Selina Narain DTSC 630 Homework 3

2023-03-25

## Question 1

**Import the pupae data, then**

# Initialize and read data  
pupae <- read.csv("pupae.csv")  
  
head(pupae)

## T\_treatment CO2\_treatment Gender PupalWeight Frass  
## 1 ambient 280 0 0.244 1.900  
## 2 ambient 280 1 0.319 2.770  
## 3 ambient 280 0 0.221 NA  
## 4 ambient 280 0 0.280 1.996  
## 5 ambient 280 0 0.257 1.069  
## 6 ambient 280 1 0.333 2.257

summary(pupae)

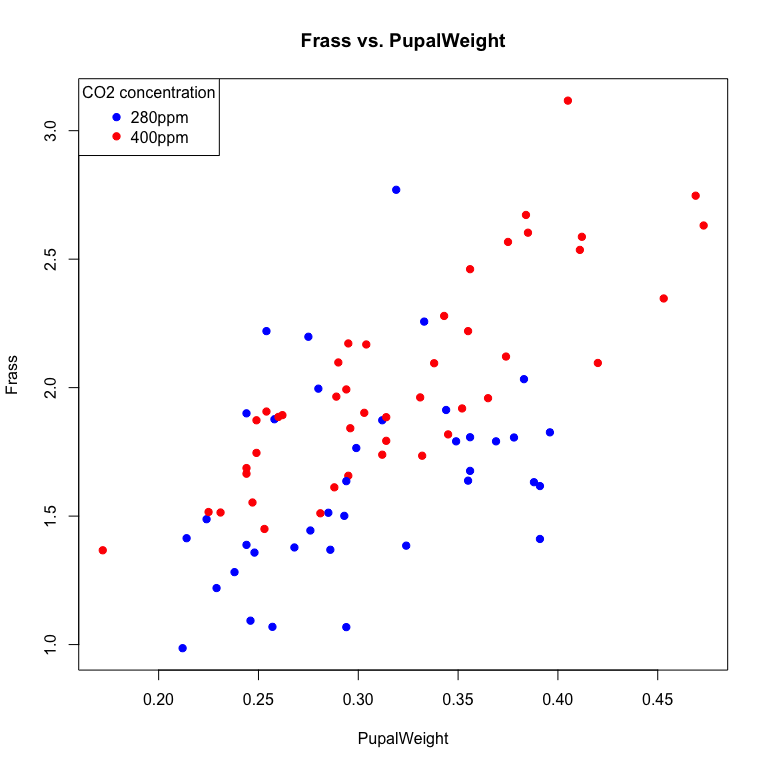
## T\_treatment CO2\_treatment Gender PupalWeight   
## Length:84 Min. :280.0 Min. :0.0000 Min. :0.1720   
## Class :character 1st Qu.:280.0 1st Qu.:0.0000 1st Qu.:0.2562   
## Mode :character Median :400.0 Median :0.0000 Median :0.2975   
## Mean :344.3 Mean :0.4487 Mean :0.3110   
## 3rd Qu.:400.0 3rd Qu.:1.0000 3rd Qu.:0.3560   
## Max. :400.0 Max. :1.0000 Max. :0.4730   
## NA's :6   
## Frass   
## Min. :0.986   
## 1st Qu.:1.515   
## Median :1.818   
## Mean :1.846   
## 3rd Qu.:2.095   
## Max. :3.117   
## NA's :1

* Convert the variable ”CO2 treatment” to a factor. Inspect the levels of this factor variable.

# Make a factor variable of CO2 treatment  
pupae$CO2\_treatment <- factor(pupae$CO2\_treatment,   
 levels=c(280, 400),   
 labels=c("280ppm", "400ppm"))

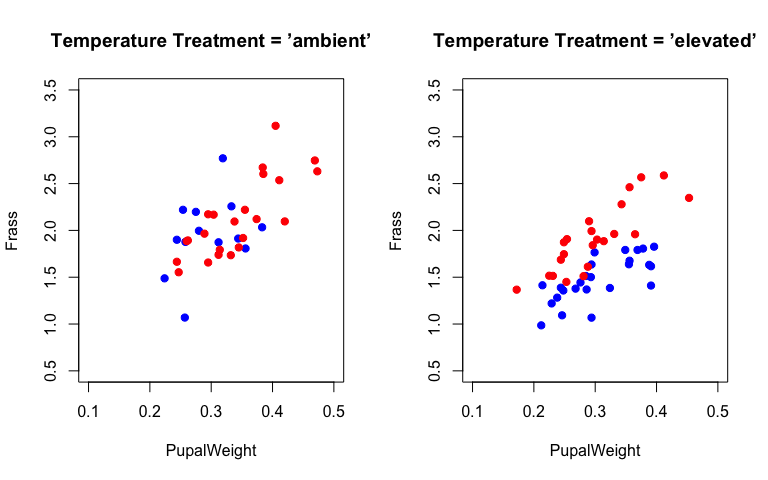
* Make a scatter plot of Frass vs. PupalWeight, with blue solid circles for a CO2 concentration of 280ppm and red for 400ppm. Also add a legend.

# Create palette to store colors to be used in a plot  
palette(c("blue", "red"))  
  
# y ~ x : y for y axis and x for x axis  
# Frass ~ PupalWeight : Frass for y axis and PupalWeight for x axis  
plot(Frass ~ PupalWeight,   
 main = "Frass vs. PupalWeight",  
 data = pupae,   
 col = CO2\_treatment,   
 pch = 19)  
  
# Add a legend  
legend("topleft",   
 title = "CO2 concentration", #title  
 levels(pupae$CO2\_treatment), #CO2\_treatment levels  
 col = palette(), #corresponding colors  
 pch = 19) #symbols



* Make side-by-side plots (similar to the above figure), one with the ’ambient’ temperature treatment, one with ’elevated’. In order to better compare data among the two treatment groups (that is the reason for drawing the side-by-side plot), make sure that the X and Y axis ranges are the same for both plots (Hint: use xlim and ylim)

#Setting the plot area to a 1 by 2 array. One row of plots, and two columns.  
par(mfrow = c(1, 2))  
  
#Takes subset of ambient data and load into dataframe  
ambient\_df <- subset(pupae, T\_treatment =="ambient")  
  
#Takes subset of elevated data and load into dataframe  
elevated\_df <- subset(pupae, T\_treatment =="elevated")  
  
#Left graph  
plot(Frass ~ PupalWeight,   
 main = "Temperature Treatment = ’ambient’",  
 data = ambient\_df,   
 col = CO2\_treatment,  
 pch = 19,  
 xlim = c(0.1, 0.5),   
 ylim = c(0.5, 3.5))  
  
#Right graph  
plot(Frass ~ PupalWeight,   
 main = "Temperature Treatment = ’elevated’",  
 data = elevated\_df,   
 col = CO2\_treatment,   
 pch = 19,  
 xlim = c(0.1, 0.5),   
 ylim = c(0.5, 3.5))



## Question 2

**Sometimes we want to combine the side-by-side plots into a single plot with combinations of symbols and colors to distinguish subgroups. Here you will generate the plot as below. Note that you can either use just plot() function with col and pch option specified for appropriate subgroups without points() functions, or use the approach we have learned (plot + points)**

# Create palette to store colors to be used in a plot  
palette(c("red", "blue"))  
  
# Combine plots for temperature treatment (’ambient’ and ’elevated’) and CO2 concentration (280ppm and 400ppm)  
plot(Frass ~ PupalWeight,  
 main = "Pupae Dataset: Treatment (CO2, Temperature)",  
 data = pupae,  
 col = CO2\_treatment,  
 pch = c(3, 15)[as.factor(T\_treatment)])  
  
# Add a legend  
legend("topleft",   
 title = "CO2 - temp",  
 c("280ppm - Tamb", "280ppm - Telev", "400ppm - Tamb", "400ppm - Telev"), #labels  
 col = c(palette()[1], palette()[1], palette()[2],palette()[2]), #corresponding colors  
 pch = c(3, 15, 3, 15)) #symbols

