Tidy Dataset

Below are the steps related to this project:

1. Merges the training and the test sets to create one data set.
2. Extracts only the measurements on the mean and standard deviation for each measurement.
3. Uses descriptive activity names to name the activities in the data set.
4. Appropriately labels the data set with descriptive variable names.
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.

Below is the detail of every step.

Merges the training and the test sets to create one data set.

Variable	File
activityLabel	activity_labels.txt
columnNames	features.txt
xTest	test/X_test.txt
yTest	test/y_test.txt
subjectTest	test/subject_test.txt
bodyaccxTest	test/Inertial Signals/body_acc_x_test.txt
bodyaccyTest	test/Inertial Signals/body_acc_y_test.txt
bodyacczTest	test/Inertial Signals/body_acc_z_test.txt
bodygyroxTest	test/Inertial Signals/body_gyro_x_test.txt
bodygyroyTest	test/Inertial Signals/body_gyro_y_test.txt
bodygyrozTest	test/Inertial Signals/body_gyro_z_test.txt
totalaccxTest	test/Inertial Signals/total_acc_x_test.txt
totalaccyTest	test/Inertial Signals/total_acc_y_test.txt
totalacczTest	test/Inertial Signals/total_acc_z_test.txt
xTrain	train/X_train.txt
yTrain	train/y_train.txt
subjectTrain	train/subject_train.txt
bodyaccxTrain	train/Inertial Signals/body_acc_x_train.txt
bodyaccyTrain	train/Inertial Signals/body_acc_y_train.txt
bodyacczTrain	train/Inertial Signals/body_acc_z_train.txt
bodygyroxTrain	train/Inertial Signals/body_gyro_x_train.txt
bodygyroyTrain	train/Inertial Signals/body_gyro_y_train.txt
bodygyrozTrain	train/Inertial Signals/body_gyro_z_train.txt
totalaccxTrain	train/Inertial Signals/total_acc_x_train.txt
totalaccyTrain	train/Inertial Signals/total_acc_y_train.txt
totalacczTrain	train/Inertial Signals/total_acc_z_train.txt

Values of column V2 of dataset columnNames that contains name of columns need to be transformed according the following table:

Value	Replaced with
[-]	_
[(]	blank value
[)]	blank value

Column names need to be generated with the following instruction

```
sapply(stri_pad_left(c(1:128), pad = "0", width = 3), function(x) paste(vname, x, sep = ""))
```

Column names start with the prefix and are generated from 1 to 128

Prefix Column Name	DataSet
bodyacc_x	bodyaccxTest
	bodyaccxTrain
bodyacc_y	bodyaccyTest
	bodyaccyTrain
bodyacc_z	bodyacczTest
	bodyacczTrain
bodygyro_x	bodygyroxTest
	bodygyroxTrain
bodygyro_y	bodygyroyTest
	bodygyroyTrain
bodygyro_z	bodygyrozTest
	bodygyrozTrain
totalacc_x	totalaccxTest
	totalaccxTrain
totalacc_y	totalaccyTest
	totalaccyTrain
totalacc_z	totalacczTest
	totalacczTrain

Test and Train dataset must be generated and then binded

```
datasetTest = cbind(xTest, yTest, subjectTest, bodyaccxTest, bodyaccyTest, bodyacczTest,
bodygyroxTest, bodygyroyTest, bodygyrozTest, totalaccxTest, totalaccyTest, totalacczTest)

datasetTrain = cbind(xTrain, yTrain, subjectTrain, bodyaccxTrain, bodyaccyTrain, bodyacczTrain,
bodygyroxTrain, bodygyroyTrain, bodygyrozTrain, totalaccxTrain, totalaccyTrain, totalacczTrain)

dataset = rbind(datasetTrain, datasetTest)
```

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

The following instruction makes the selection

```
columnNamesReport <- columnNames[ (V2 %like% "mean\\(\\)" | V2 %like% "std\\(\\)") ]
datasetReport <- dataset[,c("activitylabel", "subject", columnNamesReport$V2)]
```

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.

```
tidyDataReport <- group_by(datasetReport, activitylabel, activitydescription, subject) %>%
summarise_all(funs(mean))
```

In the final dataset all column are redefined with better names

Largest Column Name	Orignal Column Name
Actiity Label Code	activitylabel
Actiity Label Description	activitydescription
Subject	subject
Average of time domain body accelerometer signal mean axis X	tbodyacc_mean_x
Average of time domain body accelerometer signal mean axis Y	tbodyacc_mean_y
Average of time domain body accelerometer signal mean axis Z	tbodyacc_mean_z
Average of time domain body accelerometer signal standard deviation axis X	tbodyacc_std_x
Average of time domain body accelerometer signal standard deviation axis Y	tbodyacc_std_y
Average of time domain body accelerometer signal standard deviation axis Z	tbodyacc_std_z
Average of time domain gravity accelerometer signal mean axis X	tgravityacc_mean_x
Average of time domain gravity accelerometer signal mean axis Y	tgravityacc_mean_y
Average of time domain gravity accelerometer signal mean axis Z	tgravityacc_mean_z
Average of time domain gravity accelerometer signal standard deviation axis X	tgravityacc_std_x
Average of time domain gravity accelerometer signal standard deviation axis Y	tgravityacc_std_y

Largest Column Name	Orignal Column Name
Average of time domain gravity accelerometer signal standard deviation axis Z	tgravityacc_std_z
Average of time domain body accelerometer jerk signal mean axis X	tbodyaccjerk_mean_x
Average of time domain body accelerometer jerk signal mean axis Y	tbodyaccjerk_mean_y
Average of time domain body accelerometer jerk signal mean axis Z	tbodyaccjerk_mean_z
Average of time domain body accelerometer jerk signal standard deviation axis X	tbodyaccjerk_std_x
Average of time domain body accelerometer jerk signal standard deviation axis X	tbodyaccjerk_std_y
Average of time domain body accelerometer jerk signal standard deviation axis X	tbodyaccjerk_std_z
Average of time domain body gyroscope signal mean axis X	tbodygyro_mean_x
Average of time domain body gyroscope signal mean axis Y	tbodygyro_mean_y
Average of time domain body gyroscope signal mean axis Z	tbodygyro_mean_z
Average of time domain body gyroscope signal standard deviation axis X	tbodygyro_std_x
Average of time domain body gyroscope signal standard deviation axis Y	tbodygyro_std_y
Average of time domain body gyroscope signal standard deviation axis Z	tbodygyro_std_z
Average of time domain body gyroscope jerk signal mean axis X	tbodygyrojerk_mean_x
Average of time domain body gyroscope jerk signal mean axis Y	tbodygyrojerk_mean_y
Average of time domain body gyroscope jerk signal mean axis Z	tbodygyrojerk_mean_z
Average of time domain body gyroscope jerk signal standard deviation axis X	tbodygyrojerk_std_x
Average of time domain body gyroscope jerk signal standard deviation axis Y	tbodygyrojerk_std_y
Average of time domain body gyroscope jerk signal standard deviation axis Z	tbodygyrojerk_std_z
Average of time domain body accelerometer magnitude signal mean	tbodyaccmag_mean
Average of time domain body accelerometer magnitude signal standard deviation	tbodyaccmag_std
Average of time domain gravity accelerometer magnitude signal mean	tgravityaccmag_mean
Average of time domain gravity accelerometer magnitude signal standard deviation	tgravityaccmag_std

Largest Column Name	Orignal Column Name
Average of time domain body accelerometer jerk magnitude signal mean	tbodyaccjerkmag__mean
Average of time domain body accelerometer jerk magnitude signal standard deviation	tbodyaccjerkmag__std
Average of time domain body gyroscope magnitude signal mean	tbodygyromag__mean
Average of time domain body gyroscope magnitude signal standard deviation	tbodygyromag__std
Average of time domain body gyroscope jerk magnitude signal mean	tbodygyrojerkmag__mean
Average of time domain body gyroscope jerk magnitude signal standard deviation	tbodygyrojerkmag__std
Average of time domain body accelerometer signal mean axis X	fbodyacc__mean__x
Average of time domain body accelerometer signal mean axis Y	fbodyacc__mean__y
Average of time domain body accelerometer signal mean axis Z	fbodyacc__mean__z
Average of time domain body accelerometer signal standard deviation axis X	fbodyacc__std__x
Average of time domain body accelerometer signal standard deviation axis Y	fbodyacc__std__y
Average of time domain body accelerometer signal standard deviation axis Z	fbodyacc__std__z
Average of time domain body accelerometer jerk signal mean axis X	fbodyaccjerk__mean__x
Average of time domain body accelerometer jerk signal mean axis Y	fbodyaccjerk__mean__y
Average of time domain body accelerometer jerk signal mean axis Z	fbodyaccjerk__mean__z
Average of time domain body accelerometer jerk signal standard deviation axis X	fbodyaccjerk__std__x
Average of time domain body accelerometer jerk signal standard deviation axis Y	fbodyaccjerk__std__y
Average of time domain body accelerometer jerk signal standard deviation axis Z	fbodyaccjerk__std__z
Average of time domain body gyroscope signal mean axis X	fbodygyro__mean__x
Average of time domain body gyroscope signal mean axis Y	fbodygyro__mean__y
Average of time domain body gyroscope signal mean axis Z	fbodygyro__mean__z

Largest Column Name	Orignal Column Name
Average of time domain body gyroscope signal standard deviation axis X	fbodygyro_std_x
Average of time domain body gyroscope signal standard deviation axis Y	fbodygyro_std_y
Average of time domain body gyroscope signal standard deviation axis Z	fbodygyro_std_z
Average of time domain body accelerometer magnitude mean‘=	
fbodyaccmag_mean	
Average of time domain body accelerometer magnitude standard deviation	fbodyaccmag_std
Average of time domain body accelerometer jerk magnitude mean	fbodybodyaccjerkmag_mean
Average of time domain body accelerometer magnitude standard deviation	fbodybodyaccjerkmag_std
Average of time domain body gyroscope magnitude mean	fbodybodygyromag_mean
Average of time domain body gyroscope magnitude standard deviation	fbodybodygyromag_std
Average of time domain body gyroscope jerk magnitude mean	fbodybodygyrojerkmag_mean
Average of time domain body gyroscope jerk magnitude standard deviation	fbodybodygyrojerkmag_std)