <- svyglm(NGREC\_summary~v013 + v025 + v190 + v714 + v106 + v502, design= designNG, family = poisson(link = "log"), na.action = "na.omit") # Observations removed if it contains missing values  
summary(NGmodel)   
# calculate the RRs   
exp(coef(NGmodel))  
# round to make it clearer  
round(exp(coef(NGmodel)), 2)  
# RR to CI  
cbind(RR = exp(coef(NGmodel)), exp(confint(NGmodel)))  
# round to 2  
round(cbind(RR = exp(coef(NGmodel)), exp(confint(NGmodel))), 2)   
  
# Missing Data Included  
# Convert cases with NA into 0 and run a different model  
# library("imputeTS")  
#BMI  
freq(NGDHS$NGREC\_BMI, NGDHS$NGwt, na.rm= TRUE, plot = FALSE)  
NGDHS$NGREC\_BMI2 <- na.replace(NGDHS$NGREC\_BMI, 0) #gtools package  
freq(NGDHS$NGREC\_BMI2, NGDHS$NGwt, na.rm= TRUE, plot = FALSE)  
#AFB  
freq(NGDHS$NGREC\_AFB, NGDHS$NGwt, na.rm= TRUE, plot = FALSE)  
NGDHS$NGREC\_AFB2 <- na.replace(NGDHS$NGREC\_AFB, 0)  
freq(NGDHS$NGREC\_AFB2, NGDHS$NGwt, na.rm= TRUE, plot = FALSE)  
# LSP women  
freq(NGDHS$NGREC\_LSP, NGDHS$NGwt, plot = FALSE)  
NGDHS$NGREC\_LSP2 <- na.replace(NGDHS$NGREC\_LSP, 0)  
freq(NGDHS$NGREC\_LSP2, NGDHS$NGwt, na.rm= TRUE, plot = FALSE)  
# NB- Type of Contraceptive use was linked with month of usage. month of use contains NA i.e. not in the universe.   
# Variables with missing values- BMI (149), LSP woman (56) and LSP men (193)  
NGDHS$NGREC\_summary2= as.numeric(NGDHS$NGREC\_SMOKE + NGDHS$NGREC\_BMI2 + NGDHS$NGREC\_AFSI + NGDHS$NGREC\_AFB2 + NGDHS$NGREC\_Parity + NGDHS$NGyears\_FP + NGDHS$NGREC\_LSP2 + NGDHS$NGREC\_STI)  
summary(NGDHS$NGREC\_summary2)  
favstats(NGDHS$NGREC\_summary2)  
designNG <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~NGwt, data = NGDHS)  
is.numeric(NGDHS$NGREC\_summary2)  
designNG <-update(designNG, NGREC\_summary2=NGDHS$NGREC\_SMOKE + NGDHS$NGREC\_BMI2 + NGDHS$NGREC\_AFSI + NGDHS$NGREC\_AFB2 + NGDHS$NGREC\_Parity + NGDHS$NGyears\_FP + NGDHS$NGREC\_LSP2 + NGDHS$NGREC\_STI)  
  
  
# Unadjusted RR  
# Age  
NGmodelAge2 <- svyglm(NGREC\_summary2~v013, design= designNG, family = poisson(link = "log"), na.action = "na.pass")  
summary(NGmodelAge2)  
exp(coef(NGmodelAge2)) # calculate the RRs  
round(exp(coef(NGmodelAge2)), 2) # round to make it clearer  
cbind(RR = exp(coef(NGmodelAge2)), exp(confint(NGmodelAge2))) # RR to CI  
round(cbind(RR = exp(coef(NGmodelAge2)), exp(confint(NGmodelAge2))), 2) # round to 2  
# Residential Status  
NGmodelRS2 <- svyglm(NGREC\_summary2~v025, design= designNG, family = poisson(link = "log"), na.action = "na.pass")  
summary(NGmodelRS2)  
exp(coef(NGmodelRS2)) # calculate the RRs  
round(exp(coef(NGmodelRS2)), 2) # round to make it clearer  
cbind(RR = exp(coef(NGmodelRS2)), exp(confint(NGmodelRS2))) # RR to CI  
round(cbind(RR = exp(coef(NGmodelRS2)), exp(confint(NGmodelRS2))), 2) # round to 2  
# Marital Status  
NGmodelMS2 <- svyglm(NGREC\_summary2~v502, design= designNG, family = poisson(link = "log"), na.action = "na.pass")  
summary(NGmodelMS2)  
exp(coef(NGmodelMS2)) # calculate the RRs  
round(exp(coef(NGmodelMS2)), 2) # round to make it clearer  
cbind(RR = exp(coef(NGmodelMS2)), exp(confint(NGmodelMS2))) # RR to CI  
round(cbind(RR = exp(coef(NGmodelMS2)), exp(confint(NGmodelMS2))), 2) # round to 2  
# Education  
NGmodelEdu2 <- svyglm(NGREC\_summary2~v106, design= designNG, family = poisson(link = "log"), na.action = "na.pass")  
summary(NGmodelEdu2)  
exp(coef(NGmodelEdu2)) # calculate the RRs  
round(exp(coef(NGmodelEdu2)), 2) # round to make it clearer  
cbind(RR = exp(coef(NGmodelEdu2)), exp(confint(NGmodelEdu2))) # RR to CI  
round(cbind(RR = exp(coef(NGmodelEdu2)), exp(confint(NGmodelEdu2))), 2) # round to 2  
# Employment  
NGmodelEmp2 <- svyglm(NGREC\_summary2~v714, design= designNG, family = poisson(link = "log"), na.action = "na.pass")  
summary(NGmodelEmp2)  
exp(coef(NGmodelEmp2)) # calculate the RRs  
round(exp(coef(NGmodelEmp2)), 2) # round to make it clearer  
cbind(RR = exp(coef(NGmodelEmp2)), exp(confint(NGmodelEmp2))) # RR to CI  
round(cbind(RR = exp(coef(NGmodelEmp2)), exp(confint(NGmodelEmp2))), 2) # round to 2  
# Wealth Index  
NGmodelWI2 <- svyglm(NGREC\_summary2~v190, design= designNG, family = poisson(link = "log"), na.action = "na.pass")  
summary(NGmodelWI2)  
exp(coef(NGmodelWI2)) # calculate the RRs  
round(exp(coef(NGmodelWI2)), 2) # round to make it clearer  
cbind(RR = exp(coef(NGmodelWI2)), exp(confint(NGmodelWI2))) # RR to CI  
round(cbind(RR = exp(coef(NGmodelWI2)), exp(confint(NGmodelWI2))), 2) # round to 2  
  
# Adjusted RR  
# Keep NA  
NGmodel2 <- svyglm(NGREC\_summary2~v013 + v025 + v190 + v714 + v106 + v502, design= designNG, family = poisson(link = "log"), na.action = "na.pass") # keep all data, including NAs  
summary(NGmodel2)  
# calculate the RRs   
exp(coef(NGmodel2))  
# round to make it clearer  
round(exp(coef(NGmodel2)), 2)  
# RR to CI  
cbind(RR = exp(coef(NGmodel2)), exp(confint(NGmodel2)))  
# round to 2  
round(cbind(RR = exp(coef(NGmodel2)), exp(confint(NGmodel2))), 2)  
  
# Better Model  
# Comparing two models, AIC or BIC  
AIC(NGmodel, NGmodel2)  
BIC(NGmodel, NGmodel2, maximal = NGmodel2)  
plot\_summs(NGmodel, scale = TRUE, exp = TRUE)  
plot\_summs(NGmodel2, scale = TRUE, exp = TRUE)  
plot\_summs(NGmodel, NGmodel2, ci\_level = 0.95, model.names = c("Model 1"= "Model Excludes Missing", "Model 2"= "Model Includes Missing"),   
 colors = 'CUD Bright', exp = TRUE, legend.title = "Graph Comparing Models" )  
# Graph of Effect Size BJModel (Missing Data Excluded Model) Effect package  
NGeff.pres <- allEffects(NGmodel, scale = T)  
plot(NGeff.pres, ask=F)  
NGeff.pres2 <- allEffects(NGmodel2, scale = T)  
plot(NGeff.pres2, ask=F)

# Sierra Leone

# Select needed variable from women dataset in Sierra Leone  
SLDHS <- subset(SLDHSa, select = c(v001, v002, v021, v005, v025, v463z, v437, v438, v525, v212, v201,  
 v312, v337, v836, v763a, v012, v013, v502, v106, v714, v190, v130,   
 v717, v157, v158, v159, v171b, v481, v467b, v467c, v467d, v467f))  
  
# Sample weight/design for women  
SLDHS$SLwt <- SLDHS$v005/1000000  
designSL <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~SLwt, data = SLDHS)  
  
  
# Demographics  
SLDHS$SLwt <- SLDHS$v005/1000000  
freq(SLDHS$v013, SLDHS$SLwt, plot = FALSE) #Age  
freq(SLDHS$v025, SLDHS$SLwt, plot = FALSE) #Residential status  
freq(SLDHS$v502, SLDHS$SLwt, plot = FALSE) #Marital status  
freq(SLDHS$v106, SLDHS$SLwt, plot = FALSE) #Education  
freq(SLDHS$v714, SLDHS$SLwt, plot = FALSE) #Employment   
freq(SLDHS$v190, SLDHS$SLwt, plot = FALSE) #Wealth index  
designSL <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~SLwt, data = SLDHS) # Mean of Age and Confidence Interval  
SLttAge <- svyttest(v012~0, designSL)  
SLttAge  
  
# Proportion of Women With Risk Factors for Cervical Cancer  
#1. Use cigarettes and tobacco   
freq(SLDHS$v463z, SLDHS$SLwt, plot = FALSE)   
sum(is.na(SLDHS$v463z))  
#2. BMI   
favstats(SLDHS$v438, SLDHS$SLwt, plot = FALSE)  
sum(is.na(SLDHS$v438))  
is.na(SLDHS$v438)  
is.na(SLDHS$v438) <- which(SLDHS$v438 == 9996)  
is.na(SLDHS$v438) <- which(SLDHS$v438 == 9995)  
is.na(SLDHS$v438) <- which(SLDHS$v438 == 9994)  
favstats(SLDHS$v438, SLDHS$SLwt, plot = FALSE)  
SLDHS$SLHtM <- (SLDHS$v438/1000) #Height (Height is in cm, converted into meters. Also into 1 implied decimal)  
favstats(SLDHS$SLHtM, plot = FALSE)  
favstats(SLDHS$v437, SLDHS$SLwt, plot = FALSE)  
sum(is.na(SLDHS$v437))  
is.na(SLDHS$v437)  
is.na(SLDHS$v437) <- which(SLDHS$v437 == 9996)  
is.na(SLDHS$v437) <- which(SLDHS$v437 == 9995)  
is.na(SLDHS$v437) <- which(SLDHS$v437 == 9994)  
favstats(SLDHS$v437, SLDHS$SLwt, plot = FALSE)  
SLDHS$SLwtkg <- (SLDHS$v437/10) #Weight  
favstats(SLDHS$SLwtkg, plot = FALSE)  
SLDHS$BMI <- (SLDHS$SLwtkg) /(SLDHS$SLHtM \* SLDHS$SLHtM) # Calculated BMI  
favstats(SLDHS$BMI, na.rm = TRUE)  
SLDHS$BMICAT <- NA  
SLDHS$BMICAT[SLDHS$BMI < 18.50000] <-1  
SLDHS$BMICAT[SLDHS$BMI >= 18.50000 & SLDHS$BMI <= 24.9999] <-2  
SLDHS$BMICAT[SLDHS$BMI >= 25.00000 & SLDHS$BMI <= 29.9999] <-3  
SLDHS$BMICAT[SLDHS$BMI >= 30.00000] <-4  
SLDHS$BMICAT <- factor(SLDHS$BMICAT, levels = c(1,2,3,4), labels = c("Underweight", "Normal", "Overweight", "Obese"))  
SLDHS$SLwt <- SLDHS$v005/1000000  
freq(SLDHS$BMICAT, SLDHS$SLwt, plot = FALSE)  
SLBiCAT <- NA # Binary Categories of BMI   
SLBiCAT[SLDHS$BMI < 25.0000] <- "low risk bmi"  
SLBiCAT[SLDHS$BMI >= 25.0000] <- "high risk bmi"  
freq(SLBiCAT, SLDHS$SLwt, plot = FALSE)  
#3. Age at First Sexual Intercourse   
freq(SLDHS$v525, SLDHS$SLwt, plot = FALSE)   
sum(is.na(SLDHS$v525))  
SLAFSI <- subset(SLDHS, v525 >=8 & v525 <=37)  
freq(SLAFSI$v525, plot = FALSE)  
SLAFSI$SLAFSIwt <- SLAFSI$v005/1000000  
designSLAFSI <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~SLAFSIwt, data = SLAFSI)  
svymean(~v525, design = designSLAFSI, na.rm = FALSE)  
confint(svymean(~v525, designSLAFSI))  
SLAFSICAT <- NA # Binary Categories of Age at First Sexual Intercourse Combine, had sex or not   
SLAFSICAT[SLDHS$v525 <=17.9999] <- "less or equal 17 years"  
SLAFSICAT[SLDHS$v525 >= 18.0000] <- "greater than 17 years"  
freq(SLAFSICAT, SLDHS$SLwt, plot = FALSE)  
#4. Age at First Birth   
freq(SLDHS$v212, SLDHS$SLwt, plot = FALSE)  
sum(is.na(SLDHS$v212))  
favstats(SLDHS$v212, SLDHS$SLwt, plot = FALSE)  
SLAFB <- subset(SLDHS, v212 >=11 & v212 <=42)  
freq(SLAFB$v212, plot = FALSE)  
SLAFB$SLAFBwt <- SLAFB$v005/1000000  
designSLAFB <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~SLAFBwt, data = SLAFB)  
svymean(~v212, design = designSLAFB, na.rm = FALSE)  
confint(svymean(~v212, designSLAFB))  
SLAFBCAT <- NA # Binary Categories of Age at First Birth  
SLAFBCAT[SLDHS$v212 <=19.9999] <- "less or equal 19 years"  
SLAFBCAT[SLDHS$v212 >= 20.0000] <- "greater than 19 years"  
freq(SLAFBCAT, SLDHS$SLwt, plot = FALSE)  
#5. Total Children Ever Born (Parity)  
freq(SLDHS$v201, SLDHS$SLwt, plot = FALSE)   
sum(is.na(SLDHS$v201))  
SLP <- subset(SLDHS, v201 >=1 & v201 <=14)  
freq(SLP$v201, plot = FALSE)  
SLP$SLPwt <- SLP$v005/1000000  
designSLP <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~SLPwt, data = SLP)  
svymean(~v201, design = designSLP, na.rm = FALSE)  
confint(svymean(~v201, designSLP))  
SLPCAT <- NA # Binary Categories of Total Children Ever Born (Parity)  
SLPCAT[SLDHS$v201<=2.9999] <- "less or equal 2 births"  
SLPCAT[SLDHS$v201 >= 3.0000] <- "greater than 2 births"  
freq(SLPCAT, SLDHS$SLwt, plot = FALSE)  
#6. Total Lifetime Number of Sex Partners #Missing Data Excluded(430)  
freq(SLDHS$v836, SLDHS$SLwt, plot = FALSE)   
sum(is.na(SLDHS$v836))  
favstats(SLDHS$v836, SLDHS$SLwt, plot = FALSE)  
is.na(SLDHS$v836) <- which(SLDHS$v836 == 98)  
favstats(SLDHS$v836, SLDHS$SLwt, plot = FALSE)  
SLLSP <- subset(SLDHS, v836 >=1 & v836 <=95)  
freq(SLLSP$v836, plot = FALSE)  
SLLSP$SLLSPwt <- SLLSP$v005/1000000  
designSLLSP <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~SLLSPwt, data = SLLSP)  
svymean(~v836, design = designSLLSP, na.rm = FALSE)  
confint(svymean(~v836, designSLLSP))  
SLLSPCAT <- NA # Binary Categories of Total Lifetime Number of Sex Partners  
SLLSPCAT[SLDHS$v836 <=5.9999] <- "less or equal 5 partners"  
SLLSPCAT[SLDHS$v836 >= 6.0000] <- "greater than 5 partners"  
freq(SLLSPCAT, SLDHS$SLwt, plot = FALSE)  
#7. Contraceptive Usage  
freq(SLDHS$v312, SLDHS$SLwt, plot = FALSE)  
is.na(SLDHS$v312)  
sum(is.na(SLDHS$v312))  
nlevels(SLDHS$v312)  
table(SLDHS$v312)  
SLDHSContwt <- freq(SLDHS$v312, SLDHS$SLwt, plot = FALSE)  
SLDHSContwt  
SLCont\_recoded <- c("not using" = 0, "pill" =1, "iud"=1, "injections"=1, "diaphragm"= 2, "male condom"=2, "female sterilization"=2,  
 "male sterilization"=2, "periodic abstinence"=2, "withdrawal"=2, "other traditional"=2, "implants/norplant"=1,  
 "prolonged abstinence"= 2, "lactational amenorrhea (lam)"=2, "female condom"=2, "foam or jelly"= 2, "emergency contraception"= 1,  
 "other modern method"= 2, "standard days method (sdm)"=2, "specific method 1"= 2, "specific method 2"= 2)  
SLDHS$BJREC\_CONT <- SLCont\_recoded[SLDHS$v312]  
freq(SLDHS$BJREC\_CONT, SLDHS$SLwt, plot = FALSE)  
SLCont\_recoded <- c("not using" = 0, "pill" =1, "iud"=1, "injections"=1, "diaphragm"= 0, "male condom"=0, "female sterilization"=0,  
 "male sterilization"=0, "periodic abstinence"=0, "withdrawal"=0, "other traditional"=0, "implants/norplant"=1,  
 "prolonged abstinence"= 0, "lactational amenorrhea (lam)"=0, "female condom"=0, "foam or jelly"= 0, "emergency contraception"= 1,  
 "other modern method"= 0, "standard days method (sdm)"=0, "specific method 1"= 0, "specific method 2"= NA)  
SLDHS$SLREC\_CONT <- SLCont\_recoded[SLDHS$v312]  
freq(SLDHS$SLREC\_CONT, SLDHS$SLwt, plot = FALSE)  
# Months of use of current family planning method, For women who are currently using a contraceptive method  
freq(SLDHS$v337, SLDHS$SLwt, plot = FALSE)  
is.na(SLDHS$v337)  
sum(is.na(SLDHS$v337))  
favstats(SLDHS$v337, SLDHS$SLwt, plot = FALSE)  
summary(SLDHS$v337, SLDHS$SLwt, plot = FALSE)  
SLMFP <- NA  
SLMFP[SLDHS$v337 <60.0000]<- "Less than 5 years"  
SLMFP[SLDHS$v337 >=60.0000]<- "More than or equal 5 years"  
freq(SLMFP, SLDHS$SLwt, plot = FALSE)  
# Combine number hormonal and non hormonal to number of years  
SLDHS$SLcom <- paste(SLDHS$v312, SLDHS$v337)  
SLDHS$SLcom  
SLDHS$SLcom <- paste(SLDHS$v312, "-", SLDHS$v337)  
SLDHS$SLcom  
SLDHS$SLcom <- paste(SLDHS$v312, SLMFP, sep = "-" )  
SLDHS$SLcom  
freq(SLDHS$SLcom, plot = FALSE)  
SLcom\_recoded <- c("emergency contraception-Less than 5 years"=0, "emergency contraception-More than or equal 5 years"=1, "female condom-More than or equal 5 years"=0,  
 "female sterilization-Less than 5 years"=0, "female sterilization-More than or equal 5 years"=0, "implants/norplant-Less than 5 years"=0, "implants/norplant-More than or equal 5 years"=1,  
 "injections-Less than 5 years"=0, "injections-More than or equal 5 years"=1, "iud-Less than 5 years"=0, "iud-More than or equal 5 years"=1, "lactational amenorrhea (lam)-Less than 5 years"=0,   
 "male condom-Less than 5 years"=0, "male condom-More than or equal 5 years"=0, "male sterilization-Less than 5 years"=0, "not using-NA"=0, "other traditional-Less than 5 years"=0,  
 "other traditional-More than or equal 5 years"=0, "periodic abstinence-Less than 5 years"=0, "periodic abstinence-More than or equal 5 years"=0,  
 "pill-Less than 5 years"=0, "pill-More than or equal 5 years"=1, "standard days method (sdm)-Less than 5 years"=0, "standard days method (sdm)-More than or equal 5 years"=0,  
 "withdrawal-Less than 5 years"=0, "withdrawal-More than or equal 5 years"=0)  
SLDHS$SLyears\_FP <- SLcom\_recoded[SLDHS$SLcom]  
freq(SLDHS$SLyears\_FP, SLDHS$SLwt, plot = FALSE)  
#8. Sexually Transmitted Infection in last 12 months ("don't know" added to "no")  
freq(SLDHS$v763a, SLDHS$SLwt, plot = FALSE)  
sum(is.na(SLDHS$v763a))  
SLSTI\_recoded <- c("no"= 0, "yes" = 1, "don't know"= 0)  
SLDHS$SLREC\_STI <- SLSTI\_recoded[SLDHS$v763a]  
freq(SLDHS$SLREC\_STI, SLDHS$SLwt, plot = FALSE)  
  
  
#Number of Risk Factors  
# Smoking  
freq(SLDHS$v463z, SLDHS$SLwt, plot = FALSE)  
SLsmoke\_recoded <- c("no" = 1, "yes, smokes nothing" = 0)  
SLDHS$SLREC\_SMOKE <- SLsmoke\_recoded[SLDHS$v463z]  
freq(SLDHS$SLREC\_SMOKE, SLDHS$SLwt, plot = FALSE)  
#BMI  
freq(SLBiCAT, SLDHS$SLwt, plot = FALSE)  
SLbmi\_recoded <- c("high risk bmi" = 1, "low risk bmi" = 0)  
SLDHS$SLREC\_BMI <- SLbmi\_recoded[SLBiCAT]  
freq(SLDHS$SLREC\_BMI, SLDHS$SLwt, plot = FALSE)  
#AFSI  
freq(SLAFSICAT, SLDHS$SLwt, plot = FALSE)  
SLAFSI\_recoded <- c("less or equal 17 years" = 1, "greater than 17 years" = 0)  
SLDHS$SLREC\_AFSI <- SLAFSI\_recoded[SLAFSICAT]  
freq(SLDHS$SLREC\_AFSI, SLDHS$SLwt, plot = FALSE)  
#AFB  
freq(SLAFBCAT, SLDHS$SLwt, plot = FALSE)  
SLAFB\_recoded <- c("less or equal 19 years" = 1, "greater than 19 years" = 0)  
SLDHS$SLREC\_AFB <- SLAFB\_recoded[SLAFBCAT]  
freq(SLDHS$SLREC\_AFB, SLDHS$SLwt, plot = FALSE)  
#Parity  
freq(SLPCAT, SLDHS$SLwt, plot = FALSE)  
SLParity\_recoded <- c("greater than 2 births" = 1, "less or equal 2 births" = 0)  
SLDHS$SLREC\_Parity <- SLParity\_recoded[SLPCAT]  
freq(SLDHS$SLREC\_Parity, SLDHS$SLwt, plot = FALSE)  
#FP  
freq(SLDHS$SLcom, plot = FALSE)  
SLcom\_recoded <- c("emergency contraception-Less than 5 years"=0, "emergency contraception-More than or equal 5 years"=1, "female condom-More than or equal 5 years"=0,  
 "female sterilization-Less than 5 years"=0, "female sterilization-More than or equal 5 years"=0, "implants/norplant-Less than 5 years"=0, "implants/norplant-More than or equal 5 years"=1,  
 "injections-Less than 5 years"=0, "injections-More than or equal 5 years"=1, "iud-Less than 5 years"=0, "iud-More than or equal 5 years"=1, "lactational amenorrhea (lam)-Less than 5 years"=0,   
 "male condom-Less than 5 years"=0, "male condom-More than or equal 5 years"=0, "male sterilization-Less than 5 years"=0, "not using-NA"=0, "other traditional-Less than 5 years"=0,  
 "other traditional-More than or equal 5 years"=0, "periodic abstinence-Less than 5 years"=0, "periodic abstinence-More than or equal 5 years"=0,  
 "pill-Less than 5 years"=0, "pill-More than or equal 5 years"=1, "standard days method (sdm)-Less than 5 years"=0, "standard days method (sdm)-More than or equal 5 years"=0,  
 "withdrawal-Less than 5 years"=0, "withdrawal-More than or equal 5 years"=0)  
SLDHS$SLyears\_FP <- SLcom\_recoded[SLDHS$SLcom]  
freq(SLDHS$SLyears\_FP, SLDHS$SLwt, plot = FALSE)  
# LSP  
freq(SLLSPCAT, SLDHS$SLwt, plot = FALSE)  
SLLSP\_recoded <- c("greater than 5 partners" = 1, "less or equal 5 partners" = 0)  
SLDHS$SLREC\_LSP <- SLLSP\_recoded[SLLSPCAT]  
freq(SLDHS$SLREC\_LSP, SLDHS$SLwt, plot = FALSE)  
# STI  
freq(SLDHS$v763a, SLDHS$SLwt, plot = FALSE)  
SLSTI\_recoded <- c("no"= 0, "yes" = 1, "don't know"= 0)  
SLDHS$SLREC\_STI <- SLSTI\_recoded[SLDHS$v763a]  
freq(SLDHS$SLREC\_STI, SLDHS$SLwt, plot = FALSE)  
  
#Number of Risk Factors  
SLRiskcom <- data.frame(SLDHS$SLREC\_SMOKE, SLDHS$SLREC\_BMI, SLDHS$SLREC\_AFSI, SLDHS$SLREC\_AFB, SLDHS$SLREC\_Parity, SLDHS$SLyears\_FP, SLDHS$SLREC\_LSP, SLDHS$SLREC\_STI)  
SLRiskcom  
SLtotal <- cbind(SLRiskcom, total = rowSums(SLRiskcom, na.rm = TRUE))  
SLtotal <- cbind(SLRiskcom, total = rowSums(SLRiskcom, na.rm = FALSE))  
tail(SLtotal)  
head(SLtotal)  
  
# Regrouped number of Risk Factors  
SLTotalnew <- NA   
SLTotalnew[SLtotal$total <=4.9999] <- "less or equal 4"  
SLTotalnew[SLtotal$total >= 5.0000] <- "greater or equal 5"   
freq(SLTotalnew, SLDHS$SLwt, plot = FALSE)  
  
  
# Modelling  
# Missing Data Removed  
SLDHS$SLREC\_summary= as.numeric(SLDHS$SLREC\_SMOKE + SLDHS$SLREC\_BMI + SLDHS$SLREC\_AFSI + SLDHS$SLREC\_AFB + SLDHS$SLREC\_Parity + SLDHS$SLyears\_FP + SLDHS$SLREC\_LSP + SLDHS$SLREC\_STI)  
crosstab(SLDHS$SLREC\_summary, SLDHS$v013, SLDHS$SLwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(SLDHS$v013, SLDHS$SLREC\_summary, SLDHS$SLwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(SLDHS$SLREC\_summary, SLDHS$v106, SLDHS$SLwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(SLDHS$v106, SLDHS$SLREC\_summary, SLDHS$SLwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
crosstab(SLTotalnew, SLDHS$v013, SLDHS$SLwt, prop.c = TRUE, missing.include = FALSE, plot = FALSE)  
summary(SLDHS$SLREC\_summary)  
favstats(SLDHS$SLREC\_summary)  
designSL <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~SLwt, data = SLDHS)  
is.numeric(SLDHS$SLREC\_summary)  
designSL <-update(designSL, SLREC\_summary=SLDHS$SLREC\_SMOKE + SLDHS$SLREC\_BMI + SLDHS$SLREC\_AFSI + SLDHS$SLREC\_AFB + SLDHS$SLREC\_Parity + SLDHS$SLyears\_FP + SLDHS$SLREC\_LSP + SLDHS$SLREC\_STI)  
  
  
# Unadjusted RR  
# Age  
SLmodelAge <- svyglm(SLREC\_summary~v013, design= designSL, family = poisson(link = "log"), na.action = na.omit)  
summary(SLmodelAge)  
exp(coef(SLmodelAge)) # calculate the RRs  
round(exp(coef(SLmodelAge)), 2) # round to make it clearer  
cbind(RR = exp(coef(SLmodelAge)), exp(confint(SLmodelAge))) # RR to CI  
round(cbind(RR = exp(coef(SLmodelAge)), exp(confint(SLmodelAge))), 2) # round to 2  
# Residential Status  
SLmodelRS <- svyglm(SLREC\_summary~v025, design= designSL, family = poisson(link = "log"), na.action = na.omit)  
summary(SLmodelRS)  
exp(coef(SLmodelRS)) # calculate the RRs  
round(exp(coef(SLmodelRS)), 2) # round to make it clearer  
cbind(RR = exp(coef(SLmodelRS)), exp(confint(SLmodelRS))) # RR to CI  
round(cbind(RR = exp(coef(SLmodelRS)), exp(confint(SLmodelRS))), 2) # round to 2  
# Marital Status  
SLmodelMS <- svyglm(SLREC\_summary~v502, design= designSL, family = poisson(link = "log"), na.action = na.omit)  
summary(SLmodelMS)  
exp(coef(SLmodelMS)) # calculate the RRs  
round(exp(coef(SLmodelMS)), 2) # round to make it clearer  
cbind(RR = exp(coef(SLmodelMS)), exp(confint(SLmodelMS))) # RR to CI  
round(cbind(RR = exp(coef(SLmodelMS)), exp(confint(SLmodelMS))), 2) # round to 2  
# Education  
SLmodelEdu <- svyglm(SLREC\_summary~v106, design= designSL, family = poisson(link = "log"), na.action = na.omit)  
summary(SLmodelEdu)  
exp(coef(SLmodelEdu)) # calculate the RRs  
round(exp(coef(SLmodelEdu)), 2) # round to make it clearer  
cbind(RR = exp(coef(SLmodelEdu)), exp(confint(SLmodelEdu))) # RR to CI  
round(cbind(RR = exp(coef(SLmodelEdu)), exp(confint(SLmodelEdu))), 2) # round to 2  
# Employment  
SLmodelEmp <- svyglm(SLREC\_summary~v714, design= designSL, family = poisson(link = "log"), na.action = na.omit)  
summary(SLmodelEmp)  
exp(coef(SLmodelEmp)) # calculate the RRs  
round(exp(coef(SLmodelEmp)), 2) # round to make it clearer  
cbind(RR = exp(coef(SLmodelEmp)), exp(confint(SLmodelEmp))) # RR to CI  
round(cbind(RR = exp(coef(SLmodelEmp)), exp(confint(SLmodelEmp))), 2) # round to 2  
# Wealth Index  
SLmodelWI <- svyglm(SLREC\_summary~v190, design= designSL, family = poisson(link = "log"), na.action = na.omit)  
summary(SLmodelWI)  
exp(coef(SLmodelWI)) # calculate the RRs  
round(exp(coef(SLmodelWI)), 2) # round to make it clearer  
cbind(RR = exp(coef(SLmodelWI)), exp(confint(SLmodelWI))) # RR to CI  
round(cbind(RR = exp(coef(SLmodelWI)), exp(confint(SLmodelWI))), 2) # round to 2  
# Adjusted RR  
# Poisson Regression Modelling  
SLmodel <- svyglm(SLREC\_summary~v013 + v025 + v190 + v714 + v106 + v502, design= designSL, family = poisson(link = "log"), na.action = "na.omit") # Observations removed if it contains missing values  
summary(SLmodel)   
# calculate the RRs   
exp(coef(SLmodel))  
# round to make it clearer  
round(exp(coef(SLmodel)), 2)  
# RR to CI  
cbind(RR = exp(coef(SLmodel)), exp(confint(SLmodel)))  
# round to 2  
round(cbind(RR = exp(coef(SLmodel)), exp(confint(SLmodel))), 2)   
  
# Missing Data Included  
# Convert cases with NA into 0 and run a different model  
# library("imputeTS")  
#BMI  
freq(SLDHS$SLREC\_BMI, SLDHS$SLwt, na.rm= TRUE, plot = FALSE)  
SLDHS$SLREC\_BMI2 <- na.replace(SLDHS$SLREC\_BMI, 0) #gtools package  
freq(SLDHS$SLREC\_BMI2, SLDHS$SLwt, na.rm= TRUE, plot = FALSE)  
#AFB  
freq(SLDHS$SLREC\_AFB, SLDHS$SLwt, na.rm= TRUE, plot = FALSE)  
SLDHS$SLREC\_AFB2 <- na.replace(SLDHS$SLREC\_AFB, 0)  
freq(SLDHS$SLREC\_AFB2, SLDHS$SLwt, na.rm= TRUE, plot = FALSE)  
# LSP women  
freq(SLDHS$SLREC\_LSP, SLDHS$SLwt, plot = FALSE)  
SLDHS$SLREC\_LSP2 <- na.replace(SLDHS$SLREC\_LSP, 0)  
freq(SLDHS$SLREC\_LSP2, SLDHS$SLwt, na.rm= TRUE, plot = FALSE)  
# NB- Type of Contraceptive use was linked with month of usage. month of use contains NA i.e. not in the universe.   
# Variables with missing values- BMI (), LSP woman () and LSP men ()  
SLDHS$SLREC\_summary2= as.numeric(SLDHS$SLREC\_SMOKE + SLDHS$SLREC\_BMI2 + SLDHS$SLREC\_AFSI + SLDHS$SLREC\_AFB2 + SLDHS$SLREC\_Parity + SLDHS$SLyears\_FP + SLDHS$SLREC\_LSP2 + SLDHS$SLREC\_STI)  
summary(SLDHS$SLREC\_summary2)  
favstats(SLDHS$SLREC\_summary2)  
designSL <- svydesign(ids = ~v021+v002, strata = ~v025, weights = ~SLwt, data = SLDHS)  
is.numeric(SLDHS$SLREC\_summary2)  
designSL <-update(designSL, SLREC\_summary2=SLDHS$SLREC\_SMOKE + SLDHS$SLREC\_BMI2 + SLDHS$SLREC\_AFSI + SLDHS$SLREC\_AFB2 + SLDHS$SLREC\_Parity + SLDHS$SLyears\_FP + SLDHS$SLREC\_LSP2 + SLDHS$SLREC\_STI)  
  
  
# Unadjusted RR  
# Age  
SLmodelAge2 <- svyglm(SLREC\_summary2~v013, design= designSL, family = poisson(link = "log"), na.action = "na.pass")  
summary(SLmodelAge2)  
exp(coef(SLmodelAge2)) # calculate the RRs  
round(exp(coef(SLmodelAge2)), 2) # round to make it clearer  
cbind(RR = exp(coef(SLmodelAge2)), exp(confint(SLmodelAge2))) # RR to CI  
round(cbind(RR = exp(coef(SLmodelAge2)), exp(confint(SLmodelAge2))), 2) # round to 2  
# Residential Status  
SLmodelRS2 <- svyglm(SLREC\_summary2~v025, design= designSL, family = poisson(link = "log"), na.action = "na.pass")  
summary(SLmodelRS2)  
exp(coef(SLmodelRS2)) # calculate the RRs  
round(exp(coef(SLmodelRS2)), 2) # round to make it clearer  
cbind(RR = exp(coef(SLmodelRS2)), exp(confint(SLmodelRS2))) # RR to CI  
round(cbind(RR = exp(coef(SLmodelRS2)), exp(confint(SLmodelRS2))), 2) # round to 2  
# Marital Status  
SLmodelMS2 <- svyglm(SLREC\_summary2~v502, design= designSL, family = poisson(link = "log"), na.action = "na.pass")  
summary(SLmodelMS2)  
exp(coef(SLmodelMS2)) # calculate the RRs  
round(exp(coef(SLmodelMS2)), 2) # round to make it clearer  
cbind(RR = exp(coef(SLmodelMS2)), exp(confint(SLmodelMS2))) # RR to CI  
round(cbind(RR = exp(coef(SLmodelMS2)), exp(confint(SLmodelMS2))), 2) # round to 2  
# Education  
SLmodelEdu2 <- svyglm(SLREC\_summary2~v106, design= designSL, family = poisson(link = "log"), na.action = "na.pass")  
summary(SLmodelEdu2)  
exp(coef(SLmodelEdu2)) # calculate the RRs  
round(exp(coef(SLmodelEdu2)), 2) # round to make it clearer  
cbind(RR = exp(coef(SLmodelEdu2)), exp(confint(SLmodelEdu2))) # RR to CI  
round(cbind(RR = exp(coef(SLmodelEdu2)), exp(confint(SLmodelEdu2))), 2) # round to 2  
# Employment  
SLmodelEmp2 <- svyglm(SLREC\_summary2~v714, design= designSL, family = poisson(link = "log"), na.action = "na.pass")  
summary(SLmodelEmp2)  
exp(coef(SLmodelEmp2)) # calculate the RRs  
round(exp(coef(SLmodelEmp2)), 2) # round to make it clearer  
cbind(RR = exp(coef(SLmodelEmp2)), exp(confint(SLmodelEmp2))) # RR to CI  
round(cbind(RR = exp(coef(SLmodelEmp2)), exp(confint(SLmodelEmp2))), 2) # round to 2  
# Wealth Index  
SLmodelWI2 <- svyglm(SLREC\_summary2~v190, design= designSL, family = poisson(link = "log"), na.action = "na.pass")  
summary(SLmodelWI2)  
exp(coef(SLmodelWI2)) # calculate the RRs  
round(exp(coef(SLmodelWI2)), 2) # round to make it clearer  
cbind(RR = exp(coef(SLmodelWI2)), exp(confint(SLmodelWI2))) # RR to CI  
round(cbind(RR = exp(coef(SLmodelWI2)), exp(confint(SLmodelWI2))), 2) # round to 2  
  
# Adjusted RR  
# Keep NA  
SLmodel2 <- svyglm(SLREC\_summary2~v013 + v025 + v190 + v714 + v106 + v502, design= designSL, family = poisson(link = "log"), na.action = "na.pass") # keep all data, including NAs  
summary(SLmodel2)  
# calculate the RRs   
exp(coef(SLmodel2))  
# round to make it clearer  
round(exp(coef(SLmodel2)), 2)  
# RR to CI  
cbind(RR = exp(coef(SLmodel2)), exp(confint(SLmodel2)))  
# round to 2  
round(cbind(RR = exp(coef(SLmodel2)), exp(confint(SLmodel2))), 2)  
  
# Better Model  
# Comparing two models, AIC or BIC  
AIC(SLmodel, SLmodel2)  
BIC(SLmodel, SLmodel2, maximal = SLmodel2)  
plot\_summs(SLmodel, scale = TRUE, exp = TRUE)  
plot\_summs(SLmodel2, scale = TRUE, exp = TRUE)  
plot\_summs(SLmodel, SLmodel2, ci\_level = 0.95, model.names = c("Model 1"= "Model Excludes Missing", "Model 2"= "Model Includes Missing"),   
 colors = 'CUD Bright', exp = TRUE, legend.title = "Graph Comparing Models" )  
# Graph of Effect Size BJModel (Missing Data Excluded Model)  
SLeff.pres <- allEffects(SLmodel, scale = T)  
plot(SLeff.pres, ask=F)  
SLeff.pres2 <- allEffects(SLmodel2, scale = T)  
plot(SLeff.pres2, ask=F)

# Proportion of cervical cancer risk factors among women 15-49 years in West Africa  
Prop\_Risk <- c(68.8, 73.0, 63.9, 76.6, 84.1, 77.1, 68.7, 80.5, # Age at First Sexual Intercourse <= 17 years  
 52.4, 59.5, 52.5, 62.2, 67.2, 67.1, 56.2, 62.0, # Age at First Birth <= 19 years  
 48.7, 42.3, 40.9, 44.9, 42.9, 54.1, 48.5, 43.4, # Parity > 2 years  
 26.0, 37.3, 36.1, 26.4, 37.0, 27.7, 28.1, 28.1, # Overweight or Obese  
 3.0, 3.9, 1.9, 10.7, 30.9, 12.3, 6.5, 10.7, # STI in last 12 Months  
 2.8, 11.4, 0.5, 2.5, 16.2, 1.1, 2.7, 4.3, # Total number of sex partners >5  
 2.7, 0.4, 0.7, 1.4, 1.1, 1.1, 0.5, 4.6, # Smoking  
 0.5, 0.6, 0.8, 0.4, 1.8, 0.9, 0.8, 1.7) # Hormonal Contraceptive usage >= 5 years  
  
  
PropCountry\_Risk <- matrix(Prop\_Risk, nrow = 8, byrow = T)  
colnames(PropCountry\_Risk) <- c("Benin", "Cameroon", "The Gambia", "Guinea", "Liberia", "Mali", "Nigeria", "Sierra Leone")  
row.names(PropCountry\_Risk) <- c("Onset of Sexual Intercourse 17 Years or Less", "Age at First Birth 19 Years or Less", "More Than Two Children",  
 "Overweight or Obese", "STI in last 12 Months", "Lifetime Sex Partners Five or More", "Smokes", "Hormonal Contraceptives for 5 Years or More")  
PropCountry\_Risk  
color.names = c("red2", "green3", "slateblue4", "yellow2", "olivedrab2", "orange", "grey90", "black")  
barplot(PropCountry\_Risk, beside = T, xlab = "Country", ylab = "Percentage (%)", col = color.names)  
legend(1, 2300, row.names(PropCountry\_Risk), cex = 0.7, fill = color.names, bty = "n")  
  
# ggplot2 Grouped Bar Plot  
PropCountry\_Riskmelt <- melt(PropCountry\_Risk)  
head(PropCountry\_Riskmelt)  
colnames(PropCountry\_Riskmelt) <- c( "Risk", "Country", "Percentage")  
head(PropCountry\_Riskmelt)  
ggplot(PropCountry\_Riskmelt, aes(x= Country, y= Percentage, fill= Risk))  
geom\_bar(stat = "identity", position = "dodge")  
scale\_fill\_brewer(palette = "Set1")  
labs(y= "Percentage")  
theme\_bw()  
theme(panel.grid.major.x = element\_blank())  
q <- ggplot(PropCountry\_Riskmelt, aes(x= Country, y= Percentage, fill= Risk)) +  
 geom\_bar(stat = "identity", position = "dodge") +  
 scale\_fill\_brewer(palette = "Set1") +  
 labs(y= "Percentage (%)") +  
 theme\_bw() +  
 theme(panel.grid.major.x = element\_blank())  
  
q + theme(legend.position="bottom")

# Proportion of the total number of risk factors among women 15-49 years in West Africa  
Total.Riskn <- c(8.0, 5.3, 10.9, 4.2, 2.0, 4.9, 7.5, 4.4, # zero   
 20.7, 14.2, 20.9, 14.0, 8.7, 11.1, 18.0, 14.0, # One  
 29.5, 28.3, 25.5, 29.4, 23.8, 27.0, 27.8, 28.0, #Two  
 29.5, 33.2, 25.1, 35.0, 29.4, 36.7, 33.3, 33.9, # Three  
 11.6, 16.3, 16.6, 14.8, 25.4, 17.3, 12.2, 16.5, # Four  
 0.7, 2.6, 1.0, 2.7, 10.6, 3.0, 1.2, 3.2) # Greater than Five  
  
# Greater than Five number of risk factors (Basic barplot)  
Country\_Riskn <- matrix(Total.Riskn, nrow = 6, byrow = T) # 5  
colnames(Country\_Riskn) <- c("Benin", "Cameroon", "The Gambia", "Guinea", "Liberia", "Mali", "Nigeria", "Sierra Leone")  
row.names(Country\_Riskn) <- c( "Zero", "One", "Two", "Three", "Four", "Greater than Five")  
Country\_Riskn  
color.names = c("red2", "green3", "slateblue4", "yellow2", "olivedrab2", "orange")  
barplot(Country\_Riskn, beside= T, xlab= "Country", ylab= "Percentage", col= color.names)  
  
# ggplot2 Grouped Bar Plot (reshape Package)  
Country\_Riskmeltn <- melt(Country\_Riskn)  
head(Country\_Riskmeltn)  
colnames(Country\_Riskmeltn) <- c("Number", "Country", "Percentage")  
head(Country\_Riskmeltn)  
ggplot(Country\_Riskmeltn, aes(x= Country, y= Percentage, fill= Number))  
geom\_bar(stat = "identity", position = "dodge")  
scale\_fill\_brewer(palette = "Dark2")  
labs(y= "Percentage (%)")  
theme\_bw()  
theme(panel.grid.major.x = element\_blank())  
w <- ggplot(Country\_Riskmeltn, aes(x= Country, y= Percentage, fill= Number)) +  
 geom\_bar(stat = "identity", position = "dodge") +  
 scale\_fill\_brewer(palette = "Dark2") +   
 labs(y= "Percentage (%)") +  
 theme\_bw() +  
 theme(panel.grid.major.x = element\_blank())  
  
w + theme(legend.position="bottom")

# Mean total number of risk factors  
# Individual country mean and standard deviation  
favstats(BJDHS$BJREC\_summary) # Benin  
favstats(CMAGE49$CMREC\_summary) # Cameroon  
favstats(GMDHS$GMREC\_summary) # The Gambia  
favstats(GNDHS$GNREC\_summary) # Guinea  
favstats(LBDHS$LBREC\_summary) # Liberia  
favstats(MLDHS$MLREC\_summary) # Mali  
favstats(NGDHS$NGREC\_summary) # Nigeria  
favstats(SLDHS$SLREC\_summary) # Sierra Leone  
  
df <- list(BJDHS$BJREC\_summary, CMAGE49$CMREC\_summary, GMDHS$GMREC\_summary, GNDHS$GNREC\_summary,  
 LBDHS$LBREC\_summary, MLDHS$MLREC\_summary, NGDHS$NGREC\_summary, SLDHS$SLREC\_summary)  
# Creating variable  
a=BJDHS$BJREC\_summary  
b=CMAGE49$CMREC\_summary  
c=GMDHS$GMREC\_summary  
d=GNDHS$GNREC\_summary  
e=LBDHS$LBREC\_summary  
f=MLDHS$MLREC\_summary  
g=NGDHS$NGREC\_summary  
h=SLDHS$SLREC\_summary  
# Finding maximum length  
max\_ln <- max(c(length(a), length(b), length(c),length(d), length(e), length(f), length(g), length(h)))  
DHScomb <- data.frame(Benin = c(a,rep(NA, max\_ln - length(a))),  
 Cameroon = c(b,rep(NA, max\_ln - length(b))),  
 Gambia = c(c,rep(NA, max\_ln - length(c))),  
 Guinea = c(d,rep(NA, max\_ln - length(d))),  
 Liberia = c(e,rep(NA, max\_ln - length(e))),  
 Mali = c(f,rep(NA, max\_ln - length(f))),  
 Nigeria = c(g,rep(NA, max\_ln - length(g))),  
 SierraLeone = c(h,rep(NA, max\_ln - length(h))))  
DHScomb  
head(DHScomb)  
tail(DHScomb)  
colnames(DHScomb) = c("Benin","Cameroon","The Gambia","Guinea","Liberia","Mali","Nigeria","Sierra Leone")  
DHScomb\_long <- melt(DHScomb)  
DHScomb\_long  
head(DHScomb\_long)  
p <- ggplot(DHScomb\_long, aes(x=variable, y=value, fill=variable)) +   
 geom\_boxplot(outlier.colour="black", outlier.shape=16, outlier.size=2)+  
 labs(x="Country", y = "Number of Risk Factors") +  
 scale\_fill\_manual(values=c("#999999", "#E69F00", "#56B4E9", "#009E73", "#F0E442", "#0072B2", "#D55E00", "#CC79A7"),   
 name="Mean (Standard Deviation)",   
 breaks=c("Benin", "Cameroon", "The Gambia", "Guinea", "Liberia", "Mali", "Nigeria", "Sierra Leone"),  
 labels=c("2.19 (1.14)", "2.53 (1.19)", "2.24 (1.23)", "2.51 (1.12)", "3.05 (1.19)",  
 "2.60 (1.14)", "2.29 (1.16)", "2.54 (1.14)"))   
p   
p + theme(legend.position="bottom")

sessionInfo()

## R version 4.1.1 (2021-08-10)  
## Platform: x86\_64-w64-mingw32/x64 (64-bit)  
## Running under: Windows 10 x64 (build 19045)  
##   
## Matrix products: default  
##   
## locale:  
## [1] LC\_COLLATE=English\_Nigeria.1252 LC\_CTYPE=English\_Nigeria.1252   
## [3] LC\_MONETARY=English\_Nigeria.1252 LC\_NUMERIC=C   
## [5] LC\_TIME=English\_Nigeria.1252   
##   
## attached base packages:  
## [1] stats graphics grDevices utils datasets methods base   
##   
## loaded via a namespace (and not attached):  
## [1] compiler\_4.1.1 fastmap\_1.1.0 cli\_3.4.1 tools\_4.1.1   
## [5] htmltools\_0.5.2 rstudioapi\_0.15.0 yaml\_2.3.7 rmarkdown\_2.25   
## [9] knitr\_1.44 xfun\_0.40 digest\_0.6.28 rlang\_1.1.1   
## [13] evaluate\_0.22