

# HW03

```
load("hw2_list.rda")

bp_list

$treatment
  patient pre_bp post_bp diff_bp normal
Subject_1 Subject_1 120 127 -7 FALSE
Subject_2 Subject_2 151 145 6 FALSE
Subject_3 Subject_3 125 135 -10 FALSE
Subject_4 Subject_4 126 122 4 FALSE
Subject_5 Subject_5 115 115 0 TRUE
Subject_6 Subject_6 132 122 10 FALSE
Subject_7 Subject_7 132 123 9 FALSE
Subject_8 Subject_8 129 126 3 FALSE
Subject_9 Subject_9 134 126 8 FALSE
Subject_10 Subject_10 139 129 10 FALSE
Subject_11 Subject_11 127 132 -5 FALSE
Subject_12 Subject_12 122 146 -24 FALSE
Subject_13 Subject_13 127 120 7 FALSE
Subject_14 Subject_14 135 114 21 TRUE
Subject_15 Subject_15 133 121 12 FALSE
Subject_16 Subject_16 128 120 8 FALSE
Subject_17 Subject_17 147 128 19 FALSE
Subject_18 Subject_18 138 120 18 FALSE
Subject_19 Subject_19 140 133 7 FALSE
Subject_20 Subject_20 132 115 17 TRUE

$placebo
  patient pre_bp post_bp diff_bp normal
1 Subject_1 138 105 33 TRUE
2 Subject_2 135 136 -1 FALSE
3 Subject_3 147 123 24 FALSE
4 Subject_4 117 130 -13 FALSE
5 Subject_5 152 134 18 FALSE
6 Subject_6 134 143 -9 FALSE
7 Subject_7 114 135 -21 FALSE
8 Subject_8 121 139 -18 FALSE
9 Subject_9 131 120 11 FALSE
10 Subject_10 130 124 6 FALSE
```

## Task 1: Control Flow Practice

1. Suppose we want to characterize the post-treatment (or placebo) blood pressure measurement as optimal ( $\leq 120$ ), borderline ( $120 < \text{bp} \leq 130$ ), and high ( $> 130$ ). First, create a new column in each data frame from above called **status**.

```
#bp_list$status <-character(10)

bp_list$placebo$status <- ifelse(bp_list$placebo$post_bp <=120, "Optimal",
                                ifelse((bp_list$placebo$post_bp >120 & bp_list$placebo$post_bp <=130), "Borderlin
                                ifelse(bp_list$placebo$post_bp >130, "High", NA))
                                )

bp_list$placebo

  patient pre_bp post_bp diff_bp normal Status
1 Subject_1 138 105 33 TRUE Optimal
2 Subject_2 135 136 -1 FALSE High
3 Subject_3 147 123 24 FALSE Borderline
4 Subject_4 117 130 -13 FALSE Borderline
5 Subject_5 152 134 18 FALSE High
6 Subject_6 134 143 -9 FALSE High
7 Subject_7 114 135 -21 FALSE High
8 Subject_8 121 139 -18 FALSE High
9 Subject_9 131 120 11 FALSE Optimal
10 Subject_10 130 124 6 FALSE Borderline
```

2. For the non-placebo data frame (within the list), create a for loop and use if/then/else logic to create the status column's values.

```
#bp_list$treatment$status<-character(20)
for (i in -1:20){

bp_list$treatment$status <- ifelse(bp_list$treatment$post_bp <=120, "Optimal",
                                ifelse((bp_list$treatment$post_bp >120 & bp_list$treatment$post_bp <=130), "Borde
                                ifelse(bp_list$treatment$post_bp >130, "High", NA))
                                )

}

head(bp_list$treatment)

  patient pre_bp post_bp diff_bp normal Status
Subject_1 Subject_1 120 127 -7 FALSE Borderline
Subject_2 Subject_2 151 145 6 FALSE High
Subject_3 Subject_3 125 135 -10 FALSE High
Subject_4 Subject_4 126 122 4 FALSE Borderline
Subject_5 Subject_5 115 115 0 TRUE Optimal
Subject_6 Subject_6 132 122 10 FALSE Borderline
```

```
#bp_list$treatment

#bp_list$placebo$status <-character(10)
#bp_list$placebo$status
for (i in -1:20){

bp_list$placebo$status <- ifelse(bp_list$placebo$post_bp <=120, "Optimal",
                                ifelse((bp_list$placebo$post_bp >120 & bp_list$placebo$post_bp <=130), "Borderlin
                                ifelse(bp_list$placebo$post_bp >130, "High", NA))
                                )
#head(Final_data$treatment)
}

#head(bp_list$placebo)

bp_list$placebo

  patient pre_bp post_bp diff_bp normal Status
1 Subject_1 138 105 33 TRUE Optimal
2 Subject_2 135 136 -1 FALSE High
3 Subject_3 147 123 24 FALSE Borderline
4 Subject_4 117 130 -13 FALSE Borderline
5 Subject_5 152 134 18 FALSE High
6 Subject_6 134 143 -9 FALSE High
7 Subject_7 114 135 -21 FALSE High
8 Subject_8 121 139 -18 FALSE High
9 Subject_9 131 120 11 FALSE Optimal
10 Subject_10 130 124 6 FALSE Borderline
```

## Task 2 Writing function

```
# Create a function and pass 2 arguments bp_list (data frame) and statistics
patient_summary <- function(bp_list, stat = "mean") {
# Check if the list has two elements if not then give message
if(length(bp_list) != 2) {
stop("There should be exactly two data frames in the list: treatment and placebo dat
}

# Get the function for the 'stat' string
stat_fun <- get(stat)

# Define the columns to calculate the statistic for pre_bp, post_bp and diff_dp column
stat_columns <- c("pre_bp", "post_bp", "diff_bp")

# Initialize an empty vector to store the results
stat_result <- c()

# create Loop through each data frame in the list
for(df_data in bp_list) {
# Check if all the columns exist in the data frame or not if not then show message t
for (col in stat_columns) {
if (!col %in% names(df_data)) {
stop(paste("Column", col, "not found in the data frame"))
}

# Compute the statistic for the current column
stat_result <- c(stat_result, stat_fun(df_data[[col]], na.rm = TRUE))
}
}

# Create meaningful names for the result vector
final_result <- c()
data_frame_names <- c("treatment", "placebo")
for (df_name in data_frame_names) {
for (col in stat_columns) {
final_result <- c(final_result, paste(df_name, col, sep = "_"))
}
}

# Assign names to the result vector
names(stat_result) <- final_result

# Return the result vector
return(stat_result)
}
```

show “mean”, “var”, “sd”, “min”, and “max” for pre\_bp, post\_bp and diff\_dp columns

```
# Show the mean of all the values
mean_stats <- patient_summary(bp_list, stat = "mean")
#print("The mean of all the values are ")
print(mean_stats)

treatment_pre_bp treatment_post_bp treatment_diff_bp placebo_pre_bp
131.60 125.95 5.65 131.90
placebo_post_bp placebo_diff_bp
128.90 3.00

# Show the var of all the values
var_stats <- patient_summary(bp_list, stat = "var")
print(var_stats)

treatment_pre_bp treatment_post_bp treatment_diff_bp placebo_pre_bp
75.72632 78.99737 117.81842 149.87778
placebo_post_bp placebo_diff_bp
124.98889 341.33333

# Show the sd of all the values
sd_stats <- patient_summary(bp_list, stat = "sd")
print(sd_stats)

treatment_pre_bp treatment_post_bp treatment_diff_bp placebo_pre_bp
8.702087 8.888046 10.854419 12.242458
placebo_post_bp placebo_diff_bp
11.179843 18.475209

# Show the min of all the values
min_stats <- patient_summary(bp_list, stat = "min")
print(min_stats)

treatment_pre_bp treatment_post_bp treatment_diff_bp placebo_pre_bp
115 114 -24 114
placebo_post_bp placebo_diff_bp
105 -21

# Show the max of all the values
max_stats <- patient_summary(bp_list, stat = "max")
print(max_stats)

treatment_pre_bp treatment_post_bp treatment_diff_bp placebo_pre_bp
151 146 21 152
placebo_post_bp placebo_diff_bp
143 33

# Show the structure of the data frame
str(bp_list)
```

List of 2  
\$ treatment:'data.frame': 20 obs. of 6 variables:  
..\$ patient: chr [1:20] "Subject\_1" "Subject\_2" "Subject\_3" "Subject\_4" ...  
..\$ pre\_bp : num [1:20] 120 151 125 126 115 132 129 134 139 ...  
..\$ post\_bp: num [1:20] 127 145 135 122 115 122 123 126 129 ...  
..\$ diff\_bp: num [1:20] -7 6 -10 4 0 10 9 3 8 10 ...  
..\$ normal : logi [1:20] FALSE FALSE FALSE TRUE FALSE ...  
..\$ Status : chr [1:20] "Borderline" "High" "High" "Borderline" ...  
\$ placebo : 'data.frame': 10 obs. of 6 variables:  
..\$ patient: chr [1:10] "Subject\_1" "Subject\_2" "Subject\_3" "Subject\_4" ...  
..\$ pre\_bp : num [1:10] 138 135 147 117 152 134 114 121 131 130  
..\$ post\_bp: num [1:10] 105 136 123 130 134 143 135 139 120 124  
..\$ diff\_bp: num [1:10] 33 -1 24 -13 18 -9 -21 -18 11 6  
..\$ normal : logi [1:10] TRUE FALSE FALSE FALSE FALSE FALSE ...  
..\$ Status : chr [1:10] "Optimal" "High" "Borderline" "Borderline" ...