Assignment 3 - GitHub Collab

Ashley Hutchings & Tyler Yellott

2025-03-24

Box Plots and Summary Stats

```
# Set working directory to GitHub project folder
#setwd("/Users/ashleyhutchings/Desktop/R Class/Assignment 3 GitHub/Assignment3R")

# Read in CSV
txt_df<-read.csv("TextMessages.csv", header=TRUE)

# Check data
head(txt_df)
```

```
##   Group Baseline Six_months Participant
## 1     1       52         32           1
## 2     1       68         48           2
## 3     1       85         62           3
## 4     1       47         16           4
## 5     1       73         63           5
## 6     1       57         53           6
```

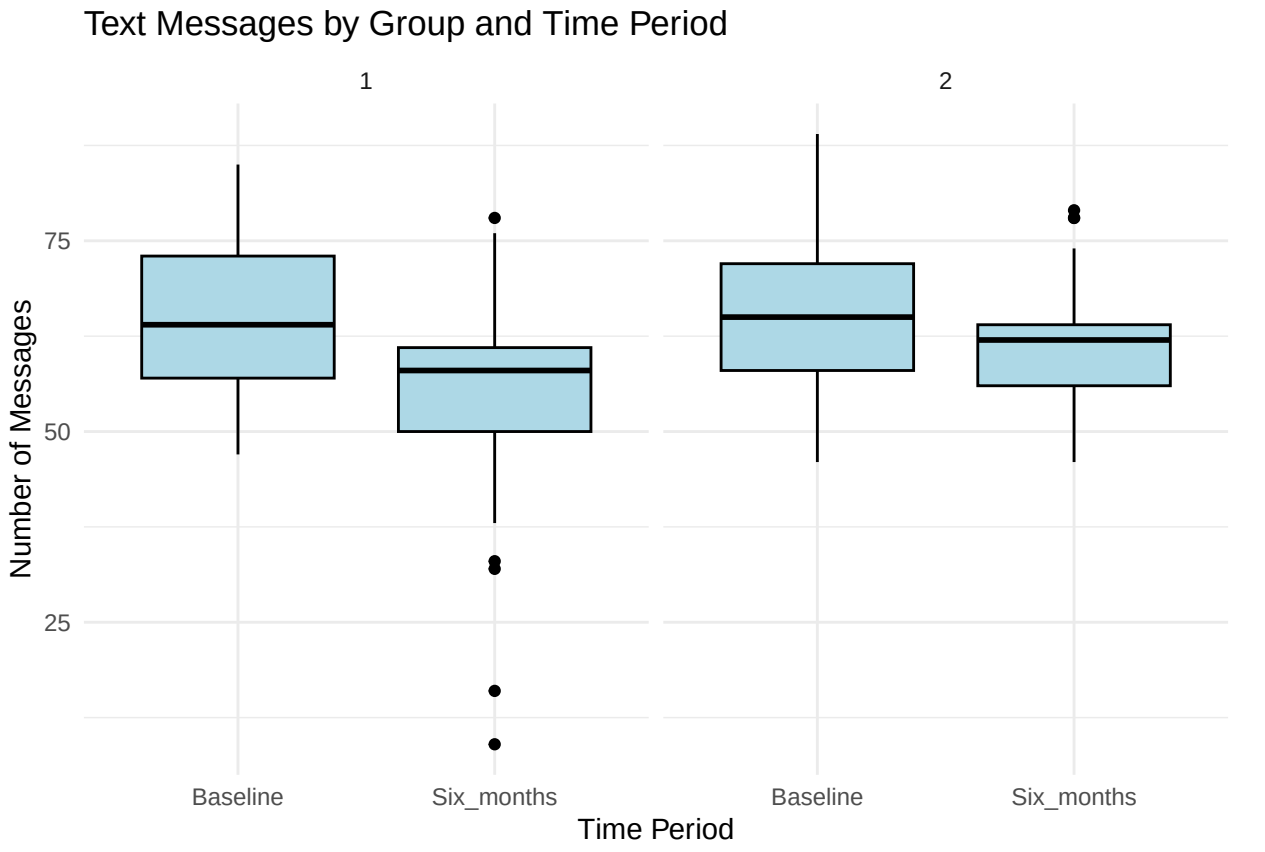
```
# Check data types
sapply(txt_df,class)
```

```
##       Group    Baseline  Six_months Participant
##   "integer"   "integer"   "integer"   "integer"
```

```
# Convert Group to Factor
txt_df$Group<-as.factor(txt_df$Group)

# Load libraries needed for functions below
library(tidyverse)
library(ggplot2)
library(tidyr)

# Pivot dataset from wide to long in order to create boxplots
txt_long <- txt_df %>% pivot_longer(cols = c(Baseline, Six_months),
               names_to = "Time",
               values_to = "Texts")

# Check data is transposed
head(txt_long)
```

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

```r
# Create boxplots using facet_wrap to group by Groupg
ggplot(txt_long, aes(x = Time, y = Texts)) +
  geom_boxplot(fill = "lightblue", color = "black") +
  facet_wrap(~ Group) +
  labs(title = "Text Messages by Group and Time Period",
       x = "Time Period",
       y = "Number of Messages") +
  theme_minimal()
```

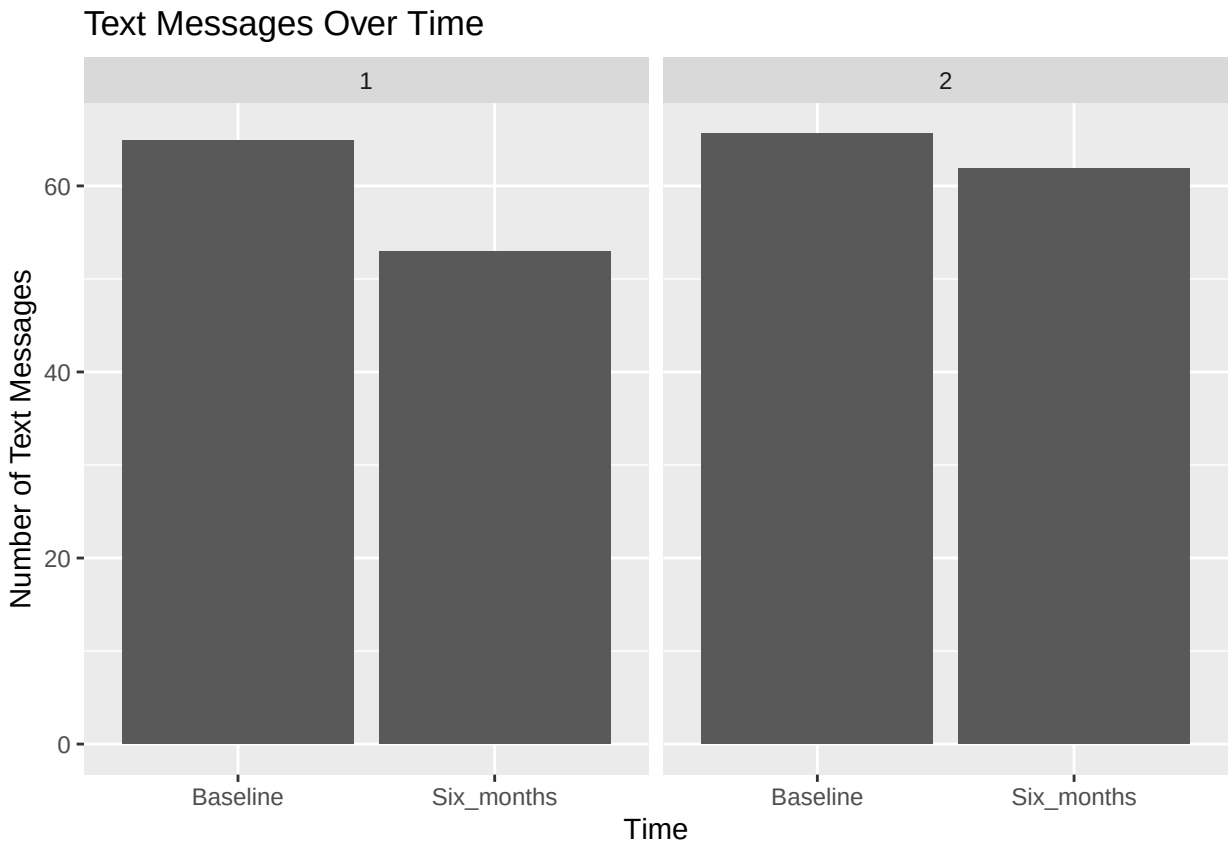



```r
# Use summarise() to define summary statistics
txt_long %>%
  group_by(Group, Time) %>%
  summarise(
    count = n(),
    mean = mean(Texts),
    median = median(Texts),
```

```
    sd = sd(Texts),
    min = min(Texts),
    max = max(Texts),
    .groups = "drop"
  )
```

```
## # A tibble: 4 x 8
##   Group Time       count  mean median    sd   min   max
##   <fct> <chr>      <int> <dbl>  <int> <dbl> <int> <int>
## 1 1     Baseline      25  64.8     64 10.7     47    85
## 2 1     Six_months    25  53.0     58 16.3      9    78
## 3 2     Baseline      25  65.6     65 10.8     46    89
## 4 2     Six_months    25  61.8     62  9.41    46    79
```

```
#Created Bar Chart

#Load libraries

library(readr)
library(tidyr)
library(ggplot2)

#Load Data

Data <- read_csv("TextMessages.csv")

#Change format

text_long <- pivot_longer(
  Data,
  cols = c(Baseline, Six_months),
  names_to = "Time",
  values_to = "TextMessages")

#Creating actual Bar Chart

ggplot(text_long, aes(x = Time, y = TextMessages)) +
  stat_summary(fun = mean, geom = "bar") +
  facet_wrap(~ Group) +
  labs(
    title = "Text Messages Over Time",
    x = "Time",
    y = "Number of Text Messages"
  )
```




```
# The bar chart compares the average number of text messages sent at baseline
# and six months for each group. It shows that the average number of messages
# decreased from baseline to six months, indicating a drop in message activity
# over time across both groups.
```