**Getting and Cleaning Data Course Project**

The data is collected from the accelerometers from the Samsung Galaxy S smartphone.

Step 1. The data set was downloaded.

MyFile <- "getdata\_projectfiles\_UCI HAR Dataset.zip"

if (!file.exists(MyFile)){

fileURL <- "https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip"

download.file(fileURL, MyFile, method="curl")

}

if (!file.exists("UCI HAR Dataset")) {

unzip(MyFile)

}

Step 2. Reading the data

We read all the .txt files in the folder, and store them in different variables.

Data\_features <- read.table("UCI HAR Dataset/features.txt", col.names = c("n","functions"))

Data\_activities <- read.table("UCI HAR Dataset/activity\_labels.txt", col.names = c("code", "activity"))

Data\_subject\_test <- read.table("UCI HAR Dataset/test/subject\_test.txt", col.names = "subject")

Data\_x\_test <- read.table("UCI HAR Dataset/test/X\_test.txt", col.names = features$functions)

Data\_y\_test <- read.table("UCI HAR Dataset/test/y\_test.txt", col.names = "code")

Data\_subject\_train <- read.table("UCI HAR Dataset/train/subject\_train.txt", col.names = "subject")

Data\_x\_train <- read.table("UCI HAR Dataset/train/X\_train.txt", col.names = features$functions)

Data\_y\_train <- read.table("UCI HAR Dataset/train/y\_train.txt", col.names = "code")

Step 3. Merge the two data sets (training and test)

X <- rbind(Data\_x\_train, Data\_x\_test)

Y <- rbind(Data\_y\_train, Data\_y\_test)

Subject <- rbind(Data\_subject\_train, Data\_subject\_test)

Data\_Merge <- cbind(Data\_Subject, Y, X)

Step 4. Extract only the measurements on the mean and standard deviation for each measurement

Data\_Only <- Data\_Merge %>% select(subject, code, contains("mean"), contains("std"))

Step 5. Use descriptive activity names to name the activities in the data set

Data\_Only$code <- activities[Data\_Only$code, 2]

Step 6. Appropriately label the data set with descriptive variable names

names(Data\_Only)[2] = "activity"

names(Data\_Only)<-gsub("Acc", "Accelerometer", names(Data\_Only))

names(Data\_Only)<-gsub("Gyro", "Gyroscope", names(Data\_Only))

names(Data\_Only)<-gsub("BodyBody", "Body", names(Data\_Only))

names(Data\_Only)<-gsub("Mag", "Magnitude", names(Data\_Only))

names(Data\_Only)<-gsub("^t", "Time", names(Data\_Only))

names(Data\_Only)<-gsub("^f", "Frequency", names(Data\_Only))

names(Data\_Only)<-gsub("tBody", "TimeBody", names(Data\_Only))

names(Data\_Only)<-gsub("-mean()", "Mean", names(Data\_Only), ignore.case = TRUE)

names(Data\_Only)<-gsub("-std()", "STD", names(Data\_Only), ignore.case = TRUE)

names(Data\_Only)<-gsub("-freq()", "Frequency", names(Data\_Only), ignore.case = TRUE)

names(Data\_Only)<-gsub("angle", "Angle", names(Data\_Only))

names(Data\_Only)<-gsub("gravity", "Gravity", names(Data\_Only))

Step 7. Create a second, independent tidy data set with the average of each variable for each activity and each subject