

# Homework3-Devore Ch10&11

Homework03Gr6

2025-03-16

**NOTE:** only those who contributes and fully participates in the work will get credit

**Scribe:**

**Moderator:**

**All contributors:**

**Attention:** If applicable, please finish all 8 problems step by step, rather than using the advance functions `aov`. However, you are encouraged to use more advance functions to check your solutions. For the problems that are difficult to implement by steps (eg. plot simultaneously CI) you can use advanced functions.

##Q1 Devore 2nd edition 10.1.6

*#Data*

```
carbonate <- c(20.5, 28.1, 27.8, 27.0, 28.0, 25.2, 25.3, 27.1, 20.5, 31.3)
```

```
silicate <- c(26.3, 24.0, 26.2, 20.2, 23.7, 34.0, 17.1, 26.8, 23.7, 24.9)
```

```
magnetite <- c(29.5, 34.0, 27.5, 29.4, 27.9, 26.2, 29.9, 29.5, 30.0, 35.6)
```

```
hematite <- c(36.5, 44.2, 34.1, 30.3, 31.4, 33.1, 34.1, 32.9, 36.3, 25.5)
```

```
total_fe <- c(carbonate, silicate, magnetite, hematite)
```

```
formation_type <- factor(rep(c("carbonate", "silicate", "magnetite", "hematite"), each = 10))
```

```
data <- data.frame(total_fe, formation_type)
```

```
mean_carbonate <- mean(carbonate)
```

```
mean_silicate <- mean(silicate)
```

```
mean_magnetite <- mean(magnetite)
```

```
mean_hematite <- mean(hematite)
```

```

grand_mean <- mean(total_fe)

#Sum of squares between groups
n_per_group <- length(carbonate) # 10 observations per group
ssb <- n_per_group * ((mean_carbonate - grand_mean)^2 +
                      (mean_silicate - grand_mean)^2 +
                      (mean_magnetite - grand_mean)^2 +
                      (mean_hematite - grand_mean)^2)

#Sum of squares within group

ssw_carbonate <- sum((carbonate - mean_carbonate)^2)
ssw_silicate <- sum((silicate - mean_silicate)^2)
ssw_magnetite <- sum((magnetite - mean_magnetite)^2)
ssw_hematite <- sum((hematite - mean_hematite)^2)

ssw <- ssw_carbonate + ssw_silicate + ssw_magnetite + ssw_hematite

#Degree of Freedom

k <- 4 # Number of groups
n <- length(total_fe) # Total number of observations

dfb <- k - 1
dfw <- n - k

#Calculate Mean Squares Between (MSB) and Mean Squares Within (MSW)
msb <- ssb / dfb
msw <- ssw / dfw

#f value
f_value <- msb / msw

#p value
p_value <- pf(f_value, dfb, dfw, lower.tail = FALSE)

#ANOVA table
cat("ANOVA Table:\n")

## ANOVA Table:

```

```

cat("Source      df      SS      MS      F      P\n")
## Source      df      SS      MS      F      P
cat("Between      ", dfb, " ", round(ssb, 2), " ", round(msb, 2), "
", round(f_value, 4), " ", format(p_value, scientific = TRUE, digits
= 3), "\n")
## Between      3      509.12      169.71      10.849      3.2e-05
cat("Within      ", dfw, " ", round(ssw, 2), " ", round(msw, 2), "
", "-\t\t -", "\n")
## Within      36      563.13      15.64      -      -
cat("Total      ", n - 1, " ", round(ssb + ssw, 2), " ", "-\t\t
-", " ", "-\t\t -", "\n")
## Total      39      1072.26      -      -      -      -

alpha <- 0.01

if (p_value < alpha) {
  cat("The p-value is less than the significance level (0.01), so we
reject the null hypothesis. There are significant differences in total
Fe among the four types of iron formations.\n")
} else {
  cat("The p-value is greater than the significance level (0.01), so
we fail to reject the null hypothesis. There is not enough evidence to
conclude that there are significant differences in total Fe among the
four types of iron formations.\n")
}

## The p-value is less than the significance level (0.01), so we
reject the null hypothesis. There are significant differences in total
Fe among the four types of iron formations.

#Given Information

#Number of mixtures (k) = 6
#Measurements per mixture (n_per_mixture) = 26
#Total observations (N) = 6 * 26 = 156
#SS (Mixture) = 5664.415
#MS (Error) = 13.929

```

```

#Degrees of Freedom (df)
#df (Mixture) = k - 1 = 6 - 1 = 5
#df (Total) = N - 1 = 156 - 1 = 155
#df (Error) = df (Total) - df (Mixture) = 155 - 5 = 150

#Sum of Squares (SS)
#MS (Mixture) = SS (Mixture) / df (Mixture)
#MS (Mixture) = 5664.415 / 5 = 1132.883
#SS (Error) = MS (Error) * df (Error)
#SS (Error) = 13.929 * 150 = 2089.35
#SS (Total) = SS (Mixture) + SS (Error)
#SS (Total) = 5664.415 + 2089.35 = 7753.765

#F-Statistic
#F = MS (Mixture) / MS (Error)
#F = 1132.883 / 13.929 = 81.332

#P value
df_mixture <- 5
df_error <- 150
f_value <- 81.332

p_value <- pf(f_value, df_mixture, df_error, lower.tail = FALSE)

cat("P-value:", p_value, "\n")
## P-value: 6.108449e-41

cat("ANOVA Table:\n")
## ANOVA Table:

cat("Source      df      SS      MS      F      P\n")
## Source      df      SS      MS      F      P

cat("Mixture      ", df_mixture, "      ", round(5664.415, 3), "      ",
round(1132.883, 3), "      ", round(81.332, 3), "      ", format(p_value,
scientific = TRUE, digits = 3), "\n")
## Mixture      5      5664.415      1132.883      81.332      6.11e-41

```

```

cat("Error      ", df_error, "      ", round(2089.35, 3), "      ",
    round(13.929, 3), "      -      -\n")

## Error      150      2089.35      13.929      -      -

cat("Total      ", 155, "      ", round(7753.765, 3), "      -      -
-\n")

## Total      155      7753.765      -      -      -

alpha <- 0.01

if (p_value < alpha) {
  cat("The P-value is less than the significance level (0.01), so we
  reject the null hypothesis.\n")
  cat("There is compelling evidence that the true average electrical
  resistivity is not the same for all six mixtures.\n")
} else {
  cat("The P-value is greater than the significance level (0.01), so
  we fail to reject the null hypothesis.\n")
  cat("There is not enough evidence to conclude that the true average
  electrical resistivity is not the same for all six mixtures.\n")
}

## The P-value is less than the significance level (0.01), so we
## reject the null hypothesis.
## There is compelling evidence that the true average electrical
## resistivity is not the same for all six mixtures.

#Data
plate_4 <- c(309.2, 409.5, 311.0, 326.5, 316.8, 349.8, 309.7)
plate_6 <- c(402.1, 347.2, 361.0, 404.5, 331.0, 348.9, 381.7)
plate_8 <- c(392.4, 366.2, 351.0, 357.1, 409.9, 367.3, 382.0)
plate_10 <- c(346.7, 452.9, 461.4, 433.1, 410.6, 384.2, 362.6)
plate_12 <- c(407.4, 441.8, 419.9, 410.7, 473.4, 441.2, 465.8)

axial_stiffness <- c(plate_4, plate_6, plate_8, plate_10, plate_12)
plate_length <- factor(rep(c("4", "6", "8", "10", "12"), each = 7))
data <- data.frame(axial_stiffness, plate_length)

mean_4 <- mean(plate_4)
mean_6 <- mean(plate_6)
mean_8 <- mean(plate_8)

```

```

mean_10 <- mean(plate_10)
mean_12 <- mean(plate_12)

grand_mean <- mean(axial_stiffness)

n_per_group <- length(plate_4) # 7 observations per group
ssb <- n_per_group * ((mean_4 - grand_mean)^2 +
                      (mean_6 - grand_mean)^2 +
                      (mean_8 - grand_mean)^2 +
                      (mean_10 - grand_mean)^2 +
                      (mean_12 - grand_mean)^2)

ssw_4 <- sum((plate_4 - mean_4)^2)
ssw_6 <- sum((plate_6 - mean_6)^2)
ssw_8 <- sum((plate_8 - mean_8)^2)
ssw_10 <- sum((plate_10 - mean_10)^2)
ssw_12 <- sum((plate_12 - mean_12)^2)

ssw <- ssw_4 + ssw_6 + ssw_8 + ssw_10 + ssw_12

k <- 5 # Number of groups
n <- length(axial_stiffness) # Total number of observations

dfb <- k - 1
dfw <- n - k

msb <- ssb / dfb
msw <- ssw / dfw

f_value <- msb / msw

p_value <- pf(f_value, dfb, dfw, lower.tail = FALSE)

cat("ANOVA Table:\n")

## ANOVA Table:

cat("Source          df      SS      MS      F      P\n")

## Source          df      SS      MS      F      P

```

```

cat("Plate Length ", dfb, " ", round(ssb, 2), " ", round(msb, 2), "
", round(f_value, 4), " ", format(p_value, scientific = TRUE, digits
= 3), "\n")

## Plate Length      4      43992.55      10998.14      10.4827      1.96e-05

cat("Error          ", dfw, " ", round(ssw, 2), " ", round(msw, 2), "
", "-\t\t -", "\n")

## Error              30      31475.03      1049.17      -          -

cat("Total          ", n - 1, " ", round(ssb + ssw, 2), " ", "-\t\t
-", " ", "-\t\t -", "\n")

## Total              34      75467.58      -          -          -          -

alpha <- 0.01

if (p_value < alpha) {
  cat("The p-value is less than the significance level (0.01), so we
reject the null hypothesis.\n")
  cat("There is evidence that variation in plate length has an effect
on true average axial stiffness.\n")
} else {
  cat("The p-value is greater than the significance level (0.01), so
we fail to reject the null hypothesis.\n")
  cat("There is not enough evidence to conclude that variation in
plate length has an effect on true average axial stiffness.\n")
}

## The p-value is less than the significance level (0.01), so we
reject the null hypothesis.
## There is evidence that variation in plate length has an effect on
true average axial stiffness.

#Data
hormone_1 <- c(13, 17, 7, 14)
hormone_2 <- c(21, 13, 20, 17)
hormone_3 <- c(18, 15, 20, 17)
hormone_4 <- c(7, 11, 18, 10)
hormone_5 <- c(6, 11, 15, 8)

plant_growth <- c(hormone_1, hormone_2, hormone_3, hormone_4,
hormone_5)

```

```

hormone_type <- factor(rep(c("1", "2", "3", "4", "5"), each = 4))
data <- data.frame(plant_growth, hormone_type)

mean_1 <- mean(hormone_1)
mean_2 <- mean(hormone_2)
mean_3 <- mean(hormone_3)
mean_4 <- mean(hormone_4)
mean_5 <- mean(hormone_5)

grand_mean <- mean(plant_growth)

n_per_group <- length(hormone_1) # 4 observations per group
ssb <- n_per_group * ((mean_1 - grand_mean)^2 +
                      (mean_2 - grand_mean)^2 +
                      (mean_3 - grand_mean)^2 +
                      (mean_4 - grand_mean)^2 +
                      (mean_5 - grand_mean)^2)

ssw_1 <- sum((hormone_1 - mean_1)^2)
ssw_2 <- sum((hormone_2 - mean_2)^2)
ssw_3 <- sum((hormone_3 - mean_3)^2)
ssw_4 <- sum((hormone_4 - mean_4)^2)
ssw_5 <- sum((hormone_5 - mean_5)^2)

ssw <- ssw_1 + ssw_2 + ssw_3 + ssw_4 + ssw_5

k <- 5 # Number of groups
n <- length(plant_growth) # Total number of observations

dfb <- k - 1
dfw <- n - k

msb <- ssb / dfb
msw <- ssw / dfw

f_value <- msb / msw

p_value <- pf(f_value, dfb, dfw, lower.tail = FALSE)

```



```

cat("ANOVA Table:\n")
## ANOVA Table:
cat("Source          df      SS      MS      F      P\n")
## Source          df      SS      MS      F      P
cat("Hormone Type  ", dfb, "    ", round(ssb, 2), "    ", round(msb, 2), "
", round(f_value, 4), "    ", format(p_value, scientific = TRUE, digits
= 3), "\n")
## Hormone Type    4      200.3    50.07    3.4855    3.34e-02
cat("Error          ", dfw, "    ", round(ssw, 2), "    ", round(msw, 2), "
", "-\t\t -", "\n")
## Error          15      215.5    14.37    -      -
cat("Total          ", n - 1, "    ", round(ssb + ssw, 2), "    ", "-\t\t
-", "    ", "-\t\t -", "\n")
## Total          19      415.8    -      -      -      -
alpha <- 0.05

if (p_value < alpha) {
  cat("The p-value is less than the significance level (0.05), so we
reject the null hypothesis.\n")
  cat("There is evidence that different hormone types affect plant
growth.\n")
} else {
  cat("The p-value is greater than the significance level (0.05), so
we fail to reject the null hypothesis.\n")
  cat("There is not enough evidence to conclude that different hormone
types affect plant growth.\n")
}

## The p-value is less than the significance level (0.05), so we
reject the null hypothesis.
## There is evidence that different hormone types affect plant growth.

q_value <- qtukey(0.95, k, dfw) # Get the studentized range statistic
msw_value <- msw

```

```

n_per_group <- length(hormone_1)
hsd <- q_value * sqrt(msw_value / n_per_group)

cat("Tukey's HSD:", hsd, "\n")

## Tukey's HSD: 8.276175

means <- c(mean_1, mean_2, mean_3, mean_4, mean_5)
hormone_names <- c("1", "2", "3", "4", "5")

cat("Pairwise Comparisons:\n")

## Pairwise Comparisons:

for (i in 1:(k - 1)) {
  for (j in (i + 1):k) {
    diff <- abs(means[i] - means[j])
    if (diff > hsd) {
      cat("Hormone", hormone_names[i], "vs Hormone", hormone_names[j],
": Significant difference (", diff, " > ", hsd, ")\n")
    } else {
      cat("Hormone", hormone_names[i], "vs Hormone", hormone_names[j],
": No significant difference (", diff, " <= ", hsd, ")\n")
    }
  }
}

## Hormone 1 vs Hormone 2 : No significant difference ( 5 <=
8.276175 )
## Hormone 1 vs Hormone 3 : No significant difference ( 4.75 <=
8.276175 )
## Hormone 1 vs Hormone 4 : No significant difference ( 1.25 <=
8.276175 )
## Hormone 1 vs Hormone 5 : No significant difference ( 2.75 <=
8.276175 )
## Hormone 2 vs Hormone 3 : No significant difference ( 0.25 <=
8.276175 )
## Hormone 2 vs Hormone 4 : No significant difference ( 6.25 <=
8.276175 )
## Hormone 2 vs Hormone 5 : No significant difference ( 7.75 <=
8.276175 )
## Hormone 3 vs Hormone 4 : No significant difference ( 6 <=
8.276175 )

```

```

## Hormone 3 vs Hormone 5 : No significant difference ( 7.5  <=
8.276175 )
## Hormone 4 vs Hormone 5 : No significant difference ( 1.5  <=
8.276175 )

#Data
brand_1 <- c(7.9, 6.2, 6.6, 8.6, 8.9, 10.1, 9.6)
brand_2 <- c(5.7, 7.5, 9.8, 6.1, 8.4)
brand_3 <- c(6.8, 7.5, 5.0, 7.4, 5.3, 6.1)
brand_4 <- c(6.4, 7.1, 7.9, 4.5, 5.0, 4.0)

folacin_content <- c(brand_1, brand_2, brand_3, brand_4)
brand <- factor(rep(c("1", "2", "3", "4"), c(length(brand_1),
length(brand_2), length(brand_3), length(brand_4))))
data <- data.frame(folacin_content, brand)

anova_result <- aov(folacin_content ~ brand, data = data)
summary(anova_result)

##              Df Sum Sq Mean Sq F value Pr(>F)
## brand          3  23.50   7.832    3.749 0.0276 *
## Residuals     20  41.78   2.089
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

alpha <- 0.05
summary_table <- summary(anova_result)
f_value <- summary_table[[1]]$`F value`[1]
p_value <- summary_table[[1]]$`Pr(>F)`[1]

cat("F-value:", f_value, "\n")

## F-value: 3.74933

cat("P-value:", p_value, "\n")

## P-value: 0.02755167

if (p_value < alpha) {
  cat("The p-value is less than the significance level (0.05), so we
reject the null hypothesis. There are significant differences in
folacin content among the brands.\n")
} else {

```

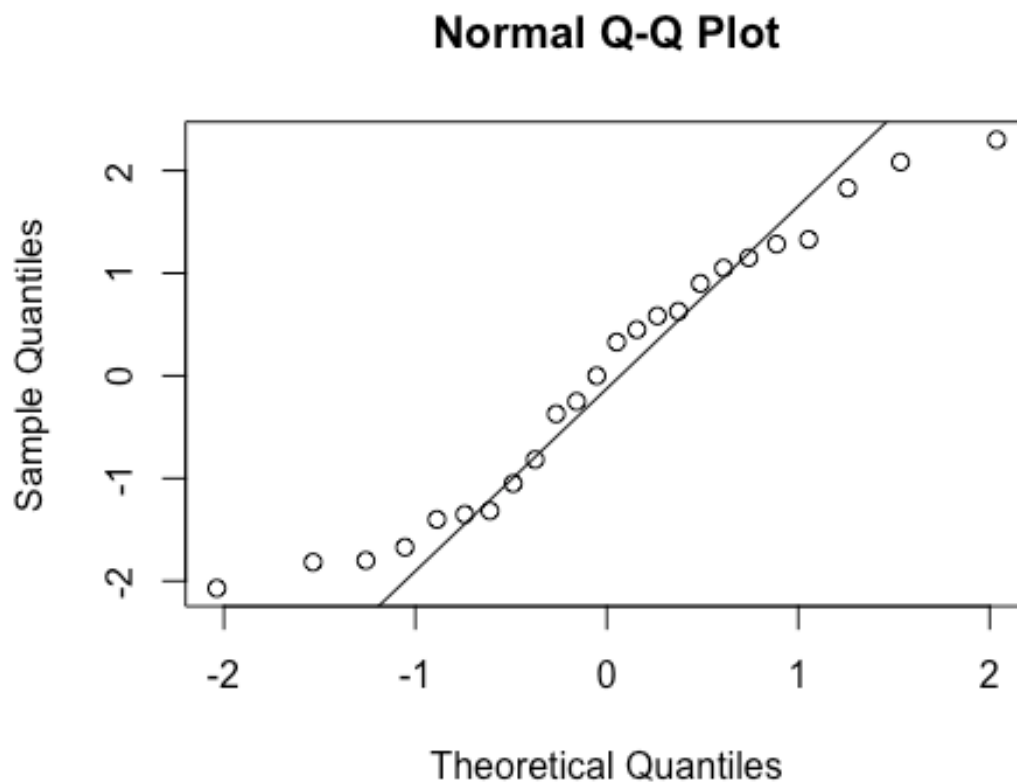
```
cat("The p-value is greater than the significance level (0.05), so  
we fail to reject the null hypothesis. There is not enough evidence to  
conclude that there are significant differences in folacin content  
among the brands.\n")  
}
```

```
## The p-value is less than the significance level (0.05), so we  
reject the null hypothesis. There are significant differences in  
folacin content among the brands.
```

```
##Assess assumptions
```

```
# Normality of Residuals
```

```
qqnorm(residuals(anova_result))  
qqline(residuals(anova_result))
```



```

# Homogeneity of Variances (Levene's Test)
library(car)

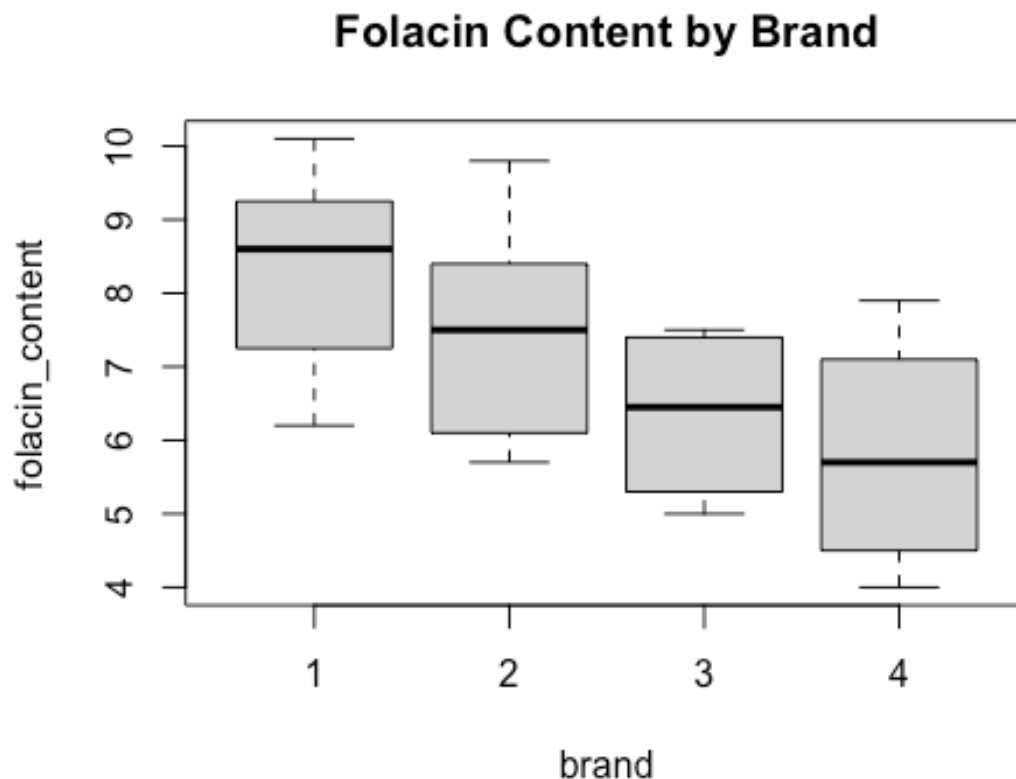
## Loading required package: carData

leveneTest(folacin_content ~ brand, data = data)

## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group  3  0.4328 0.7318
##      20

# Boxplots for visual inspection
boxplot(folacin_content ~ brand, data = data, main = "Folacin Content
by Brand")

```



```

tukey_result <- TukeyHSD(anova_result)
print(tukey_result)

##      Tukey multiple comparisons of means
##      95% family-wise confidence level
##
## Fit: aov(formula = folacin_content ~ brand, data = data)
##
## $brand
##           diff           lwr           upr           p adj
## 2-1 -0.7714286 -3.140109    1.5972517 0.7989522
## 3-1 -1.9214286 -4.172021    0.3291639 0.1115525
## 4-1 -2.4547619 -4.705354   -0.2041694 0.0294477
## 3-2 -1.1500000 -3.599546    1.2995457 0.5650524
## 4-2 -1.6833333 -4.132879    0.7662124 0.2501669
## 4-3 -0.5333333 -2.868884    1.8022169 0.9180643

cat("Tukey's HSD Test Results:\n")

## Tukey's HSD Test Results:

print(tukey_result)

##      Tukey multiple comparisons of means
##      95% family-wise confidence level
##
## Fit: aov(formula = folacin_content ~ brand, data = data)
##
## $brand
##           diff           lwr           upr           p adj
## 2-1 -0.7714286 -3.140109    1.5972517 0.7989522
## 3-1 -1.9214286 -4.172021    0.3291639 0.1115525
## 4-1 -2.4547619 -4.705354   -0.2041694 0.0294477
## 3-2 -1.1500000 -3.599546    1.2995457 0.5650524
## 4-2 -1.6833333 -4.132879    0.7662124 0.2501669
## 4-3 -0.5333333 -2.868884    1.8022169 0.9180643

cat("\nInterpretation:\n")

##
## Interpretation:

```

```

for (i in 1:nrow(tukey_result$brand)) {
  if (tukey_result$brand[i, 4] < 0.05) {
    cat(rownames(tukey_result$brand)[i], "shows a significant
difference.\n")
  } else {
    cat(rownames(tukey_result$brand)[i], "shows no significant
difference.\n")
  }
}

## 2-1 shows no significant difference.
## 3-1 shows no significant difference.
## 4-1 shows a significant difference.
## 3-2 shows no significant difference.
## 4-2 shows no significant difference.
## 4-3 shows no significant difference.

paint_data <- matrix(c(54, 46, 51, 46, 44, 47, 39, 42, 44, 44, 37,
43), nrow = 4, byrow = TRUE)
rownames(paint_data) <- c("Paint 1", "Paint 2", "Paint 3", "Paint 4")
colnames(paint_data) <- c("Roller 1", "Roller 2", "Roller 3")

#a. ANNOVA table
# Calculate totals
paint_totals <- rowSums(paint_data)
roller_totals <- colSums(paint_data)
grand_total <- sum(paint_data)

# Calculate means
paint_means <- paint_totals / ncol(paint_data)
roller_means <- roller_totals / nrow(paint_data)
grand_mean <- grand_total / (nrow(paint_data) * ncol(paint_data))

# Calculate sums of squares
ss_paint <- ncol(paint_data) * sum((paint_means - grand_mean)^2)
ss_roller <- nrow(paint_data) * sum((roller_means - grand_mean)^2)
ss_total <- sum((paint_data - grand_mean)^2)
ss_error <- ss_total - ss_paint - ss_roller

# Calculate degrees of freedom
df_paint <- nrow(paint_data) - 1

```

```

df_roller <- ncol(paint_data) - 1
df_error <- (nrow(paint_data) - 1) * (ncol(paint_data) - 1)
df_total <- nrow(paint_data) * ncol(paint_data) - 1

# Calculate mean squares
ms_paint <- ss_paint / df_paint
ms_roller <- ss_roller / df_roller
ms_error <- ss_error / df_error

# Calculate F-statistics
f_paint <- ms_paint / ms_error
f_roller <- ms_roller / ms_error

# Create ANOVA table
anova_table <- data.frame(
  Source = c("Paint Brand", "Roller Brand", "Error", "Total"),
  df = c(df_paint, df_roller, df_error, df_total),
  SS = c(ss_paint, ss_roller, ss_error, ss_total),
  MS = c(ms_paint, ms_roller, ms_error, NA),
  F = c(f_paint, f_roller, NA, NA)
)

print(anova_table)

##           Source df          SS          MS          F
## 1  Paint Brand   3 159.58333 53.194444 7.848361
## 2  Roller Brand   2  38.00000 19.000000 2.803279
## 3      Error     6  40.66667  6.777778      NA
## 4      Total    11 238.25000      NA      NA

#b) Test Hypotheses for Paint Brand
alpha <- 0.05
p_value_paint <- pf(f_paint, df_paint, df_error, lower.tail = FALSE)

cat("Paint Brand Test:\n")

## Paint Brand Test:

cat("F-statistic:", f_paint, "\n")

## F-statistic: 7.848361

```



```

cat("P-value:", p_value_paint, "\n")

## P-value: 0.01686359

if (p_value_paint < alpha) {
  cat("The p-value is less than the significance level (0.05), so we
reject the null hypothesis.\n")
  cat("There is evidence that paint brand has an effect on coverage.
\n")
} else {
  cat("The p-value is greater than the significance level (0.05), so
we fail to reject the null hypothesis.\n")
  cat("There is not enough evidence to conclude that paint brand has
an effect on coverage.\n")
}

## The p-value is less than the significance level (0.05), so we
reject the null hypothesis.
## There is evidence that paint brand has an effect on coverage.

#c) Test Hypotheses for Roller Brand
p_value_roller <- pf(f_roller, df_roller, df_error, lower.tail =
FALSE)

cat("\nRoller Brand Test:\n")

##
## Roller Brand Test:

cat("F-statistic:", f_roller, "\n")

## F-statistic: 2.803279

cat("P-value:", p_value_roller, "\n")

## P-value: 0.1381476

if (p_value_roller < alpha) {
  cat("The p-value is less than the significance level (0.05), so we
reject the null hypothesis.\n")
  cat("There is evidence that roller brand has an effect on coverage.
\n")
} else {
  cat("The p-value is greater than the significance level (0.05), so

```

```

we fail to reject the null hypothesis.\n")
  cat("There is not enough evidence to conclude that roller brand has
an effect on coverage.\n")
}

## The p-value is greater than the significance level (0.05), so we
fail to reject the null hypothesis.
## There is not enough evidence to conclude that roller brand has an
effect on coverage.

#d) Tukey's Method
q_value <- qtkey(0.95, nrow(paint_data), df_error)
hsd <- q_value * sqrt(ms_error / ncol(paint_data))

cat("\nTukey's HSD:", hsd, "\n")

##
## Tukey's HSD: 7.358493

# Pairwise comparisons
paint_means_original <- paint_means + 400
paint_names <- rownames(paint_data)

cat("\nPairwise Comparisons (Original Scale):\n")

##
## Pairwise Comparisons (Original Scale):

for (i in 1:(nrow(paint_data) - 1)) {
  for (j in (i + 1):nrow(paint_data)) {
    diff <- abs(paint_means_original[i] - paint_means_original[j])
    if (diff > hsd) {
      cat(paint_names[i], "vs", paint_names[j], ": Significant
difference (", diff, " > ", hsd, ")\n")
    } else {
      cat(paint_names[i], "vs", paint_names[j], ": No significant
difference (", diff, " <= ", hsd, ")\n")
    }
  }
}

## Paint 1 vs Paint 2 : No significant difference ( 4.666667 <=
7.358493 )

```

```
## Paint 1 vs Paint 3 : Significant difference ( 8.666667 > 7.358493
)
## Paint 1 vs Paint 4 : Significant difference ( 9 > 7.358493 )
## Paint 2 vs Paint 3 : No significant difference ( 4 <= 7.358493 )
## Paint 2 vs Paint 4 : No significant difference ( 4.333333 <=
7.358493 )
## Paint 3 vs Paint 4 : No significant difference ( 0.3333333 <=
7.358493 )
```

```
#Data
```

```
method_A <- c(30.7, 29.1, 30.0, 31.9, 30.5, 26.9, 28.2, 32.4, 26.6,
28.6)
```

```
method_B <- c(33.7, 30.6, 32.2, 34.6, 33.0, 29.3, 28.4, 32.4, 29.5,
29.4)
```

```
method_C <- c(30.5, 32.6, 30.5, 33.5, 32.4, 27.8, 30.7, 33.6, 29.2,
33.2)
```

```
compressive_strength <- c(method_A, method_B, method_C)
```

```
curing_method <- factor(rep(c("A", "B", "C"), each = 10))
```

```
data <- data.frame(compressive_strength, curing_method)
```

```
anova_result <- aov(compressive_strength ~ curing_method, data = data)
summary(anova_result)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## curing_method  2   23.23   11.614    2.829 0.0766 .
## Residuals    27  110.84    4.105
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
alpha <- 0.05
```

```
summary_table <- summary(anova_result)
```

```
f_value <- summary_table[[1]]$`F value`[1]
```

```
p_value <- summary_table[[1]]$`Pr(>F)`[1]
```

```
cat("F-value:", f_value, "\n")
```

```
## F-value: 2.829237
```

```
cat("P-value:", p_value, "\n")
```

```
## P-value: 0.07664189
```

```

if (p_value < alpha) {
  cat("The p-value is less than the significance level (0.05), so we
reject the null hypothesis. There are significant differences in
compressive strength among the three curing methods.\n")
} else {
  cat("The p-value is greater than the significance level (0.05), so
we fail to reject the null hypothesis. There is not enough evidence to
conclude that there are significant differences in compressive
strength among the three curing methods.\n")
}

```

```

## The p-value is greater than the significance level (0.05), so we
fail to reject the null hypothesis. There is not enough evidence to
conclude that there are significant differences in compressive
strength among the three curing methods.

```

*#Assess assumptions by Anova*

```

alpha <- 0.05
summary_table <- summary(anova_result)
f_value <- summary_table[[1]]$`F value`[1]
p_value <- summary_table[[1]]$`Pr(>F)`[1]

```

```

cat("F-value:", f_value, "\n")

```

```

## F-value: 2.829237

```

```

cat("P-value:", p_value, "\n")

```

```

## P-value: 0.07664189

```

```

if (p_value < alpha) {
  cat("The p-value is less than the significance level (0.05), so we
reject the null hypothesis. There are significant differences in
compressive strength among the three curing methods.\n")
} else {
  cat("The p-value is greater than the significance level (0.05), so
we fail to reject the null hypothesis. There is not enough evidence to
conclude that there are significant differences in compressive
strength among the three curing methods.\n")
}

```

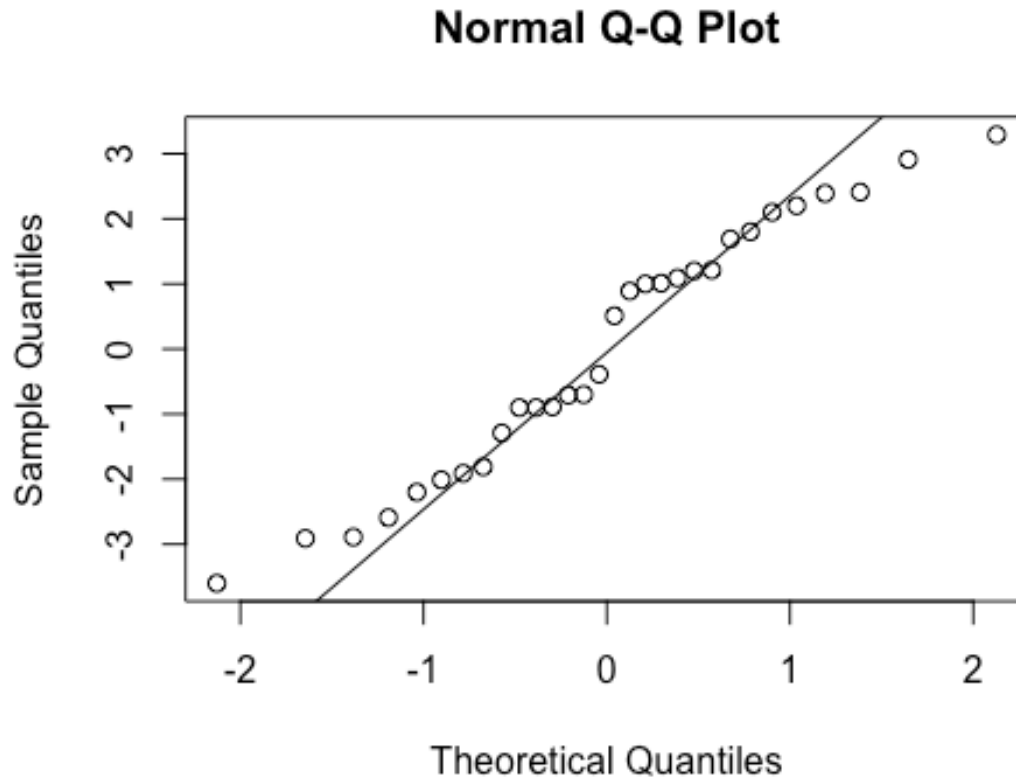
```

## The p-value is greater than the significance level (0.05), so we
fail to reject the null hypothesis. There is not enough evidence to

```

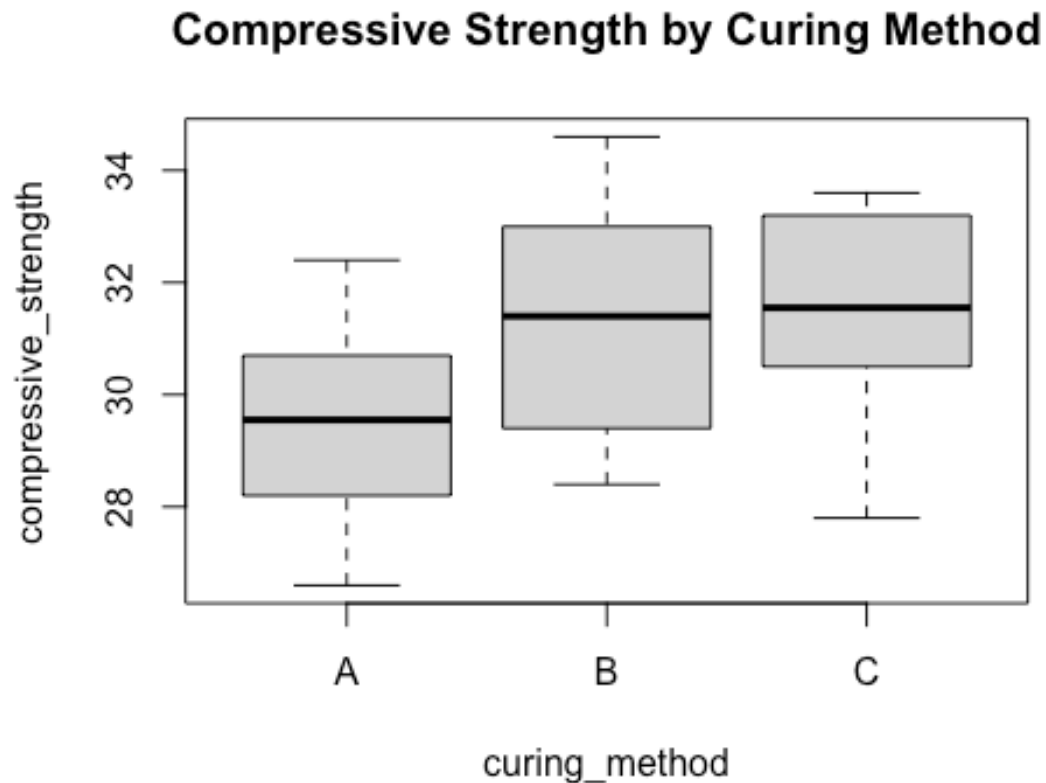
conclude that there are significant differences in compressive strength among the three curing methods.

```
# Normality of Residuals  
qqnorm(residuals(anova_result))  
qqline(residuals(anova_result))
```



```
# Homogeneity of Variances (Levene's Test)  
library(car)  
leveneTest(compressive_strength ~ curing_method, data = data)  
  
## Levene's Test for Homogeneity of Variance (center = median)  
##           Df F value Pr(>F)  
## group    2    0.226 0.7992  
##           27
```

```
# Boxplots for visual inspection
boxplot(compressive_strength ~ curing_method, data = data, main =
"Compressive Strength by Curing Method")
```



```
#Multiple comparison results
tukey_result <- TukeyHSD(anova_result)
print(tukey_result)

## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = compressive_strength ~ curing_method, data =
data)
##
## $curing_method
##      diff      lwr      upr      p adj
```

```

## B-A 1.82 -0.4266072 4.066607 0.1294154
## C-A 1.91 -0.3366072 4.156607 0.1070553
## C-B 0.09 -2.1566072 2.336607 0.9945767

cat("Tukey's HSD Test Results:\n")

## Tukey's HSD Test Results:

print(tukey_result)

## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = compressive_strength ~ curing_method, data =
data)
##
## $curing_method
##      diff      lwr      upr      p adj
## B-A 1.82 -0.4266072 4.066607 0.1294154
## C-A 1.91 -0.3366072 4.156607 0.1070553
## C-B 0.09 -2.1566072 2.336607 0.9945767

cat("\nInterpretation:\n")

##
## Interpretation:

for (i in 1:nrow(tukey_result$curing_method)) {
  if (tukey_result$curing_method[i, 4] < 0.05) {
    cat(rownames(tukey_result$curing_method)[i], "shows a significant
difference.\n")
  } else {
    cat(rownames(tukey_result$curing_method)[i], "shows no significant
difference.\n")
  }
}

## B-A shows no significant difference.
## C-A shows no significant difference.
## C-B shows no significant difference.

#given data

#Factor A: Curing Time (3 levels)

```

```

#Factor B: Type of Mix (4 levels)
#Observations per combination: 3
SSA <- 30763.0
SSB <- 34185.6
SSE <- 97436.8
SST <- 205966.6
# $\bar{x}_{1..} = 4010.88$ ,  $\bar{x}_{2..} = 4029.10$ ,  $\bar{x}_{3..} = 3960.$ 

#Degrees of Freedom (df)

# $df_A = a - 1 = 3 - 1 = 2$ 
# $df_B = b - 1 = 4 - 1 = 3$ 
# $df_{AB} = df_A * df_B = 2 * 3 = 6$ 
# $df_E = N - ab = 36 - 12 = 24$ 
# $df_T = N - 1 = 36 - 1 = 35$ 
#Sum of Squares Interaction (SSAB)
# $SSAB = SST - SSA - SSB - SSE$ 
# $SSAB = 205,966.6 - 30,763.0 - 34,185.6 - 97,436.8$ 
SSAB <- 43581.2

#Mean Squares (MS)
# $MSA = SSA / df_A$ 
MSA <- 15381.5
# $MSB = SSB / df_B$ 
MSB <- 11395.2
# $MSAB = SSAB / df_{AB}$ 
MSAB <- 7263.53
# $MSE = SSE / df_E$ 
MSE <- 4059.87

#F-Statistics

# $FA = MSA / MSE$ 
FA <- 3.79
# $FB = MSB / MSE$ 
FB <- 2.81
# $FAB = MSAB / MSE$ 
FAB <- 1.79

dfA <- 2

```



```

dfB <- 3
dfAB <- 6
dfT <- 35
dfE <- 24
FA <- 3.79
FB <- 2.81
FAB <- 1.79

p_value_A <- pf(FA, dfA, dfE, lower.tail = FALSE)
p_value_B <- pf(FB, dfB, dfE, lower.tail = FALSE)
p_value_AB <- pf(FAB, dfAB, dfE, lower.tail = FALSE)

cat("P-value for Factor A:", p_value_A, "\n")
## P-value for Factor A: 0.03711841

cat("P-value for Factor B:", p_value_B, "\n")
## P-value for Factor B: 0.06102411

cat("P-value for Interaction:", p_value_AB, "\n")
## P-value for Interaction: 0.1437743

#a. ANNOVA TABLE

anova_table <- data.frame(
  Source = c("Curing Time (A)", "Type of Mix (B)", "Interaction (AB)",
"Error", "Total"),
  df = c(dfA, dfB, dfAB, dfE, dfT),
  SS = c(SSA, SSB, SSAB, SSE, SST),
  MS = c(MSA, MSB, MSAB, MSE, NA),
  F = c(FA, FB, FAB, NA, NA),
  P = c(p_value_A, p_value_B, p_value_AB, NA, NA)
)

print(anova_table)

##           Source df          SS          MS      F          P
## 1 Curing Time (A)  2  30763.0 15381.50  3.79 0.03711841
## 2 Type of Mix (B)  3   34185.6 11395.20  2.81 0.06102411
## 3 Interaction (AB)  6   43581.2   7263.53  1.79 0.14377427

```

```
## 4          Error 24  97436.8  4059.87   NA        NA
## 5          Total 35 205966.6        NA   NA        NA
```

#### *#b. Test interaction (AB)*

```
alpha <- 0.05
if (p_value_AB < alpha) {
  cat("The p-value for interaction is less than 0.05. Reject H0AB.
There is a significant interaction.\n")
} else {
  cat("The p-value for interaction is greater than 0.05. Fail to
reject H0AB. There is no significant interaction.\n")
}
```

```
## The p-value for interaction is greater than 0.05. Fail to reject
H0AB. There is no significant interaction.
```

#### *#c. Test factor A (curing time)*

```
if (p_value_A < alpha) {
  cat("The p-value for Factor A is less than 0.05. Reject H0A. There
is a significant effect of curing time.\n")
} else {
  cat("The p-value for Factor A is greater than 0.05. Fail to reject
H0A. There is no significant effect of curing time.\n")
}
```

```
## The p-value for Factor A is less than 0.05. Reject H0A. There is a
significant effect of curing time.
```

#### *#d. Test Factor B (Type of Mix)*

```
if (p_value_B < alpha) {
  cat("The p-value for Factor B is less than 0.05. Reject H0B. There
is a significant effect of type of mix.\n")
} else {
  cat("The p-value for Factor B is greater than 0.05. Fail to reject
H0B. There is no significant effect of type of mix.\n")
}
```

```
## The p-value for Factor B is greater than 0.05. Fail to reject H0B.
There is no significant effect of type of mix.
```

#### *#e. Tukey's Procedure for Curing Times*

```

means <- c(4010.88, 4029.10, 3960.02)
names <- c("Curing Time 1", "Curing Time 2", "Curing Time 3")
k <- 3
n <- 4 * 3 # observations per curing time

q_value <- qtukey(0.95, k, dfE)
hsd <- q_value * sqrt(MSE / n)

cat("Tukey's HSD:", hsd, "\n")

## Tukey's HSD: 64.96043

# Pairwise comparisons
cat("\nPairwise Comparisons:\n")

##
## Pairwise Comparisons:

for (i in 1:(k - 1)) {
  for (j in (i + 1):k) {
    diff <- abs(means[i] - means[j])
    if (diff > hsd) {
      cat(names[i], "vs", names[j], ": Significant difference (",
diff, " > ", hsd, ")\n")
    } else {
      cat(names[i], "vs", names[j], ": No significant difference (",
diff, " <= ", hsd, ")\n")
    }
  }
}

## Curing Time 1 vs Curing Time 2 : No significant difference ( 18.22
<= 64.96043 )
## Curing Time 1 vs Curing Time 3 : No significant difference ( 50.86
<= 64.96043 )
## Curing Time 2 vs Curing Time 3 : Significant difference ( 69.08 >
64.96043 )

```