######################################~15 mins

#functions that begin with "fn." are custom r functions

##########################

#numerators section;

# creates new R dataframes with "numerators" of per share columns, multiplying per-share data by appropriate share count (primary shares or fully-diluted shares)

#a newer and maybe better alternative to R dataframes are R datatables  
#if i was starting over from scratch in R, I might use datatables rather than dataframes  
#I am unsure what is best to use in Python, datatable or dataframe or something else

#ch <- odbcConnect("xpressfeed", uid = "xpressfeed", pwd = "xpressfeed") #old code used to access SQL

setwd(r\_ri.num) #sets the directory on my computer that data is accessed from and saved to

######################

rm(ch) #removes the ch “vector”# this code is probably not needed

#inputs

mypers.i = c(-1:0) #pers means “periods”; in this case,a period is a year

mypers <- ifelse(mypers.i >=0, mypers.i, paste('m', abs(mypers.i), sep='') ) #"m" mean "minus"; m1 means "minus1" or "-1"

#get fully diluted per share and primary pershare colnums using mydata\_af.0

fn.flga('mydata\_af.0')  
fn.flga('myshsinfo') #this is an R list

#loads data files using the custom R-function “fn.flga”

#fn.flga('mycols\_fd') #columns with "fully diluted" per share data

#fn.flga('mycols\_pr') #columns with "primary" per share data

#fn.flga('myunsure\_ps') #columns with per share data where we are uncertain if the data is "primary AKA primary" or "fully diluted" per share data

mycols\_fd <- match(myshsinfo$co\_afnd\_fd$columnname,colnames(mydata\_af.0) )  
mycols\_pr <- match (c(myshsinfo$co\_afnd\_pr$columnname,myshsinfo$co\_afnd\_unsure$columnname) , colnames(mydata\_af.0) )

#Create df with total assets in non share cols, assets per primary share in per primary share cols, and assets per fully diluted share in fully diluted share cols?

#Which cols are per primary share cols? mycols\_pr

#Which cols are per fully diluted share cols? mycols\_fd

#Which cols are not share cols? c(-mycol\_fdshs, -mycol\_prshs)

coltype <- 1:ncol(mydata\_af.0)

coltype [mycols\_pr] <- 'p' #primary columns

coltype [mycols\_fd] <- 'f' #fully diluted columns

coltype [c(-mycols\_fd, -mycols\_pr)] <- 't' #t means "total", rather than per-share

dim(mydata\_af.0)

#checks

summary(is.na(coltype))

summary(coltype=='p'|coltype=='f'|coltype=='t')

summary(coltype=='p')

summary(coltype=='f')

summary(coltype=='t')

fn.asr('coltype') #use this later for calculating denominator of roa df's (that is, a df with: assets per primary share, assets per fully diluted share, or total assets)

#cols for fully diluted shares

mycol\_fdshs <- match( c('cshfd','cshpri'), colnames(mydata\_af.0) ) # match( c('cshfd','cshpri', 'csho'), colnames(mydata\_af.0) )

mycol\_prshs <- match( c('cshpri'), colnames(mydata\_af.0) ) #match( c('cshpri', 'csho'), colnames(mydata\_af.0) )

#df of fully diluted numerators (fully diluted per share data \* coalesce(cshfd, cshpri, csho) for each period i

for(p in 1: length(mypers)) {

i <- mypers[p]

tempfilename <- paste( "mydata\_af.", i, sep='')

my\_co\_afnd\_df <- fn.flga(tempfilename) #load file

#df of fd shares to multiply fd data by

shsfd <- coalesce( my\_co\_afnd\_df[mycol\_fdshs[1] ][,] , my\_co\_afnd\_df[mycol\_fdshs[2]][,])

shsfd\_df <- data.frame(matrix(data=shsfd, nrow = nrow(my\_co\_afnd\_df) , ncol = length(mycols\_fd) ))

#head(shsfd\_df)

#df of fd numerators (fd per share data \* shsfd\_df)

nums\_fd <- (my\_co\_afnd\_df [mycols\_fd] \* (1000000\*shsfd\_df))/1000000

colnames(nums\_fd) <- paste(colnames(mydata\_af.0)[mycols\_fd], '\_nfd', sep='')

########################

#df of primary shares to multiply primary data by

shspr <- coalesce( my\_co\_afnd\_df[mycol\_fdshs[1] ][,])

shspr\_df <- data.frame(matrix(data=shspr, nrow = nrow(my\_co\_afnd\_df) , ncol = length(mycols\_pr) ))

#df of primary numerators (primary per share data \* shspr\_df)

nums\_pr <- (my\_co\_afnd\_df [mycols\_pr] \* (1000000\*shspr\_df))/1000000

colnames(nums\_pr) <- paste(colnames(mydata\_af.0)[mycols\_pr], '\_npr', sep='') #add colname with numerator & fd or primary info

#head(my\_co\_afnd\_df [mycols\_pr])

#head(shspr\_df)

#head(nums\_pr)

#cbind df of numerator dfs above #maybe cbind df to co\_afnd df

my\_co\_afnd\_nums\_df <- cbind (my\_co\_afnd\_df,nums\_fd ,nums\_pr)

tempfilename <- paste( "mydata\_af\_nums.", i, sep='')

fn.asr(tempfilename , my\_co\_afnd\_nums\_df)

#rm(mydata\_af1, mydata\_af2, mydata\_af,iname, temp, tempfilename) ; gc();gc()

} #end i

p

i

###end numerators section