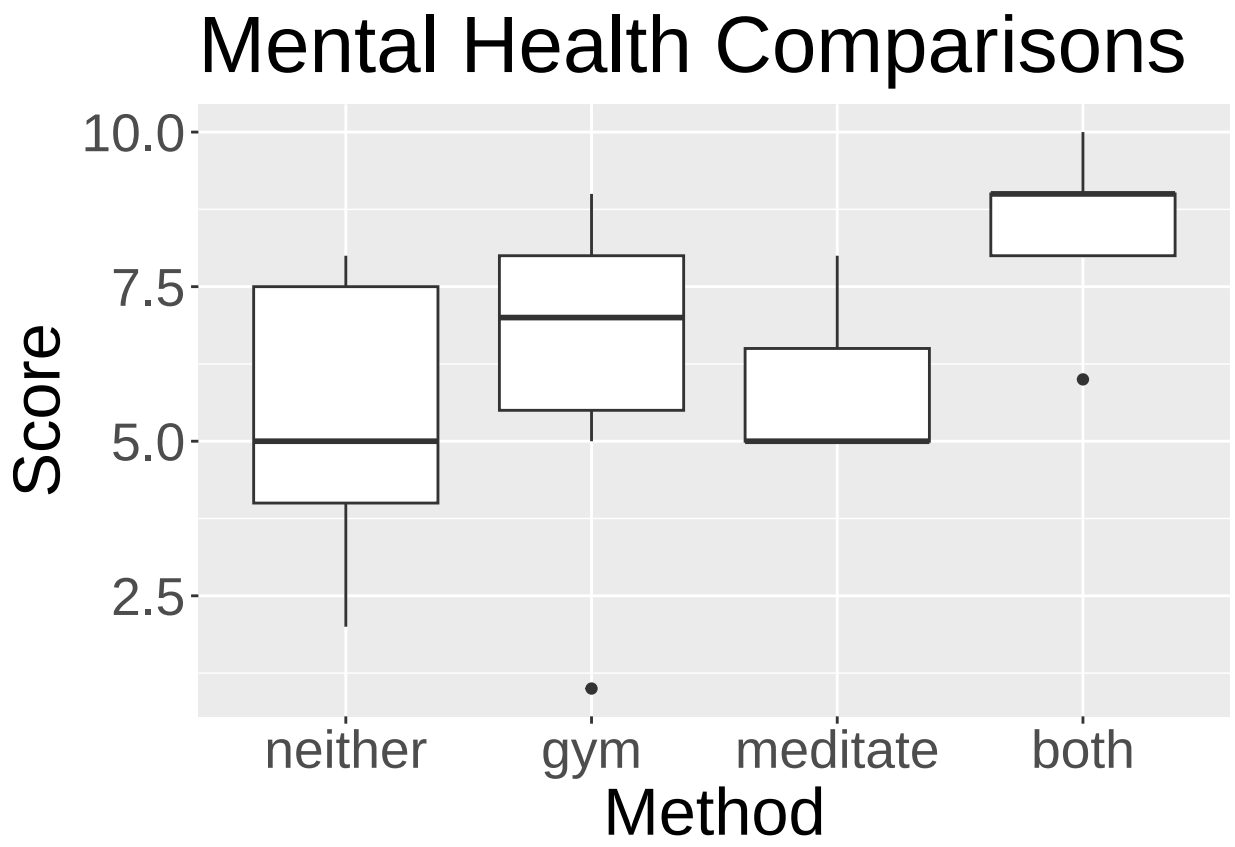
DailyCheck#12

2025-05-07

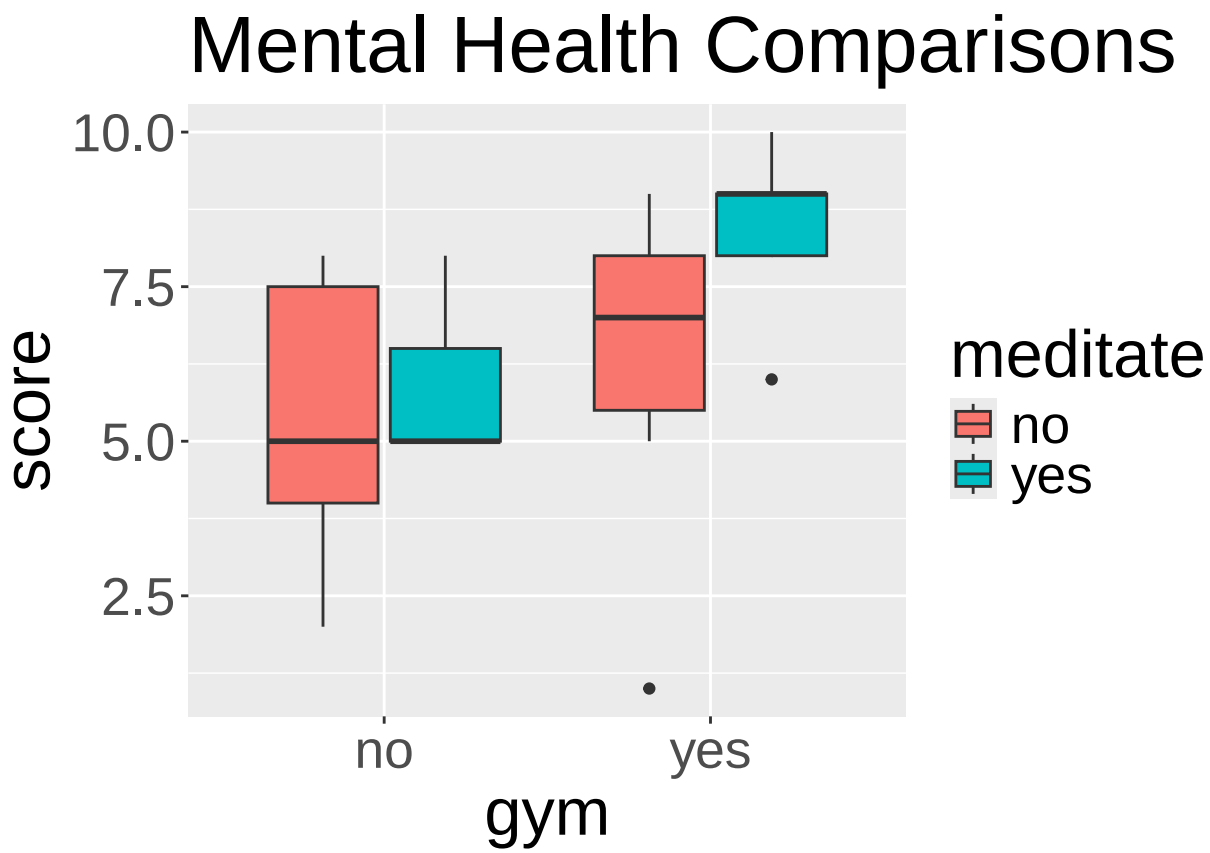
```
neither <- c(7, 4, 2, 8, 4, 5, 8)
gym <- c(5, 1, 8, 8, 9, 7, 6)
meditate <- c(7, 5, 5, 8, 5, 6, 5)
both <- c(9, 6, 8, 9, 9, 8, 10)
mh_df <- data.frame(
score = c(neither, gym, meditate, both),
gym = c(rep("no", 7), rep("yes", 7),
rep("no", 7), rep("yes", 7)),
meditate = c(rep("no", 14), rep("yes",14))
)
```

```
library(ggplot2)
theme_update(text = element_text(size = 25))
ggplot() +
geom_boxplot(aes(x="neither", y=neither)) +
geom_boxplot(aes(x="gym", y=gym)) +
geom_boxplot(aes(x="meditate", y=meditate)) +
geom_boxplot(aes(x="both", y=both)) +
scale_x_discrete(limits=c("neither", "gym", "meditate", "both")) + ggtitle("Mental Health Comparisons")
ylab("Score") +
xlab("Method")
```



Description: This boxplot compares mental health scores across four daily routines: doing nothing, going to the gym, meditating, or doing both. It's clear that the days with both gym and meditation had the highest and most consistent scores, suggesting a strong positive effect when the two activities are combined. Gym-only days also showed relatively high scores, while meditation alone had a smaller impact. The lowest scores came from days with neither activity, showing more ups and downs. Overall, it seems that combining physical activity with mindfulness may be the most helpful for boosting mental health.

```
ggplot(data=mh_df, mapping=aes(x=gym, y=score, fill=meditate)) +
geom_boxplot() + ggtitle("Mental Health Comparisons")
```



This boxplot shows how daily mental health scores vary based on whether the person went to the gym and whether they meditated. Overall, scores were highest when both activities were done together—gym and meditation days had the best outcomes. Going to the gym alone already gave a noticeable boost compared to doing nothing, and meditation helped more when added on top of that. Without gym, meditation didn't have as strong of an effect on its own. This suggests that physical activity may be more impactful, and combining it with meditation could bring the biggest benefit to mental well-being.

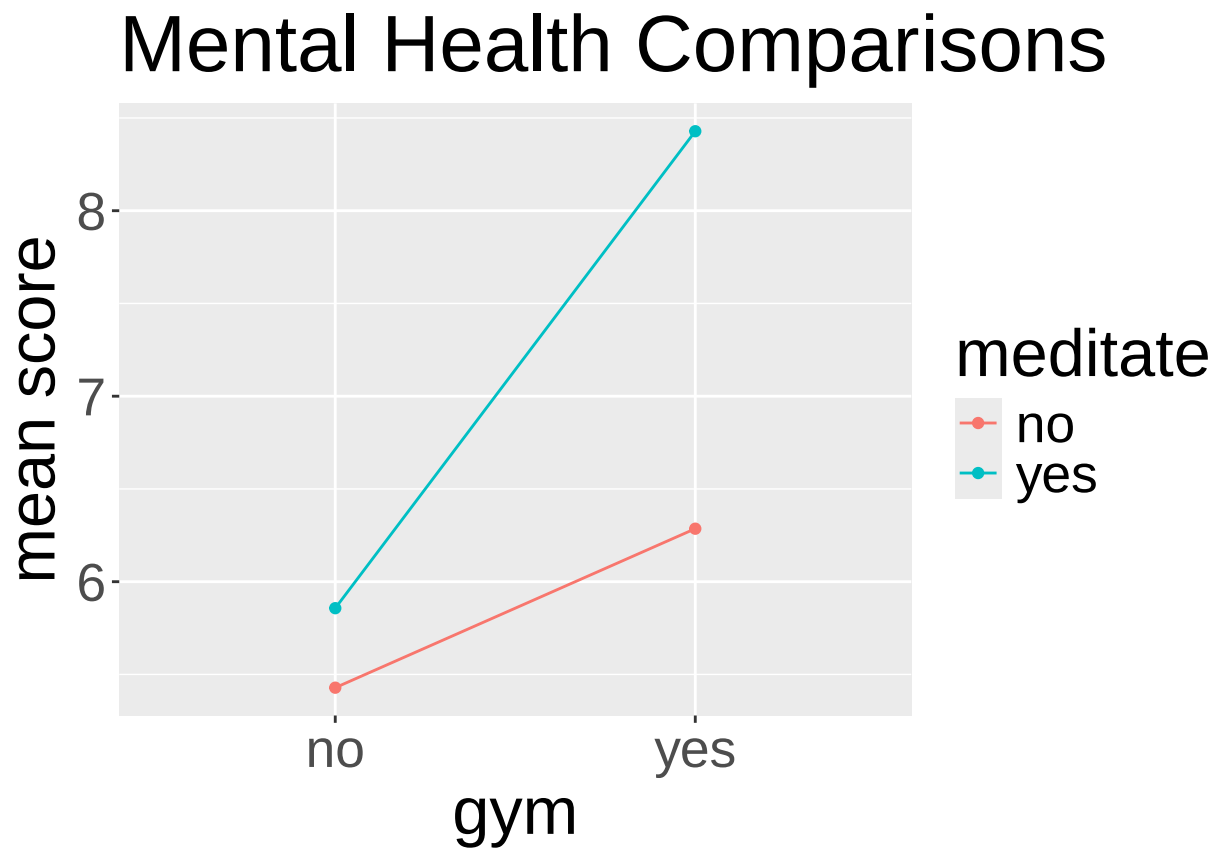
```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
tmp <- mh_df %>%
group_by(gym, meditate) %>%
summarize(mean=mean(score))
```

```
## `summarise()` has grouped output by 'gym'. You can override using the `.groups`
## argument.
```

```
ggplot(data=tmp, aes(y=mean, x=gym,
group=meditate, color=meditate)) +
geom_point() + geom_line() +
ggtitle("Mental Health Comparisons") +
ylab("mean score")
```



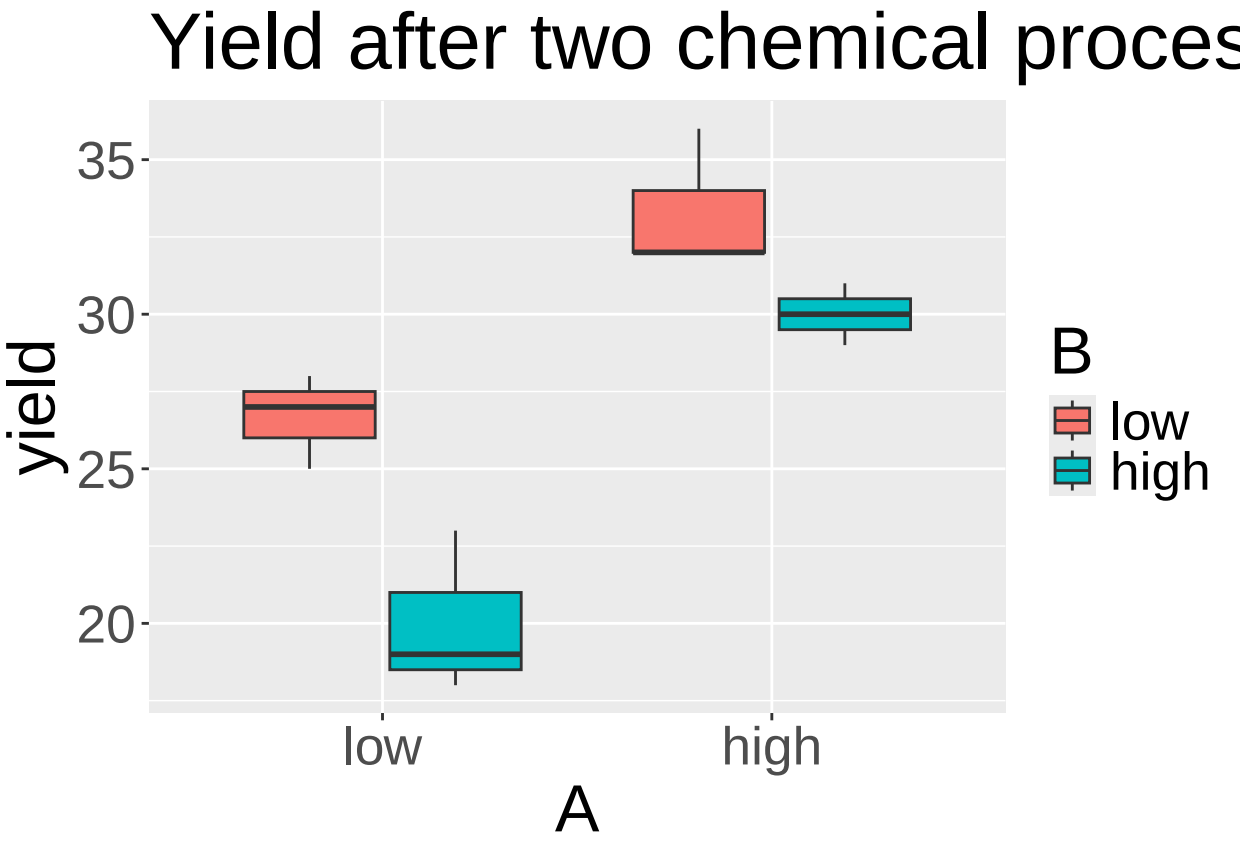
This line plot shows the average mental health scores for each combination of gym and meditation routines. We can see that going to the gym consistently improved average scores, but the improvement was especially strong when combined with meditation. The steeper slope for the meditation group ("yes") suggests an interaction effect—gym and meditation together boost mental health more than either one alone. This pattern supports the idea that combining physical activity with mindfulness practices can lead to greater mental well-being.

Your turn #1

```
library(readxl)
yield <- read_excel("yield.xlsx")
yield$A <- factor(yield$A, levels = c("low", "high"))
yield$B <- factor(yield$B, levels = c("low", "high"))
```

First graph:

```
library(ggplot2)
ggplot(data = yield, mapping = aes(x = A, y = yield, fill = B)) + geom_boxplot() + ggtitle("Yield after
```



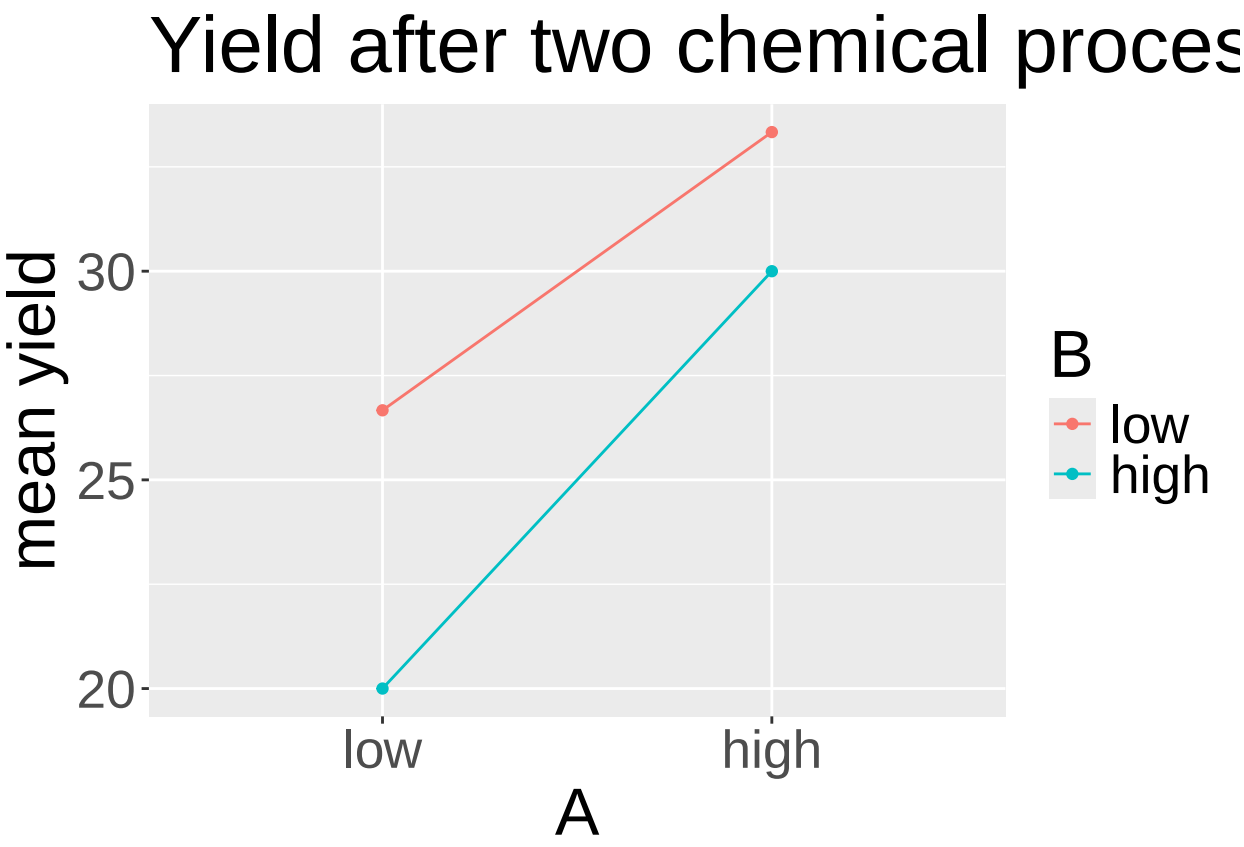
This boxplot shows how different combinations of two chemical processes (A and B) affect yield. When both A and B are set to low, the yield is somewhere in the middle. But when A is high and B is low, the yield is the highest out of all the groups. On the other hand, setting A to low and B to high gives the lowest yield. It's interesting that when both A and B are high, the yield isn't as strong as when only A is high. This suggests that the best result comes from turning up process A while keeping process B low—there seems to be some interaction between the two that affects the outcome.

Second graph:

```
library(dplyr)
tmp <- yield %>%
  group_by(A,B) %>%
  summarize(mean = mean(yield))
```

```
## `summarise()` has grouped output by 'A'. You can override using the `.groups`
## argument.
```

```
ggplot(data = tmp, mapping = aes(y = mean, x = A, group = B, color = B)) +
  geom_point() +
  geom_line() +
  ggtitle("Yield after two chemical processes") +
  ylab("mean yield")
```



This line plot shows the average yield based on different combinations of chemical processes A and B. When both A and B are low, the yield is moderate. As we increase factor A from low to high, yield increases in both cases—but more sharply when B is set to high. The lines aren't perfectly parallel, which suggests that there might be some interaction between A and B. In other words, the effect of changing A depends somewhat on the level of B. This kind of plot helps us see how the two processes might work together to influence the outcome.

Your Turn #2

```
model1 <- lm(yield ~ A + B + A*B, data = yield)
summary(model1)
```

```
##
## Call:
## lm(formula = yield ~ A + B + A * B, data = yield)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -2.000 -1.333 -0.500  1.083  3.000
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   26.667      1.143  23.338 1.21e-08 ***
## Ahigh          6.667      1.616   4.126  0.00332 **
```

```
## Bhigh         -6.667      1.616  -4.126  0.00332 **
## Ahigh:Bhigh    3.333      2.285   1.459  0.18278
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.979 on 8 degrees of freedom
## Multiple R-squared:  0.903,  Adjusted R-squared:  0.8666
## F-statistic: 24.82 on 3 and 8 DF,  p-value: 0.0002093
```

The output from the linear model shows a strong overall effect of the factors on yield, with a very small p-value for the model ($p = 0.0002093$). This means we have statistically significant evidence that at least one of the factors—A, B, or their interaction—impacts the yield.

When both process A and B are set to low, the average yield is around 26.67. If we increase A to high while keeping B low, the yield goes up by about 6.67 units—a pretty strong and statistically significant improvement. But if we increase B instead, the yield actually drops by the same amount, which is also significant. So A helps, and B hurts. The interaction between A and B isn't statistically meaningful, which suggests that the way they affect yield doesn't really depend on each other—their effects just add up.

There is strong evidence that both A and B individually affect the yield, but no strong evidence of a significant interaction between them. Increasing A helps, while increasing B hurts the yield.