MA 2611 Lab 2

Xin Yu Wu

2024-09-09

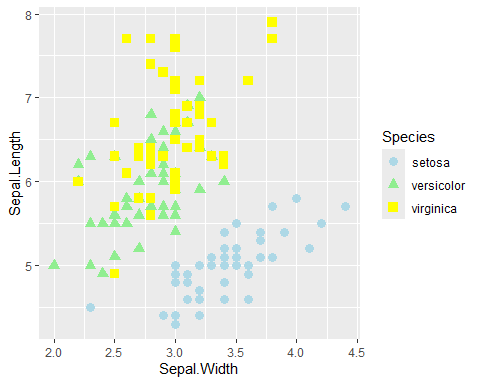
if(!require("ggplot2")){  
install.packages("ggplot2")  
require ("ggplot2")  
}

## Loading required package: ggplot2

library(ggplot2)

**Problem L.5**

ggplot(data = iris, aes(x = Sepal.Width, y = Sepal.Length, shape = Species, colour = Species)) +   
 geom\_point(size=3) +  
 scale\_colour\_manual(values=c("lightblue","palegreen2","yellow"))

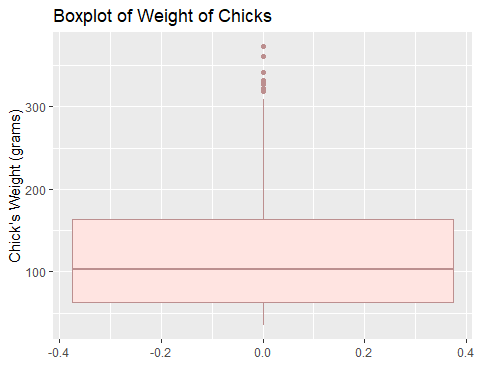


labs(title="Boxplot of Sepal Width vs Sepal Length by Species",   
 x = "Sepal Width (cm)",   
 y = "Sepal Length (cm)" )

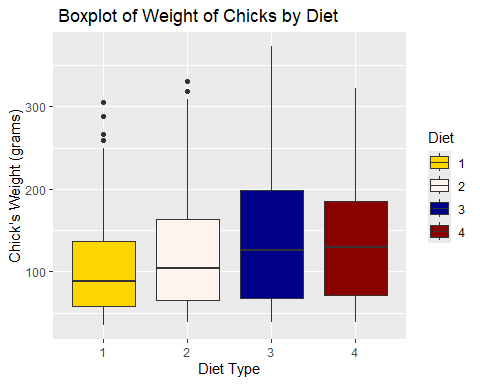
## $x  
## [1] "Sepal Width (cm)"  
##   
## $y  
## [1] "Sepal Length (cm)"  
##   
## $title  
## [1] "Boxplot of Sepal Width vs Sepal Length by Species"  
##   
## attr(,"class")  
## [1] "labels"

**Problem L.6 & L.7**

#a)  
  
ggplot(data = ChickWeight, aes(x=NULL, y= weight)) +  
 geom\_boxplot(fill = "mistyrose", color = "rosybrown") +  
 labs(title="Boxplot of Weight of Chicks",   
 x ="",  
 y = "Chick's Weight (grams)")



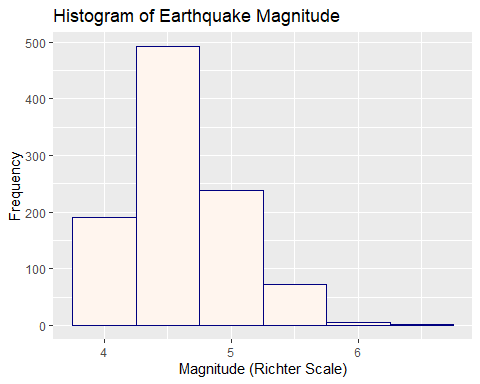
#b)  
ggplot(data = ChickWeight, aes(x=Diet, y=weight, fill = Diet)) +  
 geom\_boxplot() +  
 scale\_fill\_manual(values=c("gold","seashell","blue4", "red4")) +  
 labs(title=" Boxplot of Weight of Chicks by Diet",   
 x ="Diet Type",  
 y = "Chick's Weight (grams)")



1. The box plot from part A indicates that most chicks weigh on the lower end, between 65-160 grams with many outliers. The box plot from part B provides a more detailed illustration than the one from part A as it provides a detailed view of the weight distribution of chicks categorized by diet style. It demonstrates that varying diets appear to influence the weight distribution among the chicks in grams. The median weight of the chicks varies by approximately 25 g across different diets, while the third quartile can differ by up to about 65 g depending on the diet. It also shows more outliers in diets 1 and 2. Moreover, part B shows a clearer picture of how diet affects weight variation.

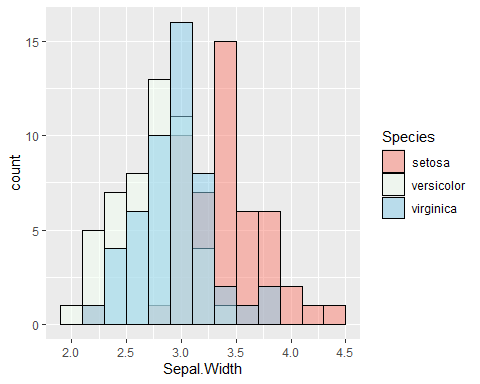
**Problem L.8**

ggplot(data=quakes,aes(x=mag)) +  
geom\_histogram(binwidth= 0.5, fill="seashell", colour="navy") +  
labs(title="Histogram of Earthquake Magnitude",x="Magnitude (Richter Scale)",y="Frequency")

 Interpretation: The histogram is positively skewed.This means that majority of the earthquakes have magnitues that are lower on the Richter scale (4-5.5).

**Problem L.9**

#bin size  
sepal\_bin <- diff(range(iris$Sepal.Width))/12  
  
ggplot(data=iris,aes(x= Sepal.Width, fill = Species)) +  
 geom\_histogram(binwidth = sepal\_bin, position= "identity", alpha= 0.5, color = "black") +  
 scale\_fill\_manual(values = c("salmon", "honeydew", "skyblue"))



labs(title="Histogram of Iris Sepal Width",  
 x="Sepal Width (cm)",  
 y="Frequency")

## $x  
## [1] "Sepal Width (cm)"  
##   
## $y  
## [1] "Frequency"  
##   
## $title  
## [1] "Histogram of Iris Sepal Width"  
##   
## attr(,"class")  
## [1] "labels"

Interpretation: The histogram is bimodal and is slightly positively skewed. This means that majority of the data are slightly on the lower side of the range of widths.