

Project 1

201805070 → Emrah Fidan → Group 35

This is a report I created so that the answer to all questions can appear in a single file. You can look at other folders to review codes and outputs.

Project 1.1

```
# Read Data
data <- read.table("DatasetNA.txt", header = TRUE)

# a. The function should have options of main title, x and y labels, x and y limits, color options.
# b. Group and Gender categorical variables should be barplotted in a separate page.

# Create My Barplot
my_barplot_group <- function() {
  # Group Person Count
  group_freq <- table(data$Group)

  heights <- as.numeric(group_freq)

  bar_names <- names(group_freq)

  plot_height <- max(heights) + 2
  plot_width <- length(heights) * 1.5
  bar_width <- 0.8

  # Graphic
  plot(0, 0, xlim = c(0, plot_width-1), ylim = c(0, plot_height), type = "h", xlab = "Group", ylab = "Person", main = "Groups")

  # Bars
  rect(1 - bar_width / 2, 0, 1 + bar_width / 2, heights[1], col = "skyblue")
  rect(2 - bar_width / 2, 0, 2 + bar_width / 2, heights[2], col = "yellow")
  rect(3 - bar_width / 2, 0, 3 + bar_width / 2, heights[3], col = "pink")
  rect(4 - bar_width / 2, 0, 4 + bar_width / 2, heights[4], col = "green")

  # Bars Name
  text(1:length(heights), par("usr")[2] - 0, labels = bar_names, srt = 0, xpd = TRUE)
}

my_barplot_gender <- function() {
  # Gender Person Count
  gender_freq <- table(data$Gender)

  heights <- as.numeric(gender_freq)

  bar_names <- names(gender_freq)

  plot_height <- max(heights) + 2
  plot_width <- length(heights) * 3
  bar_width <- 0.8

  # Graphic
  plot(0, 0, xlim = c(0, plot_width-1), ylim = c(0, plot_height), type = "h", xlab = "Genders", ylab = "Frekans", main = "Gender Dağılımı")

  # Bars
  rect(1 - bar_width / 2, 0, 1 + bar_width / 2, heights[1], col = "red")
  rect(2 - bar_width / 2, 0, 2 + bar_width / 2, heights[2], col = "blue")

  # Bars Name
  text(1:length(heights), par("usr")[2] - 0, labels = bar_names, srt = 0, xpd = TRUE)
}

my_barplot_group()
my_barplot_gender()

# c. Group and Gender categorical variables should be barplotted horizontally or vertically in the same page together :

# Create My Barplot Function
my_barplot_group_gender <- function() {
```

```

# Group and Gender Person Count
group_freq <- table(data$Group)
gender_freq <- table(data$Gender)

heights <- as.numeric(group_freq)
heights2 <- as.numeric(gender_freq)

bar_names <- names(group_freq)
bar_names2 <- names(gender_freq)

plot_height <- max(heights) + 2
plot_width <- length(heights) * 1.5
bar_width <- 0.8

plot_height2 <- max(heights2) + 2
plot_width2 <- length(heights2) * 3
bar_width2 <- 0.8

# Graphic
par(mfrow = c(1, 2)) # same page
plot(0, 0, xlim = c(0, plot_width-1), ylim = c(0, plot_height), type = "h", xlab = "Group", ylab = "Person", main = "Groups")

rect(1 - bar_width / 2, 0, 1 + bar_width / 2, heights[1], col = "skyblue")
rect(2 - bar_width / 2, 0, 2 + bar_width / 2, heights[2], col = "yellow")
rect(3 - bar_width / 2, 0, 3 + bar_width / 2, heights[3], col = "pink")
rect(4 - bar_width / 2, 0, 4 + bar_width / 2, heights[4], col = "green")

text(1:length(heights), par("usr")[2] - 0, labels = bar_names, srt = 90, xpd = TRUE)

plot(0, 0, xlim = c(0, plot_width2-3), ylim = c(0, plot_height2), type = "h", xlab = "Gender", ylab = "Person", main = "Genders")

rect(1 - bar_width2 / 2, 0, 1 + bar_width2 / 2, heights2[1], col = "red")
rect(2 - bar_width2 / 2, 0, 2 + bar_width2 / 2, heights2[2], col = "blue")

text(1:length(heights2), par("usr")[3] + 9, labels = bar_names2, srt = 90, xpd = TRUE)

}
my_barplot_group_gender()

```

Project 1.2

```

# Read File
veri <- read.table("DatasetNA.txt", header = TRUE, sep = " ", dec = ",")

# Create my histogram()
custom_histogram <- function(data, bins = 10, main_title = "Histogram", x_label = "X", y_label = "Frequency", color = "blue") {
  # ignore <NA>
  data <- data[!is.na(data)]

  # Calculate histogram
  hist_counts <- rep(0, bins)
  data_min <- min(data)
  data_max <- max(data)
  range_width <- data_max - data_min
  bin_width <- range_width / bins

  for (i in 1:length(data)) {
    bin_index <- floor((data[i] - data_min) / bin_width) + 1
    if (bin_index > 0 && bin_index <= bins) {
      hist_counts[bin_index] <- hist_counts[bin_index] + 1
    }
  }

  # Use plot()
  breaks <- seq(data_min, data_max, length.out = bins + 1)
  bar_centers <- breaks[-(length(breaks))] + bin_width / 2
  plot(bar_centers, hist_counts, type = "h", lwd = 10, col = color, xlab = x_label, ylab = y_label, main = main_title)
}

# Select Var1-Var8 colums
veri_alkume <- veri[, c("Var1", "Var2", "Var3", "Var4", "Var5", "Var6", "Var7", "Var8")]

# String -> numeric
veri_numeric <- apply(veri_alkume, 2, function(x) as.numeric(as.character(gsub(",", ".", x))))

# a: The function should have options of the number of bins, main title, x and y labels, x and y limits, color options.
# b: Each histogram should be plotted in a separate page (output screen).
# histograms : separte page
for (i in 1:ncol(veri_numeric)) {
  custom_histogram(veri_numeric[, i], main_title = paste("Histogram of", colnames(veri_numeric)[i]), x_label = colnames(veri_numeric)[

```

```

}

# c : More than one histogram could be plotted in the same page (output screen) together.
# Create a PDF file
pdf("all_histograms.pdf")

# Create histograms and save in the PDF file
for (i in 1:ncol(veri_numeric)) {
  custom_histogram(veri_numeric[, i], main_title = paste("Histogram of", colnames(veri_numeric)[i]), x_label = colnames(veri_numeric)[i])
}

# Close the PDF file
dev.off()

```

Project 1.3

```

# Read Data
veri <- read.table("DatasetNA.txt", header = TRUE, sep = " ", dec = ",")

# Create my Boxplot
custom_boxplot <- function(data, main_title = "Boxplot", x_label = "X", y_label = "Y", color = "lightblue") {
  # Calculate boxplot
  stats <- fivenum(data)

  # Generate x values
  x <- rep(1, length(data))

  # Plot boxplot components
  plot(x, data, type = "n", xlim = c(0.5, 1.5), ylim = range(data, na.rm = TRUE, finite = TRUE),
       xlab = x_label, ylab = y_label, main = main_title)

  rect(0.9, stats[2], 1.1, stats[4], col = color)
  segments(0.9, stats[2], 1.1, stats[2])
  segments(0.9, stats[4], 1.1, stats[4])
  segments(1, stats[2], 1, stats[1])
  segments(1, stats[4], 1, stats[5])
  segments(0.8, stats[3], 1.2, stats[3], lwd = 2) # Horizontal line
}

# Select Var1-Var8
veri_alkume <- veri[, c("Var1", "Var2", "Var3", "Var4", "Var5", "Var6", "Var7", "Var8")]

# String -> numeric
veri_numeric <- apply(veri_alkume, 2, function(x) as.numeric(as.character(gsub(",", ".", x))))

# Draw boxplot
for (i in 1:ncol(veri_numeric)) {
  custom_boxplot(veri_numeric[, i], main_title = paste("Boxplot of", colnames(veri_numeric)[i]), x_label = colnames(veri_numeric)[i], color = "lightblue")
}

# A page
# Create a PDF file
pdf("all_boxplots.pdf")

# Create boxplots and save in the PDF file
for (i in 1:ncol(veri_numeric)) {
  custom_boxplot(veri_numeric[, i], main_title = paste("Boxplot - ", colnames(veri_numeric)[i]), x_label = colnames(veri_numeric)[i], color = "lightblue")
}

# Close the PDF file
dev.off()

```

I separated the outputs according to their folders and put them in the Project1 file in pdf format