CodingChallenge7

Olivia Brown

2025-04-03

[Link to Github](https://github.com/olixiabrown/PLPA5820)

#### Q1 - Load in required packages and dataset

getwd()

## [1] "/Users/oliviabrown/Desktop/R directories and code/PLPA5820"

plantemergence <- read.csv("CodingChallenge7/PlantEmergence.csv")  
  
library(tidyverse)

## Warning: package 'tidyr' was built under R version 4.4.1

## Warning: package 'purrr' was built under R version 4.4.1

## Warning: package 'lubridate' was built under R version 4.4.1

## ── 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 1.9.4 ✔ tidyr 1.3.1  
## ✔ purrr 1.0.4   
## ── 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 become errors

library(lme4)

## Loading required package: Matrix  
##   
## Attaching package: 'Matrix'  
##   
## The following objects are masked from 'package:tidyr':  
##   
## expand, pack, unpack

library(emmeans)

## Warning: package 'emmeans' was built under R version 4.4.1

## Welcome to emmeans.  
## Caution: You lose important information if you filter this package's results.  
## See '? untidy'

library(multcompView)

## Warning: package 'multcompView' was built under R version 4.4.1

library(multcomp)

## Warning: package 'multcomp' was built under R version 4.4.1

## Loading required package: mvtnorm

## Warning: package 'mvtnorm' was built under R version 4.4.1

## Loading required package: survival

## Warning: package 'survival' was built under R version 4.4.1

## Loading required package: TH.data

## Warning: package 'TH.data' was built under R version 4.4.1

## Loading required package: MASS  
##   
## Attaching package: 'MASS'  
##   
## The following object is masked from 'package:dplyr':  
##   
## select  
##   
##   
## Attaching package: 'TH.data'  
##   
## The following object is masked from 'package:MASS':  
##   
## geyser

#turn df vectors into factors  
str(plantemergence)

## 'data.frame': 144 obs. of 7 variables:  
## $ Plot : int 101 102 103 104 105 106 107 108 109 201 ...  
## $ Treatment : int 1 2 3 4 5 6 7 8 9 6 ...  
## $ Rep : int 1 1 1 1 1 1 1 1 1 2 ...  
## $ Emergence : num 180.5 54.5 195 198.5 202 ...  
## $ DatePlanted : chr "9-May-22" "9-May-22" "9-May-22" "9-May-22" ...  
## $ DateCounted : chr "16-May-22" "16-May-22" "16-May-22" "16-May-22" ...  
## $ DaysAfterPlanting: int 7 7 7 7 7 7 7 7 7 7 ...

plantemergence$Treatment <- as.factor(plantemergence$Treatment)  
plantemergence$DaysAfterPlanting <- as.factor(plantemergence$DaysAfterPlanting)  
plantemergence$Rep <- as.factor(plantemergence$Rep)

####Q2 - lm with interaction

emergence.lm <- lm(Emergence ~ Treatment\*DaysAfterPlanting, data = plantemergence)  
  
summary(emergence.lm)

##   
## Call:  
## lm(formula = Emergence ~ Treatment \* DaysAfterPlanting, data = plantemergence)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -21.250 -6.062 -0.875 6.750 21.875   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.823e+02 5.324e+00 34.229 <2e-16 \*\*\*  
## Treatment2 -1.365e+02 7.530e+00 -18.128 <2e-16 \*\*\*  
## Treatment3 1.112e+01 7.530e+00 1.477 0.142   
## Treatment4 2.500e+00 7.530e+00 0.332 0.741   
## Treatment5 8.750e+00 7.530e+00 1.162 0.248   
## Treatment6 7.000e+00 7.530e+00 0.930 0.355   
## Treatment7 -1.250e-01 7.530e+00 -0.017 0.987   
## Treatment8 9.125e+00 7.530e+00 1.212 0.228   
## Treatment9 2.375e+00 7.530e+00 0.315 0.753   
## DaysAfterPlanting14 1.000e+01 7.530e+00 1.328 0.187   
## DaysAfterPlanting21 1.062e+01 7.530e+00 1.411 0.161   
## DaysAfterPlanting28 1.100e+01 7.530e+00 1.461 0.147   
## Treatment2:DaysAfterPlanting14 1.625e+00 1.065e+01 0.153 0.879   
## Treatment3:DaysAfterPlanting14 -2.625e+00 1.065e+01 -0.247 0.806   
## Treatment4:DaysAfterPlanting14 -6.250e-01 1.065e+01 -0.059 0.953   
## Treatment5:DaysAfterPlanting14 2.500e+00 1.065e+01 0.235 0.815   
## Treatment6:DaysAfterPlanting14 1.000e+00 1.065e+01 0.094 0.925   
## Treatment7:DaysAfterPlanting14 -2.500e+00 1.065e+01 -0.235 0.815   
## Treatment8:DaysAfterPlanting14 -2.500e+00 1.065e+01 -0.235 0.815   
## Treatment9:DaysAfterPlanting14 6.250e-01 1.065e+01 0.059 0.953   
## Treatment2:DaysAfterPlanting21 3.500e+00 1.065e+01 0.329 0.743   
## Treatment3:DaysAfterPlanting21 -1.000e+00 1.065e+01 -0.094 0.925   
## Treatment4:DaysAfterPlanting21 1.500e+00 1.065e+01 0.141 0.888   
## Treatment5:DaysAfterPlanting21 2.875e+00 1.065e+01 0.270 0.788   
## Treatment6:DaysAfterPlanting21 4.125e+00 1.065e+01 0.387 0.699   
## Treatment7:DaysAfterPlanting21 -2.125e+00 1.065e+01 -0.200 0.842   
## Treatment8:DaysAfterPlanting21 -1.500e+00 1.065e+01 -0.141 0.888   
## Treatment9:DaysAfterPlanting21 -1.250e+00 1.065e+01 -0.117 0.907   
## Treatment2:DaysAfterPlanting28 2.750e+00 1.065e+01 0.258 0.797   
## Treatment3:DaysAfterPlanting28 -1.875e+00 1.065e+01 -0.176 0.861   
## Treatment4:DaysAfterPlanting28 3.123e-13 1.065e+01 0.000 1.000   
## Treatment5:DaysAfterPlanting28 2.500e+00 1.065e+01 0.235 0.815   
## Treatment6:DaysAfterPlanting28 2.125e+00 1.065e+01 0.200 0.842   
## Treatment7:DaysAfterPlanting28 -3.625e+00 1.065e+01 -0.340 0.734   
## Treatment8:DaysAfterPlanting28 -1.500e+00 1.065e+01 -0.141 0.888   
## Treatment9:DaysAfterPlanting28 -8.750e-01 1.065e+01 -0.082 0.935   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 10.65 on 108 degrees of freedom  
## Multiple R-squared: 0.9585, Adjusted R-squared: 0.945   
## F-statistic: 71.21 on 35 and 108 DF, p-value: < 2.2e-16

anova(emergence.lm)

## Analysis of Variance Table  
##   
## Response: Emergence  
## Df Sum Sq Mean Sq F value Pr(>F)   
## Treatment 8 279366 34921 307.9516 < 2.2e-16 \*\*\*  
## DaysAfterPlanting 3 3116 1039 9.1603 1.877e-05 \*\*\*  
## Treatment:DaysAfterPlanting 24 142 6 0.0522 1   
## Residuals 108 12247 113   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

####Q3 - lm without interaction We do not need to fit the interaction term because the interactions are not statistically significant, so will use the following model for the rest of the assignment.

emergence.lm2 <- lm(Emergence ~ Treatment + DaysAfterPlanting, data = plantemergence)  
summary(emergence.lm2)

##   
## Call:  
## lm(formula = Emergence ~ Treatment + DaysAfterPlanting, data = plantemergence)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -21.1632 -6.1536 -0.8542 6.1823 21.3958   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 182.163 2.797 65.136 < 2e-16 \*\*\*  
## Treatment2 -134.531 3.425 -39.277 < 2e-16 \*\*\*  
## Treatment3 9.750 3.425 2.847 0.00513 \*\*   
## Treatment4 2.719 3.425 0.794 0.42876   
## Treatment5 10.719 3.425 3.129 0.00216 \*\*   
## Treatment6 8.812 3.425 2.573 0.01119 \*   
## Treatment7 -2.188 3.425 -0.639 0.52416   
## Treatment8 7.750 3.425 2.263 0.02529 \*   
## Treatment9 2.000 3.425 0.584 0.56028   
## DaysAfterPlanting14 9.722 2.283 4.258 3.89e-05 \*\*\*  
## DaysAfterPlanting21 11.306 2.283 4.951 2.21e-06 \*\*\*  
## DaysAfterPlanting28 10.944 2.283 4.793 4.36e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 9.688 on 132 degrees of freedom  
## Multiple R-squared: 0.958, Adjusted R-squared: 0.9545   
## F-statistic: 273.6 on 11 and 132 DF, p-value: < 2.2e-16

anova(emergence.lm2)

## Analysis of Variance Table  
##   
## Response: Emergence  
## Df Sum Sq Mean Sq F value Pr(>F)   
## Treatment 8 279366 34921 372.070 < 2.2e-16 \*\*\*  
## DaysAfterPlanting 3 3116 1039 11.068 1.575e-06 \*\*\*  
## Residuals 132 12389 94   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The estimate of treatment 2 deviates from the intercept (Treatment 1) by 134.531 which is more than the estimates for all other treatments.

####Q4 - emmeans least squared test

lsmeans <- emmeans(emergence.lm2, ~Treatment)  
results\_lsmeans <- cld(lsmeans, alpha = 0.05, details = TRUE)

####Q5 significance of letters: treatments that share a letter are not significantly different from each other and treatments that do not share a letter are significantly different from each other. treatment two is the only treatment that is significantly different from every other treatment.

plot\_cldbars\_onefactor <- function(emergence.lm2, factor) {  
data <- emergence.lm2$model  
variables <- colnames(emergence.lm2$model)  
dependent\_var <- variables[1]  
independent\_var <- variables[2:length(variables)]  
  
#estimate lsmeans  
lsmeans2 <- emmeans(emergence.lm2, as.formula(paste("~", factor)))   
# contrast with Tukey adjustment by default  
Results\_lsmeans2 <- cld(lsmeans2, alpha = 0.05, reversed = TRUE, details =  
TRUE, Letters = letters)   
  
# Extracting the letters for the bars  
sig.diff.letters <- data.frame(Results\_lsmeans2$emmeans[,1],  
str\_trim(Results\_lsmeans2$emmeans[,7]))  
colnames(sig.diff.letters) <- c(factor, "Letters")  
  
  
# for plotting with letters from significance test  
ave\_stand2 <- emergence.lm2$model %>%  
group\_by(!!sym(factor)) %>%  
dplyr::summarize(  
ave.emerge = mean(.data[[dependent\_var]], na.rm = TRUE),  
se = sd(.data[[dependent\_var]]) / sqrt(n())  
) %>%  
left\_join(sig.diff.letters, by = factor) %>%  
mutate(letter\_position = ave.emerge + 10 \* se)  
  
  
plot <- ggplot(data, aes(x = !! sym(factor), y = !! sym(dependent\_var))) +  
stat\_summary(fun = mean, geom = "bar") +  
stat\_summary(fun.data = mean\_se, geom = "errorbar", width = 0.5) +  
ylab("Number of emerged plants") +  
geom\_jitter(width = 0.02, alpha = 0.5) +  
geom\_text(data = ave\_stand2, aes(label = Letters, y = letter\_position),  
size = 5) +  
xlab(as.character(factor)) +  
theme\_classic()  
  
return(plot)  
}  
  
plot\_cldbars\_onefactor(emergence.lm2, "Treatment")

