Ulcerative Colitis Weight Rehabilitation

Peyton Hall

04/25/2024

```
library(ggplot2)
```

```
# Create a matrix UC
UC <- matrix(NA, nrow = 14, ncol = 4)

# Assign the data to the matrix
UC <- matrix(
  c("02/24/2024", "148.6", "4 Pills (40MG)", NA,
    "02/25/2024", "150.2", "4 Pills (40MG)", NA,
    "02/26/2024", "151.6", "4 Pills (40MG)", "Visited Gym",
    "02/27/2024", "152.4", "4 Pills (40MG)", NA,
    "02/28/2024", "154.2", "4 Pills (40MG)", NA,
    "02/29/2024", "156.6", "4 Pills (40MG)", NA,
    "03/01/2024", "156.4", "4 Pills (40MG)", "Visited Gym",
    "03/02/2024", "158.0", "3 Pills (30MG)", NA,
    "03/03/2024", "158.2", "3 Pills (30MG)", NA,
    "03/04/2024", "158.8", "3 Pills (30MG)", "Visited Gym",
    "03/05/2024", "158.8", "3 Pills (30MG)", NA,
    "03/06/2024", "157.4", "3 Pills (30MG)", NA,
    "03/07/2024", "160.6", "3 Pills (30MG)", "Visited Gym",
    "03/08/2024", "158.4", "3 Pills (30MG)", "Visited Gym"),
  nrow = 14,
  byrow = TRUE
)

# Assign column names
colnames(UC) <- c("Date", "Weight", "Prednisone Dosage", "Activity")

# Print the matrix
UC
```

```
##       Date         Weight  Prednisone Dosage Activity
##  [1,] "02/24/2024" "148.6" "4 Pills (40MG)"  NA
##  [2,] "02/25/2024" "150.2" "4 Pills (40MG)"  NA
##  [3,] "02/26/2024" "151.6" "4 Pills (40MG)"  "Visited Gym"
##  [4,] "02/27/2024" "152.4" "4 Pills (40MG)"  NA
##  [5,] "02/28/2024" "154.2" "4 Pills (40MG)"  NA
##  [6,] "02/29/2024" "156.6" "4 Pills (40MG)"  NA
##  [7,] "03/01/2024" "156.4" "4 Pills (40MG)"  "Visited Gym"
##  [8,] "03/02/2024" "158.0" "3 Pills (30MG)"  NA
##  [9,] "03/03/2024" "158.2" "3 Pills (30MG)"  NA
```

```
## [10,] "03/04/2024" "158.8" "3 Pills (30MG)"  "Visited Gym"
## [11,] "03/05/2024" "158.8" "3 Pills (30MG)"  NA
## [12,] "03/06/2024" "157.4" "3 Pills (30MG)"  NA
## [13,] "03/07/2024" "160.6" "3 Pills (30MG)"  "Visited Gym"
## [14,] "03/08/2024" "158.4" "3 Pills (30MG)"  "Visited Gym"
```

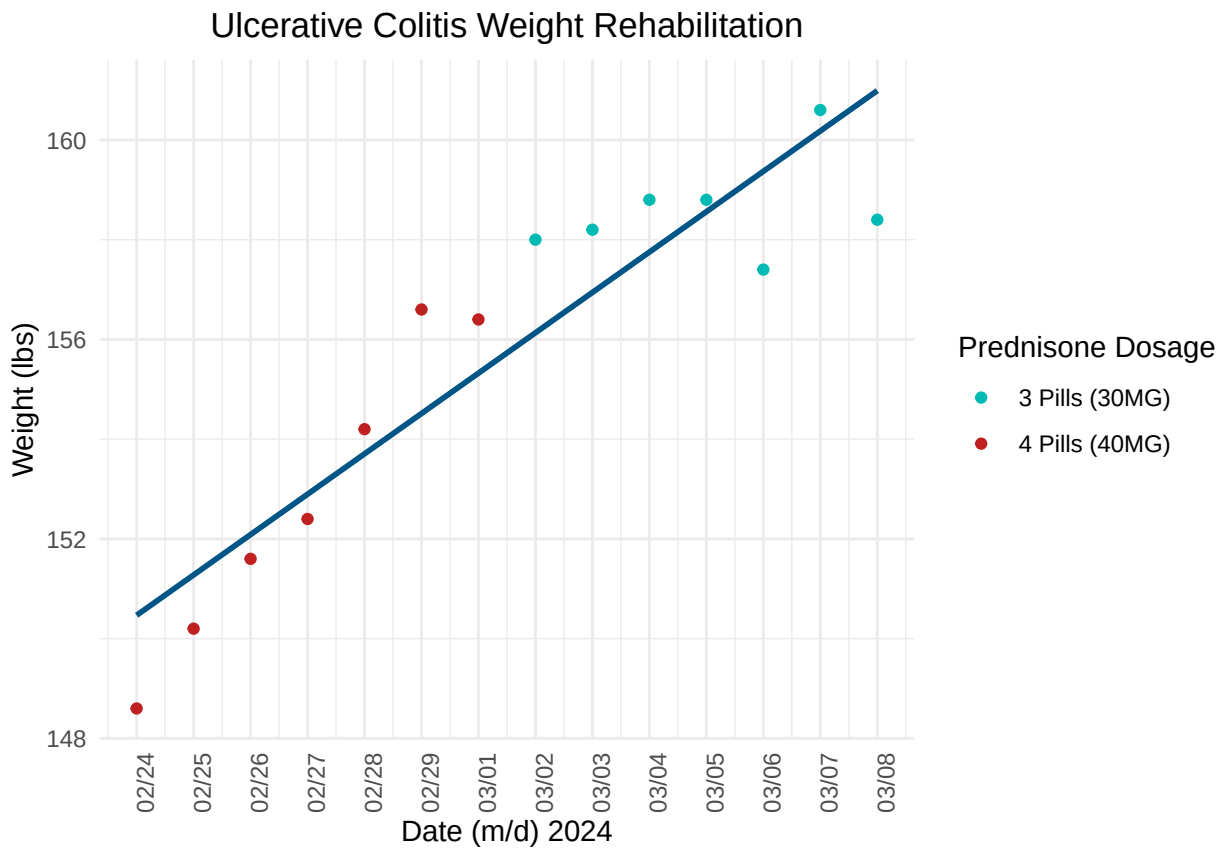
Note: Prednisone Chemical Formular = C_21 H_26 O_5

```
# Convert matrix UC to data frame
UC_df <- as.data.frame(UC)

# Convert weight column to numeric, handling non-numeric entries
UC_df$Weight <- as.numeric(ifelse(grepl("[0-9.]+", UC_df$Weight),
                                  UC_df$Weight, NA))

# Create the scatter plot with regression line using the data frame UC_df
ggplot(UC_df, aes(x = as.Date(Date, "%m/%d/%Y"), y = Weight,
                  color = as.factor(`Prednisone Dosage`))) +
  geom_point() +  # Scatter plot points
  guides(color = guide_legend(title = "Prednisone Dosage")) +
  scale_color_manual(values=c("#01BAB3", "#BF2020")) +
  geom_smooth(method = "lm", se = FALSE, color = "#005587") +  # Regression line
  labs(title = "Ulcerative Colitis Weight Rehabilitation",
       x = "Date (m/d) 2024", y = "Weight (lbs)") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 90, hjust = 1),
        plot.title = element_text(hjust = 0.5)) +  # Center the title
  scale_x_date(date_breaks = "1 day", date_labels = "%m/%d")
```

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



Hypothesis: The average weight gain rate was higher on 4 pills rather than 3.

$H_0 : \mu f = \mu t$ vs $H_a : \mu f > \mu t$ Let f = average rate of change of weight on four pills Let t = average rate of change of weight on three pills

```
UC_df$days <- c(1:14) # create a vector 1 through 14 and save it into the UC_df data frame

model.4days = lm(Weight ~ days, data = UC_df[1:7,])
model.3days = lm(Weight ~ days, data = UC_df[8:14,])

summary(model.4days)
```

```
##
## Call:
## lm(formula = Weight ~ days, data = UC_df[1:7, ])
##
## Residuals:
##        1        2        3        4        5        6        7
## -0.10000  0.11429  0.12857 -0.45714 -0.04286  0.97143 -0.61429
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 147.3143     0.4738  310.92 6.53e-12 ***
## days          1.3857     0.1059   13.08 4.66e-05 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5606 on 5 degrees of freedom
## Multiple R-squared:  0.9716, Adjusted R-squared:  0.9659
## F-statistic: 171.1 on 1 and 5 DF,  p-value: 4.662e-05
```

```
summary(model.3days)
```

```
##
## Call:
## lm(formula = Weight ~ days, data = UC_df[8:14, ])
##
## Residuals:
##        8        9       10       11       12       13       14
## -0.10714 -0.07143  0.36429  0.20000 -1.36429  1.67143 -0.69286
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 156.7929     2.1803  71.913 9.85e-09 ***
## days          0.1643     0.1950   0.842    0.438
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.032 on 5 degrees of freedom
## Multiple R-squared:  0.1243, Adjusted R-squared:  -0.05085
## F-statistic: 0.7097 on 1 and 5 DF,  p-value: 0.438
```

```
# average rate of change formula = (y2 - y1) / (x2 - x1)
# average rate of change = (weight2 - weight 1) / (time2 - time1)
f = c(1.6, 1.4, 0.8, 1.8, 2.4, -0.2)
t = c(0.2, 0.6, 0.0, -1.4, 3.2, -2.2)

# use paired t-test because there is a difference in medication per week
t.test(f, t, alternative = "greater", paired = TRUE) # (x,n,p)
```

```
##
##  Paired t-test
##
## data:  f and t
## t = 2.2526, df = 5, p-value = 0.03702
## alternative hypothesis: true mean difference is greater than 0
## 95 percent confidence interval:
##  0.1300547       Inf
## sample estimates:
## mean difference
##        1.233333
```

The probability value is 0.03702. The significance level can be defined as alpha = 0.05. The mean rate of change for the first set is greater than the mean rate of change for the second set. That is, the mean difference equals 1.233333. The H_0 should be rejected, and the H_a should be accepted. There is significant evidence that the average rate of change for four pills is greater than the average rate of change for three pills.