simon-5502-11-solution

## File details

This program was written by Steve Simon on 2025-03-30 and is placed in the public domain. You can use this program any way you please.

* Data used in this program
  + [arthritis-treatments.txt](https://github.com/pmean/data/blob/main/files/arthritis-treatments.txt)
  + [data dictionary](https://github.com/pmean/data/blob/main/files/arthritis-treatments.yaml)
* Tibbles created by this program
  + oa: Original data from arthritis-treatments.txt
  + oa\_1: Renamed variables, exclude ROM variables
  + oa\_2: Pivot to a wide format
* Models created by this program
  + m1: Random effects model

library(broom)  
library(foreign)  
library(lme4)  
library(tidyverse)  
  
R.version.string

[1] "R version 4.5.0 (2025-04-11 ucrt)"

Sys.Date()

[1] "2025-04-30"

## Read oa

oa <- read\_table(  
 file="../data/arthritis-treatments.txt",  
 col\_names=TRUE,  
 col\_types="nnnnnnn")  
  
glimpse(oa)

Rows: 10  
Columns: 7  
$ Subject <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10  
$ NoROM <dbl> 35, 110, 101, 99, 126, 118, 117, 73, 95, 110  
$ NoVAS <dbl> 5.3, 2.0, 1.1, 6.3, 4.0, 0.9, 2.0, 6.1, 5.2, 2.2  
$ TENSROM <dbl> 50, 90, 110, 103, 137, 89, 70, 68, 38, 87  
$ TENSVAS <dbl> 3.8, 7.3, 3.6, 4.0, 1.9, 5.6, 6.6, 4.1, 7.7, 4.8  
$ SWDROM <dbl> 64, 120, 116, 135, 150, 100, 74, 93, 100, 73  
$ SWDVAS <dbl> 7.0, 1.6, 2.4, 0.8, 1.0, 2.0, 8.0, 4.5, 2.3, 4.0

## Question 2 (done out of order!)

Select the pain variables (those ending in VAS) and pivot the data to a wider format. Include a glimpse of the original data and the pivoted data to show that the restructuring was done properly. The original dataset should have 10 rows and 4 columns (after removing the range of motion measurements). The restructured dataset should have 30 rows and 3 columns.

oa |>  
 select(  
 Subject,  
 NoVAS,  
 TENSVAS,  
 SWDVAS) |>  
 rename(  
 no=NoVAS,  
 tens=TENSVAS,  
 swd=SWDVAS) -> oa\_1

oa\_1 |>  
 pivot\_longer(  
 cols=no:swd,  
 names\_to="treatment",  
 values\_to="vas") -> oa\_2  
  
glimpse(oa\_2)

Rows: 30  
Columns: 3  
$ Subject <dbl> 1, 1, 1, 2, 2, 2, 3, 3, 3, 4, 4, 4, 5, 5, 5, 6, 6, 6, 7, 7, …  
$ treatment <chr> "no", "tens", "swd", "no", "tens", "swd", "no", "tens", "swd…  
$ vas <dbl> 5.3, 3.8, 7.0, 2.0, 7.3, 1.6, 1.1, 3.6, 2.4, 6.3, 4.0, 0.8, …

## Question 1

You will not be graded on this, but get in the habit of drawing a few graphs and computing a few statistics that will help you better understand the dataset you are working with. Only examine variables that will be part of the further analyses.

It would also be fine to wait until after restructuring to calculate descriptive statistics.

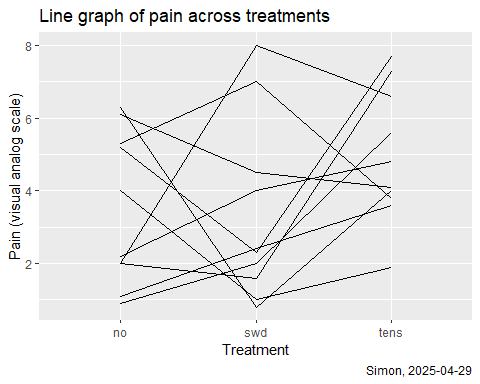
oa\_2 |>  
 group\_by(treatment) |>  
 summarize(  
 vas\_mean=mean(vas),  
 vas\_sd=sd(vas)) -> oa\_means  
  
oa\_means

# A tibble: 3 × 3  
 treatment vas\_mean vas\_sd  
 <chr> <dbl> <dbl>  
1 no 3.51 2.10  
2 swd 3.36 2.49  
3 tens 4.94 1.84

## Question 3

Draw a line graph. Do the data show a consistent pattern (e.g., patients with large values on one measurement tend to have large values on the other measurements)?

oa\_2 |>  
 ggplot() +  
 aes(  
 x = treatment,   
 y = vas,   
 group = Subject) +  
 geom\_line() +  
 labs(  
 x="Treatment",  
 y="Pain (visual analog scale)",  
 title="Line graph of pain across treatments",  
 caption="Simon, 2025-04-29") -> oa\_plot  
  
oa\_plot



## Question 4

Fit a mixed model with pain score (vas) as the dependent variable, treatment as the independent variable, and Subject as a random effect. Interpret the fixed effects. Are the t-statistics for TENS and SWD close enough to zero to conclude that there is no effect of the two treatments on pain, compared to the control group?

m1 <- lmer(  
 vas ~ treatment + (1 | Subject),  
 data=oa\_2)

boundary (singular) fit: see help('isSingular')

summary(m1)

Linear mixed model fit by REML ['lmerMod']  
Formula: vas ~ treatment + (1 | Subject)  
 Data: oa\_2  
  
REML criterion at convergence: 125  
  
Scaled residuals:   
 Min 1Q Median 3Q Max   
-1.4094 -0.6827 -0.4126 0.7800 2.1512   
  
Random effects:  
 Groups Name Variance Std.Dev.   
 Subject (Intercept) 4.133e-15 6.429e-08  
 Residual 4.652e+00 2.157e+00  
Number of obs: 30, groups: Subject, 10  
  
Fixed effects:  
 Estimate Std. Error t value  
(Intercept) 3.5100 0.6821 5.146  
treatmentswd -0.1500 0.9646 -0.156  
treatmenttens 1.4300 0.9646 1.482  
  
Correlation of Fixed Effects:  
 (Intr) trtmnts  
treatmntswd -0.707   
treatmnttns -0.707 0.500   
optimizer (nloptwrap) convergence code: 0 (OK)  
boundary (singular) fit: see help('isSingular')

## Question 5

Calculate the intraclass correlation. It is very small. What does this tell you?

icc <- 4.133e-15 /(4.133e-15 + 4.652e+00)  
  
icc

[1] 8.884351e-16