Phase 1.2: Statistical Analysis – ANOVA & Tukey HSD

Isaac Appiah-Asare Koomson

2025-06-13

## Load Dataset

# Load data  
data <- read\_csv("D:/myR/seedscience/data/seed\_germination\_stress.csv")  
head(data)

## # A tibble: 6 × 4  
## Treatment Replicate Germinated\_Seeds Total\_Seeds  
## <chr> <dbl> <dbl> <dbl>  
## 1 Salinity\_0.0 1 29 30  
## 2 Salinity\_0.0 2 28 30  
## 3 Salinity\_0.0 3 30 30  
## 4 Salinity\_0.0 4 29 30  
## 5 Salinity\_0.0 5 30 30  
## 6 Salinity\_0.5 1 26 30

## Calculate Germination Percentage

data <- data %>%   
 mutate(Germination\_Percentage = (Germinated\_Seeds / Total\_Seeds) \* 100)

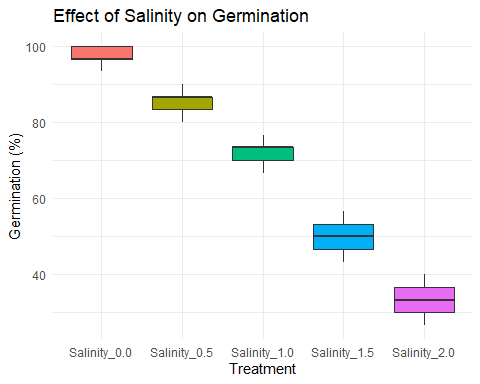
## Summary Statistics per Treatment

summary\_stats <- data %>%  
 group\_by(Treatment) %>%  
 summarise(  
 Mean = mean(Germination\_Percentage),  
 SD = sd(Germination\_Percentage),  
 SE = SD / sqrt(n()),  
 CV = (SD / Mean) \* 100  
 )  
summary\_stats

## # A tibble: 5 × 5  
## Treatment Mean SD SE CV  
## <chr> <dbl> <dbl> <dbl> <dbl>  
## 1 Salinity\_0.0 97.3 2.79 1.25 2.87  
## 2 Salinity\_0.5 85.3 3.80 1.70 4.45  
## 3 Salinity\_1.0 72 3.80 1.70 5.28  
## 4 Salinity\_1.5 50 5.27 2.36 10.5   
## 5 Salinity\_2.0 33.3 5.27 2.36 15.8

## Boxplot: Germination Percentage by Salinity Treatment

ggplot(data, aes(x = Treatment, y = Germination\_Percentage, fill = Treatment)) +  
 geom\_boxplot() +  
 labs(title = "Effect of Salinity on Germination", y = "Germination (%)") +  
 theme\_minimal() +  
 theme(legend.position = "none")



## Treat Treatment as Factor

data$Treatment <- as.factor(data$Treatment)

## Run ANOVA

anova\_model <- aov(Germination\_Percentage ~ Treatment, data = data)  
summary(anova\_model)

## Df Sum Sq Mean Sq F value Pr(>F)   
## Treatment 4 13509 3377 183.1 1.89e-15 \*\*\*  
## Residuals 20 369 18   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Tukey HSD Test

tukey\_result <- TukeyHSD(anova\_model)  
tukey\_result

## Tukey multiple comparisons of means  
## 95% family-wise confidence level  
##   
## Fit: aov(formula = Germination\_Percentage ~ Treatment, data = data)  
##   
## $Treatment  
## diff lwr upr p adj  
## Salinity\_0.5-Salinity\_0.0 -12.00000 -20.12791 -3.872093 0.0021957  
## Salinity\_1.0-Salinity\_0.0 -25.33333 -33.46124 -17.205426 0.0000001  
## Salinity\_1.5-Salinity\_0.0 -47.33333 -55.46124 -39.205426 0.0000000  
## Salinity\_2.0-Salinity\_0.0 -64.00000 -72.12791 -55.872093 0.0000000  
## Salinity\_1.0-Salinity\_0.5 -13.33333 -21.46124 -5.205426 0.0007239  
## Salinity\_1.5-Salinity\_0.5 -35.33333 -43.46124 -27.205426 0.0000000  
## Salinity\_2.0-Salinity\_0.5 -52.00000 -60.12791 -43.872093 0.0000000  
## Salinity\_1.5-Salinity\_1.0 -22.00000 -30.12791 -13.872093 0.0000009  
## Salinity\_2.0-Salinity\_1.0 -38.66667 -46.79457 -30.538759 0.0000000  
## Salinity\_2.0-Salinity\_1.5 -16.66667 -24.79457 -8.538759 0.0000480

## Extract Tukey Group Letters

tukey\_letters <- multcompLetters4(anova\_model, tukey\_result)  
  
# Create data frame of group letters  
group\_letters <- as.data.frame.list(tukey\_letters$Treatment) %>%  
 rownames\_to\_column("Treatment") %>%  
 rename(group = Letters)  
  
# Merge with summary stats  
summary\_stats <- merge(summary\_stats, group\_letters, by = "Treatment")  
summary\_stats

## Treatment Mean SD SE CV group monospacedLetters  
## 1 Salinity\_0.0 97.33333 2.788867 1.247219 2.865274 a a   
## 2 Salinity\_0.5 85.33333 3.800585 1.699673 4.453810 b b   
## 3 Salinity\_1.0 72.00000 3.800585 1.699673 5.278590 c c   
## 4 Salinity\_1.5 50.00000 5.270463 2.357023 10.540926 d d   
## 5 Salinity\_2.0 33.33333 5.270463 2.357023 15.811388 e e  
## LetterMatrix.a LetterMatrix.b LetterMatrix.c LetterMatrix.d LetterMatrix.e  
## 1 TRUE FALSE FALSE FALSE FALSE  
## 2 FALSE TRUE FALSE FALSE FALSE  
## 3 FALSE FALSE TRUE FALSE FALSE  
## 4 FALSE FALSE FALSE TRUE FALSE  
## 5 FALSE FALSE FALSE FALSE TRUE

## Final Clean Table

summary\_stats <- summary\_stats %>%  
 select(Treatment, Mean, SD, SE, CV, group)  
summary\_stats

## Treatment Mean SD SE CV group  
## 1 Salinity\_0.0 97.33333 2.788867 1.247219 2.865274 a  
## 2 Salinity\_0.5 85.33333 3.800585 1.699673 4.453810 b  
## 3 Salinity\_1.0 72.00000 3.800585 1.699673 5.278590 c  
## 4 Salinity\_1.5 50.00000 5.270463 2.357023 10.540926 d  
## 5 Salinity\_2.0 33.33333 5.270463 2.357023 15.811388 e