run\_analysis.R

LouisG

# create training tables from training files  
tbl\_xtrain <- read.table("x\_train.txt")  
tbl\_ytrain <- read.table("y\_train.txt")  
tbl\_subject\_train <- read.table("subject\_train.txt")  
colnames(tbl\_subject\_train) <- ("SubjectID")  
  
# create test tables from test files   
tbl\_xtest <- read.table("x\_test.txt")  
tbl\_ytest <- read.table("y\_test.txt")  
tbl\_subject\_test <- read.table("subject\_test.txt")  
colnames(tbl\_subject\_test) <- ("SubjectID")  
  
# create features tables and assign column names  
tbl\_features <- read.table("features.txt")  
colnames(tbl\_features) <- c("FeatureID", "Feature")  
  
# create activity Lables and assign column names   
tbl\_activity\_labels <- read.table("activity\_labels.txt")  
colnames(tbl\_activity\_labels) <- c("ActivityID", "Activity")  
  
  
# create table of rows of mean and std only.   
required\_features <- grep( ".\*mean.\*|.\*std.\*",tbl\_features[,2])  
  
# Assign descriptive variable names for training data  
colnames(tbl\_xtrain) <- tbl\_features[,2]  
colnames(tbl\_ytrain) <- "ActivityID"  
#tbl\_ytrain <- cbind(tbl\_ytrain, "Activity")  
  
# add descriptive activity to training - prior to merge  
tbl\_ytrain$Activity <- tbl\_activity\_labels$Activity[match(tbl\_ytrain$ActivityID,tbl\_activity\_labels$ActivityID)]  
  
# Assign descriptive variable names for test data  
colnames(tbl\_xtest) <- tbl\_features[,2]  
colnames(tbl\_ytest) <- "ActivityID"  
#tbl\_ytest <- cbind(tbl\_ytest, "Activity")  
  
# add descriptive activity to test- prior to merge  
tbl\_ytest$Activity <- tbl\_activity\_labels$Activity[match(tbl\_ytest$ActivityID,tbl\_activity\_labels$ActivityID)]  
  
  
# merge test and traing data  
tbl\_xtrain\_xtest <- rbind(tbl\_xtrain,tbl\_xtest)  
tbl\_ytrain\_ytest <- rbind(tbl\_ytrain,tbl\_ytest)  
  
# Select ony required measurement (mean and std)  
tbl\_xtrain\_xtest <- tbl\_xtrain\_xtest[,required\_features]  
  
# Complete merge process  
tbl\_subject <- rbind(tbl\_subject\_train,tbl\_subject\_test)  
tbl\_allReqData <- cbind(tbl\_subject, tbl\_ytrain\_ytest, tbl\_xtrain\_xtest)  
   
# creates a second, independent tidy data set with the average of each variable   
tbl\_TidyData <- aggregate(. ~SubjectID + Activity,tbl\_allReqData, mean)  
tbl\_TidyData <- tbl\_TidyData[order(tbl\_TidyData$SubjectID, tbl\_TidyData$Activity),]  
  
write.table(tbl\_TidyData, "tidyData.txt")