You should create one R script called run_analysis.R that does the following.

1. Merges the training and the test sets to create one data set.
[JH] What the assignment is telling you to do is to create a single data set from the files in test and train folders of your data. The "test" and "train" folders each have three .txt files and a directory. You are to take the three .txt files from each of the "test" and "train" folders and combine them into one, single data frame.
Open question: "What to do with the Inertial Signals data?" I have not been able to discover this yet.

2. Extracts only the measurements on the mean and standard deviation for each measurement.
[JH] In other words, you are to filter for column names that include strings "mean" and "std" and discard the rest.

3. Uses descriptive activity names to name the activities in the data set
[JH] This is another gem of a statement. What they really want you to do is to apply the names given in "activity_labels.txt" to your merged data. To think they would have just said so!

4. Appropriately labels the data set with descriptive variable names.

[JH] This one continues in the spirit of the previous statement. The variable names you want are given in "features.txt". I would do this ***before*** *Step 2 to avoid all sorts of pain and grief. Otherwise, you will have to rely on a fairy godmother, a unicorn, a leperchaun or another magical creature of similar nature to tell you what numbers represent "mean" and "std".*

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.

[JH] This is the grand-daddy of confusing statements and quite rightly reserved for the final requirement.

This is a multi-part requirement.

What they mean by tidy is that you should assign "lowercasenamesthatnoonecanreadorunderstand" to variable names that you obtained from "features.txt".

Additionally, they want you to calculate the "means" of columns in your merged data (the data with non-Mean and non-Std columns removed). In other words, you are to calculate the "mean" of each variable (aka feature) for each activity performed by each subject. For example, if there were 10 observations for activity 1 peformed by subject 1, you are to calculate the mean of those 10 observations.