Tidyverse Create Assignment

Mohammad Zahid Chowdhury

2025-04-13

Overview of the dataset: This dataset is collected from kaggle datasets and this dataset provides a detailed overview of gym members' exercise routines, physical attributes, and fitness metrics. It contains 973 samples of gym data, including key performance indicators such as heart rate, calories burned, and workout duration. Each entry also includes demographic data and experience levels, allowing for comprehensive analysis of fitness patterns, athlete progression, and health trends.

Load libraries:

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.1
## v purrr     1.0.2
## -- 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
```

```
library(ggplot2)
library(dplyr)
```

Read the CSV data set:

```
gymdata<-read.csv("https://raw.githubusercontent.com/zahid607/Tidyverse/refs/heads/main/gym_members_exe

head(gymdata)
```

```
##   Age Gender Weight..kg. Height..m. Max_BPM Avg_BPM Resting_BPM
## 1  56   Male        88.3       1.71     180     157          60
## 2  46 Female        74.9       1.53     179     151          66
## 3  32 Female        68.1       1.66     167     122          54
## 4  25   Male        53.2       1.70     190     164          56
## 5  38   Male        46.1       1.79     188     158          68
## 6  56 Female        58.0       1.68     168     156          74
##   Session_Duration..hours. Calories_Burned Workout_Type Fat_Percentage
## 1                     1.69            1313         Yoga           12.6
## 2                     1.30             883         HIIT           33.9
## 3                     1.11             677       Cardio           33.4
## 4                     0.59             532     Strength           28.8
## 5                     0.64             556     Strength           29.2
## 6                     1.59            1116         HIIT           15.5
##   Water_Intake..liters. Workout_Frequency..days.week. Experience_Level   BMI
## 1                   3.5                             4                3 30.20
## 2                   2.1                             4                2 32.00
## 3                   2.3                             4                2 24.71
## 4                   2.1                             3                1 18.41
## 5                   2.8                             3                1 14.39
## 6                   2.7                             5                3 20.55
```

Name of Columns of the data set.

```
colnames(gymdata)
```

```
##  [1] "Age"                           "Gender"
##  [3] "Weight..kg."                   "Height..m."
##  [5] "Max_BPM"                       "Avg_BPM"
##  [7] "Resting_BPM"                   "Session_Duration..hours."
##  [9] "Calories_Burned"               "Workout_Type"
## [11] "Fat_Percentage"                "Water_Intake..liters."
## [13] "Workout_Frequency..days.week." "Experience_Level"
## [15] "BMI"
```

Summary statistics of gymdata

```
summary(gymdata)
```

```
##       Age           Gender           Weight..kg.       Height..m.
##  Min.   :18.00   Length:973         Min.   : 40.00   Min.   :1.500
##  1st Qu.:28.00   Class :character   1st Qu.: 58.10   1st Qu.:1.620
##  Median :40.00   Mode  :character   Median : 70.00   Median :1.710
##  Mean   :38.68                      Mean   : 73.85   Mean   :1.723
##  3rd Qu.:49.00                      3rd Qu.: 86.00   3rd Qu.:1.800
##  Max.   :59.00                      Max.   :129.90   Max.   :2.000
##     Max_BPM         Avg_BPM       Resting_BPM    Session_Duration..hours.
##  Min.   :160.0   Min.   :120.0   Min.   :50.00   Min.   :0.500
```

```
##  1st Qu.:170.0   1st Qu.:131.0   1st Qu.:56.00   1st Qu.:1.040
##  Median :180.0   Median :143.0   Median :62.00   Median :1.260
##  Mean   :179.9   Mean   :143.8   Mean   :62.22   Mean   :1.256
##  3rd Qu.:190.0   3rd Qu.:156.0   3rd Qu.:68.00   3rd Qu.:1.460
##  Max.   :199.0   Max.   :169.0   Max.   :74.00   Max.   :2.000
##  Calories_Burned  Workout_Type       Fat_Percentage  Water_Intake..liters.
##  Min.   : 303.0   Length:973         Min.   :10.00   Min.   :1.500
##  1st Qu.: 720.0   Class :character   1st Qu.:21.30   1st Qu.:2.200
##  Median : 893.0   Mode  :character   Median :26.20   Median :2.600
##  Mean   : 905.4                      Mean   :24.98   Mean   :2.627
##  3rd Qu.:1076.0                      3rd Qu.:29.30   3rd Qu.:3.100
##  Max.   :1783.0                      Max.   :35.00   Max.   :3.700
##  Workout_Frequency..days.week. Experience_Level      BMI
##  Min.   :2.000                 Min.   :1.00     Min.   :12.32
##  1st Qu.:3.000                 1st Qu.:1.00     1st Qu.:20.11
##  Median :3.000                 Median :2.00     Median :24.16
##  Mean   :3.322                 Mean   :1.81     Mean   :24.91
##  3rd Qu.:4.000                 3rd Qu.:2.00     3rd Qu.:28.56
##  Max.   :5.000                 Max.   :3.00     Max.   :49.84
```

Tyepes of Workout and their frequency:

```
gymdata %>%
  count(Workout_Type, sort = TRUE)
```

```
##   Workout_Type   n
## 1     Strength 258
## 2       Cardio 255
## 3         Yoga 239
## 4         HIIT 221
```

Total Calories burned by Gender

```
gymdata %>%
  group_by(Gender) %>%
  summarize(Total_Calories = sum(Calories_Burned, na.rm = TRUE))
```

```
## # A tibble: 2 x 2
##   Gender Total_Calories
##   <chr>           <dbl>
## 1 Female         398359
## 2 Male           482617
```

Gender Differences in Workout Habits:

```r
gymdata %>%
  group_by(Gender) %>%
  summarize(
    avg_duration = mean(Session_Duration..hours., na.rm = TRUE),
    avg_calories = mean(Calories_Burned, na.rm = TRUE)
  )
```

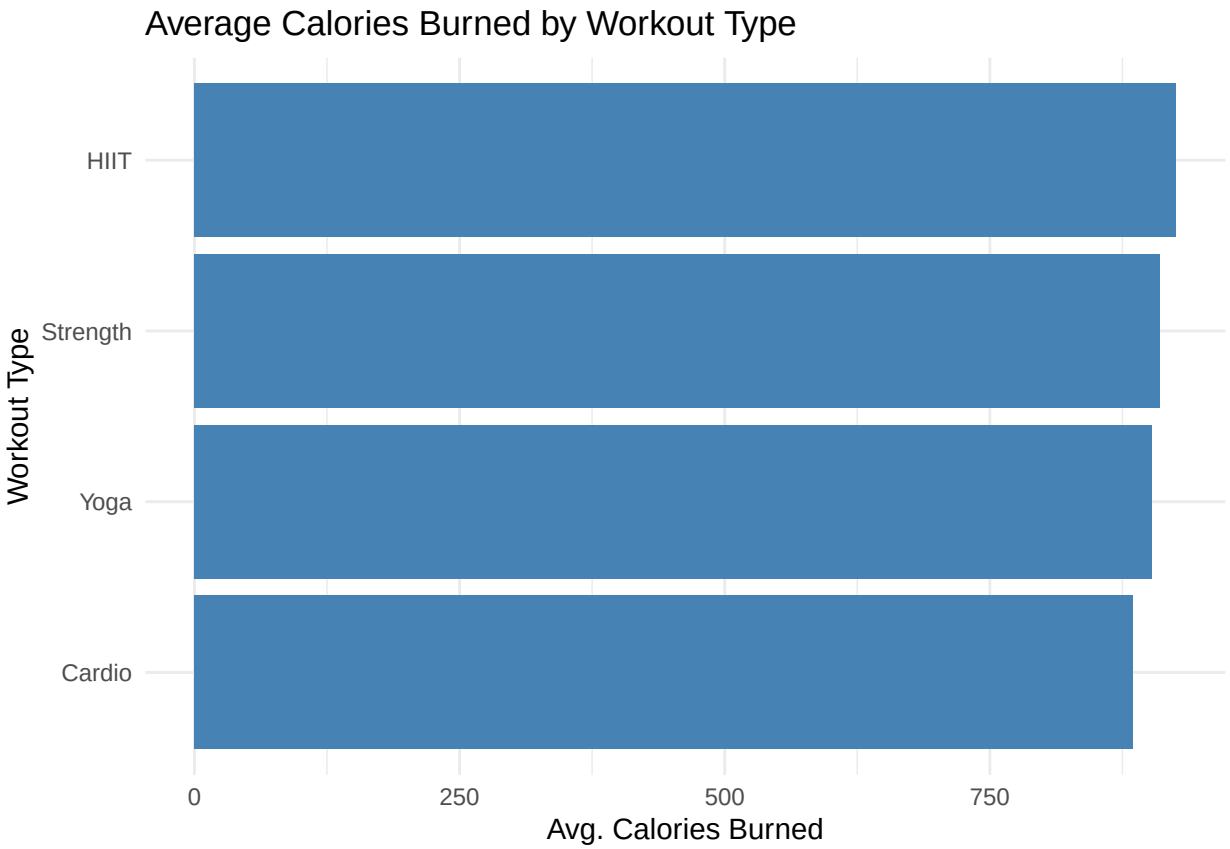
```
## # A tibble: 2 x 3
##   Gender avg_duration avg_calories
##   <chr>         <dbl>        <dbl>
## 1 Female         1.26         862.
## 2 Male           1.25         944.
```

Comment: Females speding average time in gym is slightly more than Males but males are burned more calories than females.

Bar diagram of Average Calories Burned by Workout Type

```r
# Clean + group data: average calories burned per exercise type
calories_by_exercise <- gymdata %>%
  group_by(Workout_Type) %>%
  summarize(avg_calories = mean(Calories_Burned, na.rm = TRUE)) %>%
  arrange(desc(avg_calories))

# Plot it
ggplot(calories_by_exercise, aes(x = reorder(Workout_Type, avg_calories), y = avg_calories)) +
  geom_col(fill = "steelblue") +
  coord_flip() +
  labs(
    title = "Average Calories Burned by Workout Type",
    x = "Workout Type",
    y = "Avg. Calories Burned"
  ) +
  theme_minimal()
```



Conclusion: This represents a basic analysis of the gym members' exercise tracking data.