Final Project

Kevin Dodd

September 25, 2018

knitr::opts\_chunk$set(echo = TRUE, message=FALSE, warning=FALSE)

# Introduction

For this report the data that will be explored is from a study of two species of *Tamalia* aphids; *Tamalia coweni* and *Tamalia inquilinus* and their population distribution on two species of Manzanita, the White-Leaf Manzanita and the Common Manzanita. The data is a sample of the most recent growing year, as the full study has been taking place over many years.

library(readxl)  
PopEcoLab\_data <- read\_excel("C:/Users/Kevin/Documents/college/Ecology/PopEcoLab-data.xlsx", sheet=1, col\_names=TRUE)  
names(PopEcoLab\_data) <- c("Day", "ID", "species\_M", "height", "width", "radius", "galls", "hyp", "surface\_area", "gall\_density", "coweni", "inquilinus", "total\_aphids")

## Variables of Interest

The variables that we are interested in are the species of Manzanita, which is either White-Leaf Manzanita or Common Manzanita, and the total population of *Tamalia* aphids found on these Manzanita.

# Univariate Description

1. Manzanita Species

PopEcoLab\_data$species\_M[PopEcoLab\_data$species\_M == "AMZ"] <- "Common"  
PopEcoLab\_data$species\_M[PopEcoLab\_data$species\_M == "AV"] <- "White-leaf"

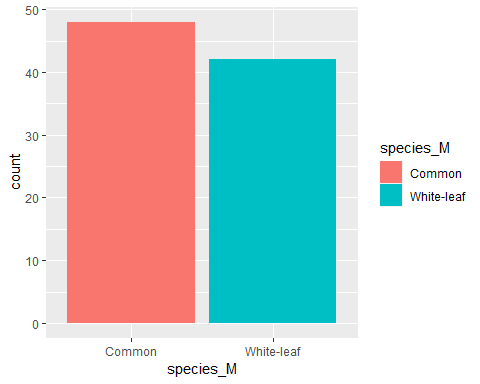
First, the number of each Manzanita species in the sample must be determined.

table(PopEcoLab\_data$species\_M)

##   
## Common White-leaf   
## 48 42

This table does display the appropriate information, however it is not easy to read, so lets graph it.

library(ggplot2)  
ggplot(PopEcoLab\_data, aes(x=species\_M, fill=species\_M)) + geom\_bar()



1. Total aphid population, both *T. coweni* and *T. inquilinus*.

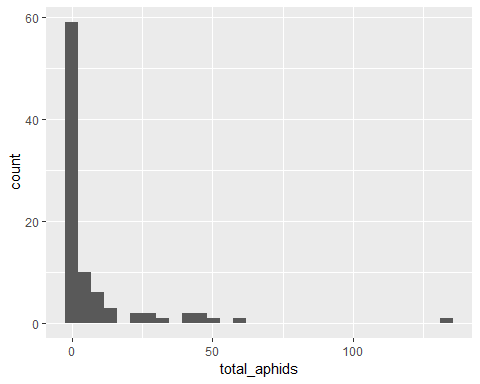
This variable shows the estimated total number of *Tamalia* aphids per Manzanita plant and is based on the number of galls that plant posessed.

summary(PopEcoLab\_data$total\_aphids)

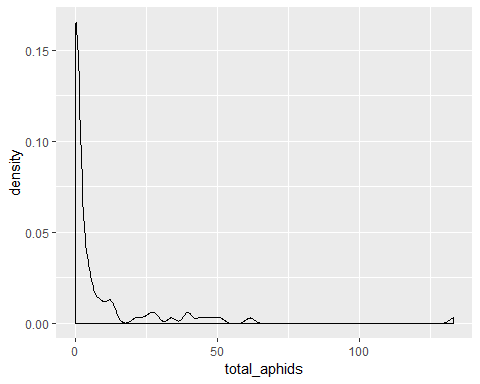
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.000 0.000 0.000 7.877 5.367 133.340

Distribution of the estimated total aphid population per Manzanita.

ggplot(PopEcoLab\_data, aes(x=total\_aphids, fill=total\_aphids)) + geom\_histogram()



ggplot(PopEcoLab\_data, aes(x=total\_aphids)) + geom\_density()



These graphs display that our average is skewed by the outlier located at about 130, that Manzanita apparently had a large number of galls. The majority of the Manzanita plants in this sample have an estimated aphid population of zero, likely because they had no galls on this season’s growth.

Total estimated *Tamalia* aphid population in this sample.

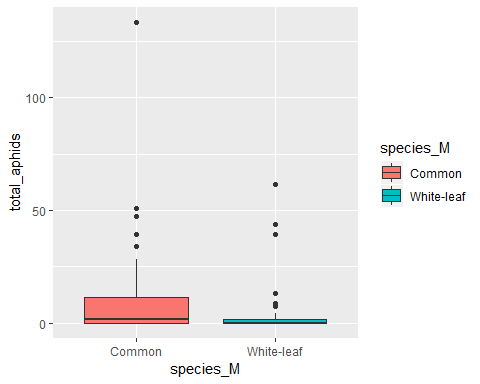
sum(PopEcoLab\_data$total\_aphids)

## [1] 708.91

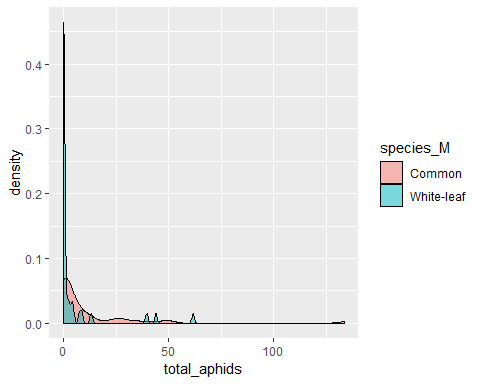
# Bivariate

Next, we will compare the *Tamalia* aphid populations on each Manzanita species to see if one has a greater aphid population than the other.

ggplot(PopEcoLab\_data, aes(x=species\_M, y=total\_aphids, fill=species\_M)) + geom\_boxplot()



ggplot(PopEcoLab\_data, aes(x=total\_aphids, fill=species\_M)) + geom\_density(alpha=.5)



The boxplot and the density plot indicate that Common Manzanita tends to have more aphids per plant than White-leaf for this sample. It was seen earlier that a majority of the Manzanita plants had an aphid population of zero, and now these graphs indicate that most of those Manzanitas with a zero population were White-leaf Manzanita.