Hw3Final

Jenny, Syed, Kevin

2025-10-23

```r
# Setting Working Directory
setwd("/cloud/project")

#Read the Data and call it TextMessages
TextMessages <- read.csv("TextMessages.csv", header=TRUE)

# Create the dataset
text_data <- data.frame(Participant = 1:50, Group = c(rep(1, 25), rep(2, 25)),
                        Baseline = c(52,68,85,47,73,57,63,50,66,60,51,72,77,57,
                                     79,75,53,72,62,71,53,64,79,75,60,65,57,66,
                                     71,75,61,80,66,53,62,61,77,66,52,60,58,54,
                                     72,71,87,75,57,59,46,89),
                        Six_months = c(32,48,62,16,63,53,59,58,59,57,60,56,61,
                                       52,9,76,38,63,53,61,50, 78,33,68,59,62,
                                       50,62,61,70,64,64,55,47,61,56,64,62,47,
                                       56,78,74, 61,61,78,62,71,55,46,79))


# *** RESHAPE DATE FROM WIDE TO LONG ***
library(tidyr)
library(dplyr)

text_long <- text_data %>% pivot_longer(cols = c(Baseline, Six_months),
                            names_to = "Timepoint", values_to = "TextMessages")


##############
## BOX PLOT ##
##############

library(ggplot2)

# quick peek to confirm the data is ready
str(text_long)
```

```
## tibble [100 x 4] (S3: tbl_df/tbl/data.frame)
##  $ Participant : int [1:100] 1 1 2 2 3 3 4 4 5 5 ...
##  $ Group       : num [1:100] 1 1 1 1 1 1 1 1 1 1 ...
##  $ Timepoint   : chr [1:100] "Baseline" "Six_months" "Baseline" "Six_months" ...
##  $ TextMessages: num [1:100] 52 32 68 48 85 62 47 16 73 63 ...
```

```r
head(text_long)
```

```
## # A tibble: 6 x 4
##   Participant Group Timepoint   TextMessages
##         <int> <dbl> <chr>              <dbl>
## 1           1     1 Baseline              52
## 2           1     1 Six_months            32
## 3           2     1 Baseline              68
## 4           2     1 Six_months            48
## 5           3     1 Baseline              85
## 6           3     1 Six_months            62
```

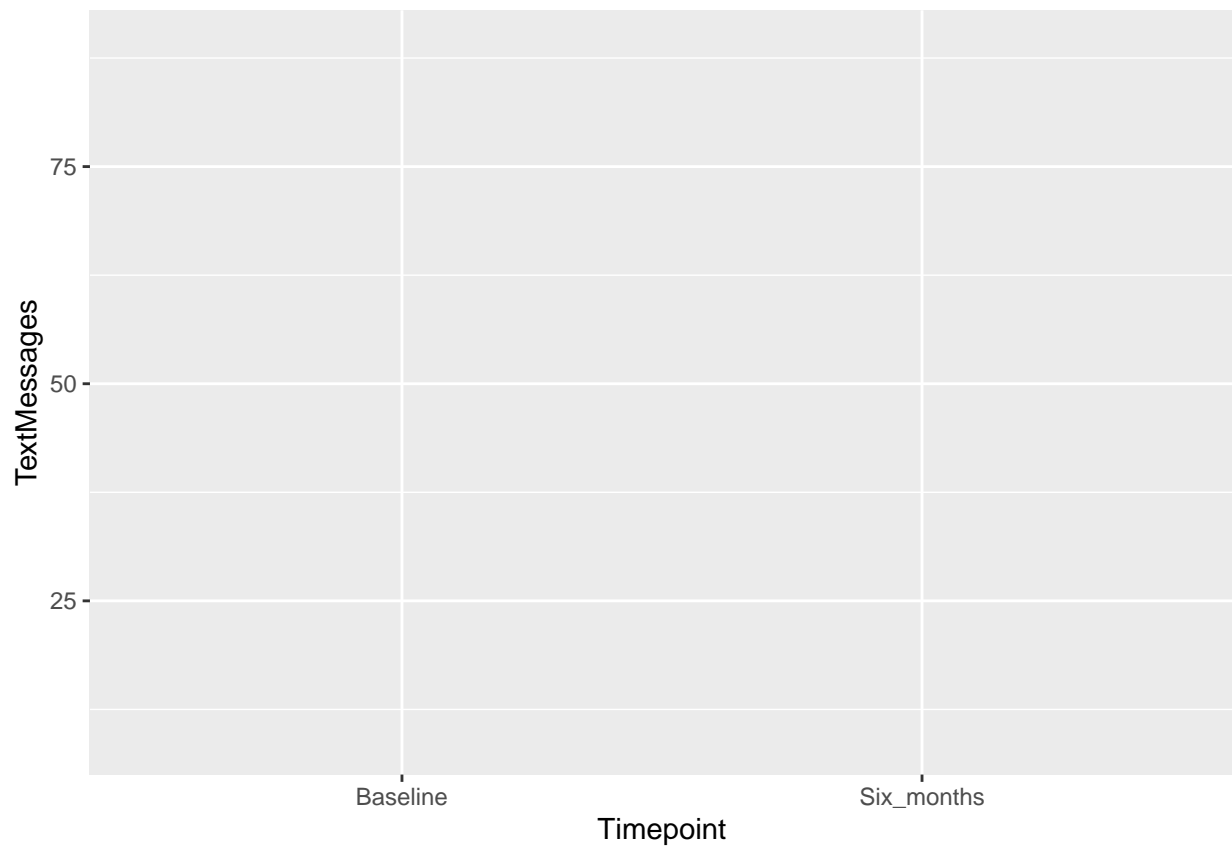```r
table(text_long$Timepoint)
```

```
##
##   Baseline Six_months
##         50         50
```

```r
table(text_long$Group)
```
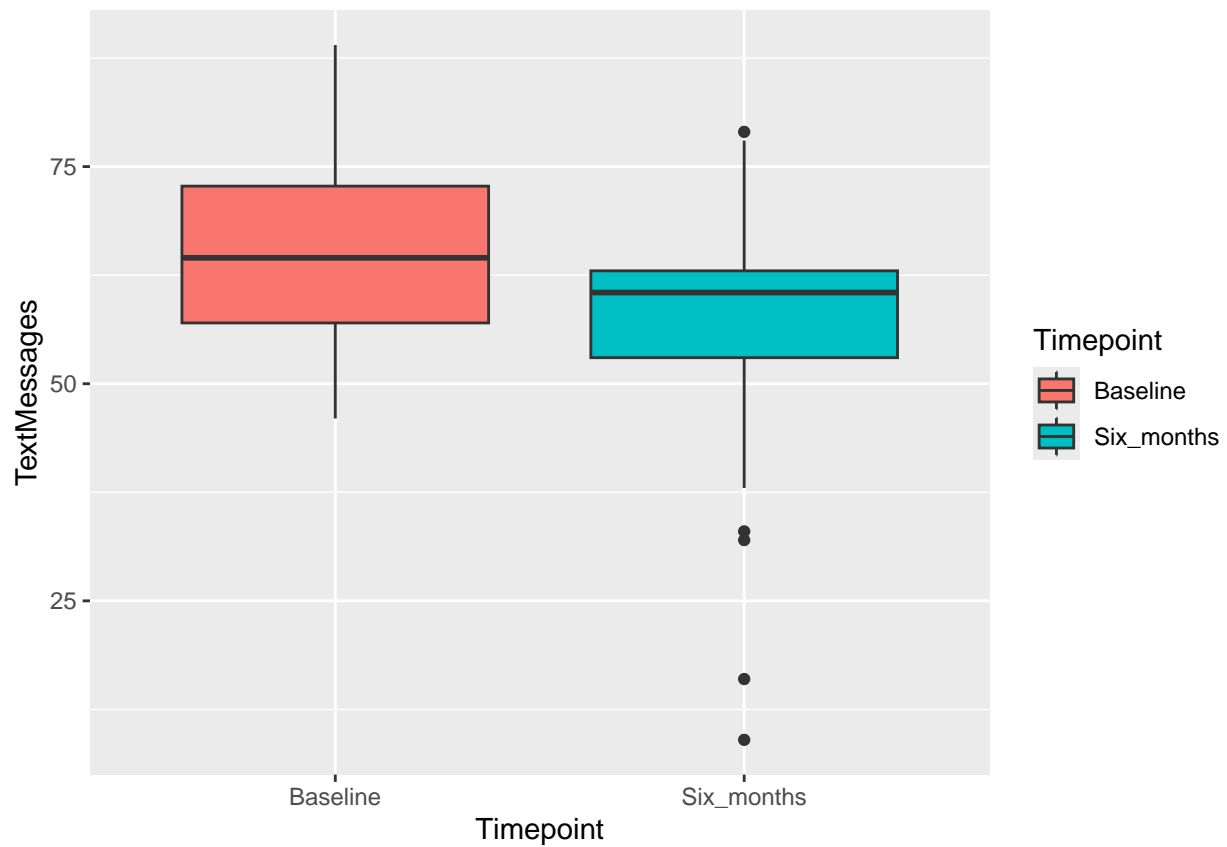
```
##
##  1  2
## 50 50
```

```r
# Make sure Categorical Variables exist
text_long$Group <- as.factor(text_long$Group)
text_long$Timepoint <- factor(text_long$Timepoint, levels =
                          c("Baseline", "Six_months"))
# Create the boxplot
Boxplot <- ggplot(text_long, aes(x = Timepoint, y = TextMessages,
                          fill = Timepoint))
Boxplot
```
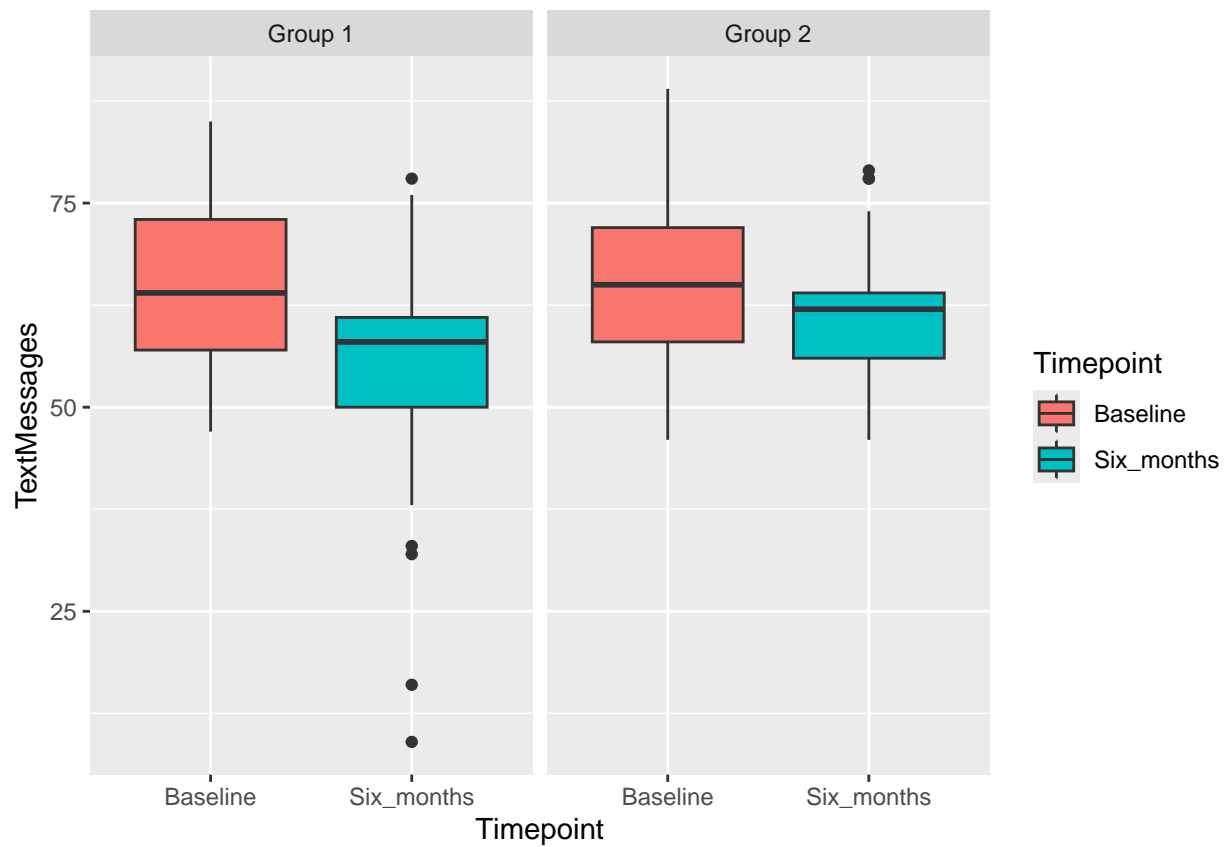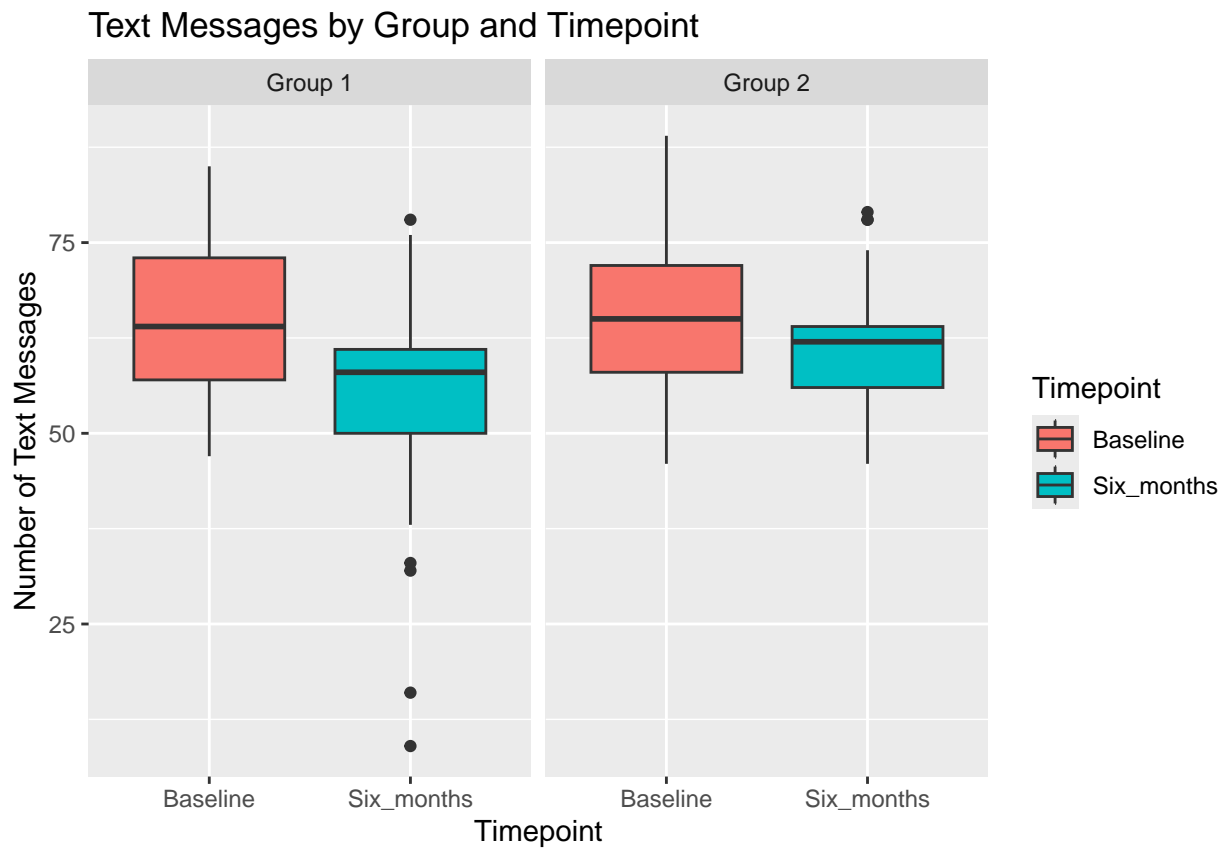
```
# Add box layer
Boxplot + geom_boxplot()
```

```r
# Stratify into 2 seperate groups
Boxplot + geom_boxplot() + facet_wrap(~ Group, labeller = labeller(Group =
                                        c(`1` = "Group 1", `2` = "Group 2")))
```

```
#add user-defined labels
Boxplot + geom_boxplot() + facet_wrap(~ Group, labeller = labeller(Group =
                  c(`1` = "Group 1", `2` = "Group 2"))) + labs(title =
                "Text Messages by Group and Timepoint", x = "Timepoint",
                y = "Number of Text Messages")
```

Text Messages by Group and Timepoint


```
# Make it look cleaner
Boxplot + geom_boxplot() + facet_wrap(~ Group, labeller = labeller(Group =
                    c(`1` = "Group 1", `2` = "Group 2"))) + labs(title =
                "Text Messages by Group and Timepoint", x = "Timepoint",
                y = "Number of Text Messages") + theme_minimal() +
        theme(legend.position = "none", plot.title =  element_text(hjust = 0.5))
```

Text Messages by Group and Timepoint
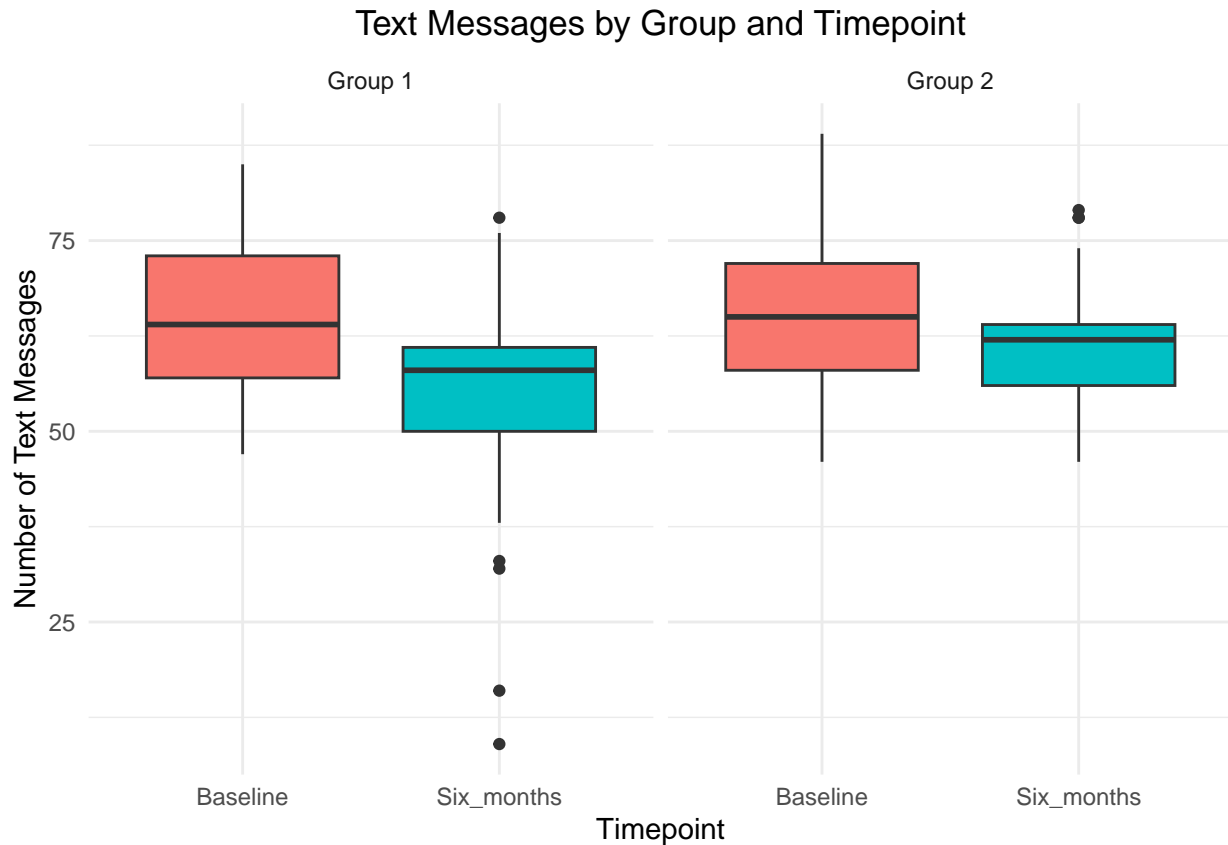
```
#SYED'S PART


TextMessages <- read.csv("TextMessages.csv", header=TRUE)


#install.packages("pastecs")

library(pastecs)

TextMessages$Group <- as.factor(TextMessages$Group)
by(TextMessages$Baseline,TextMessages$Group,stat.desc)
```

```
## TextMessages$Group: 1
##      nbr.val      nbr.null       nbr.na          min          max          range
##    25.000000      0.000000     0.000000    47.000000    85.000000     38.000000
##          sum        median         mean      SE.mean CI.mean.0.95          var
##  1621.000000     64.000000    64.840000     2.135946     4.408377   114.056667
##      std.dev      coef.var
##    10.679732      0.164709
## -------------------------------------------------------------
## TextMessages$Group: 2
##      nbr.val      nbr.null       nbr.na          min          max          range
##    25.0000000     0.0000000    0.0000000    46.0000000   89.0000000    43.0000000
##          sum        median         mean      SE.mean CI.mean.0.95          var
## 1640.0000000    65.0000000    65.6000000     2.1671794    4.4728385  117.4166667
##      std.dev      coef.var
```

```
##      10.8358971      0.1651814
```