140_FinalProject_Code

Qianping Wu

2024-12-05

```r
library(readxl)
library(dplyr)
library(ggplot2)
library(tidyr)
library(tidyverse)

# Load the data
data <- read_excel("NASA_Astronaut2-6-2020.xlsx", sheet = 2)
sheet1 <- read_excel("NASA_Astronaut2-6-2020.xlsx", sheet =1)
# head(data)
# colnames(data)
```

Paired t-Tests: Katherine Jin

```r
# Perform paired t-tests
t_test_MRD1_R <- t.test(data$`MRD1- R Avg (E)`, data$`MRD1- R Avg (S)`, paired = TRUE)
t_test_MRD1_L <- t.test(data$`MRD1- L Avg (E)`, data$`MRD1- L Avg (S)`, paired = TRUE)
t_test_PTB_R <- t.test(data$`PTB R (E) Avg`, data$`PTB R (S) Avg`, paired = TRUE)
t_test_PTB_L <- t.test(data$`PTB L (E) Avg`, data$`PTB L (S) Avg`, paired = TRUE)
# Print the results
print(t_test_MRD1_R)
```

```
##
##  Paired t-test
##
## data:  data$‘MRD1- R Avg (E)‘ and data$‘MRD1- R Avg (S)‘
## t = 1.0667, df = 17, p-value = 0.301
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
##  -0.1330550  0.4051994
## sample estimates:
## mean difference
##       0.1360722
```

```r
print(t_test_MRD1_L)
```

```
##
##  Paired t-test
##
## data:  data$‘MRD1- L Avg (E)‘ and data$‘MRD1- L Avg (S)‘
## t = 0.96056, df = 17, p-value = 0.3502
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
##  -0.1634864  0.4367734
## sample estimates:
## mean difference
##       0.1366435
```

```r
print(t_test_PTB_R)
```

```
##
##  Paired t-test
##
## data:  data$‘PTB R (E) Avg‘ and data$‘PTB R (S) Avg‘
## t = -12.718, df = 17, p-value = 4.114e-10
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
##  -4.467065 -3.195880
## sample estimates:
## mean difference
##       -3.831472
```

```r
print(t_test_PTB_L)
```

```
##
##  Paired t-test
##
## data:  data$‘PTB L (E) Avg‘ and data$‘PTB L (S) Avg‘
```

```
## t = -11.419, df = 17, p-value = 2.14e-09
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
##  -4.374424 -3.010085
## sample estimates:
## mean difference
##       -3.692255
```

Correlation Analysis: Jamie Tian

```
# Extract MRD and PTB columns
earth_data <- data[, c("MRD1- R Avg (E)", "MRD1- L Avg (E)", "PTB R (E) Avg", "PTB L (E) Avg")]
space_data <- data[, c("MRD1- R Avg (S)", "MRD1- L Avg (S)", "PTB R (S) Avg", "PTB L (S) Avg")]
colnames(earth_data) <- c("MRD_R", "MRD_L", "PTB_R", "PTB_L")
colnames(space_data) <- c("MRD_R", "MRD_L", "PTB_R", "PTB_L")
# Pearson correlations
cor_earth <- cor(earth_data, use = "complete.obs", method = "pearson")
cor_space <- cor(space_data, use = "complete.obs", method = "pearson")
# Correlation matrices
cat("Correlation matrix for Earth:\n")
```

```
## Correlation matrix for Earth:
```

```
print(cor_earth)
```

```
##             MRD_R      MRD_L       PTB_R       PTB_L
## MRD_R  1.00000000  0.9285134 -0.04988999 -0.05827674
## MRD_L  0.92851343  1.0000000 -0.13799337 -0.15270541
## PTB_R -0.04988999 -0.1379934  1.00000000  0.87534571
## PTB_L -0.05827674 -0.1527054  0.87534571  1.00000000
```

```
cat("\nCorrelation matrix for Space:\n")
```

```
##
## Correlation matrix for Space:
```

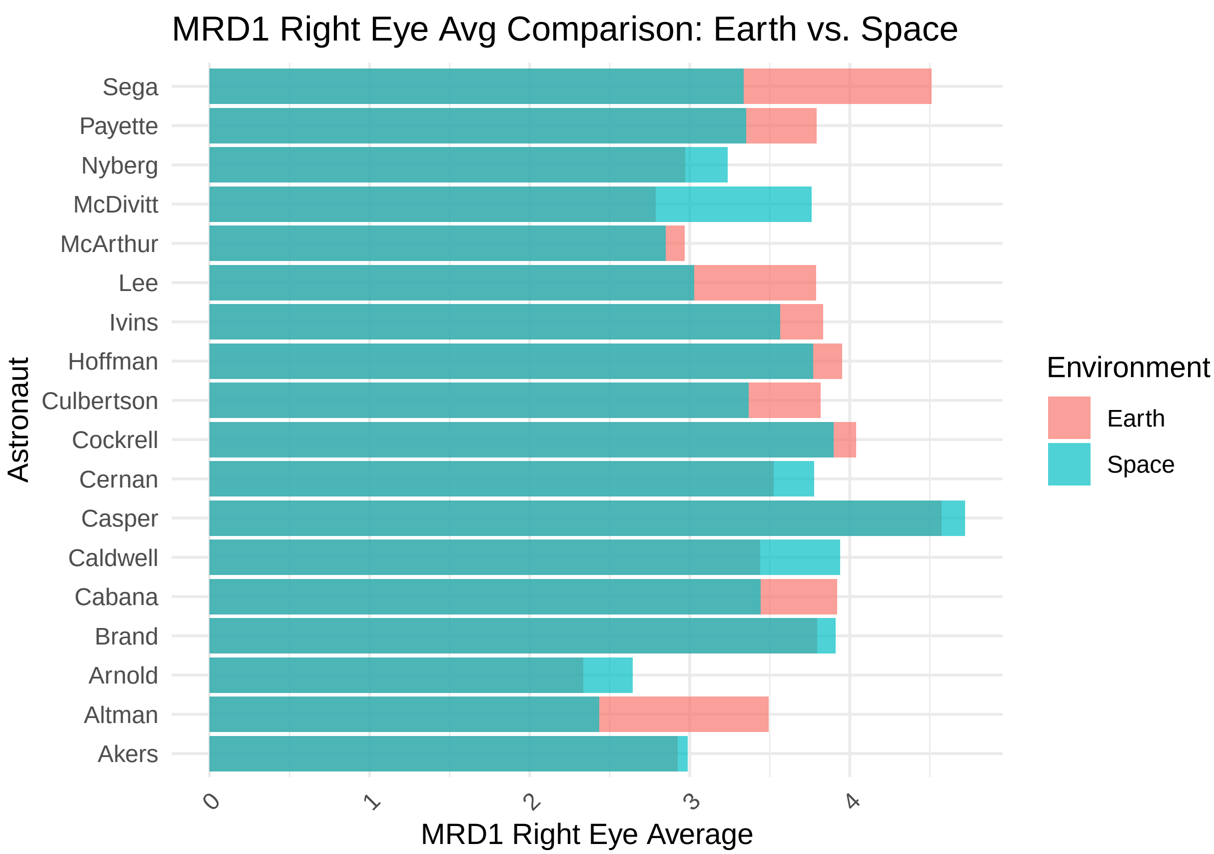
```
print(cor_space)
```

```
##              MRD_R       MRD_L        PTB_R      PTB_L
## MRD_R  1.000000000 0.838483702 -0.006570915 0.05327018
## MRD_L  0.838483702 1.000000000  0.009479456 0.10747914
## PTB_R -0.006570915 0.009479456  1.000000000 0.81881684
## PTB_L  0.053270175 0.107479142  0.818816837 1.00000000
```

Visualization

Barplot, scatterplot: Qianping Wu

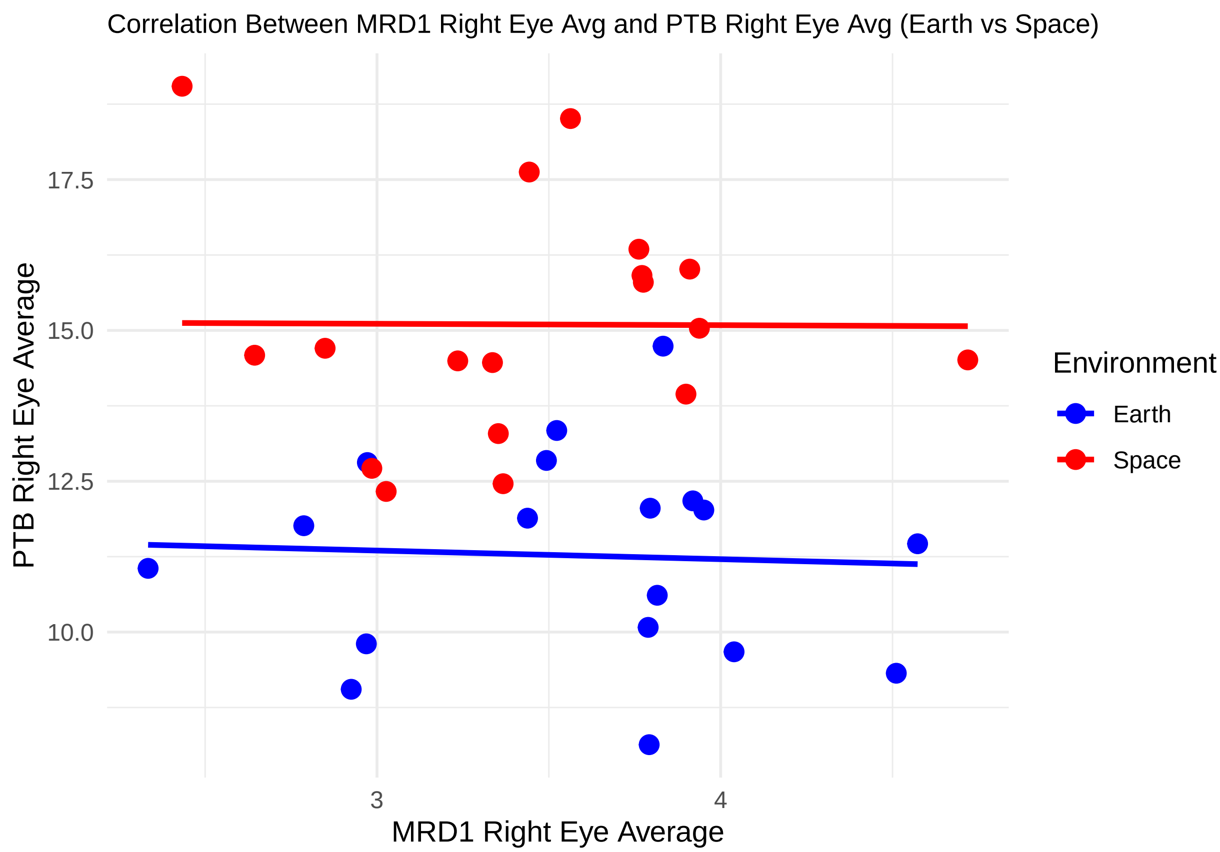
```
# Bar Plot for MRD1 Right Eye Avg (Earth vs. Space)
ggplot(data) +
  geom_bar(aes(x = Astronaut, y = `MRD1- R Avg (E)`, fill = "Earth"),
           stat = "identity", position = "dodge", alpha = 0.7) +
  geom_bar(aes(x = Astronaut, y = `MRD1- R Avg (S)`, fill = "Space"),
           stat = "identity", position = "dodge", alpha = 0.7) +
  labs(title = "MRD1 Right Eye Avg Comparison: Earth vs. Space",
       x = "Astronaut",
       y = "MRD1 Right Eye Average",
       fill = "Environment") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  coord_flip()
```




```
# Scatter Plot for Correlation Between MRD1 Right Eye Avg and PTB Right Eye Avg (Earth vs Space)
data_earth <- data %>%
  select(Astronaut, `MRD1- R Avg (E)`, `PTB R (E) Avg`) %>%
  rename(MRD1_R = `MRD1- R Avg (E)`, PTB_R = `PTB R (E) Avg`) %>%
  mutate(Environment = "Earth")

data_space <- data %>%
  select(Astronaut, `MRD1- R Avg (S)`, `PTB R (S) Avg`) %>%
```

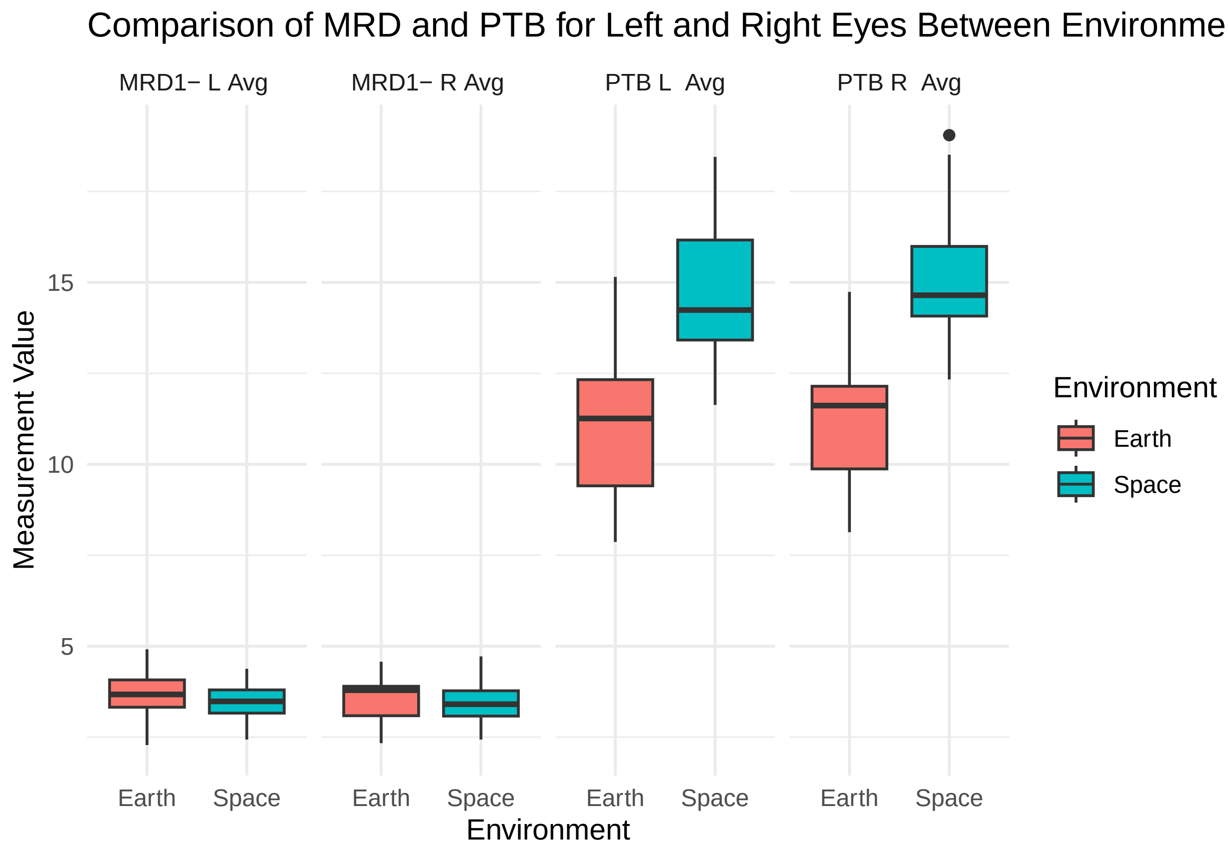
```
  rename(MRD1_R = `MRD1- R Avg (S)`, PTB_R = `PTB R (S) Avg`) %>%
  mutate(Environment = "Space")

data_combined <- bind_rows(data_earth, data_space)

ggplot(data_combined, aes(x = MRD1_R, y = PTB_R, color = Environment)) +
  geom_point(size = 3) +
  geom_smooth(method = "lm", se = FALSE, aes(color = Environment)) +
  labs(title = "Correlation Between MRD1 Right Eye Avg and PTB Right Eye Avg (Earth vs Space)",
       x = "MRD1 Right Eye Average",
       y = "PTB Right Eye Average") +
  theme_minimal() +
  theme(
    plot.title = element_text(size = 10)
  ) +
  scale_color_manual(values = c("Earth" = "blue", "Space" = "red"))
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



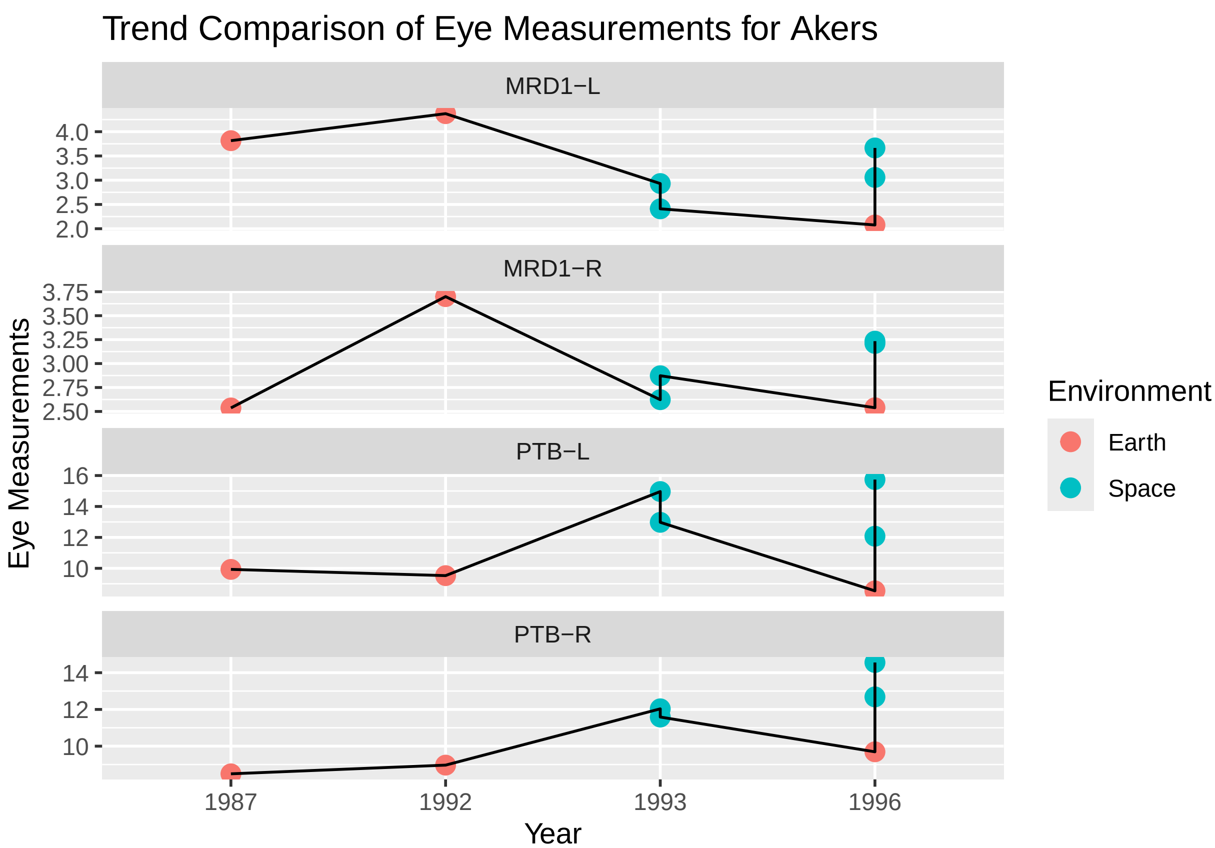
Boxplot, line plot: Evelyn Yi Tsing Ng

```
long_data <- data %>%
  pivot_longer(cols = c("MRD1- R Avg (E)", "MRD1- R Avg (S)", "MRD1- L Avg (E)", "MRD1- L Avg (S)", "PTB
               names_to = "Measurement", values_to = "Value") %>%
  mutate(
    Eye = case_when(grepl("R", Measurement) ~ "Right Eye",
                    grepl("L", Measurement) ~ "Left Eye"),
    Environment = case_when(grepl("E", Measurement) ~ "Earth",
                            grepl("S", Measurement) ~ "Space")) %>%
  mutate(Measurement = gsub("\\(E\\)|\\(S\\)", "", Measurement))

ggplot(long_data, aes(x = Environment, y = Value, fill = Environment)) +
  geom_boxplot() +
  facet_grid(. ~ Measurement) +
  labs(title = "Comparison of MRD and PTB for Left and Right Eyes Between Environments",
       x = "Environment", y = "Measurement Value") +
  theme_minimal()
```




```
data2 <- sheet1 %>%
  separate(col = "Name/Date", into = c("Name", "Year"), sep = "/'") %>%
  filter(Name == "Akers") %>%
  mutate(
    Year = paste0("19", Year),
    X = gsub("[0-9]", "", ...1),
     Environment = case_when(
      grepl("Earth", X, ignore.case = TRUE) ~ "Earth",
```

```
      grepl("Space", X, ignore.case = TRUE) ~ "Space"
    )
    ) %>%
    pivot_longer(
    cols = c("MRD1-R", "PTB-R", "MRD1-L", "PTB-L"),
    names_to = "Measurement",
    values_to = "Value"
  )

ggplot(data2, aes(x = Year, y = Value, group = Measurement)) +
  geom_point(aes(color = Environment), size = 3) +
  geom_line(color = "black") +
  facet_wrap(~ Measurement, ncol = 1, scales = "free_y") +
  labs(
    title = "Trend Comparison of Eye Measurements for Akers",
    x = "Year",
    y = "Eye Measurements"
  )
```



Variability Analysis: Xuanang Li

```r
# Calculate standard deviation and range for MRD and PTB columns
mrd_columns <- c("MRD1-R", "MRD1-L")
ptb_columns <- c("PTB-R", "PTB-L")

# Define a function to calculate range
calc_range <- function(x) {
  return(max(x, na.rm = TRUE) - min(x, na.rm = TRUE))
}

# Calculations for MRD
mrd_std <- sapply(sheet1[mrd_columns], sd, na.rm = TRUE)
mrd_range <- sapply(sheet1[mrd_columns], calc_range)

# Calculations for PTB
ptb_std <- sapply(sheet1[ptb_columns], sd, na.rm = TRUE)
ptb_range <- sapply(sheet1[ptb_columns], calc_range)

# Combine results into a data frame
variability_stats <- data.frame(
  Metric = c("Standard Deviation", "Range"),
  MRD_R = c(mrd_std["MRD1-R"], mrd_range["MRD1-R"]),
  MRD_L = c(mrd_std["MRD1-L"], mrd_range["MRD1-L"]),
  PTB_R = c(ptb_std["PTB-R"], ptb_range["PTB-R"]),
  PTB_L = c(ptb_std["PTB-L"], ptb_range["PTB-L"])
)

# Display the results
print(variability_stats)
```

```
##               Metric     MRD_R     MRD_L    PTB_R     PTB_L
## 1 Standard Deviation 0.6671125 0.7681143  2.84173  2.893673
## 2              Range 3.1300000 3.4070000 12.37700 12.733000
```

Effect Size Calculation: Yangsheng Xu

```r
t_test_MRD1_R <- t.test(data$`MRD1- R Avg (E)`, data$`MRD1- R Avg (S)`, paired = TRUE)
t_test_MRD1_L <- t.test(data$`MRD1- L Avg (E)`, data$`MRD1- L Avg (S)`, paired = TRUE)
t_test_PTB_R <- t.test(data$`PTB R (E) Avg`, data$`PTB R (S) Avg`, paired = TRUE)
t_test_PTB_L <- t.test(data$`PTB L (E) Avg`, data$`PTB L (S) Avg`, paired = TRUE)

#install.packages("effectsize")
library(effectsize)

effect_size_MRD1_R <- cohens_d(data$`MRD1- R Avg (E)`, data$`MRD1- R Avg (S)`, paired = TRUE)
```

```
## For paired samples, 'repeated_measures_d()' provides more options.
```

```r
effect_size_MRD1_L <- cohens_d(data$`MRD1- L Avg (E)`, data$`MRD1- L Avg (S)`, paired = TRUE)
```

```
## For paired samples, 'repeated_measures_d()' provides more options.
```

```r
effect_size_PTB_R <- cohens_d(data$`PTB R (E) Avg`, data$`PTB R (S) Avg`, paired = TRUE)
```

```
## For paired samples, 'repeated_measures_d()' provides more options.
```

```r
effect_size_PTB_L <- cohens_d(data$`PTB L (E) Avg`, data$`PTB L (S) Avg`, paired = TRUE)
```

```
## For paired samples, 'repeated_measures_d()' provides more options.
```

```r
print(effect_size_MRD1_R)
```

```
## Cohen's d |         95% CI
## -------------------------
## 0.25      | [-0.22, 0.72]
```

```r
print(effect_size_MRD1_L)
```

```
## Cohen's d |         95% CI
## -------------------------
## 0.23      | [-0.24, 0.69]
```

```r
print(effect_size_PTB_R)
```

```
## Cohen's d |          95% CI
## --------------------------
## -3.00     | [-4.09, -1.89]
```

```r
print(effect_size_PTB_L)
```

```
## Cohen's d |          95% CI
## --------------------------
## -2.69     | [-3.69, -1.67]
```