run\_analysis.R

gog

2021-02-20

#Loading of the required packages  
library(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.3 v purrr 0.3.4  
## v tibble 3.0.6 v dplyr 1.0.4  
## v tidyr 1.1.2 v stringr 1.4.0  
## v readr 1.4.0 v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

library(dplyr)  
library(readr)  
library(reshape2)

##   
## Attaching package: 'reshape2'

## The following object is masked from 'package:tidyr':  
##   
## smiths

# Cleaning up  
rm(list=ls())  
 # Establish paths for reading the files that corresponds to the project.  
setwd("~/R/COURSERA/Getting and cleaning data/getdata\_projectfiles\_UCI HAR Dataset/UCI HAR Dataset")  
 # contents of the working directory  
(dir1<-dir())

## [1] "activity\_labels.txt" "features.txt" "features\_info.txt"   
## [4] "ReadMe-Gabriel.html" "ReadMe Gabriel.Rmd" "README.txt"   
## [7] "run\_analysis.R" "run\_analysis.spin.R" "run\_analysis.spin.Rmd"  
## [10] "run\_analysis\_1.R" "test" "tidy\_means"   
## [13] "tidy\_means.txt" "trabajo.RData" "train"

# Getting the activity labels and variable names that are in the main directory  
act\_labels<-read.table(dir1[1])  
colnames(act\_labels)[2]<-"activity"  
features<-read.table(dir1[2])  
path\_test<-"C:/Users/gog/OneDrive/Documentos/R/COURSERA/Getting and cleaning data/getdata\_projectfiles\_UCI HAR Dataset/UCI HAR Dataset/test"  
path\_train<-"C:/Users/gog/OneDrive/Documentos/R/COURSERA/Getting and cleaning data/getdata\_projectfiles\_UCI HAR Dataset/UCI HAR Dataset/train"  
 # Reading test files  
setwd(path\_test)  
test\_files<-dir(path\_test)  
dir\_test<-dir(test\_files)  
sub\_test<-read.table(test\_files[2],quote = "\"",comment.char = "")  
colnames(sub\_test)<-"id"  
x\_test<-as\_tibble(read.table(test\_files[3],quote = "\"",comment.char = ""))  
y\_test<-as\_tibble(read.table(test\_files[4],quote = "\"",comment.char = ""))  
 # Code to assign features and Activity Labels to the test and train files  
colnames(x\_test)<-features[,2]  
colnames(y\_test)<-"tlabels"  
 # reading train files  
setwd(path\_train)  
train\_files<-dir(path\_train)  
dir\_train<-dir(train\_files)  
sub\_train<-as\_tibble(read.table(train\_files[2], quote = "\"",comment.char = ""))  
colnames(sub\_train)<-"id"  
x\_train<-as\_tibble(read.table(train\_files[3],quote = "\"",comment.char = ""))  
colnames(x\_train)<-features[,2]  
y\_train<-as\_tibble(read.table(train\_files[4],quote = "\"",comment.char = ""))  
colnames(y\_train)<-"tlabels"  
path\_data\_test<-"~/R/COURSERA/Getting and cleaning data/getdata\_projectfiles\_UCI HAR Dataset/UCI HAR Dataset/test/Inertial Signals/"  
path\_data\_train<-"~/R/COURSERA/Getting and cleaning data/getdata\_projectfiles\_UCI HAR Dataset/UCI HAR Dataset/train/Inertial Signals/"  
   
 # joining activity names corresponding to the activity file labels (act\_labels)  
  
y\_test<-merge(y\_test, act\_labels,by.x="tlabels",by.y="V1",sort=FALSE)  
y\_train<-merge(y\_train, act\_labels,by.x="tlabels",by.y="V1",sort=FALSE)  
x\_test<-cbind.data.frame(sub\_test,y\_test,x\_test)  
x\_train<-cbind.data.frame(sub\_train,y\_train,x\_train)  
  
 # merge the x\_train and x\_test sets into one "data\_base"  
data\_base<-as\_tibble(rbind.data.frame(x\_train,x\_test),.name\_repair = make.names)  
  
#Extraction of the measurements of the mean and standard deviation for each measurement  
  
mean\_std<-as\_tibble(select(data\_base,id,tlabels,activity,contains(c("mean","std"))),.name\_repair = make.names)  
  
 # Independent tidy data set with the average of each variable for each activity and each subject.  
 # data\_of\_means<-select(mean\_std,id,tlabels,activity,contains("mean"))  
  
means\_by\_id<-aggregate.data.frame(mean\_std[,4:89],mean\_std[,1],mean)  
means\_by\_activity<-aggregate.data.frame(mean\_std[,4:89],mean\_std[,c(2,3)],mean)  
tidy\_data\_means<-merge.data.frame(means\_by\_id,means\_by\_activity,all=TRUE)  
  
# WRITTING THE DATA FILE FOR STEP 5 OF THE SUBMISSION.  
  
setwd("~/R/COURSERA/Getting and cleaning data/getdata\_projectfiles\_UCI HAR Dataset/UCI HAR Dataset")  
write.table(tidy\_data\_means,file="tidy\_means.txt",row.name=FALSE)