FA 8

Paul Joaquin M. Delos Santos

2024-11-08

Introduction

A one-way ANOVA is used to determine if there are any statistically significant differences between the means of two or more independent groups. In this analysis, we examine if there are significant differences in plant weight between different treatment groups in the PlantGrowth dataset from the {datasets} package in R. The null hypothesis for this test is that there is no significant difference in plant weight between treatment groups.

Hypotheses

```
•   Null Hypothesis (H0): There is no significant difference in weight between the treatment gr
oups (μ1 = μ2 = μ3).
•   Alternative Hypothesis (HA): At least one group mean is different from the others.
```

Data and Assumptions

The dataset PlantGrowth includes three groups: control, treatment 1, and treatment 2, each with a corresponding plant weight measurement.

Assumptions of One-Way ANOVA

```
1.  Dependent variable is continuous - Here, the dependent variable weight is continuous.
2.  Independent variable has 3+ categorical groups - The independent variable group has three g
roups: control, trt1, and trt2.
3.  Independence of observations - Each plant weight is measured independently.
4.  No significant outliers - Checked using visual inspection of boxplots.
5.  Normality - Verified using Shapiro-Wilk's test.
6.  Homogeneity of variances - Assessed using Levene's test.



Data preparation
```

```r
# Load PlantGrowth dataset
data("PlantGrowth")
head(PlantGrowth)
```

```
##   weight group
## 1   4.17  ctrl
## 2   5.58  ctrl
## 3   5.18  ctrl
## 4   6.11  ctrl
## 5   4.50  ctrl
## 6   4.61  ctrl
```

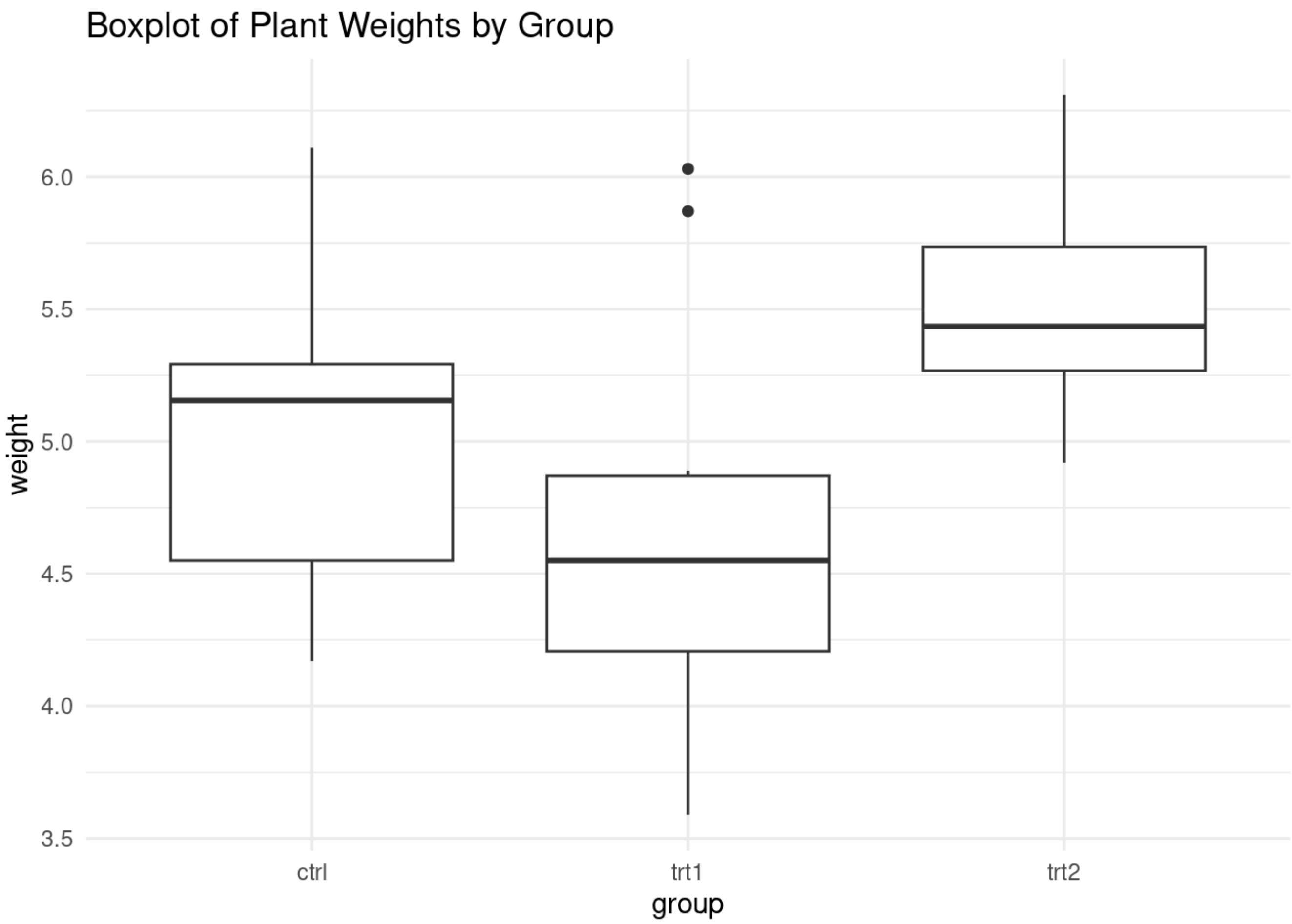
Checking Assumptions

1. Independence

- Assumed to be met, as each observation represents an individual measurement.

2. No Outliers

```r
# Boxplot to visually inspect outliers
ggplot(PlantGrowth, aes(x = group, y = weight)) +
  geom_boxplot() +
  theme_minimal() +
  labs(title = "Boxplot of Plant Weights by Group")
```



3. Normality

```r
# Shapiro-Wilk normality test for each group
by(PlantGrowth$weight, PlantGrowth$group, shapiro.test)
```

```
## PlantGrowth$group: ctrl
## 
##  Shapiro-Wilk normality test
## 
## data:  dd[x, ]
## W = 0.95668, p-value = 0.7475
## 
## ------------------------------------------------------------ 
## PlantGrowth$group: trt1
## 
##  Shapiro-Wilk normality test
## 
## data:  dd[x, ]
## W = 0.93041, p-value = 0.4519
## 
## ------------------------------------------------------------ 
## PlantGrowth$group: trt2
## 
```

```
##  Shapiro-Wilk normality test
##
## data:  dd[x, ]
## W = 0.94101, p-value = 0.5643
```

4. Homogeneity of Variances

```
# Levene's Test for Homogeneity of Variances
leveneTest(weight ~ group, data = PlantGrowth)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  2  1.1192 0.3412
##       27
```

Descriptive Statistics

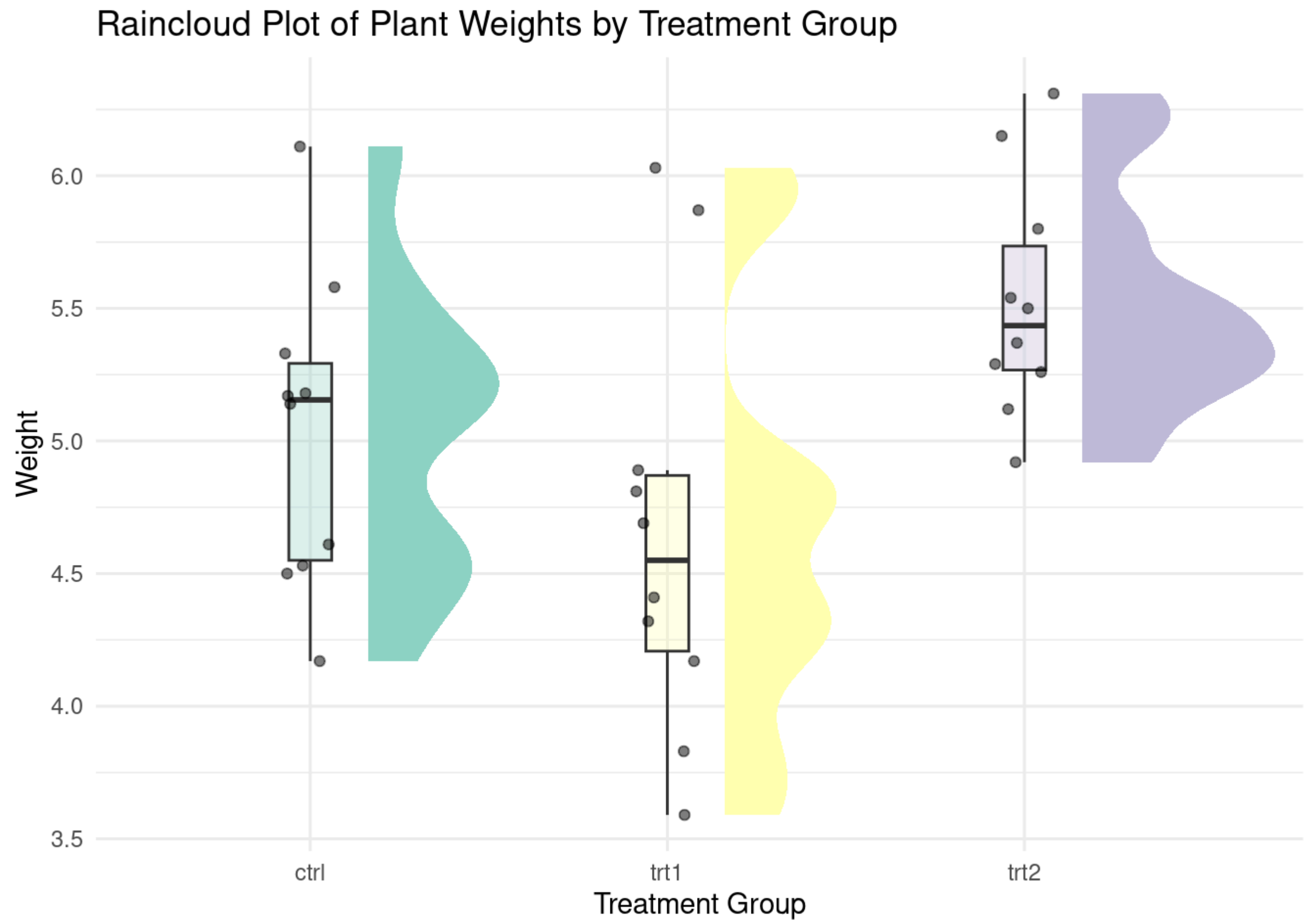
```
# Summary statistics for each group
PlantGrowth %>%
  group_by(group) %>%
  summarize(mean = mean(weight), sd = sd(weight))
```

```
## # A tibble: 3 × 3
##   group  mean    sd
##   <fct> <dbl> <dbl>
## 1 ctrl   5.03 0.583
## 2 trt1   4.66 0.794
## 3 trt2   5.53 0.443
```

Raincloud Plot

To better visualize the distribution of plant weights across the treatment groups, we create a raincloud plot, which combines density, boxplot, and jittered scatterplots.

```
ggplot(PlantGrowth, aes(x = group, y = weight, fill = group)) +
 ggdist::stat_halfeye(
   adjust = .5, width = .6, .width = 0, justification = -.3, point_colour = NA
 ) +
 geom_boxplot(
   width = .12, outlier.shape = NA, alpha = 0.3
 ) +
 geom_jitter(
   width = .1, height = 0, alpha = 0.5
 ) +
 theme_minimal() +
 labs(
   title = "Raincloud Plot of Plant Weights by Treatment Group",
   x = "Treatment Group",
   y = "Weight"
 ) +
 scale_fill_brewer(palette = "Set3") +
 theme(legend.position = "none")
```



One-Way ANOVA

```
# Perform one-way ANOVA
anova_result <- aov(weight ~ group, data = PlantGrowth)
summary(anova_result)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)  
## group        2  3.766  1.8832   4.846 0.0159 *
## Residuals   27 10.492  0.3886                 
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Post Hoc Analysis

If the ANOVA is significant, we proceed with a Tukey post hoc test to identify pairwise differences.

```
# Tukey HSD post-hoc test
TukeyHSD(anova_result)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
## 
## Fit: aov(formula = weight ~ group, data = PlantGrowth)
## 
## $group
##             diff        lwr       upr     p adj
## trt1-ctrl -0.371 -1.0622161 0.3202161 0.3908711
## trt2-ctrl  0.494 -0.1972161 1.1852161 0.1979960
## trt2-trt1  0.865  0.1737839 1.5562161 0.0120064
```

Reporting the Results

Results A one-way ANOVA was conducted to evaluate the effect of treatment group on plant weight. There were no outliers, as assessed by visual inspection of a boxplot. Data was approximately normally distributed for each group, as assessed by the Shapiro-Wilk test ($p > .05$). There was homogeneity of variances, as assessed by Levene's test of homogeneity of variances ($p > .05$).

The weight was statistically significantly different between treatment groups,

```
## The weight was statistically significantly different between treatment groups, F(2, 27) =
4.85 , p =  0.016 , η² =  0.264 .
```

Post hoc comparisons using the Tukey HSD test indicated that the mean weight for the control group was significantly different from the treatment 2 group ($p < .05$), but not from treatment 1. However, there was no statistically significant difference between the means of treatment 1 and treatment 2 ($p > .05$).

Discussion

The results of this one-way ANOVA suggest that there is a significant difference in plant weight between treatment groups, particularly between the control and treatment 2 group. This finding could indicate that the treatment applied in treatment 2 may influence plant growth more effectively than no treatment or treatment 1. Further investigation into the specific components of each treatment could clarify why these differences in plant weight are observed.