

Lab Exercises Week 09

Alison Lawyer

2024-11-26

```
# keep this chunk in all your RMarkdown scripts
```

```
knitr::opts_chunk$set(echo = TRUE)
```

```
knitr::opts_chunk$set(tidy.opts = list(width.cutoff = 60), tidy = TRUE)
```

```
# List required packages
```

```
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  1.9.3      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 to become errors
```

```
library(openxlsx)
```

LAB EXERCISES

For this week's lab you will re-use the vigilance dataset available in the "vigilance.xlsx" file. As a reminder, this file contains data on vigilant behavior of several European song bird species (expressed as a rate per hour). The study aimed to address the question of whether birds show higher vigilance behavior in urban areas with more noise, which could make it more difficult to use auditory cues for avoiding predators. For each species, data were collected at several sites (cities), each having an "urban noise" and "control" condition. You will re-analyze this data using ANOVA.

Exercise 1

Your first objective is to assess species differences in vigilance behavior in the control condition, using all species together in your analyses instead of one pair of species at a time as you did in the last lab.

Part 1. Exploratory analyses and visualizations

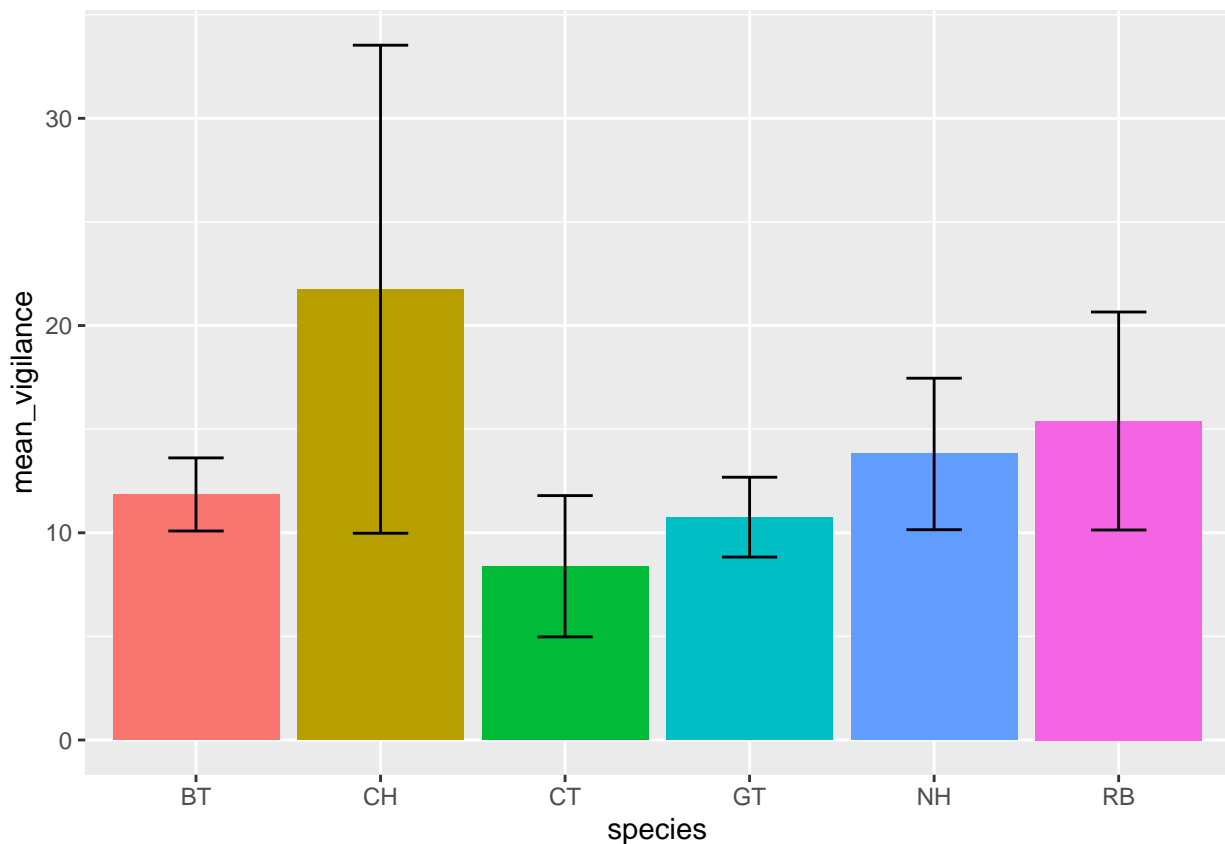
Filter the dataset to records from the control condition only. Create a summary table that contains all species and their means, medians, standard error of the mean, and sample size. Plot the means by species for all species (species on x-axis) using 95% CI for error bars.

```

vigilance <- read.xlsx("vigilance.xlsx")
colnames(vigilance) <- sub(pattern = "\\.(.*)\\.xlsx", replacement = "",
  colnames(vigilance))
# Filter the dataset to the control condition only
control <- vigilance %>%
  filter(treatment == "control")

# Plot the mean rates per species along with 95% confidence
# intervals
control %>%
  group_by(species) %>%
  summarize(mean_vigilance = mean(vigilance.rate), sd_vigilance = sd(vigilance.rate),
    median_vigilance = median(vigilance.rate), iqr = IQR(vigilance.rate),
    var = var(vigilance.rate), se = sd_vigilance/sqrt(n()),
    CI_low = mean_vigilance - 1.96 * se, CI_high = mean_vigilance +
      1.96 * se) %>%
  ggplot(aes(x = species, y = mean_vigilance, fill = species)) +
  geom_col(show.legend = FALSE) + geom_errorbar(aes(ymin = CI_low,
    ymax = CI_high), width = 0.3, color = "black")

```



```

# Make the plot

```

Part 2. Conduct an anova test

Create a one-way ANOVA of vigilance rate by species Print the summary Interpret the result.

```

vig_aov <- aov(vigilance.rate ~ species, data = control)
vig_aov

```

```

## Call:
## aov(formula = vigilance.rate ~ species, data = control)
##
## Terms:
##              species Residuals
## Sum of Squares 1257.979  9515.751
## Deg. of Freedom      5      139
##
## Residual standard error: 8.273974
## Estimated effects may be unbalanced

```

```

summary(vig_aov)

```

```

##              Df Sum Sq Mean Sq F value  Pr(>F)
## species        5   1258   251.60   3.675 0.00373 **
## Residuals     139   9516    68.46
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

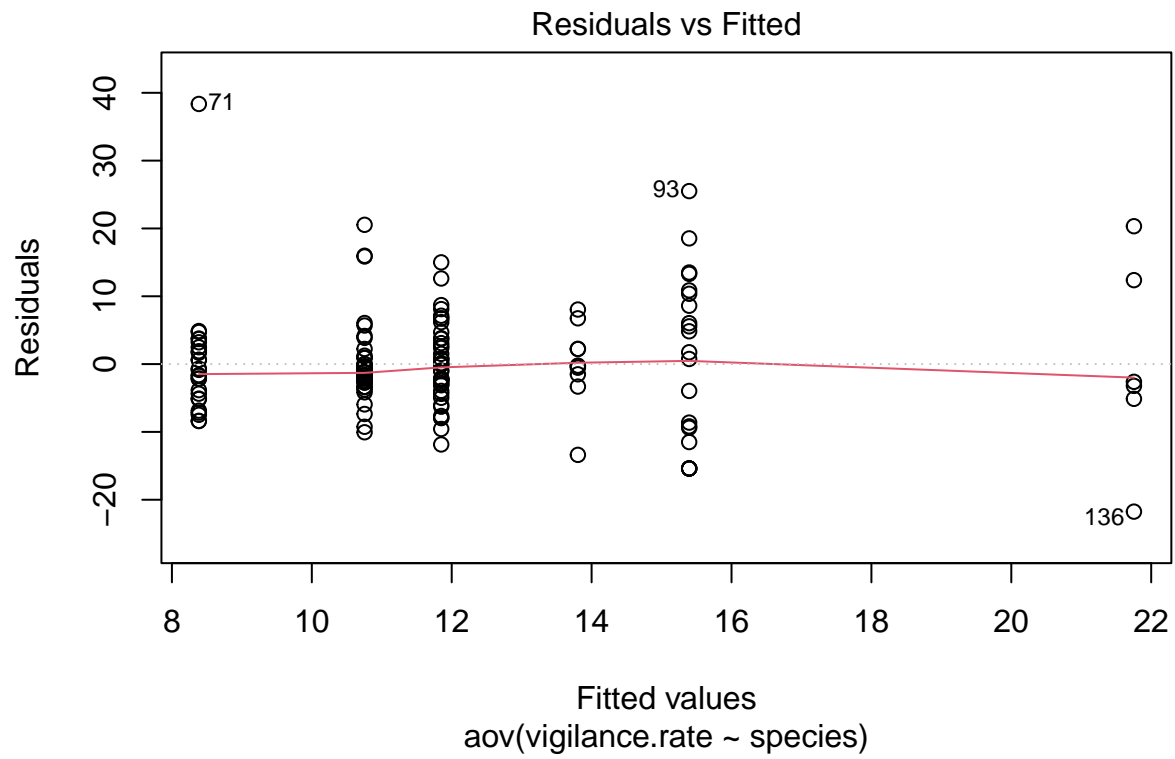
Part 3. Verify that assumptions are met

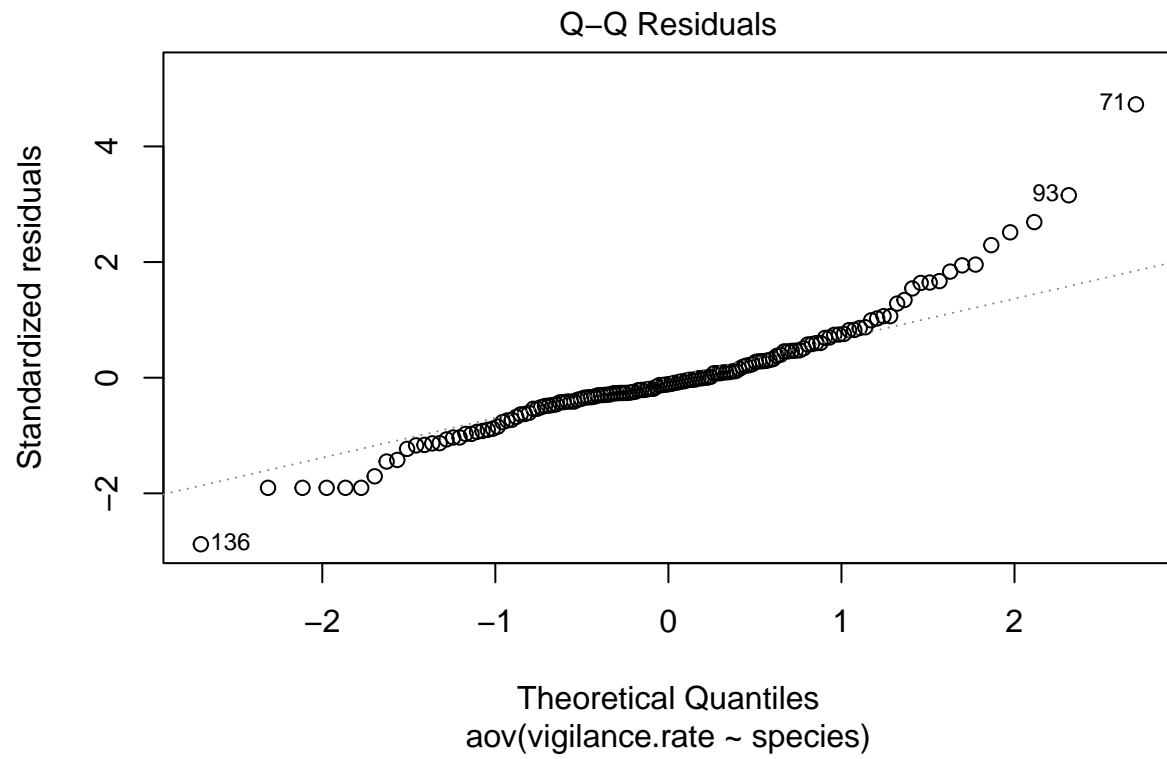
Verify the assumption of homogeneity of variances and normal distribution. Interpret your findings. Is the ANOVA result valid?

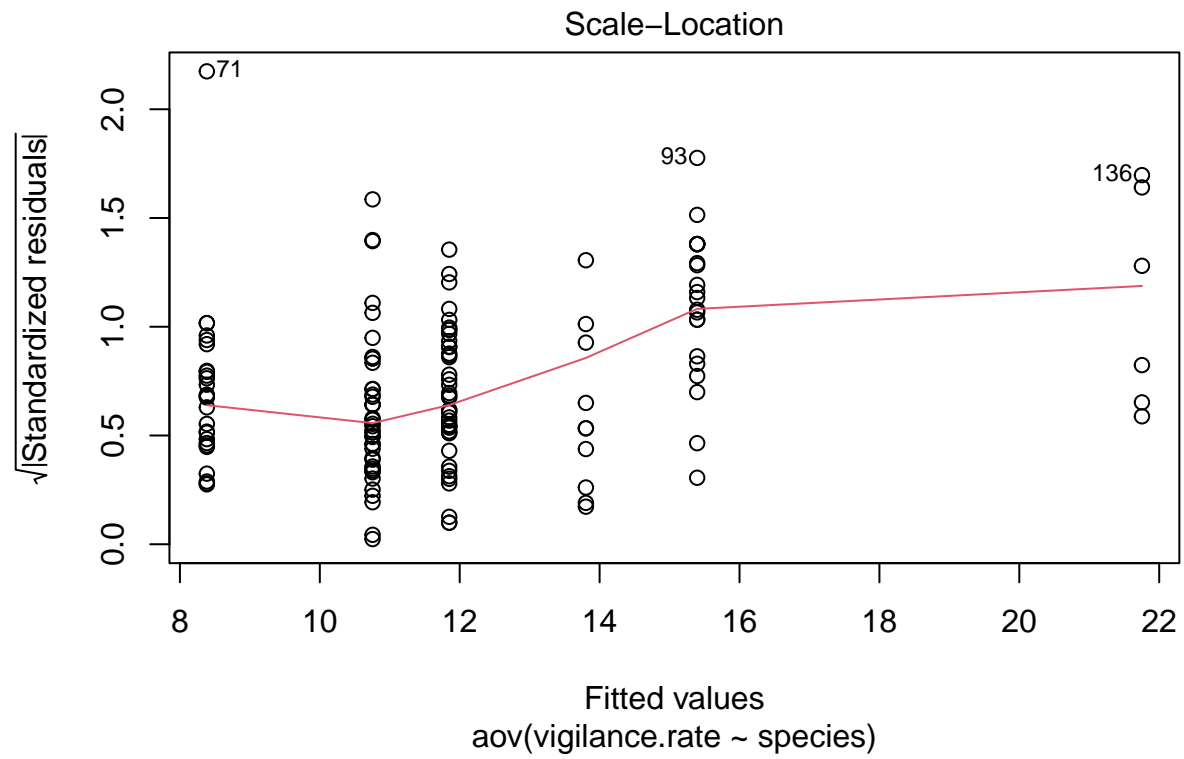
```

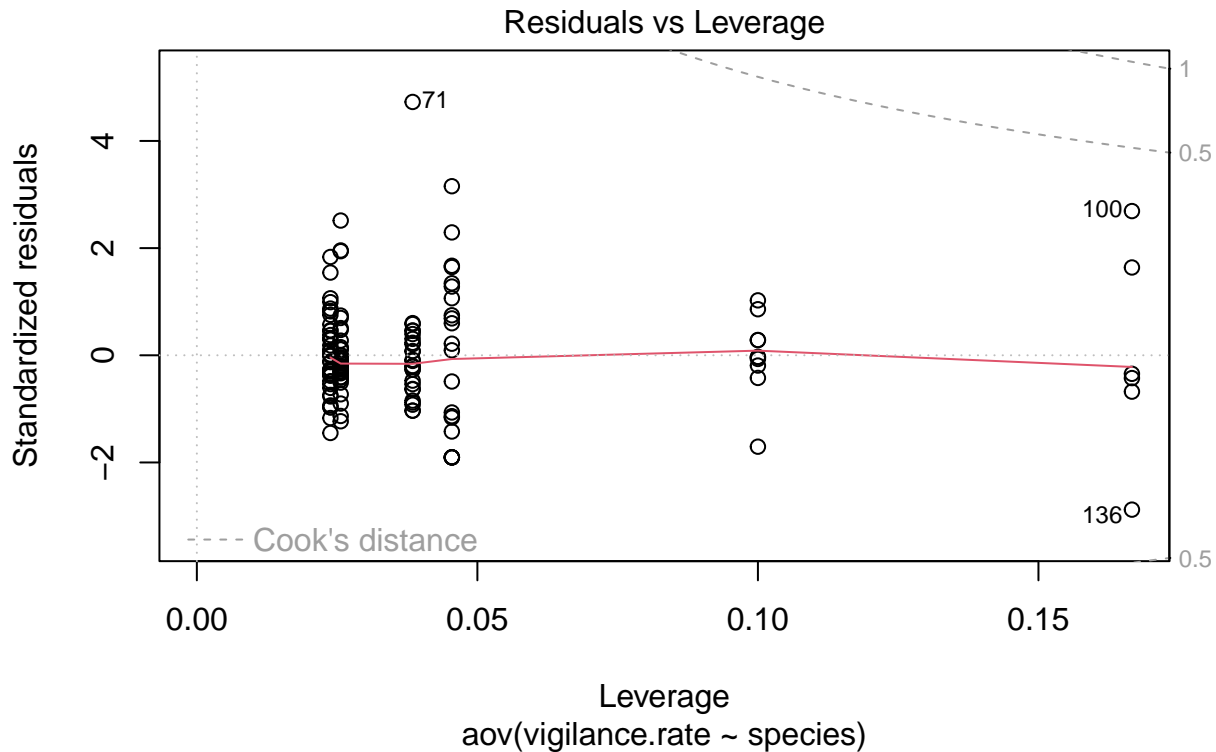
vig_aov %>%
  plot()

```









Part 4. Conduct post-hoc analyses

Regardless of the answer to the validity question in part 3, conduct a Tukey HSD post-hoc to assess the pairwise differences in vigilance rates between species. Which species differ significantly from each other?

```
TukeyHSD(vig_aov)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = vigilance.rate ~ species, data = control)
##
## $species
##          diff          lwr          upr      p adj
## CH-BT   9.903095 -0.53242030 20.3386108 0.0734512
## CT-BT  -3.465238 -9.43197238  2.5014962 0.5483893
## GT-BT  -1.095623 -6.41277527  4.2215299 0.9911946
## NH-BT   1.953762 -6.45961970 10.3671435 0.9848233
## RB-BT   3.544307 -2.74854528  9.8371600 0.5817662
## CT-CH -13.368333 -24.19777960 -2.5388871 0.0064577
## GT-CH -10.998718 -21.48428421 -0.5131517 0.0338399
## NH-CH   -7.949333 -20.29680184  4.3981352 0.4308866
## RB-CH   -6.358788 -17.37128000  4.6537042 0.5547524
## GT-CT   2.369615 -3.68422912  8.4234599 0.8676157
## NH-CT   5.419000 -3.47829801 14.3162980 0.4948595
```

```
## RB-CT    7.009545    0.08300835 13.9360826 0.0455284
## NH-GT    3.049385   -5.42599769 11.5247669 0.9036474
## RB-GT    4.639930   -1.73557851 11.0154386 0.2917462
## RB-NH    1.590545   -7.52866439 10.7097553 0.9959515
```

Exercise 2

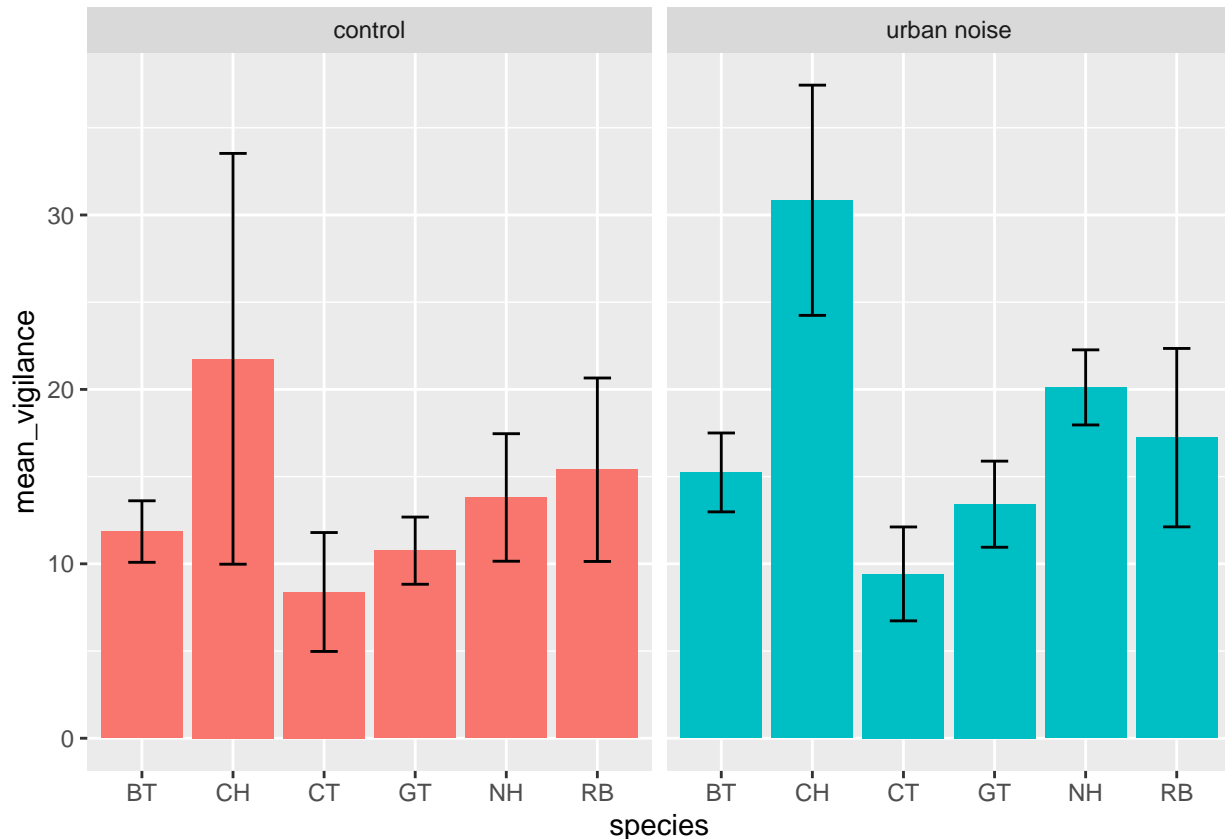
In this exercise you will conduct a two-way ANOVA and interpret its findings.

Part 1.

Expand your summary table from above to include the noise condition. Create a new figure showing the mean and 95% confidence interval of vigilance rates for control and noise condition for each species. Interpret the plot.

```
vigilance %>%
  group_by(species, treatment) %>%
  summarize(mean_vigilance = mean(vigilance.rate), sd_vigilance = sd(vigilance.rate),
            median_vigilance = median(vigilance.rate), iqr = IQR(vigilance.rate),
            var = var(vigilance.rate), se = sd_vigilance/sqrt(n()),
            CI_low = mean_vigilance - 1.96 * se, CI_high = mean_vigilance +
              1.96 * se) %>%
  ggplot(aes(x = species, y = mean_vigilance, fill = treatment)) +
  geom_col(show.legend = FALSE) + geom_errorbar(aes(ymin = CI_low,
            ymax = CI_high), width = 0.3, color = "black") + facet_wrap(~treatment)
```

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

Part 2. Conduct the two-way ANOVA

Conduct the two-way ANOVA using both species and condition as your predictor variables. Interpret the result. If your ANOVA shows significant differences in mean vigilance rates or a significant interaction, conduct a post-hoc test to identify where the differences are located.

```
two_way_vig <- aov(vigilance.rate ~ species * treatment, data = vigilance)
two_way_vig
```

```
## Call:
## aov(formula = vigilance.rate ~ species * treatment, data = vigilance)
##
## Terms:
##              species treatment species:treatment Residuals
## Sum of Squares  3807.855    641.417           236.333 18978.823
## Deg. of Freedom      5         1             5         278
##
## Residual standard error: 8.262515
## Estimated effects may be unbalanced
```

```
summary(two_way_vig)
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## species        5   3808   761.6    11.155 8.24e-10 ***
```

```
## treatment      1    641    641.4    9.395  0.00239 **
## species:treatment  5    236    47.3    0.692  0.62961
## Residuals      278 18979    68.3
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Part 3. Verify assumptions

Verify the assumption of homogeneity of variances and normal distribution. Interpret your findings. Is the ANOVA result valid?

```
two_way_vig %>%
  plot()
```

