Coding challenge 4 - R\_markdown

Bibek

2025-02-27

Table of Contents

[Link 1](#_Toc191555508)

[First level header 2](#_Toc191555509)

[Second level header 3](#_Toc191555510)

[Third level header 5](#_Toc191555511)

[Fourth level header 6](#_Toc191555512)

# Link

**Link to the manuscript where the data are published**

[**Noel, Z.A., Roze, L.V., Breunig, M., Trail, F. 2022. Endophytic fungi as promising biocontrol agent to protect wheat from Fusarium graminearum head blight. Plant Disease. https://doi.org/10.1094/PDIS-06-21-1253-RE**](https://doi.org/10.1094/PDIS-06-21-1253-RE)

**Questions 1**

1. YAML header

In R markdown, YAML header is a section at the top of the document that contains metadata about the file. It is enclosed between “—” lines and helps to tell how to format the output file.

For example, in this markdown i have YAML hear with infromation like title, author, date which will appearin the rendered files. I have also specified output format to generate markdown file using gfm. Other than this i have also other output options which i can choose while knitting.

1. Literate programming

Literate programming means writing both the code and explanation in the same document whihch make analysis easy to understand and reproduce.

**Question 2**

Here i will use the relative path of the Mycotoxin.csv data file with na.strings option set to “na”. I will be using the codes from the coding challenge 3 assignment.

#Read .csv  
#using relative path  
DON.database <- read.csv("Sample\_data/MycotoxinData.csv", na.strings = "na")  
str(DON.database) #visualize the structure of the data frame.

## 'data.frame': 375 obs. of 6 variables:  
## $ Treatment : chr "Fg" "Fg" "Fg" "Fg" ...  
## $ Cultivar : chr "Wheaton" "Wheaton" "Wheaton" "Wheaton" ...  
## $ BioRep : int 2 2 2 2 2 2 2 2 2 3 ...  
## $ MassperSeed\_mg: num 10.29 12.8 2.85 6.5 10.18 ...  
## $ DON : num 107.3 32.6 416 211.9 124 ...  
## $ X15ADON : num 3 0.85 3.5 3.1 4.8 3.3 6.9 2.9 2.1 0.71 ...

Now, I will separate code chunk for the figures plotting the DON data, 15ADON, and Seedmass, and one for the three combined using ggarrange.

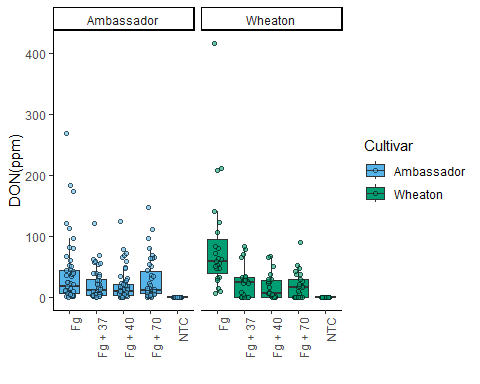
# First level header

This is for DON data.

library(ggplot2) # loading the ggplot2 package  
library(ggpubr) ##integrating multiple statistics and plots  
  
cbbPalette <-c("#000000", "#E69F00", "#56B4E9", "#009E73", "#F0E442", "#0072B2", "#D55E00", "#CC79A7")  
myCols <- cbbPalette[c(3, 4)] ## choosing 3rd and 4th color palette from the cbbPalette.  
  
##Question 1####  
#Create a boxplot for DON (deoxynivalenol) values  
DON <- ggplot(DON.database, aes(x = Treatment, y = DON, fill = Cultivar)) +  
 geom\_boxplot(position = position\_dodge(0.85), outlier.shape = NA ) + # Create boxplots with dodge positioning; remove default outlier points.  
 geom\_point(position = position\_jitterdodge(0.3), shape = 21, color ="black", alpha =0.6) + #jittered points to show individual observations  
 scale\_fill\_manual(values = myCols) + ## Use custom fill colors for the Cultivar groups.  
 ylab("DON(ppm)") + # Label the y-axis and remove x-axis label.  
 xlab("") +  
 theme\_classic() +  
 theme(axis.text.x = element\_text(angle = 90, hjust = 1)) +  
 facet\_wrap(~Cultivar) # Create separate panels for each Cultivar.  
  
DON #Visual

## Warning: Removed 8 rows containing non-finite outside the scale range  
## (`stat\_boxplot()`).

## Warning: Removed 8 rows containing missing values or values outside the scale range  
## (`geom\_point()`).



#Reorder the factor order level  
DON.database$Treatment <- factor(DON.database$Treatment, levels = c("NTC", "Fg", "Fg + 37", "Fg + 40", "Fg + 70"))  
  
# Re-displaying the plot will now use the new Treatment order.  
DON <- ggplot(DON.database, aes(x = Treatment, y = DON, fill = Cultivar)) +  
 geom\_boxplot(position = position\_dodge(0.85), outlier.shape = NA ) + # Create boxplots with dodge positioning; remove default outlier points.  
 geom\_point(position = position\_jitterdodge(0.3), shape = 21, color ="black", alpha =0.6) + #jittered points to show individual observations  
 scale\_fill\_manual(values = myCols) + ## Use custom fill colors for the Cultivar groups.  
 ylab("DON(ppm)") + # Label the y-axis and remove x-axis label.  
 xlab("") +  
 theme\_classic() +  
 theme(axis.text.x = element\_text(angle = 90, hjust = 1)) +  
 facet\_wrap(~Cultivar) # Create separate panels for each Cultivar.

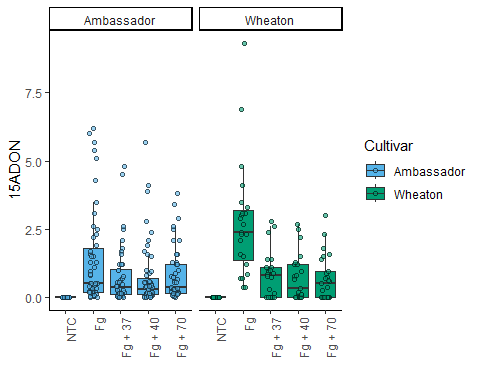
# Second level header

This id for 15ADON plot.

#Change the y-variable to plot X15ADON and MassperSeed\_mg.   
  
# Boxplot for X15ADON values  
X15ADON <- ggplot(DON.database, aes(x = Treatment, y = X15ADON, fill = Cultivar)) +  
 geom\_boxplot(position = position\_dodge(0.85), outlier.shape = NA) +  
 geom\_point(shape = 21, color = "black", position = position\_jitterdodge(0.3), alpha = 0.6) +  
 scale\_fill\_manual(values = myCols) +  
 ylab("15ADON") + # y-axis label for X15ADON  
 xlab("") +  
 theme\_classic() +  
 theme(axis.text.x = element\_text(angle = 90, hjust = 1)) +  
 facet\_wrap(~Cultivar)  
   
X15ADON # Display the X15ADON plot

## Warning: Removed 10 rows containing non-finite outside the scale range  
## (`stat\_boxplot()`).

## Warning: Removed 10 rows containing missing values or values outside the scale range  
## (`geom\_point()`).



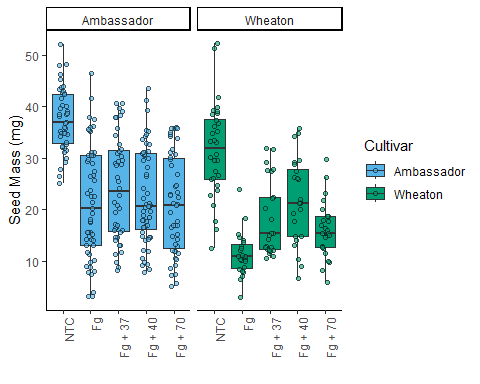
# Third level header

This id for Seedmass plot.

# Boxplot for MassperSeed\_mg (Seed Mass in mg)  
  
MassperSeed\_mg <- ggplot(DON.database, aes(x = Treatment, y = MassperSeed\_mg, fill = Cultivar)) +  
 geom\_boxplot(position = position\_dodge(0.85), outlier.shape = NA) +  
 geom\_point(shape = 21, color = "black", position = position\_jitterdodge(0.3), alpha = 0.6) +  
 scale\_fill\_manual(values = myCols) +  
 ylab("Seed Mass (mg)") + # y-axis label for Seed Mass  
 xlab("") +  
 theme\_classic() +  
 theme(axis.text.x = element\_text(angle = 90, hjust = 1)) +  
 facet\_wrap(~Cultivar)  
  
MassperSeed\_mg #Display the MassperSeed\_mg plot

## Warning: Removed 2 rows containing non-finite outside the scale range  
## (`stat\_boxplot()`).

## Warning: Removed 2 rows containing missing values or values outside the scale range  
## (`geom\_point()`).



# Fourth level header

This is combined 3 plots using ggarrange.

# Combining all plots together  
  
Fig\_combo<-ggarrange( ## Arrange the three plots (DON, X15ADON, MassperSeed\_mg) side by side in one row  
 DON,  
 X15ADON,  
 MassperSeed\_mg,  
 labels = "auto", # Add labels (a, b, c) to each plot  
 nrow = 1, # Arrange in one row  
 ncol = 3, # Three columns  
 legend = FALSE, # Do not display the legend in the combined plot  
 common.legend = T #If TRUE, a common unique legend will be created for arranged plots.  
)

## Warning: Removed 8 rows containing non-finite outside the scale range  
## (`stat\_boxplot()`).

## Warning: Removed 8 rows containing missing values or values outside the scale range  
## (`geom\_point()`).

## Warning: Removed 10 rows containing non-finite outside the scale range  
## (`stat\_boxplot()`).

## Warning: Removed 10 rows containing missing values or values outside the scale range  
## (`geom\_point()`).

## Warning: Removed 2 rows containing non-finite outside the scale range  
## (`stat\_boxplot()`).

## Warning: Removed 2 rows containing missing values or values outside the scale range  
## (`geom\_point()`).

## Warning: Removed 8 rows containing non-finite outside the scale range  
## (`stat\_boxplot()`).

## Warning: Removed 8 rows containing missing values or values outside the scale range  
## (`geom\_point()`).

## Warning: Removed 10 rows containing non-finite outside the scale range  
## (`stat\_boxplot()`).

## Warning: Removed 10 rows containing missing values or values outside the scale range  
## (`geom\_point()`).

## Warning: Removed 2 rows containing non-finite outside the scale range  
## (`stat\_boxplot()`).

## Warning: Removed 2 rows containing missing values or values outside the scale range  
## (`geom\_point()`).

Fig\_combo

