```
---
title: "Human Activity Recognition Using Smartphones"
author: "Bassel okasha"
date: "`r Sys.Date()`"
output: html_document
---
```

Introduction

This project analyzes human activity data collected from smartphone sensors. The goal is to create a tidy dataset that summarizes the average of each measurement for each activity and each subject. The original dataset, provided by the UCI Machine Learning Repository, includes accelerometer and gyroscope readings from participants performing activities such as walking, sitting, and standing.

Analysis Steps

The following steps were performed using R:

1. Merged the training and test datasets into one unified dataset.
2. Extracted only the measurements related to mean and standard deviation.
3. Replaced activity codes with descriptive activity names.
4. Labeled the dataset with clear, descriptive variable names.
5. Created a second, independent tidy dataset with the average of each variable for each activity and each subject.

Script Execution

````
```{r setup, include=TRUE}
data_path <- "C:/Users/TSC/Desktop/my_br_in_R/UCI HAR Dataset"

features <- read.table(file.path(data_path, "features.txt"), col.names = c("index", "feature"))
activities <- read.table(file.path(data_path, "activity_labels.txt"), col.names = c("code", "activity"))

valid_feature_names <- make.names(features$feature)

subject_train <- read.table(file.path(data_path, "train", "subject_train.txt"), col.names = "subject")
x_train <- read.table(file.path(data_path, "train", "X_train.txt"), col.names = valid_feature_names)
y_train <- read.table(file.path(data_path, "train", "y_train.txt"), col.names = "activity_code")

subject_test <- read.table(file.path(data_path, "test", "subject_test.txt"), col.names = "subject")
x_test <- read.table(file.path(data_path, "test", "X_test.txt"), col.names = valid_feature_names)
y_test <- read.table(file.path(data_path, "test", "y_test.txt"), col.names = "activity_code")

train_data <- cbind(subject_train, y_train, x_train)
test_data <- cbind(subject_test, y_test, x_test)
full_data <- rbind(train_data, test_data)

full_data$activity <- activities[full_data$activity_code, "activity"]
full_data$activity_code <- NULL

selected_columns <- grep("mean\\(\\)|std\\(\\)", features$feature, value = TRUE)
selected_columns <- make.names(selected_columns)

selected_data <- full_data[, c("subject", "activity", selected_columns)]

library(dplyr)

tidy_data <- selected_data %>%
  group_by(subject, activity) %>%
  summarise_all(mean)
````

```
write.table(tidy_data, "tidy_dataset.txt", row.name = FALSE)
```

#my cod
This R script processes the UCI Human Activity Recognition (HAR) dataset by merging, cleaning, and
summarizing sensor data collected from smartphones. The final output is a tidy dataset containing
the average of selected measurements for each subject and activity.
1. Set Data Path
data_path <- "C:/Users/TSC/Desktop/my_br_in_R/UCI HAR Dataset"
2. Load Metadata
features <- read.table(...); activities <- read.table(...)
3. Sanitize Feature Names
valid_feature_names <- make.names(features$feature)
4. Load Training and Test Data
subject_train, x_train, y_train
subject_test, x_test, y_test
5. Merge Datasets
train_data <- cbind(...); test_data <- cbind(...)
full_data <- rbind(train_data, test_data)
6. Add Descriptive Activity Labels
full_data$activity <- activities[full_data$activity_code, "activity"]
full_data$activity_code <- NULL
7. Select Mean and Standard Deviation Features
selected_columns <- grep("mean\\(\\)|std\\(\\)", features$feature, value = TRUE)
selected_columns <- make.names(selected_columns)
selected_data <- full_data[, c("subject", "activity", selected_columns)]
8. Create Tidy Dataset
tidy_data <- selected_data %>%
  group_by(subject, activity) %>%
  summarise_all(mean)
9.Export Final Dataset
write.table(tidy_data, "tidy_dataset.txt", row.name = FALSE)