**Student Name: Chan Tsz Kwan, ID: 56346872**

**BMS2901 Project: R code Progress**

library(dplyr)

downloadNHANES <- function(fileprefix){

print (fileprefix)

outdf <- data.frame(NULL)

urlstring <- paste('https://wwwn.cdc.gov/Nchs/Nhanes/',yrs,'/',fileprefix,letters,'.XPT', sep='')

download.file(urlstring, tf <- tempfile(), mode="wb")

tmpframe <- foreign::read.xport(tf)

outdf <- bind\_rows(outdf, tmpframe)

return(outdf)}

yrs <- c('2015-2016')

letters <- c('\_I')

library(nhanesA)

**# Sleep Disorders**

sd <- downloadNHANES('SLQ')

sd.vars <- c("SEQN","SLQ030","SLQ040")

sd.data <- nhanesTranslate('SLQ\_I', sd.vars, data=sd[,sd.vars])

names(sd.data) <- c("id", "snore.freq", "snort.stopbreath.freq")

**# Kidney condition-urology**

kd <- downloadNHANES('KIQ\_U')

kd.vars <- c("SEQN", "KIQ480")

kd.data <- nhanesTranslate('KIQ\_U\_I', kd.vars, data=kd[,kd.vars])

names(kd.data) <- c("id", "night.urinate.freq")

**# merge by id**

merged.data <- merge(kd.data, sd.data, by = c("id"), all=TRUE)

dim(merged.data)

save(merged.data, file="NHANES78.RData")

**#Recode the variables values**

int.count<-sapply(merged.data,function(x) is.integer(x))

int.count[int.count==TRUE]

**#Check variables individually**

summary(merged.data$snore.freq)

*Never Rarely - 1-2 nights a week Occasionally - 3-4 nights a week*

*1820 1475 1067*

*Frequently - 5 or more nights a Refused Don't know*

*1513 2 450*

*summary(merged.data$snort.stopbreath.freq)*

*Never Rarely - 1-2 nights a week Occasionally - 3-4 nights a week*

*4535 773 351*

*Frequently - 5 or more nights a Refused Don't know*

286 2 380

summary(merged.data$night.urinate.freq)

*0 1 2 3 4 5 or more Refused Don't know NA's*

*1407 1850 928 486 165 108 2 5 1376*

**#Exclude the cases with missing values**

invalid.exclude <- function(x, exclude=c("Refused", "Don't know", "Never")) {

x <- as.character(x)

x[x %in% exclude] <- NA

return(x)}

require(data.table)

DT2 <- data.table(merged.data)

table(DT2$snore.freq)

DT2$snore.freq <- invalid.exclude(DT2$snore.freq)

table(DT2$snore.freq)

*Frequently - 5 or more nights a Occasionally - 3-4 nights a week Rarely - 1-2 nights a week*

*1513 1067 1475*

table(DT2$snort.stopbreath.freq)

DT2$snort.stopbreath.freq <- invalid.exclude(DT2$snort.stopbreath.freq)

table(DT2$snort.stopbreath.freq)

*Frequently - 5 or more nights a Occasionally - 3-4 nights a week Rarely - 1-2 nights a week*

*286 351 773*

table(DT2$night.urinate.freq)

DT2$night.urinate.freq <- invalid.exclude(DT2$night.urinate.freq)

table(DT2$night.urinate.freq)

*0 1 2 3 4 5 or more*

*1407 1850 928 486 165 108*

**#Omit incomplete NA cases**

DT3 <- na.omit(DT2, cols=c("snore.freq","snort.stopbreath.freq","night.urinate.freq"))

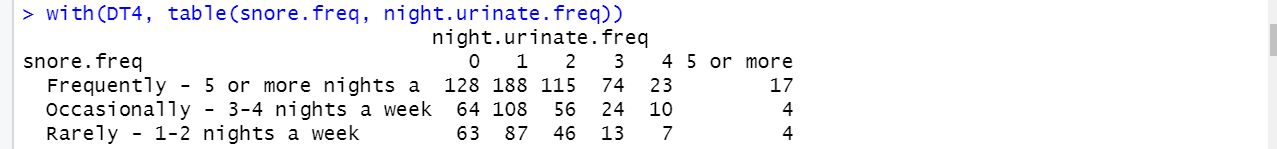
dim(DT3)

**#Exclude cases where outcome happened before exposure**

DT4 <- DT3[(DT3$snore.freq <= DT3$snort.stopbreath.freq)]

dim(DT4)

with(DT4, table(snore.freq, night.urinate.freq))



with(DT4, table(snort.stopbreath.freq, night.urinate.freq))

Text

Description automatically generated with low confidence

**#Recoding covariates**

table(DT4$night.urinate.freq)

##

DT4$night.urinate.freq <- car:::recode(DT4$night.urinate.freq,

"'0'='Never';

'1'='Rarely';

c('2', '3')= 'Ocassionally';

c('4', '5 or more')='Frequently';

else=NA")

table(DT4$night.urinate.freq)

DT4$night.urinate.freq <- invalid.exclude(DT4$night.urinate.freq)

table(DT4$night.urinate.freq)

**#Determine what to do with missing values**

DT6 <- DT4[,c("snore.freq", "snort.stopbreath.freq", "night.urinate.freq")]

sort(sapply(DT6,function(x) sum(is.na (x))), decreasing = TRUE)

require(DataExplorer)

plot\_missing(DT6)

DT7 <- na.omit(DT6)

dim(DT7)

**#Bivariate Analysis**

require(tableone)

tab1 <- CreateTableOne(data = DT7, includeNA = TRUE,

strata = "night.urinate.freq", test = FALSE,

var = c("snore.freq", "snort.stopbreath.freq"))

print(tab1,showAllLevels = FALSE)

Graphical user interface, text, application

Description automatically generated

table(DT7$snore.freq)

*Frequently - 5 or more nights a Occasionally - 3-4 nights a week Rarely - 1-2 nights a week*

*417 202 157*

table(DT7$night.urinate.freq)

*Frequently Ocassionally Rarely*

*65 328 383*

with(DT7, table(snore.freq, night.urinate.freq))

*night.urinate.freq*

*snore.freq Frequently Ocassionally Rarely*

*Frequently - 5 or more nights a 40 189 188*

*Occasionally - 3-4 nights a week 14 80 108*

*Rarely - 1-2 nights a week 11 59 87*

with(DT7, table(snort.stopbreath.freq, night.urinate.freq))

*night.urinate.freq*

*snort.stopbreath.freq Frequently Ocassionally Rarely*

*Frequently - 5 or more nights a 17 78 74*

*Occasionally - 3-4 nights a week 19 84 85*

*Rarely - 1-2 nights a week 29 166 224*

**#Stratified by exposure**

tab3 <- CreateTableOne(data = DT7, includeNA = TRUE,

strata = "night.urinate.freq", test = FALSE,

var = c("snore.freq", "snort.stopbreath.freq"))

print(tab3,showAllLevels = FALSE, smd = TRUE)

Text

Description automatically generated

cross.tab <- with(DT7, table(snort.stopbreath.freq, night.urinate.freq))

cross.tab

Text

Description automatically generated

save(DT6, DT7, file="analyticNHANES78.RData")

dim(DT7)

**#Summarize data**

load("analyticNHANES78.RData")

analytic.data <- DT7

require(rms)

describe(analytic.data)

analytic.data

Table

Description automatically generated

Table

Description automatically generated

**#Table One-p-value**

require(tableone)

tab1 <- CreateTableOne(data = analytic.data, includeNA = TRUE,

strata = "night.urinate.freq", test = FALSE,

var = c("snore.freq", "snort.stopbreath.freq"))

print(tab1,showAllLevels = FALSE, smd = TRUE)

Text, application

Description automatically generated

**#Fit a crude regression**

library(jtools)

fit1 <- glm(I(night.urinate.freq=="Frequently")~snort.stopbreath.freq,

data = analytic.data)

require(jtools)

summ(fit1, conf = TRUE, digits = 2, exp = TRUE)

A picture containing table

Description automatically generated

**#Fit an adjusted regression**

fit2 <- glm(I(night.urinate.freq=="Frequently")~snore.freq+snort.stopbreath.freq

, data = analytic.data)

require(ggstance)

summ(fit2, conf = TRUE, digits = 2, vif = TRUE, exp = TRUE)

plot\_summs(fit2, exp = TRUE)

Table

Description automatically generated