2025_4_3_CodingChallengeSeven_mer0127

Madeline Redd

2025 - 04 - 03

Question One

Read in the data called "PlantEmergence.csv" using a relative file path and load the following libraries. tidyverse, lme4, emmeans, multcomp, and multcompView. Turn the Treatment , DaysAfterPlanting and Rep into factors using the function as.factor

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.4     v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts
```

```
library(lme4)
```

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

```r
library(emmeans)
```

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

```r
library(multcompView)
library(multcomp)
```

```
## Loading required package: mvtnorm
## Loading required package: survival
## Loading required package: TH.data
## 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
```

```r
PlantEmerg <- read.csv("PlantEmergence.csv")

head(PlantEmerg)
```

```
##   Plot Treatment Rep Emergence DatePlanted DateCounted DaysAfterPlanting
## 1  101         1   1     180.5    9-May-22   16-May-22                 7
## 2  102         2   1      54.5    9-May-22   16-May-22                 7
## 3  103         3   1     195.0    9-May-22   16-May-22                 7
## 4  104         4   1     198.5    9-May-22   16-May-22                 7
## 5  105         5   1     202.0    9-May-22   16-May-22                 7
## 6  106         6   1     184.0    9-May-22   16-May-22                 7
```

Question Two

Fit a linear model to predict Emergence using Treatment and DaysAfterPlanting along with the interaction. Provide the summary of the linear model and ANOVA results.

```r
lm_Q2<- lm(DaysAfterPlanting ~ Emergence, data = PlantEmerg)

summary(lm_Q2)
```

```
##
## Call:
## lm(formula = DaysAfterPlanting ~ Emergence, data = PlantEmerg)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.9250  -5.1621   0.7038   6.6266  12.5531
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 14.85867    2.67842   5.548 1.37e-07 ***
## Emergence    0.01471    0.01446   1.017    0.311
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.853 on 142 degrees of freedom
## Multiple R-squared:  0.007231,   Adjusted R-squared:  0.0002393
## F-statistic: 1.034 on 1 and 142 DF,  p-value: 0.3109
```

Question Three

Based on the results of the linear model in question 2, do you need to fit the interaction term? Provide a simplified linear model without the interaction term but still testing both main effects. Provide the summary and ANOVA results. Then, interpret the intercept and the coefficient for Treatment 2.

```r
lm_Q2<- lm(DaysAfterPlanting ~ Emergence, data = PlantEmerg)

summary(lm_Q2)
```

```
##
## Call:
## lm(formula = DaysAfterPlanting ~ Emergence, data = PlantEmerg)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.9250  -5.1621   0.7038   6.6266  12.5531
##
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 14.85867    2.67842   5.548 1.37e-07 ***
## Emergence    0.01471    0.01446   1.017    0.311
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.853 on 142 degrees of freedom
## Multiple R-squared:  0.007231,   Adjusted R-squared:  0.0002393
## F-statistic: 1.034 on 1 and 142 DF,  p-value: 0.3109
```

Question Four

Calculate the least square means for Treatment using the emmeans package and perform a Tukey separation with the compact letter display using the cld function. Interpret the results.

```r
lm_Q2<- lm(DaysAfterPlanting ~ Emergence, data = PlantEmerg)

summary(lm_Q2)
```

```
##
## Call:
## lm(formula = DaysAfterPlanting ~ Emergence, data = PlantEmerg)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.9250  -5.1621   0.7038   6.6266  12.5531
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 14.85867    2.67842   5.548 1.37e-07 ***
## Emergence    0.01471    0.01446   1.017    0.311
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.853 on 142 degrees of freedom
## Multiple R-squared:  0.007231,   Adjusted R-squared:  0.0002393
## F-statistic: 1.034 on 1 and 142 DF,  p-value: 0.3109
```

Question Five

The provided function lets you dynamically add a linear model plus one factor from that model and plots a bar chart with letters denoting treatment differences. Use this model to generate the plot shown below. Explain the significance of the letters.

```
plot_cldbars_onefactor <- function(lm_model, factor) {
  data <- lm_model$model
  variables <- colnames(lm_model$model)
  dependent_var <- variables[1]
  independent_var <- variables[2:length(variables)]

  lsmeans <- emmeans(lm_model, as.formula(paste("~", factor))) #Estimate lsmeans
  Results_lsmeans <- cld(lsmeans, alpha = 0.05, reversed = TRUE, details = TRUE,
                Letters = letters) # contrast with Tukey adjustment by default.

    # Extracting the letters for the bars
  sig.diff.letters <- data.frame(Results_lsmeans$emmeans[,1],
                                 str_trim(Results_lsmeans$emmeans[,7]))
  colnames(sig.diff.letters) <- c(factor, "Letters")

  # for plotting with letters from significance test
  ave_stand2 <- lm_model$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)
}
```

Question Six

Generate the gfm .md file along with a .html, .docx, or .pdf. Commit, and push the .md file to github and turn in the .html, .docx, or .pdf to Canvas. Provide me a link here to your github.

[Coding Challenge Seven Link] ()