**Overview**

This repository contains an R script, run\_analysis.R, that processes and cleans the UCI Human Activity Recognition (HAR) Using Smartphones Dataset. The script performs the following operations:

1. Merges the training and test sets to create one dataset.
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 dataset.
4. Appropriately labels the dataset with descriptive variable names.
5. Creates a second, independent tidy dataset with the average of each variable for each activity and each subject.

**📁 Dataset Source**

Original dataset:  
UCI HAR Dataset

Make sure to unzip the dataset into your working directory so that the folder UCI HAR Dataset is accessible.

**📜 Script: run\_analysis.R**

This R script performs the following:

**🔹 Step 1: Merge Training and Test Sets**

* Reads the training and test datasets (X\_train.txt, X\_test.txt) and merges them into one.
* Merges the activity labels and subject identifiers for both sets.

**🔹 Step 2: Extract Mean and Standard Deviation**

* Extracts only columns containing -mean() or -std() using the features.txt file.

**🔹 Step 3: Descriptive Activity Names**

* Maps the numeric activity codes to descriptive names (e.g., "walking", "sitting").

**🔹 Step 4: Label Dataset with Descriptive Variable Names**

* Cleans and formats variable names (removes special characters and standardizes casing).

**🔹 Step 5: Create a Tidy Data Set**

* Computes the average of each variable for each activity and each subject.
* Outputs the result to tidy\_average\_data.txt.

**📦 Output**

The final tidy dataset is written to:

txt

CopyEdit

tidy\_average\_data.txt

* Format: Tab-separated text
* Dimensions: 180 rows × 68 columns
* Each row is the average of each variable for a subject-activity pair.

**🧑‍💻 How to Run**

1. Download and unzip the UCI HAR dataset into your working directory.
2. Open R or RStudio.
3. Source the script:

r

CopyEdit

source("run\_analysis.R")

run\_analysis()

1. The output file tidy\_average\_data.txt will be saved in your working directory.

**📚 Dependencies**

The script requires the following R packages:

dplyr

reshape2

Install them if needed:

install.packages("dplyr")

install.packages("reshape2")