CodingChallenge4

Muhtarin Khayer Brohee

2025-02-27

## **YAML header**

The YAML (Yet Another Markup Language) header is a metadata block at the beginning of an R Markdown (.Rmd) document. It defines settings for the document, such as the title, author, output format, and table of contents. It is enclosed in triple dashes (—).

## **Literate programming**

Literate programming is an approach where code and natural language explanations are combined in a single document. The goal is to create a human-readable and reproducible document that integrates analysis with documentation.

How does R Markdown support literate programming? - Code chunks (`{r}) allow embedding of executable R code. - Markdown text enables writing explanations, interpretation, and formatting. - Knit output produces reports in HTML, PDF, or Word, making them easy to share.

[Article Link](https://pubmed.ncbi.nlm.nih.gov/34587775/)

DON\_data <- read.csv("MycotoxinData.csv", na.strings = "na")  
str(DON\_data)

## '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 ...

# Print column names to check for typos or formatting issues  
print(colnames(DON\_data))

## [1] "Treatment" "Cultivar" "BioRep" "MassperSeed\_mg"  
## [5] "DON" "X15ADON"

# Show the first few rows of the dataset  
head(DON\_data)

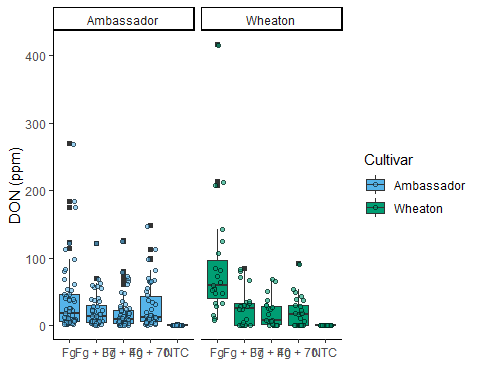
## Treatment Cultivar BioRep MassperSeed\_mg DON X15ADON  
## 1 Fg Wheaton 2 10.291304 107.3 3.00  
## 2 Fg Wheaton 2 12.803226 32.6 0.85  
## 3 Fg Wheaton 2 2.846667 416.0 3.50  
## 4 Fg Wheaton 2 6.500000 211.9 3.10  
## 5 Fg Wheaton 2 10.179167 124.0 4.80  
## 6 Fg Wheaton 2 12.044444 73.1 3.30

# Check the structure of the data  
str(DON\_data)

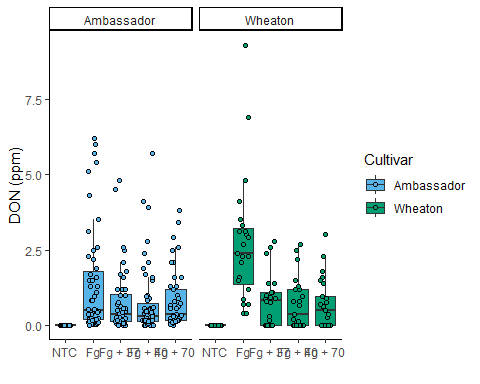
## 'data.frame': 375 obs. of 6 variables:  
## $ Treatment : chr "Fg" "Fg" "Fg" "Fg" ...  
## $ Cultivar : chr "Wheaton" "Wheaton" "Wheaton" "Wheaton" ...  
## $ BioRep : int 2 2 2 2 2 2 2 2 2 3 ...  
## $ MassperSeed\_mg: num 10.29 12.8 2.85 6.5 10.18 ...  
## $ DON : num 107.3 32.6 416 211.9 124 ...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

## **Plot**

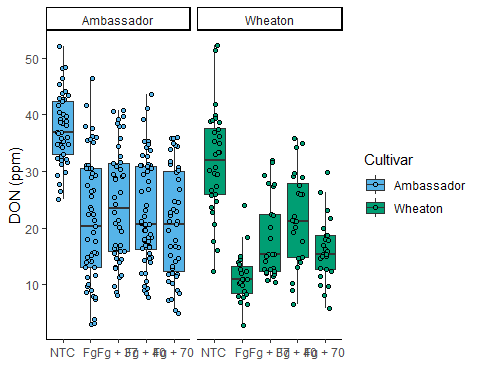
cbbPalette <- c("#000000", "#E69F00", "#56B4E9", "#009E73", "#F0E442", "#0072B2", "#D55E00", "#CC79A7")  
  
  
DON\_plot <- ggplot(DON\_data, aes(x = Treatment, y = DON, fill = Cultivar)) +  
 geom\_boxplot(outliers = T, outlier.shape = 15) +  
 ylab("DON (ppm)") +  
 xlab("") +  
 geom\_jitter(pch = 21, position = position\_jitterdodge(), color = "black", alpha = 0.6) +   
 scale\_fill\_manual(values = c(cbbPalette[[3]], cbbPalette[[4]])) +  
 theme\_classic() +  
 facet\_wrap(~Cultivar)  
DON\_plot



DON\_data$Treatment <- factor(DON\_data$Treatment, levels = c("NTC", "Fg", "Fg + 37", "Fg + 40", "Fg + 70"))  
  
DON\_15 <- ggplot(DON\_data, aes(x = Treatment, y = X15ADON, fill = Cultivar)) +  
 geom\_boxplot(outliers = F) +  
 ylab("DON (ppm)") +  
 xlab("") +  
 geom\_jitter(pch = 21, position = position\_jitterdodge(), color = "black") +   
 scale\_fill\_manual(values = c(cbbPalette[[3]], cbbPalette[[4]])) +  
 theme\_classic() +  
 facet\_wrap(~Cultivar)  
DON\_15

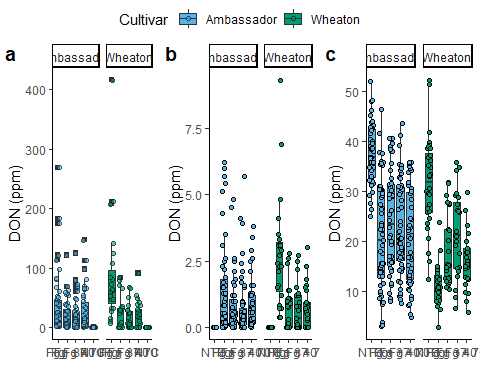


seedmass <- ggplot(DON\_data, aes(x = Treatment, y = MassperSeed\_mg, fill = Cultivar)) +  
 geom\_boxplot(outliers = F) +  
 ylab("DON (ppm)") +  
 xlab("") +  
 geom\_jitter(pch = 21, position = position\_jitterdodge(), color = "black") +   
 scale\_fill\_manual(values = c(cbbPalette[[3]], cbbPalette[[4]])) +  
 theme\_classic() +  
 facet\_wrap(~Cultivar)  
seedmass



### **Combined plot**

ggarrange(DON\_plot, DON\_15, seedmass, labels = "auto", ncol = 3, nrow = 1, common.legend = T)



## **Statistical comparison**

stats\_donplot <- DON\_plot +   
 geom\_pwc(aes(group = Treatment), method = "t\_test", label = "{p.adj.format}{p.adj.signif}", p.adjust.method = "fdr")  
  
stats\_DON\_15 <- DON\_15 +   
 geom\_pwc(aes(group = Treatment), method = "t\_test", label = "{p.adj.format}{p.adj.signif}", p.adjust.method = "fdr")  
  
stats\_seedmass <- seedmass +   
 geom\_pwc(aes(group = Treatment), method = "t\_test", label = "{p.adj.format}{p.adj.signif}", p.adjust.method = "fdr")