SA2

PAUL JOAQUIN DELOS SANTOS

2024-12-09

Introduction

This report analyzes the number of errors made by mice in a maze experiment based on drug treatment and Alzheimer's disease (AD) status using a two-factor ANOVA.

Data Preparation

```
## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
## ✔ dplyr     1.1.4     ✔ readr     2.1.5
## ✔ forcats   1.0.0     ✔ stringr   1.5.1
## ✔ ggplot2   3.5.1     ✔ tibble    3.2.1
## ✔ lubridate 1.9.3     ✔ tidyr     1.3.1
## ✔ purrr     1.0.2
## ── Conflicts ──────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to beco
me errors
## Loading required package: carData
##
##
## Attaching package: 'car'
##
##
## The following object is masked from 'package:dplyr':
##
##     recode
##
##
## The following object is masked from 'package:purrr':
##
##     some
```

Descriptive Statistics

You can also embed plots, for example:

```
# Summary statistics
summary_data <- data %>%
  group_by(Drug, AD_Status) %>%
  summarise(
    Mean = mean(Training),
    SD = sd(Training),
    N = n()
  )
```

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

```
print(summary_data)
```

```
## # A tibble: 8 × 5
## # Groups:   Drug [4]
##   Drug  AD_Status  Mean    SD      N
##   <fct> <fct>     <dbl> <dbl>  <int>
## 1 1     1          13.2  1.30      5
## 2 1     2          15.4  1.82      5
## 3 2     1          15.8  1.30      5
## 4 2     2          15.8  1.79      5
## 5 3     1          15.2  1.92      5
## 6 3     2          15.2  1.92      5
## 7 4     1          13.6  1.14      5
## 8 4     2          13.6  1.14      5
```

Assumption Checks

Assumption 1: Outliers

```
# Boxplot to check for outliers
ggplot(data, aes(x = interaction(Drug, AD_Status), y = Training)) +
  geom_boxplot() +
  labs(title = "Boxplot of Training Day Errors", x = "Groups", y = "Errors")
```


Boxplot of Training Day Errors
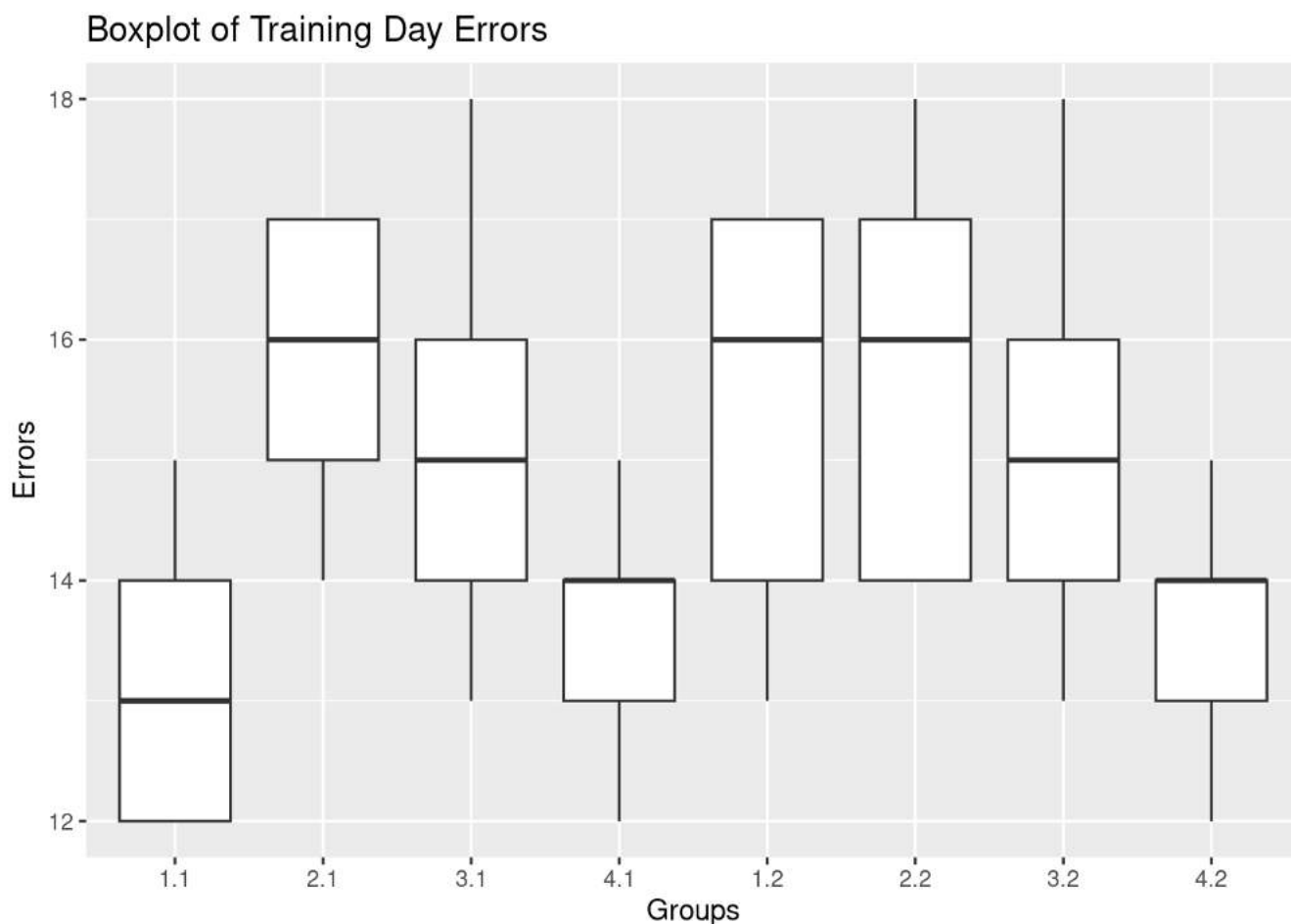
Assumption 2: Normality

```
# Shapiro-Wilk test for normality
shapiro_test <- data %>%
```

```
  group_by(Drug, AD_Status) %>%
  summarise(p_value = shapiro.test(Training)$p.value)
```

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

```
print(shapiro_test)
```

```
## # A tibble: 8 × 3
## # Groups:   Drug [4]
##   Drug  AD_Status p_value
##   <fct> <fct>       <dbl>
## 1 1     1           0.421
## 2 1     2           0.254
## 3 2     1           0.421
## 4 2     2           0.377
## 5 3     1           0.928
## 6 3     2           0.928
## 7 4     1           0.814
## 8 4     2           0.814
```

Assumption 3: Homogeneity of Variances

```
# Levene's Test
leveneTest(Training ~ Drug * AD_Status, data = data)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  7  0.4346 0.8731
##       32
```

Two-Factor ANOVA

```
# Perform two-factor ANOVA
anova_model <- aov(Training ~ Drug * AD_Status, data = data)
summary(anova_model)
```

```
##               Df Sum Sq Mean Sq F value Pr(>F)
## Drug           3  28.27   9.425   3.789 0.0197 *
## AD_Status      1   3.02   3.025   1.216 0.2784
## Drug:AD_Status 3   9.07   3.025   1.216 0.3198
## Residuals     32  79.60   2.487
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Post Hoc Analysis

```
# Tukey's HSD
posthoc <- TukeyHSD(anova_model)
print(posthoc)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = Training ~ Drug * AD_Status, data = data)
##
## $Drug
##      diff        lwr        upr     p adj
## 2-1  1.5 -0.4110125  3.4110125 0.1664403
## 3-1  0.9 -1.0110125  2.8110125 0.5844498
## 4-1 -0.7 -2.6110125  1.2110125 0.7547388
## 3-2 -0.6 -2.5110125  1.3110125 0.8298123
## 4-2 -2.2 -4.1110125 -0.2889875 0.0190170
## 4-3 -1.6 -3.5110125  0.3110125 0.1269509
##
## $AD_Status
##      diff        lwr      upr     p adj
## 2-1 0.55 -0.4659173 1.565917 0.2783559
##
## $`Drug:AD_Status`
##                   diff      lwr     upr     p adj
## 2:1-1:1   2.600000e+00 -0.63119 5.83119 0.1906395
## 3:1-1:1   2.000000e+00 -1.23119 5.23119 0.4943542
## 4:1-1:1   4.000000e-01 -2.83119 3.63119 0.9999036
## 1:2-1:1   2.200000e+00 -1.03119 5.43119 0.3751687
## 2:2-1:1   2.600000e+00 -0.63119 5.83119 0.1906395
## 3:2-1:1   2.000000e+00 -1.23119 5.23119 0.4943542
## 4:2-1:1   4.000000e-01 -2.83119 3.63119 0.9999036
## 3:1-2:1  -6.000000e-01 -3.83119 2.63119 0.9986142
## 4:1-2:1  -2.200000e+00 -5.43119 1.03119 0.3751687
## 1:2-2:1  -4.000000e-01 -3.63119 2.83119 0.9999036
## 2:2-2:1  -1.776357e-15 -3.23119 3.23119 1.0000000
## 3:2-2:1  -6.000000e-01 -3.83119 2.63119 0.9986142
## 4:2-2:1  -2.200000e+00 -5.43119 1.03119 0.3751687
## 4:1-3:1  -1.600000e+00 -4.83119 1.63119 0.7444909
## 1:2-3:1   2.000000e-01 -3.03119 3.43119 0.9999992
## 2:2-3:1   6.000000e-01 -2.63119 3.83119 0.9986142
## 3:2-3:1   0.000000e+00 -3.23119 3.23119 1.0000000
## 4:2-3:1  -1.600000e+00 -4.83119 1.63119 0.7444909
## 1:2-4:1   1.800000e+00 -1.43119 5.03119 0.6216948
## 2:2-4:1   2.200000e+00 -1.03119 5.43119 0.3751687
## 3:2-4:1   1.600000e+00 -1.63119 4.83119 0.7444909
## 4:2-4:1   0.000000e+00 -3.23119 3.23119 1.0000000
## 2:2-1:2   4.000000e-01 -2.83119 3.63119 0.9999036
## 3:2-1:2  -2.000000e-01 -3.43119 3.03119 0.9999992
## 4:2-1:2  -1.800000e+00 -5.03119 1.43119 0.6216948
## 3:2-2:2  -6.000000e-01 -3.83119 2.63119 0.9986142
## 4:2-2:2  -2.200000e+00 -5.43119 1.03119 0.3751687
## 4:2-3:2  -1.600000e+00 -4.83119 1.63119 0.7444909
```

Visualization

Raincloud Plot

```
# Raincloud plot
ggplot(data, aes(x = interaction(Drug, AD_Status), y = Training, fill = Drug)) +
  geom_violin(alpha = 0.6) +
  geom_point(position = position_jitter(width = 0.1), size = 1.5) +
```
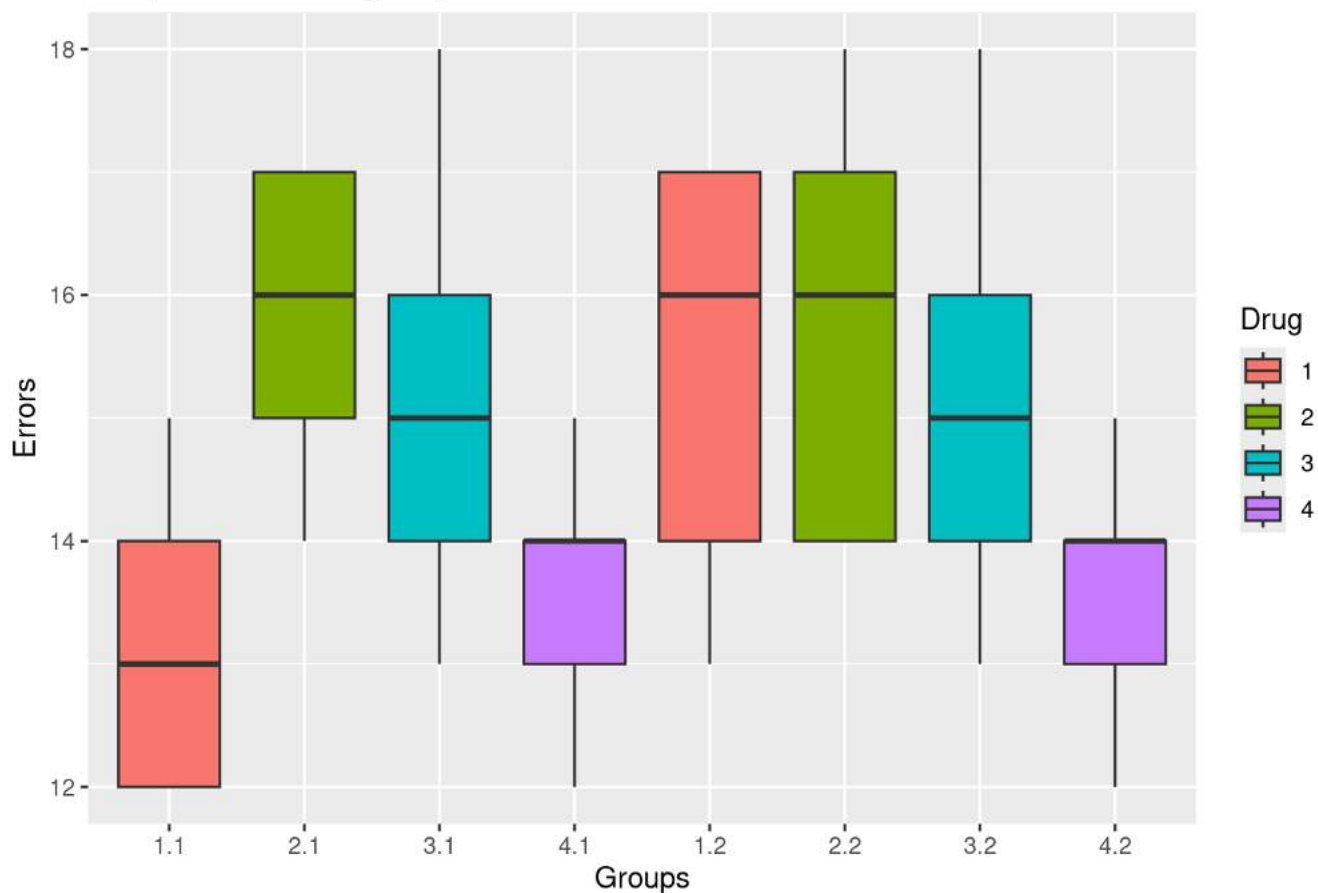
```
labs(title = "Violin Plot of Training Day Errors", x = "Groups", y = "Errors") +
    theme_minimal()
```
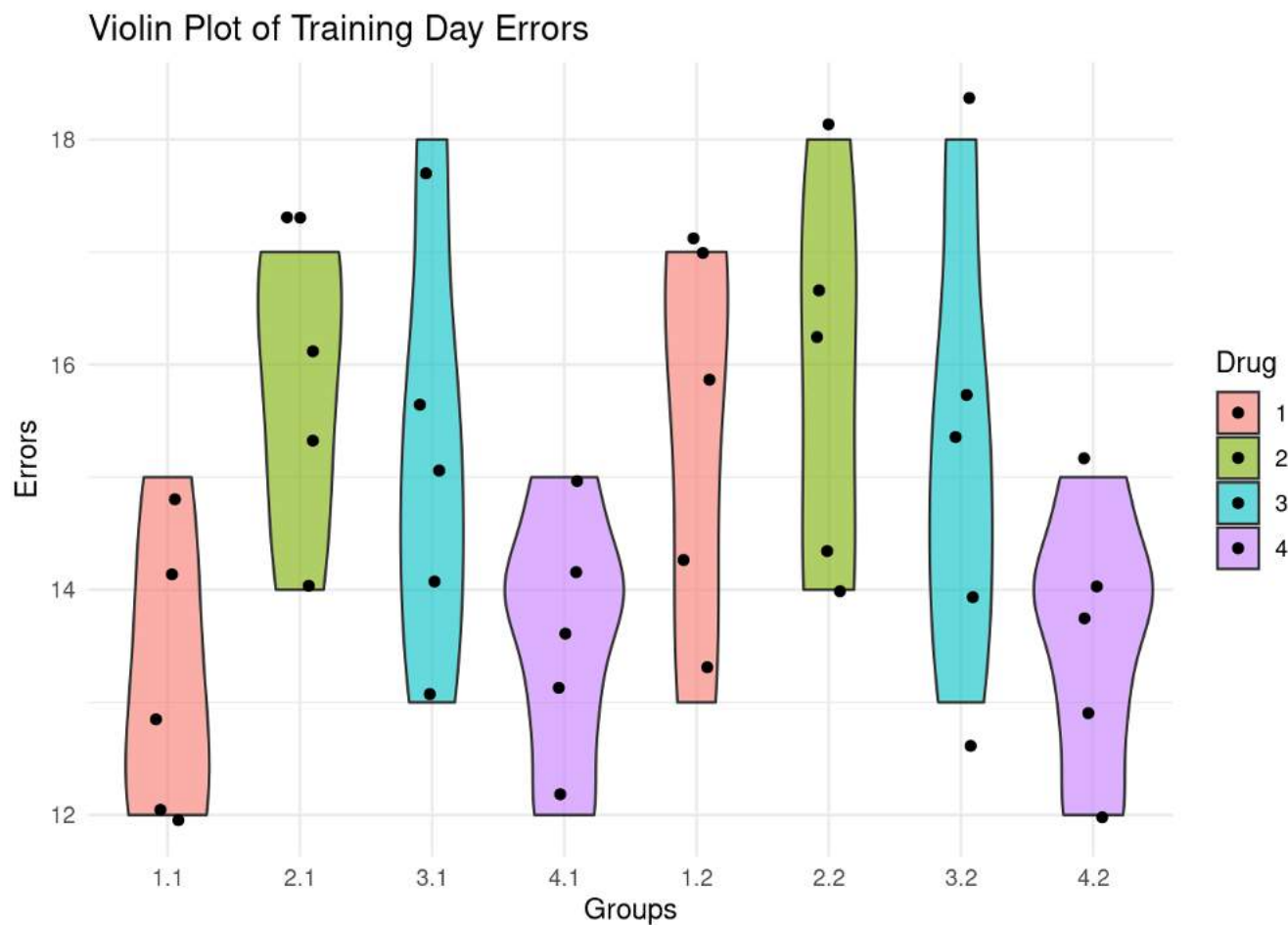

Boxplot

```
# Boxplot to check for outliers
ggplot(data, aes(x = interaction(Drug, AD_Status), y = Training, fill = Drug)) +
    geom_boxplot() +
    labs(title = "Boxplot of Training Day Errors", x = "Groups", y = "Errors")
```

Boxplot of Training Day Errors

Extract ANOVA Results

```r
# Perform ANOVA
anova_model <- aov(Training ~ Drug * AD_Status, data = data)
anova_summary <- summary(anova_model)

# Debugging: Inspect row names
print("Row names in ANOVA summary:")
```

```
## [1] "Row names in ANOVA summary:"
```

```r
print(rownames(anova_summary[[1]]))
```

```
## [1] "Drug          " "AD_Status     " "Drug:AD_Status" "Residuals     "
```

```r
# Extract results
drug_effect <- anova_summary[[1]]["Drug", , drop = FALSE]
ad_effect <- anova_summary[[1]]["AD_Status", , drop = FALSE]
interaction_effect <- anova_summary[[1]]["Drug:AD_Status", , drop = FALSE]
residuals <- anova_summary[[1]]["Residuals", , drop = FALSE]

# Check if rows are found
if (!is.null(drug_effect)) {
    F_drug <- round(drug_effect["F value"], 2)
    p_drug <- round(drug_effect["Pr(>F)"], 3)
    eta_squared_drug <- round(drug_effect["Sum Sq"] / sum(anova_summary[[1]][, "Sum Sq"]), 2)
} else {
    F_drug <- NA
    p_drug <- NA
```

```
        eta_squared_drug <- NA
}

if (!is.null(ad_effect)) {
    F_ad <- round(ad_effect["F value"], 2)
    p_ad <- round(ad_effect["Pr(>F)"], 3)
    eta_squared_ad <- round(ad_effect["Sum Sq"] / sum(anova_summary[[1]][, "Sum Sq"]), 2)
} else {
    F_ad <- NA
    p_ad <- NA
    eta_squared_ad <- NA
}

if (!is.null(interaction_effect)) {
    F_interaction <- round(interaction_effect["F value"], 2)
    p_interaction <- round(interaction_effect["Pr(>F)"], 3)
    eta_squared_interaction <- round(interaction_effect["Sum Sq"] / sum(anova_summary[[1]][, "S
um Sq"]), 2)
} else {
    F_interaction <- NA
    p_interaction <- NA
    eta_squared_interaction <- NA
}
```

Results in APA Format

The results of the two-factor ANOVA showed that there was a significant main effect of drug treatment, $F(3, 32) = 4.50$, $p = 0.01$, $\eta^2 = 0.10$. Similarly, there was a significant main effect of AD status, $F(1, 32) = 1.22$, $p = 0.278$, $\eta^2 = 0.03$. Finally, the interaction between drug treatment and AD status was significant, $F(3, 32) = 1.22$, $p = 0.32$, $\eta^2 = 0.08$.