Graduate Research Preprocessing

Sangeeta Mondal

November 2017

```
require(MASS)
require(dplyr)
require(reshape2)
require(formatR)
```

```
sc_new<-read.csv('C:/Users/PC/Documents/Bath GSR/sc_revised_new.csv', sep = ',', header = TRUE, stringsAsFactors
 = FALSE)
#drop all cases that started in 2017
sc_new<-sc_new[!sc_new$id %in% c('MJ23815', 'NJ28668', 'NJ29047', 'NJ29092', 'PJ52282', 'TJ22529', 'TJ22594', 'VJ
45650', 'YJ38874', 'YJ38882', 'yj38884'), ]
```

```
#changing MH vars with 1/2 indicator to 0/1 indicator as needed for future rowSum calculations
mh_vars<-c('suicide', 'angerco', 'instimhc', 'othermht', 'mhco', 'famco', 'mhc', 'mhc_new')
t_med2<-c("depressionmed","anxietymed", "admed", "pdmed", "bipolarmed", "mdmed", "adhdmed", "cdmed", "simed", "mh
cothermed", "disruptdisordermed", "fasmed", "genderdysmed", "intellectualdismed", "impulsmed", "psychoticdismed",
 "reactiveattachmed", "sleepprobmed", "traumaticdismed")
cc_t1<-grep('tmp.*_past|tmp.*now', names(sc_new), value = TRUE)

sc_new[4201:4417]<-lapply(sc_new[c(mh_vars, t_med2, cc_t1)], function(x) {    y<-ifelse(x == 1, 1, 0) })
names(sc_new)[4201:4417]<-lapply(c(mh_vars,t_med2, cc_t1), function(x) {    y<-paste0(x, "_2") })

#table(sc_new$suicide,sc_new$suicide2)- make sure conversions correct
```

```
#convert date vars to date formats
sdates<-grep('date', names(sc_new), value = TRUE)
sc_new[,dates]<-lapply(sc_new[,dates], as.Date, format = '%Y-%m-%d')
```

```
#calculate bench warrant durations
time_d<-function(x,y){as.numeric(x-y, units = 'days')}

btim<-mapply(time_d,x= sc_new[,grep('benchrelease',names(sc_new), value = TRUE)],y = sc_new[,grep('benchissue', n
ames(sc_new), value = TRUE)])

colnames(btim)<-paste0('benchtime',seq(1:6))
sc_new<-cbind(sc_new, btim)
```