Hobbies and Emotional Regulation (DERS)

Krzysztof Oleksiak

2025-11-19

Introduction

Research Question: Is there a correlation between the Hobbies and the responses to the Difficulties in Emotional Regulation Scale (DERS) Questionnaire?

Data:

- Hobbies importance
  Hobbies_Imp_1 to Hobbies_Imp_8, also together as Imp_overall

  (Likert: 1 = Not at all important, . . . , 5 = Extremely important)
- Hobbies time
  Hobbies_Time_1 to Hobbies_Time_8, also together as Time_overall

  (Likert: 1 = < 1 hour/week, . . . , 7 = more than 20 hours/week)
- Mental health metric DERS_1 to DERS_16, summarized as DERS_mean.

  ("indicate how often difficulties in emotional regulation": Almost never (0-10%) (1), . . . , Almost always (91-100%) (5))

To answer the research question it is needed to:

- Compare all 16 hobby items individually along with Imp_overall and Time_overall, to the DERS
- Then summary, anova, and graph the most significant items (these are Hobbies_Time_1 and Hobbies_Imp_4).

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.0     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(mosaic)
```

```
## Registered S3 method overwritten by 'mosaic':
##   method                           from
##   fortify.SpatialPolygonsDataFrame ggplot2
```

```
##
## The 'mosaic' package masks several functions from core packages in order to add
## additional features.  The original behavior of these functions should not be affected by this.
##
## Attaching package: 'mosaic'
##
## The following object is masked from 'package:Matrix':
##
##     mean
##
## The following objects are masked from 'package:dplyr':
##
##     count, do, tally
##
## The following object is masked from 'package:purrr':
##
##     cross
##
## The following object is masked from 'package:ggplot2':
##
##     stat
##
## The following objects are masked from 'package:stats':
##
##     binom.test, cor, cor.test, cov, fivenum, IQR, median, prop.test,
##     quantile, sd, t.test, var
##
## The following objects are masked from 'package:base':
##
##     max, mean, min, prod, range, sample, sum
```

```
library(knitr)
```

Data Preparation

```
Data <- read.csv("Data/student_mental_health.csv")[1:1193, ]

ders_items <- paste0("DERS_", 1:16)
imp_items  <- paste0("Hobbies_Imp_", 1:8)
time_items <- paste0("Hobbies_Time_", 1:8)

Data_filtered <- Data %>%
  filter(Catch_question != "NA") %>%
  mutate(
    DERS_mean = rowMeans(select(., all_of(ders_items)), na.rm = TRUE),
    Imp_overall  = rowMeans(select(., all_of(imp_items)),  na.rm = TRUE),
    Time_overall = rowMeans(select(., all_of(time_items)), na.rm = TRUE)
  )

Data_analysis <- Data_filtered %>% filter(!is.na(DERS_mean))

dim(Data_analysis)
```

```
## [1] 748 151
```

Table of Hobby Means

```
all_hobby_vars <- c(imp_items, time_items, "Imp_overall", "Time_overall")

hobby_table <- tibble(
  Variable = all_hobby_vars,
  Mean     = sapply(Data_analysis[all_hobby_vars], mean, na.rm = TRUE),
  SD       = sapply(Data_analysis[all_hobby_vars], sd,  na.rm = TRUE),
  n        = sapply(Data_analysis[all_hobby_vars], function(x) sum(!is.na(x)))
)

kable(hobby_table, digits = 3,
      caption = "Means and Standard Deviations for All Hobby Variables")
```

Table 1: Means and Standard Deviations for All Hobby Variables

Variable	Mean	SD	n
Hobbies_Imp_1	2.267	1.307	748
Hobbies_Imp_2	2.116	1.091	748
Hobbies_Imp_3	2.586	1.111	748
Hobbies_Imp_4	3.229	1.017	748
Hobbies_Imp_5	2.961	1.115	748
Hobbies_Imp_6	4.167	0.798	748
Hobbies_Imp_7	2.418	0.989	748
Hobbies_Imp_8	3.270	1.063	748
Hobbies_Time_1	1.717	1.131	748
Hobbies_Time_2	1.509	0.925	748
Hobbies_Time_3	2.028	1.293	748
Hobbies_Time_4	4.162	1.388	748
Hobbies_Time_5	1.999	1.336	748
Hobbies_Time_6	5.352	1.501	748
Hobbies_Time_7	1.488	0.840	748
Hobbies_Time_8	2.485	1.370	748
Imp_overall	2.877	0.468	748
Time_overall	2.592	0.495	748

Variables Relation to DERS_mean

What is important is which hobby variables are most related to DERS_mean.
Computing a correlation between DERS_mean and each hobby variable, then look for variables with a correlation with statistically significant p-values.

```
screen_results <- tibble(
  Variable = all_hobby_vars,
  r        = NA_real_,
  p_value  = NA_real_,
  n        = NA_integer_
)

for (i in seq_along(all_hobby_vars)) {
```

```
  v   <- all_hobby_vars[i]
  tmp <- Data_analysis[, c("DERS_mean", v)]
  tmp <- tmp[complete.cases(tmp), ]
  test <- cor.test(tmp$DERS_mean, tmp[[v]])

  screen_results$r[i]       <- unname(test$estimate)
  screen_results$p_value[i] <- test$p.value
  screen_results$n[i]       <- nrow(tmp)
}

screen_results <- screen_results %>%
  mutate(abs_r = abs(r)) %>%
  arrange(desc(abs_r))

kable(screen_results,
      digits  = 3,
      caption = "Correlations Between DERS_mean and All Hobby Variables (Screening Table)")
```

Table 2: Correlations Between DERS_mean and All Hobby Variables (Screening Table)

Variable	r	p_value	n	abs_r
Hobbies_Time_1	-0.120	0.001	748	0.120
Hobbies_Imp_4	0.108	0.003	748	0.108
Hobbies_Imp_1	-0.067	0.065	748	0.067
Hobbies_Imp_8	0.067	0.066	748	0.067
Hobbies_Time_4	0.063	0.084	748	0.063
Imp_overall	0.050	0.174	748	0.050
Hobbies_Imp_3	0.046	0.206	748	0.046
Time_overall	-0.034	0.357	748	0.034
Hobbies_Time_2	-0.027	0.466	748	0.027
Hobbies_Time_5	-0.025	0.488	748	0.025
Hobbies_Time_6	-0.021	0.564	748	0.021
Hobbies_Imp_2	0.019	0.608	748	0.019
Hobbies_Imp_6	0.017	0.643	748	0.017
Hobbies_Imp_5	0.011	0.767	748	0.011
Hobbies_Time_7	0.011	0.770	748	0.011
Hobbies_Time_8	-0.005	0.884	748	0.005
Hobbies_Imp_7	-0.005	0.896	748	0.005
Hobbies_Time_3	0.003	0.944	748	0.003

To decide what to focus on, variables with a correlation (abs_r >= 0.10) and p_value < 0.05 are designated significant:

```
screen_focus <- screen_results %>%
  filter(!is.na(r),
         abs_r >= 0.10,
         p_value < 0.05)

kable(screen_focus,
      digits  = 3,
      caption = "Hobby Variables With A Correlation (abs_r >= 0.10) With DERS_mean")
```

Table 3: Hobby Variables With A Correlation (abs_r >= 0.10) With DERS_mean

Variable	r	p_value	n	abs_r
Hobbies_Time_1	-0.120	0.001	748	0.120
Hobbies_Imp_4	0.108	0.003	748	0.108

Correlations and Linear Models

- Hobbies_Time_1 - "How many hours per week do you spend participating in athletics, such as varsity sports or intramurals?"
- Hobbies_Imp_4 - "How important is watching online recreational content such as on Netflix or Youtube to you?"

Correlation Tests

```
cor_time1 <- cor.test(~ DERS_mean + Hobbies_Time_1, data = Data_analysis)
cor_imp4  <- cor.test(~ DERS_mean + Hobbies_Imp_4, data = Data_analysis)

cor_time1
```

```
##
##  Pearson's product-moment correlation
##
## data:  DERS_mean and Hobbies_Time_1
## t = -3.3048, df = 746, p-value = 0.0009959
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.1901676 -0.0488571
## sample estimates:
##        cor
## -0.1201208
```

```
cor_imp4
```

```
##
##  Pearson's product-moment correlation
##
## data:  DERS_mean and Hobbies_Imp_4
## t = 2.969, df = 746, p-value = 0.003083
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.03666589 0.17836881
## sample estimates:
##       cor
## 0.1080662
```

Simple Linear Models

```
lm_time1 <- lm(DERS_mean ~ Hobbies_Time_1, data = Data_analysis)
lm_imp4  <- lm(DERS_mean ~ Hobbies_Imp_4, data = Data_analysis)

summary(lm_time1)
```

```
##
## Call:
## lm(formula = DERS_mean ~ Hobbies_Time_1, data = Data_analysis)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.9011 -0.7136 -0.0261  0.6964  2.1840
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     3.06112    0.06065  50.471  < 2e-16 ***
## Hobbies_Time_1 -0.09752    0.02951  -3.305 0.000996 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9123 on 746 degrees of freedom
## Multiple R-squared:  0.01443,    Adjusted R-squared:  0.01311
## F-statistic: 10.92 on 1 and 746 DF,  p-value: 0.0009959
```

```
anova(lm_time1)
```

```
## Analysis of Variance Table
##
## Response: DERS_mean
##                 Df Sum Sq Mean Sq F value    Pr(>F)
## Hobbies_Time_1   1   9.09  9.0907  10.922 0.0009959 ***
## Residuals      746 620.94  0.8324
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(lm_imp4)
```

```
##
## Call:
## lm(formula = DERS_mean ~ Hobbies_Imp_4, data = Data_analysis)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.00416 -0.71903 -0.03153  0.69110  2.22623
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.57850    0.11130  23.167  < 2e-16 ***
## Hobbies_Imp_4  0.09763    0.03288   2.969  0.00308 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9136 on 746 degrees of freedom
## Multiple R-squared:  0.01168,    Adjusted R-squared:  0.01035
## F-statistic: 8.815 on 1 and 746 DF,  p-value: 0.003083
```

```
anova(lm_imp4)
```

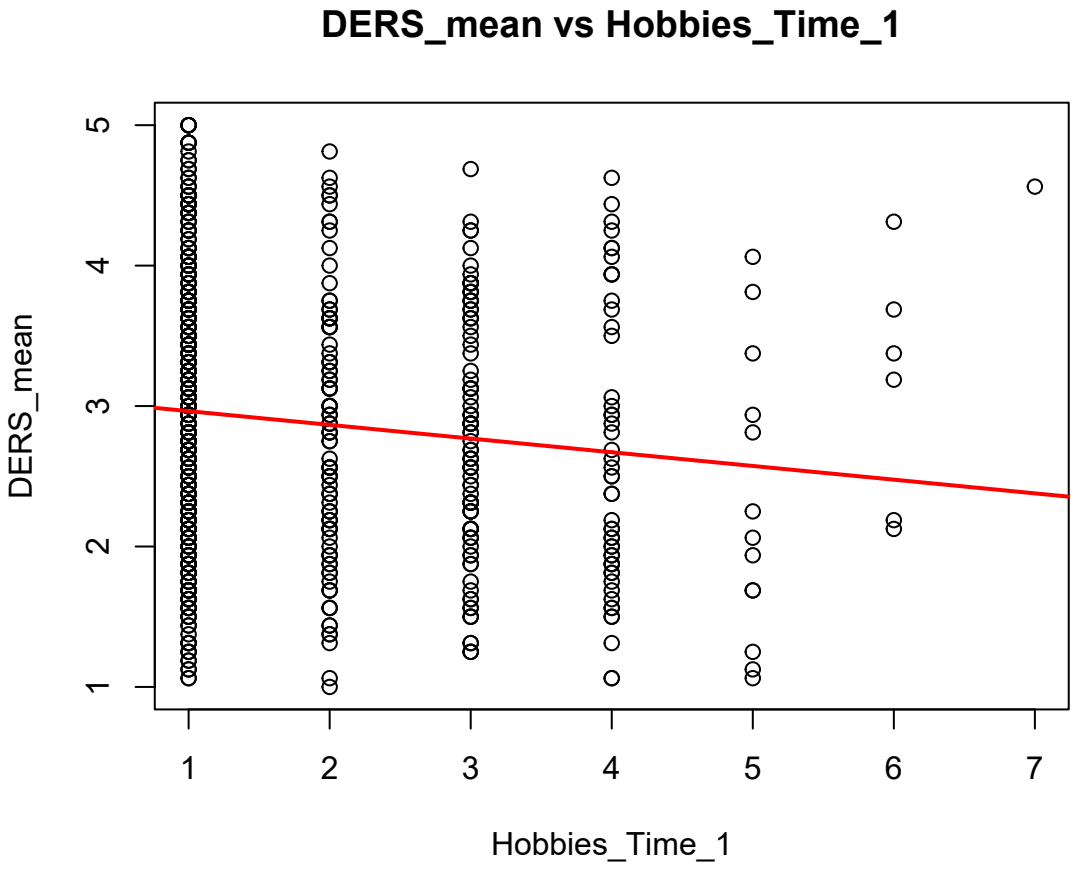
```
## Analysis of Variance Table
##
## Response: DERS_mean
```

```
##                Df Sum Sq Mean Sq F value   Pr(>F)
## Hobbies_Imp_4   1   7.36  7.3576   8.815 0.003083 **
## Residuals     746 622.67  0.8347
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Graphs

DERS_mean vs Hobbies_Time_1

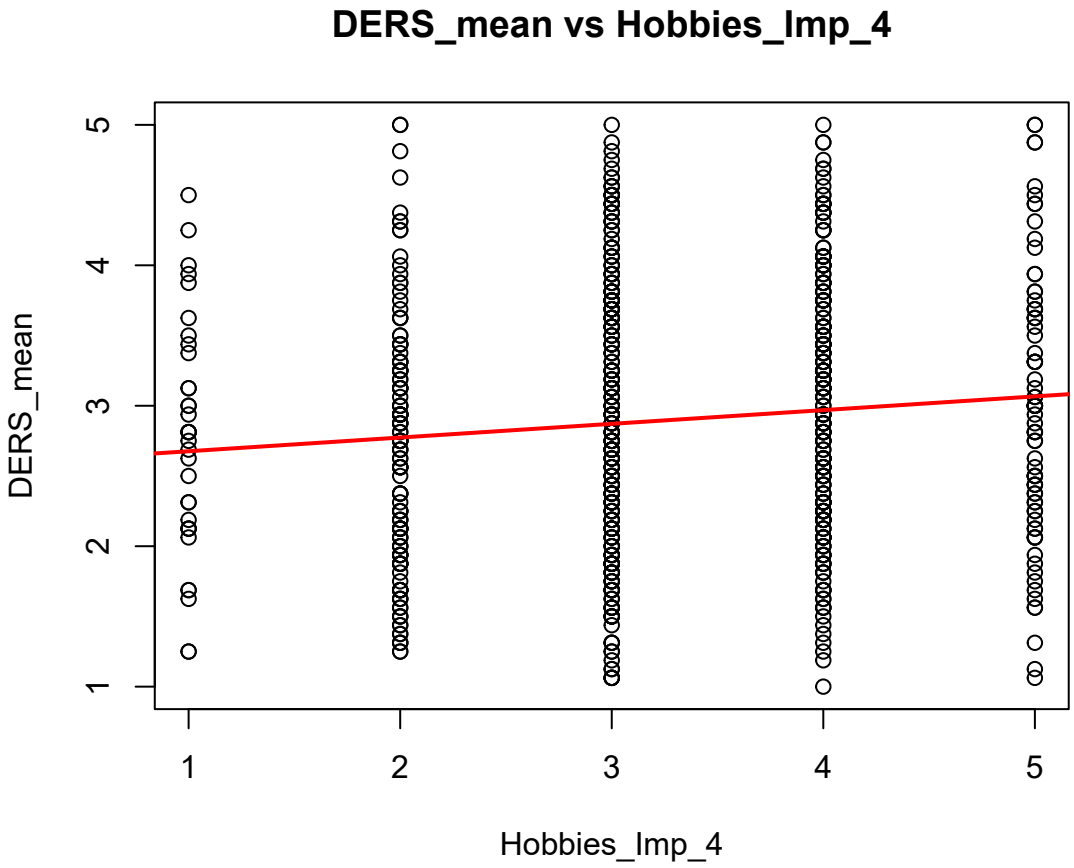
```
plot(DERS_mean ~ Hobbies_Time_1,
     data = Data_analysis,
     xlab = "Hobbies_Time_1",
     ylab = "DERS_mean",
     main = "DERS_mean vs Hobbies_Time_1")

abline(lm_time1, col = "red", lwd = 2)
```



DERS_mean vs Hobbies_Imp_4

```
plot(DERS_mean ~ Hobbies_Imp_4,
     data = Data_analysis,
     xlab = "Hobbies_Imp_4",
     ylab = "DERS_mean",
     main = "DERS_mean vs Hobbies_Imp_4")
```

```r
abline(lm_imp4, col = "red", lwd = 2)
```



Conclusion

From comparing all 16 hobby items individually along with Imp_overall and Time_overall to the DERS, two variables stood out as having small but statistically significant relationships to Difficulties in Emotional Regulation Scale (DERS). These are:

Time spent on Hobbies_Time_1, which asked "How many hours per week do you spend participating in athletics, such as varsity sports or intramurals?", showed a slight negative association with DERS. This suggests that students who spend more time in athletic activities tend to less often have difficulties in emotional regulation.

In contrast, the perceived importance of Hobbies_Imp_4, which asked "How important is watching online recreational content such as on Netflix or YouTube to you?", showed a small positive association. This suggests that students who view watching online recreational content as more important tend to more often have difficulties in emotional regulation.

While these correlations are not large, they suggest that both the amount of time invested in certain hobbies and the personal value assigned to them contribute to mental well-being.