

Reproducible Research: Peer Assessment 1

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.2      v readr      2.1.4
## v forcats    1.0.0      v stringr   1.5.0
## v ggplot2    3.4.2      v tibble    3.2.1
## v lubridate  1.9.2      v tidyr     1.3.0
## v purrr      1.0.1
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

Loading and preprocessing the data

This section describes the process to unzip dataset file and read it into a dataframe called “**activity**.” The first two lines of code which have been commented out shows the steps to download the dataset if unavailable in the project folder.

```
# file.URL <- "https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2Factivity.zip"
# download.file(file.URL, file.dir, method = "auto")

file.name <- "activity.csv"
file.dir <- "./activity.zip"
unzip(file.dir)
activity <- read.csv(file.name, na.strings = "NA")
file.remove(file.name)
```

Next, we will check the data and perform any pre-processing as required. ‘data.frame’: 17568 obs. of 3 variables: \$ steps : int NA NA NA NA NA NA NA NA NA NA NA NA ... \$ date : chr “2012-10-01” “2012-10-01” “2012-10-01” “2012-10-01” ... \$ interval: int 0 5 10 15 20 25 30 35 40 45 ... steps date interval 2304 0 0 steps date interval 1 NA 2012-10-01 0 2 NA 2012-10-01 5 3 NA 2012-10-01 10 4 NA 2012-10-01 15 5 NA 2012-10-01 20 6 NA 2012-10-01 25 ‘data.frame’: 17568 obs. of 3 variables: \$ steps : int NA NA NA NA NA NA NA NA NA NA NA NA ... \$ date : Date, format: “2012-10-01” “2012-10-01” ... \$ interval: int 0 5 10 15 20 25 30 35 40 45 ...

What is mean total number of steps taken per day?

What is the average daily activity pattern?

Imputing missing values

Are there differences in activity patterns between weekdays and weekends?