This step shows the script in RSTUDIO

```
run_analysis.R ×
Source on Save                                                          Run    Source

1   ## Create one R script called run_analysis.R that does the following:
2   ## 1. Merges the training and the test sets to create one data set.
3   ## 2. Extracts only the measurements on the mean and standard deviation for each measurement.
4   ## 3. Uses descriptive activity names to name the activities in the data set
5   ## 4. Appropriately labels the data set with descriptive activity names.
6   ## 5. Creates a second, independent tidy data set with the average of each variable for each activity and each subject.
7
8   # Verificación e instalación de paquete data.table
9 ▾ if (!require("data.table")) {
10      install.packages("data.table")
11 ▴ }
12  # Verificación e instalación de paquete reshape2
13 ▾ if (!require("reshape2")) {
14      install.packages("reshape2")
15 ▴ }
16
17  require("data.table")
18  require("reshape2")
19
20  # Load: activity labels - carga de etiquetas
21  activity_labels <- read.table("./UCI HAR Dataset/activity_labels.txt")[,2]
22
23  # Load: data column names - carga de nombres de columnas
24  features <- read.table("./UCI HAR Dataset/features.txt")[,2]
25
26  # Extract only the measurements on the mean and standard deviation for each measurement., realizando el punto 2
27  extract_features <- grepl("mean|std", features)
28
29  # Load and process X_test & y_test data. Cargar y procesar la data
30  X_test <- read.table("./UCI HAR Dataset/test/X_test.txt")
31  y_test <- read.table("./UCI HAR Dataset/test/y_test.txt")
32  subject_test <- read.table("./UCI HAR Dataset/test/subject_test.txt")
33
34  names(X_test) = features
35
36
```

```
run_analysis.R ×
Source on Save                                                          Run    Source

35
36  # Extract only the measurements on the mean and standard deviation for each measurement.
37  X_test = X_test[,extract_features]
38
39  # Load activity labels
40  y_test[,2] = activity_labels[y_test[,1]]
41  names(y_test) = c("Activity_ID", "Activity_Label")
42  names(subject_test) = "subject"
43
44  # Bind data
45  test_data <- cbind(as.data.table(subject_test), y_test, X_test)
46
47  # Load and process X_train & y_train data. , carga y procesamiento de data
48  X_train <- read.table("./UCI HAR Dataset/train/X_train.txt")
49  y_train <- read.table("./UCI HAR Dataset/train/y_train.txt")
50
51  subject_train <- read.table("./UCI HAR Dataset/train/subject_train.txt")
52
53  names(X_train) = features
54
55  # Extract only the measurements on the mean and standard deviation for each measurement.
56  X_train = X_train[,extract_features]
57
58  # Load activity data para el punto 4
59  y_train[,2] = activity_labels[y_train[,1]]
60  names(y_train) = c("Activity_ID", "Activity_Label")
61  names(subject_train) = "subject"
62
63  # Bind data
64  train_data <- cbind(as.data.table(subject_train), y_train, X_train)
65
66  # Merge test and train data,  --- realizando el punto 1
67  data = rbind(test_data, train_data)
68
69  id_labels   = c("subject", "Activity_ID", "Activity_Label")
70
```

```r
43
44    # Bind data
45    test_data <- cbind(as.data.table(subject_test), y_test, X_test)
46
47    # Load and process X_train & y_train data. , carga y procesamiento de data
48    X_train <- read.table("./UCI HAR Dataset/train/X_train.txt")
49    y_train <- read.table("./UCI HAR Dataset/train/y_train.txt")
50
51    subject_train <- read.table("./UCI HAR Dataset/train/subject_train.txt")
52
53    names(X_train) = features
54
55    # Extract only the measurements on the mean and standard deviation for each measurement.
56    X_train = X_train[,extract_features]
57
58    # Load activity data para el punto 4
59    y_train[,2] = activity_labels[y_train[,1]]
60    names(y_train) = c("Activity_ID", "Activity_Label")
61    names(subject_train) = "subject"
62
63    # Bind data
64    train_data <- cbind(as.data.table(subject_train), y_train, X_train)
65
66    # Merge test and train data,  --- realizando el punto 1
67    data = rbind(test_data, train_data)
68
69    id_labels    = c("subject", "Activity_ID", "Activity_Label")
70    data_labels = setdiff(colnames(data), id_labels)
71    melt_data       = melt(data, id = id_labels, measure.vars = data_labels)
72
73    # Apply mean function to dataset using dcast function, realizando el punto 3,5
74    tidy_data    = dcast(melt_data, subject + Activity_Label ~ variable, mean)
75
76    write.table(tidy_data, file = "./Data_Ordenada.txt")
```

Then I execute the script

```r
43
44    # Bind data
45    test_data <- cbind(as.data.table(subject_test), y_test, X_test)
46
47    # Load and process X_train & y_train data. , carga y procesamiento de data
48    X_train <- read.table("./UCI HAR Dataset/train/X_train.txt")
49    y_train <- read.table("./UCI HAR Dataset/train/y_train.txt")
50
51
```

76:53   (Top Level)   R Script

Console   Terminal ×   Jobs ×

```r
>
> names(X_test) = features
>
> # Extract only the measurements on the mean and standard deviation for each measurement.
> X_test = X_test[,extract_features]
>
> # Load activity labels
> y_test[,2] = activity_labels[y_test[,1]]
> names(y_test) = c("Activity_ID", "Activity_Label")
> names(subject_test) = "subject"
>
> # Bind data
> test_data <- cbind(as.data.table(subject_test), y_test, X_test)
>
> # Load and process X_train & y_train data. , carga y procesamiento de data
> X_train <- read.table("./UCI HAR Dataset/train/X_train.txt")
> y_train <- read.table("./UCI HAR Dataset/train/y_train.txt")
>
> subject_train <- read.table("./UCI HAR Dataset/train/subject_train.txt")
>
> names(X_train) = features
>
> # Extract only the measurements on the mean and standard deviation for each measurement.
> X_train = X_train[,extract_features]
>
> # Load activity data para el punto 4
> y_train[,2] = activity_labels[y_train[,1]]
> names(y_train) = c("Activity_ID", "Activity_Label")
> names(subject_train) = "subject"
>
> # Bind data
> train_data <- cbind(as.data.table(subject_train), y_train, X_train)
>
> # Merge test and train data,  --- realizando el punto 1
> data = rbind(test_data, train_data)
>
> id_labels    = c("subject", "Activity_ID", "Activity_Label")
> data_labels = setdiff(colnames(data), id_labels)
> melt_data       = melt(data, id = id_labels, measure.vars = data_labels)
>
> # Apply mean function to dataset using dcast function, realizando el punto 3,5
> tidy_data    = dcast(melt_data, subject + Activity_Label ~ variable, mean)
>
> write.table(tidy_data, file = "./Data_Ordenada.txt")
> |
```

FINALLY WE CAN SEE THE GENERATE FILE