Github Collaboration

Tingwei Adeck, Kaity Trinidad, Noor Aayla

March 22, 2025

Import Data

```
library(readr)
TextMessages <- read_csv("TextMessages.csv") #1-use read_csv (parent df)
#View(TextMessages)
#head(TextMessages)
```

```
kable(head(TextMessages,5), format = "markdown", #
      caption = "Text Message Data Preview",
      table.envir = "table", align = "c")
```

Table 1: Text Message Data Preview

Group	Baseline	Six_months	Participant
1	52	32	1
1	68	48	2
1	85	62	3
1	47	16	4
1	73	63	5

Factorization (Data Wrangling)

```
str(TextMessages)
```

```
## spc_tbl_ [50 x 4] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
##  $ Group      : num [1:50] 1 1 1 1 1 1 1 1 1 1 ...
##  $ Baseline   : num [1:50] 52 68 85 47 73 57 63 50 66 60 ...
##  $ Six_months : num [1:50] 32 48 62 16 63 53 59 58 59 57 ...
##  $ Participant: num [1:50] 1 2 3 4 5 6 7 8 9 10 ...
##  - attr(*, "spec")=
##   .. cols(
##   ..   Group = col_double(),
##   ..   Baseline = col_double(),
##   ..   Six_months = col_double(),
##   ..   Participant = col_double()
##   .. )
##  - attr(*, "problems")=<externalptr>
```

```
#factor a variable en place - risky
make_factor <- function(data, var){
  refactored <- as.factor(data[[var]]) #make factor
  data[[var]] <- refactored #assign factor back to var
  data #return
}

#factor with labels scalably
#!!! to unpack the vectors and setnames to map values to labels
make_factor_scale <- function(data, var, newvarName, fact_num_vect,
                                    fact_char_vect) {
  data[[newvarName]] <- dplyr::recode(data[[var]],
                                      !!!setNames(fact_char_vect,
                                                  fact_num_vect))
  refactored <- as.factor(data[[newvarName]])
  data[[newvarName]] <- refactored
  data
}

#check the groups
unique(TextMessages$Group)
```

```
## [1] 1 2
```

```
TextMessages <- make_factor_scale(TextMessages,
                                      "Group",
                                      "Group_factor", c(1,2),
                                      c("Group 1",
                                        "Group 2"))

TextMessages <- TextMessages %>%
  dplyr::relocate("Group_factor", .after = "Group")
```

Summary Statistics

```
summary_stats_long <- TextMessages %>%
  group_by(Group_factor) %>%
  summarise(
    Baseline_Mean = mean(Baseline, na.rm = TRUE),
    Baseline_SD = sd(Baseline, na.rm = TRUE),
    Baseline_Min = min(Baseline, na.rm = TRUE),
    Baseline_Max = max(Baseline, na.rm = TRUE),
    SixMonths_Mean = mean(Six_months, na.rm = TRUE),
    SixMonths_SD = sd(Six_months, na.rm = TRUE),
    SixMonths_Min = min(Six_months, na.rm = TRUE),
    SixMonths_Max = max(Six_months, na.rm = TRUE)
  ) %>%
  pivot_longer(
    cols = -Group_factor,
    names_to = c("Timepoint", "Statistic"),
    names_sep = "_"
  )
```

Table 2: Statistical Summary for Baseline and Six Months

Group_factor	Timepoint	Statistic	value
Group 1	Baseline	Mean	64.840000
Group 1	Baseline	SD	10.679732
Group 1	Baseline	Min	47.000000
Group 1	Baseline	Max	85.000000
Group 1	SixMonths	Mean	52.960000
Group 1	SixMonths	SD	16.331156
Group 1	SixMonths	Min	9.000000
Group 1	SixMonths	Max	78.000000
Group 2	Baseline	Mean	65.600000
Group 2	Baseline	SD	10.835897
Group 2	Baseline	Min	46.000000
Group 2	Baseline	Max	89.000000
Group 2	SixMonths	Mean	61.840000
Group 2	SixMonths	SD	9.410455
Group 2	SixMonths	Min	46.000000
Group 2	SixMonths	Max	79.000000

Breakdown

At the Baseline (Initial Time Point), participants in Group 1 sent an average of 64.8 text messages, with a standard deviation of 10.7, indicating moderate variability in texting behavior. The minimum number of messages sent was 47, while the maximum was 85. Similarly, participants in Group 2 had a slightly higher mean of 65.6 messages, with a standard deviation of 10.8, suggesting a similar spread of values. The minimum number of messages in this group was 46, and the maximum was 89.

After six months, both groups experienced a decline in the number of messages sent. Group 1 saw a more substantial reduction, with a new mean of 53.0 messages (a decrease of approximately 11.8 messages from baseline). Additionally, the standard deviation increased to 16.3, suggesting that individuals in this group displayed more varied texting behavior over time. The minimum and maximum values were not explicitly provided but are likely within the range of 9 to 78 messages. In contrast, Group 2 maintained a relatively higher number of messages, with a mean of 61.8 messages (a smaller decrease of about 3.8 messages from baseline). The standard deviation was 9.41, indicating that texting behavior remained more consistent within this group. The minimum and maximum values were at least 46 and 79 messages, respectively.

Overall Interpretation

Both groups showed a decline in the number of text messages sent over time. However, Group 1 experienced a greater decrease, suggesting that participants in this group reduced their texting behavior more substantially. Additionally,the increase in standard deviation for Group 1 at six months indicates greater variability in how much texting behavior changed within this group—some participants may have significantly reduced their texting, while others maintained a more consistent level. On the other hand, Group 2 exhibited more stable texting behavior, with a smaller overall decline and less variability over time.

References

ggplot2 facet : split a plot into a matrix of panels