

```
In [2]: library(dplyr)
```

```
In [4]: filename <- "uci.zip"
if (!file.exists(filename)) {
  fileURL <- "https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.z
  download.file(fileURL, filename, method="curl")
}
# Checking if folder exists
if (!file.exists("UCI HAR Dataset")) {
  unzip(filename)
}
```

```
In [51]: features<- read.table("UCI HAR Dataset/features.txt", col.names = c("n", "functions"))
activities<-read.table("UCI HAR Dataset/activity_labels.txt", col.names = c("label", "activity"))
subjet_test<-read.table("UCI HAR Dataset/test/subject_test.txt", col.names = "subject")
```

```
In [103]: head(activities)
```

A data.frame: 6 × 2

	label	activity
	<int>	<chr>
1	1	WALKING
2	2	WALKING_UPSTAIRS
3	3	WALKING_DOWNSTAIRS
4	4	SITTING
5	5	STANDING
6	6	LAYING

```
In [104]: x_test<-read.table("UCI HAR Dataset/test/X_test.txt", col.names = features$functions)
y_test<-read.table("UCI HAR Dataset/test/y_test.txt", col.names = "label")
```

```
In [54]: head(x_test)
```

A data.frame: 6 × 561

	tBodyAcc.mean...X	tBodyAcc.mean...Y	tBodyAcc.mean...Z	tBodyAcc.std...X	tBodyAcc.std...Y	tBodyAcc.stc
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<c
1	0.2571778	-0.02328523	-0.01465376	-0.9384040	-0.9200908	-0.6676
2	0.2860267	-0.01316336	-0.11908252	-0.9754147	-0.9674579	-0.9449
3	0.2754848	-0.02605042	-0.11815167	-0.9938190	-0.9699255	-0.9627
4	0.2702982	-0.03261387	-0.11752018	-0.9947428	-0.9732676	-0.9670
5	0.2748330	-0.02784779	-0.12952716	-0.9938525	-0.9674455	-0.9782
6	0.2792199	-0.01862040	-0.11390197	-0.9944552	-0.9704169	-0.9653

```
In [105]: subject_train <- read.table("UCI HAR Dataset/train/subject_train.txt", col.names = "subject")
x_train <- read.table("UCI HAR Dataset/train/X_train.txt", col.names = features$functions)
y_train <- read.table("UCI HAR Dataset/train/y_train.txt", col.names = "label")
```

```
In [106]: head(subject_train)
```

A data.frame:

6 × 1

subject	
<int>	
1	1
2	1
3	1
4	1
5	1
6	1

```
In [107]: head(y_train)
           tail(y_train)
```

A

data.frame:

6 × 1

label	
<int>	
1	5
2	5
3	5
4	5
5	5
6	5

A data.frame:

6 × 1

label	
<int>	
7347	2
7348	2
7349	2
7350	2
7351	2
7352	2

1 Merge Dataset:

```
In [108]: X<-rbind(x_train,x_test)
           Y<-rbind(y_train,y_test)
           Subject<-rbind(subject_train,subjet_test)
           Merge_Data<-cbind(Subject,Y,X)
```

In [109]: `head(Merge_Data)`

A data.frame: 6 × 563

	subject	label	tBodyAcc.mean...X	tBodyAcc.mean...Y	tBodyAcc.mean...Z	tBodyAcc.std...X	tBodyAcc.std...
	<int>	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	1	5	0.2885845	-0.02029417	-0.1329051	-0.9952786	-0.983110
2	1	5	0.2784188	-0.01641057	-0.1235202	-0.9982453	-0.975300
3	1	5	0.2796531	-0.01946716	-0.1134617	-0.9953796	-0.967187
4	1	5	0.2791739	-0.02620065	-0.1232826	-0.9960915	-0.983402
5	1	5	0.2766288	-0.01656965	-0.1153619	-0.9981386	-0.980817
6	1	5	0.2771988	-0.01009785	-0.1051373	-0.9973350	-0.990486

2 Extracts only the measurements on the mean and standard deviation for each measurement.

In [110]: `TidyData<-Merge_Data%>%select(subject, label, contains("mean"), contains("std"))`

In [111]: `head(TidyData)`

A data.frame: 6 × 88

	subject	label	tBodyAcc.mean...X	tBodyAcc.mean...Y	tBodyAcc.mean...Z	tGravityAcc.mean...X	tGravityAc
	<int>	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	1	5	0.2885845	-0.02029417	-0.1329051	0.9633961	.
2	1	5	0.2784188	-0.01641057	-0.1235202	0.9665611	.
3	1	5	0.2796531	-0.01946716	-0.1134617	0.9668781	.
4	1	5	0.2791739	-0.02620065	-0.1232826	0.9676152	.
5	1	5	0.2766288	-0.01656965	-0.1153619	0.9682244	.
6	1	5	0.2771988	-0.01009785	-0.1051373	0.9679482	.

3 Uses descriptive activity names to name the activities in the data set.

In [112]:

In []:

In [113]: `tail(TidyData)`

A data.frame: 6 × 88

	subject	label	tBodyAcc.mean...X	tBodyAcc.mean...Y	tBodyAcc.mean...Z	tGravityAcc.mean...X	tGravi
	<int>	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
10294	24	2	0.1922746	-0.03364257	-0.10594911	0.8891608	
10295	24	2	0.3101546	-0.05339125	-0.09910872	0.8909677	
10296	24	2	0.3633846	-0.03921402	-0.10591509	0.8912809	
10297	24	2	0.3499661	0.03007744	-0.11578796	0.8908972	
10298	24	2	0.2375938	0.01846687	-0.09649893	0.8927357	
10299	24	2	0.1536272	-0.01843651	-0.13701846	0.8943309	

4 Appropriately labels the data set with descriptive variable names.

```
In [115]: names(TidyData)[2] = "activity"
names(TidyData) <- gsub("Acc", "Accelerometer", names(TidyData))
names(TidyData) <- gsub("Gyro", "Gyroscope", names(TidyData))
names(TidyData) <- gsub("BodyBody", "Body", names(TidyData))
names(TidyData) <- gsub("Mag", "Magnitude", names(TidyData))
names(TidyData) <- gsub("^t", "Time", names(TidyData))
names(TidyData) <- gsub("^f", "Frequency", names(TidyData))
names(TidyData) <- gsub("tBody", "TimeBody", names(TidyData))
names(TidyData) <- gsub("-mean()", "Mean", names(TidyData), ignore.case = TRUE)
names(TidyData) <- gsub("-std()", "STD", names(TidyData), ignore.case = TRUE)
names(TidyData) <- gsub("-freq()", "Frequency", names(TidyData), ignore.case = TRUE)
names(TidyData) <- gsub("angle", "Angle", names(TidyData))
names(TidyData) <- gsub("gravity", "Gravity", names(TidyData))
```

In [117]: `tail(TidyData)`

A data.frame: 6 × 88

	subject	activity	TimeBodyAccelerometerelerometer.mean...X	TimeBodyAccelerometerelerometer.mean...Y	TimeBodyAccelerometerelerometer.mean...Z	TimeBodyGravityAccelerometerelerometer.mean...X	TimeBodyGravityAccelerometerelerometer.mean...Y	TimeBodyGravityAccelerometerelerometer.mean...Z
	<int>	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
10294	24	2	0.1922746	-0.03364257	-0.10594911	0.8891608	-0.11578796	0.8908972
10295	24	2	0.3101546	-0.05339125	-0.09910872	0.8909677	-0.10591509	0.8912809
10296	24	2	0.3633846	-0.03921402	-0.10591509	0.8912809	-0.11578796	0.8908972
10297	24	2	0.3499661	0.03007744	-0.11578796	0.8908972	-0.10591509	0.8912809
10298	24	2	0.2375938	0.01846687	-0.09649893	0.8927357	-0.10591509	0.8912809
10299	24	2	0.1536272	-0.01843651	-0.13701846	0.8943309	-0.10591509	0.8912809

5 From the data set in step 4, creates a second, independent tidy data set with the average of each variable for each activity and each subject.

```
In [130]: FinalData<-TidyData%>%  
          group_by(subject, activity)%>%  
          summarize_all(list(mean))
```

```
In [131]: write.table(FinalData, "FinalData.txt", row.name=FALSE)
```

```
In [128]: head(FinalData)
```

A grouped_df: 6 × 88

subject	activity	TimeBodyAccelerometerelerometer.mean...X	TimeBodyAccelerometerelerometer.mean...Y	Time
<int>	<int>	<dbl>	<dbl>	
1	1	0.2773308	-0.017383819	
1	2	0.2554617	-0.023953149	
1	3	0.2891883	-0.009918505	
1	4	0.2612376	-0.001308288	
1	5	0.2789176	-0.016137590	
1	6	0.2215982	-0.040513953	

In [129]: FinalData

A grouped_df: 180 × 88

subject	activity	TimeBodyAccelerometerelerometer.mean...X	TimeBodyAccelerometerelerometer.mean...Y	Ti
<int>	<int>	<dbl>	<dbl>	
1	1	0.2773308	-0.017383819	
1	2	0.2554617	-0.023953149	
1	3	0.2891883	-0.009918505	
1	4	0.2612376	-0.001308288	
1	5	0.2789176	-0.016137590	
1	6	0.2215982	-0.040513953	
2	1	0.2764266	-0.018594920	
2	2	0.2471648	-0.021412113	
2	3	0.2776153	-0.022661416	
2	4	0.2770874	-0.015687994	
2	5	0.2779115	-0.018420827	
2	6	0.2813734	-0.018158740	
3	1	0.2755675	-0.017176784	
3	2	0.2608199	-0.032410941	
3	3	0.2924235	-0.019355408	
3	4	0.2571976	-0.003502998	
3	5	0.2800465	-0.014337656	
3	6	0.2755169	-0.018955679	
4	1	0.2785820	-0.014839948	
4	2	0.2708767	-0.031980430	
4	3	0.2799653	-0.009802009	
4	4	0.2715383	-0.007163065	
4	5	0.2804997	-0.009489111	
4	6	0.2635592	-0.015003184	
5	1	0.2778423	-0.017285032	
5	2	0.2684595	-0.032526976	
5	3	0.2935439	-0.008501075	
5	4	0.2736941	-0.009900835	
5	5	0.2825444	-0.007004186	
5	6	0.2783343	-0.018304212	
...	
26	1	0.2792644	-0.015425959	
26	2	0.2726914	-0.028163380	
26	3	0.2792846	-0.012625810	
26	4	0.2582435	-0.007133645	
26	5	0.2811270	-0.016660831	
26	6	0.2716459	-0.019189573	
27	1	0.2768495	-0.016647878	
27	2	0.2657703	-0.020095332	

subject	activity	TimeBodyAccelerometerelerometer.mean...X	TimeBodyAccelerometerelerometer.mean...Y	Ti
<int>	<int>	<dbl>	<dbl>	
27	3	0.2975442	-0.013556447	
27	4	0.2739413	-0.015526980	
27	5	0.2795669	-0.016593236	
27	6	0.2741025	-0.017986761	
28	1	0.2812282	-0.015675972	
28	2	0.2620058	-0.027944394	
28	3	0.2936421	-0.022023047	
28	4	0.2769776	-0.018540444	
28	5	0.2777951	-0.017263512	
28	6	0.2759135	-0.016753786	
29	1	0.2719999	-0.016291560	
29	2	0.2654231	-0.029946531	
29	3	0.2931404	-0.014941215	
29	4	0.2771800	-0.016630680	
29	5	0.2779651	-0.017260587	
29	6	0.2872952	-0.017196548	
30	1	0.2764068	-0.017588039	
30	2	0.2714156	-0.025331170	
30	3	0.2831906	-0.017438390	
30	4	0.2683361	-0.008047313	
30	5	0.2771127	-0.017016389	
30	6	0.2810339	-0.019449410	