Data Visualization

Maria Leticia

2023-02-27

## Ex.1

### a. What is a geom?

A ‘geom’ (geometric object) is used within ggplot to create layers. For example, you can use geom to create scatter plot, box plots, and line diagrams.

### b. What is a facet?

A facet partitions a plot into a matrix of panels. Faceting separate the data up into subsets, according to a categorical variable, then generates separate plots for each subset.

### c. Explain the concept of layering.

The ‘ggplot’ works in layers. A layer is a combination of data, stat and/or geom. Layering is the act of adding layers in a graph. The ‘+’ is used to add layers and it is placed at the end of each line containing a layer.

### d. Where do you add x and y variables and map different shapes, colors, and other attributes to the data?

We use aesthetic mapping (aes) to add x and y variables and map different shapes, colors, and other attributes to the data.

### e. Explain what ‘jittering’ does and why it’s important.

The ‘jitter’ option is a tool to help better visualize the data. By ‘jittering’ we add random variation to the observations before generating the scatter plot. It is a useful way of handling overplotting, resulting in a better visual sense of how many observations have each pair of x and y values.

## Ex.2

library (ggplot2)  
  
mycotoxin <- read.csv ("MycotoxinData.csv", na.string = "NA")  
  
str (mycotoxin)

## '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: chr "10.29130435" "12.80322581" "2.846666667" "6.5" ...  
## $ DON : chr "107.3" "32.6" "416" "211.9" ...  
## $ X15ADON : chr "3" "0.85" "3.5" "3.1" ...

mycotoxin$DON1 <- as.numeric(mycotoxin$DON)

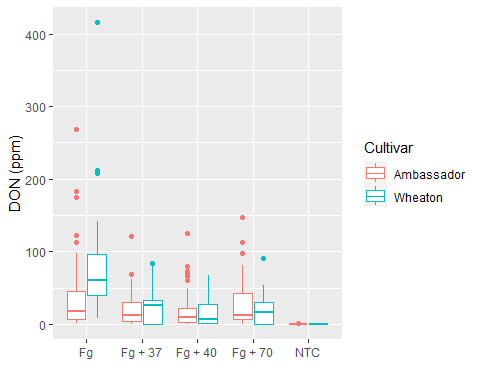
## Warning: NAs introduced by coercion

str (mycotoxin)

## 'data.frame': 375 obs. of 7 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: chr "10.29130435" "12.80322581" "2.846666667" "6.5" ...  
## $ DON : chr "107.3" "32.6" "416" "211.9" ...  
## $ X15ADON : chr "3" "0.85" "3.5" "3.1" ...  
## $ DON1 : num 107.3 32.6 416 211.9 124 ...

ggplot(mycotoxin, aes (x= Treatment, y= DON1, color= Cultivar)) +  
 geom\_boxplot () +  
 xlab ("") +  
 ylab ("DON (ppm)")

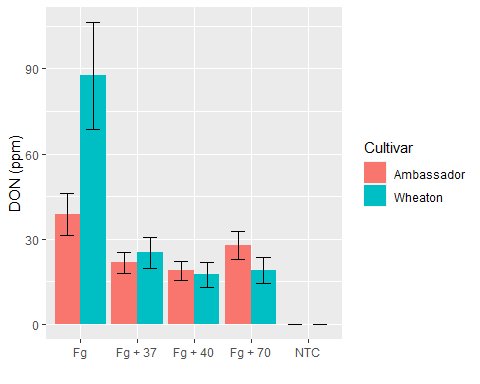
## Warning: Removed 8 rows containing non-finite values (`stat\_boxplot()`).



## Ex.3

ggplot(mycotoxin, aes (x= Treatment, y= DON1, fill=Cultivar)) +  
stat\_summary (fun = mean, geom = "bar", position = "dodge") +  
stat\_summary (fun.data = mean\_se, geom = "errorbar", position = position\_dodge(width=0.9), width = 0.5) +  
xlab ("") +  
ylab ("DON (ppm)")

## Warning: Removed 8 rows containing non-finite values (`stat\_summary()`).  
## Removed 8 rows containing non-finite values (`stat\_summary()`).

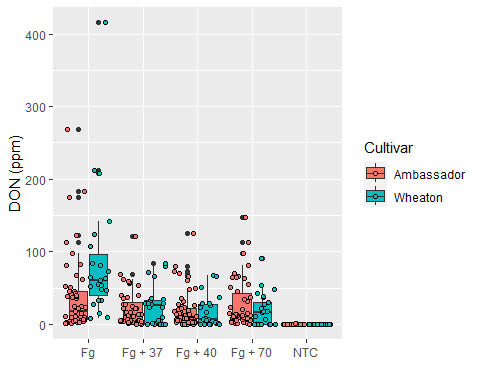


## Ex.4

ggplot(mycotoxin, aes(x=Treatment, y=DON1, fill=Cultivar)) +  
 geom\_boxplot() +  
 geom\_point(shape=21, color= "black", position=position\_jitterdodge(dodge.width=0.9)) +  
 xlab (" ") +  
 ylab ("DON (ppm)")

## Warning: Removed 8 rows containing non-finite values (`stat\_boxplot()`).

## Warning: Removed 8 rows containing missing values (`geom\_point()`).

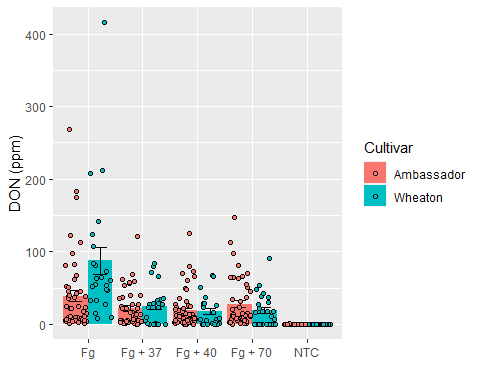


ggplot(mycotoxin, aes (x= Treatment, y= DON1, fill=Cultivar)) +  
 stat\_summary (fun = mean, geom = "bar", position = "dodge") +  
 stat\_summary (fun.data = mean\_se, geom = "errorbar", position = position\_dodge(width=0.9), width = 0.5) +  
 geom\_point(shape=21, color= "black", position=position\_jitterdodge(dodge.width=0.9)) +  
 xlab ("") +  
 ylab ("DON (ppm)")

## Warning: Removed 8 rows containing non-finite values (`stat\_summary()`).

## Warning: Removed 8 rows containing non-finite values (`stat\_summary()`).

## Warning: Removed 8 rows containing missing values (`geom\_point()`).

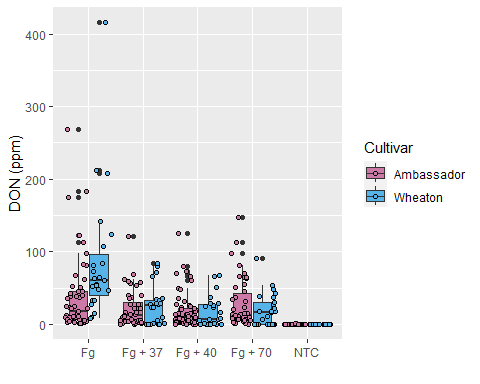


## Ex.5

cbbPalette <- c("#000000", "#E69F00", "#56B4E9", "#009E73", "#F0E442", "#0072B2",   
"#D55E00", "#CC79A7")  
  
  
ggplot(mycotoxin, aes(x=Treatment, y=DON1, fill=Cultivar)) +  
 geom\_boxplot() +  
 scale\_fill\_manual(values=c("#CC79A7", "#56B4E9")) +  
 geom\_point(shape=21, color= "#000000", position=position\_jitterdodge(dodge.width=0.9)) +  
 xlab (" ") +  
 ylab ("DON (ppm)")

## Warning: Removed 8 rows containing non-finite values (`stat\_boxplot()`).

## Warning: Removed 8 rows containing missing values (`geom\_point()`).

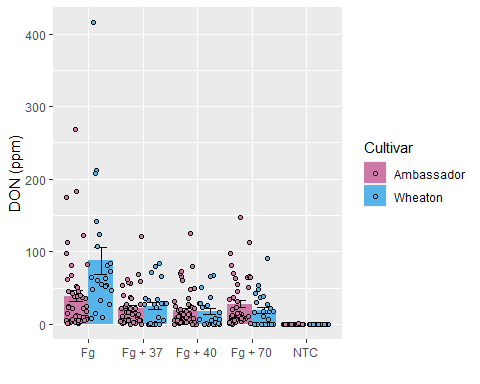


ggplot(mycotoxin, aes (x= Treatment, y= DON1, fill=Cultivar)) +  
 stat\_summary (fun = mean, geom = "bar", position = "dodge") +  
 stat\_summary (fun.data = mean\_se, geom = "errorbar", position = position\_dodge(width=0.9), width = 0.5) +  
 scale\_fill\_manual(values=c("#CC79A7", "#56B4E9")) +  
 geom\_point(shape=21, color= "#000000", position=position\_jitterdodge(dodge.width=0.9)) +  
 xlab ("") +  
 ylab ("DON (ppm)")

## Warning: Removed 8 rows containing non-finite values (`stat\_summary()`).

## Warning: Removed 8 rows containing non-finite values (`stat\_summary()`).

## Warning: Removed 8 rows containing missing values (`geom\_point()`).

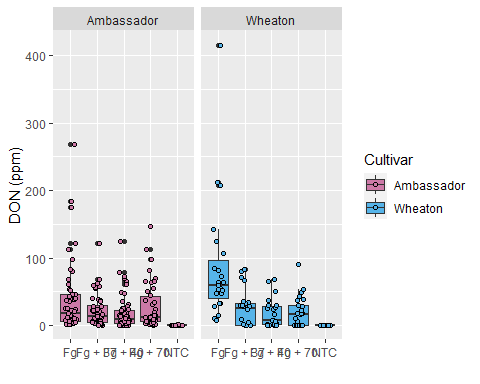


## Ex.6

ggplot(mycotoxin, aes(x=Treatment, y=DON1, fill=Cultivar)) +  
 geom\_boxplot() +  
 scale\_fill\_manual(values=c("#CC79A7", "#56B4E9")) +  
 geom\_point(shape=21, color= "#000000", position=position\_jitterdodge(dodge.width=0.9)) +  
 xlab (" ") +  
 ylab ("DON (ppm)") +  
 facet\_wrap(~Cultivar)

## Warning: Removed 8 rows containing non-finite values (`stat\_boxplot()`).

## Warning: Removed 8 rows containing missing values (`geom\_point()`).

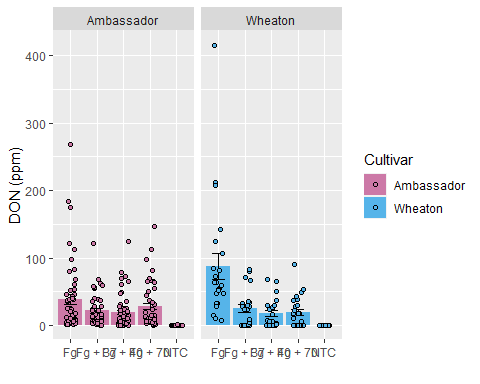


ggplot(mycotoxin, aes (x= Treatment, y= DON1, fill=Cultivar)) +  
 stat\_summary (fun = mean, geom = "bar", position = "dodge") +  
 stat\_summary (fun.data = mean\_se, geom = "errorbar", position = position\_dodge(width=0.9), width = 0.5) +  
 scale\_fill\_manual(values=c("#CC79A7", "#56B4E9")) +  
 geom\_point(shape=21, color= "#000000", position=position\_jitterdodge(dodge.width=0.9)) +  
 xlab ("") +  
 ylab ("DON (ppm)") +  
 facet\_wrap(~Cultivar)

## Warning: Removed 8 rows containing non-finite values (`stat\_summary()`).

## Warning: Removed 8 rows containing non-finite values (`stat\_summary()`).

## Warning: Removed 8 rows containing missing values (`geom\_point()`).

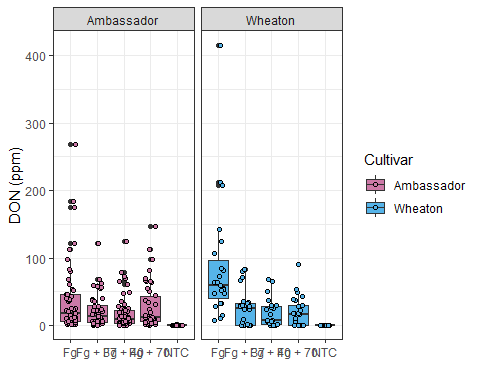


## Ex. 7

ggplot(mycotoxin, aes(x=Treatment, y=DON1, fill=Cultivar)) +  
 geom\_boxplot() +  
 scale\_fill\_manual(values=c("#CC79A7", "#56B4E9")) +  
 geom\_point(shape=21, color= "#000000", position=position\_jitterdodge(dodge.width=0.9)) +  
 xlab (" ") +  
 ylab ("DON (ppm)") +  
 facet\_wrap(~Cultivar) +  
 theme\_bw()

## Warning: Removed 8 rows containing non-finite values (`stat\_boxplot()`).

## Warning: Removed 8 rows containing missing values (`geom\_point()`).

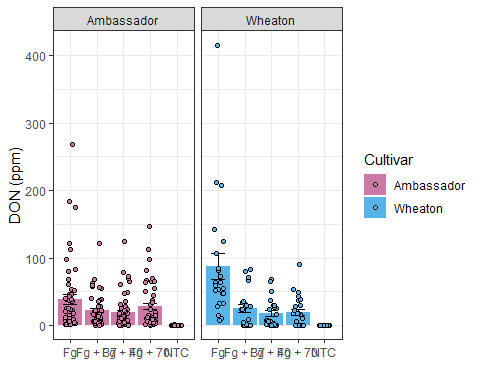


ggplot(mycotoxin, aes (x= Treatment, y= DON1, fill=Cultivar)) +  
 stat\_summary (fun = mean, geom = "bar", position = "dodge") +  
 stat\_summary (fun.data = mean\_se, geom = "errorbar", position = position\_dodge(width=0.9), width = 0.5) +  
 scale\_fill\_manual(values=c("#CC79A7", "#56B4E9")) +  
 geom\_point(shape=21, color= "#000000", position=position\_jitterdodge(dodge.width=0.9)) +  
 xlab ("") +  
 ylab ("DON (ppm)") +  
 facet\_wrap(~Cultivar) +  
 theme\_bw()

## Warning: Removed 8 rows containing non-finite values (`stat\_summary()`).

## Warning: Removed 8 rows containing non-finite values (`stat\_summary()`).

## Warning: Removed 8 rows containing missing values (`geom\_point()`).

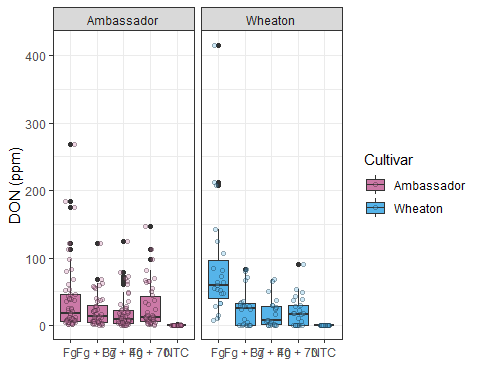


## Ex. 8

ggplot(mycotoxin, aes(x=Treatment, y=DON1, fill=Cultivar)) +  
 geom\_boxplot() +  
 scale\_fill\_manual(values=c("#CC79A7", "#56B4E9")) +  
 geom\_point(shape=21, color= "#000000", position=position\_jitterdodge(dodge.width=0.9), alpha = 0.3) +  
 xlab (" ") +  
 ylab ("DON (ppm)") +  
 facet\_wrap(~Cultivar) +  
 theme\_bw()

## Warning: Removed 8 rows containing non-finite values (`stat\_boxplot()`).

## Warning: Removed 8 rows containing missing values (`geom\_point()`).

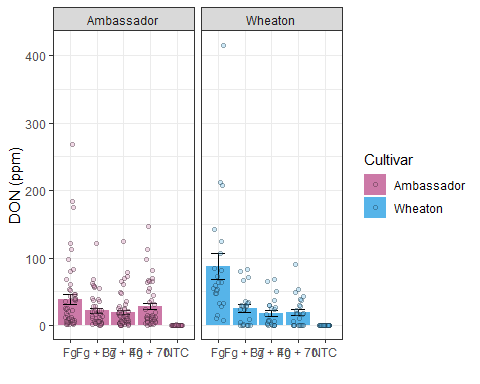


ggplot(mycotoxin, aes (x= Treatment, y= DON1, fill=Cultivar)) +  
 stat\_summary (fun = mean, geom = "bar", position = "dodge") +  
 stat\_summary (fun.data = mean\_se, geom = "errorbar", position = position\_dodge(width=0.9), width = 0.5) +  
 scale\_fill\_manual(values=c("#CC79A7", "#56B4E9")) +  
 geom\_point(shape=21, color= "#000000", position=position\_jitterdodge(dodge.width=0.9), alpha =0.3) +  
 xlab ("") +  
 ylab ("DON (ppm)") +  
 facet\_wrap(~Cultivar) +  
 theme\_bw()

## Warning: Removed 8 rows containing non-finite values (`stat\_summary()`).

## Warning: Removed 8 rows containing non-finite values (`stat\_summary()`).

## Warning: Removed 8 rows containing missing values (`geom\_point()`).



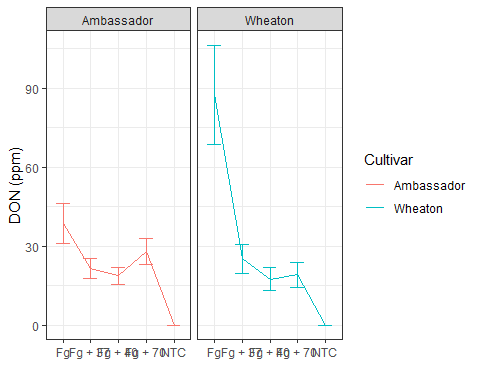
## Ex. 9

The boxplot and the barchart are showing that the inoculation of Fg with the fungal endophytes (37, 40, and 70) resulted in a lower level of Deoxynivalenol mycotoxin in the wheat grains for both cultivars tested compared to the Fg alone. However, cultivar Wheaton showed better response in the reduction of mycotoxin when inoculated with Fg + fungal endophytes than cultivar Ambassador. Therefore, we can conclude that inoculation with fungal endophytes are efficient to protect the wheat plants againt *Fusarium graminearum*, reducing the mycotoxin DON in the grains. For me, the barchat is better to represent this data because it gives a better visualization of the whole data and it shows better the relationship between the treatments (inoculations) and the mycotoxin content (DON).

## Ex. 10

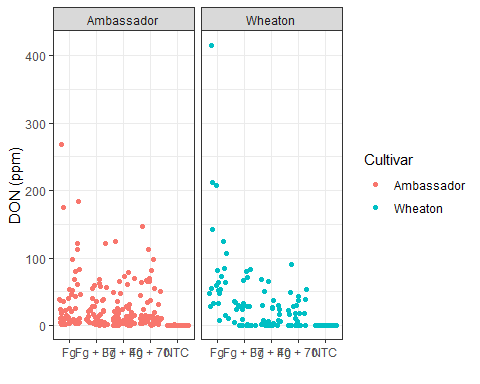
ggplot(mycotoxin, aes (x= Treatment, y= DON1, group=Cultivar, color=Cultivar)) +  
 stat\_summary (fun = mean, geom = "line") +  
 stat\_summary (fun.data = mean\_se, geom = "errorbar", width = 0.5) +  
 xlab ("") +  
 ylab ("DON (ppm)") +  
 facet\_wrap(~Cultivar) +  
 theme\_bw()

## Warning: Removed 8 rows containing non-finite values (`stat\_summary()`).  
## Removed 8 rows containing non-finite values (`stat\_summary()`).



ggplot(mycotoxin, aes (x= Treatment, y= DON1, group = Cultivar, color=Cultivar)) +  
 geom\_jitter()+  
 xlab ("") +  
 ylab ("DON (ppm)") +  
 facet\_wrap(~Cultivar) +  
 theme\_bw()

## Warning: Removed 8 rows containing missing values (`geom\_point()`).



I would choose the first one (the lines connecting the mean) because it is better to visualize the data than the second one (Jitter points) and to show the relationship between the treatment and DON. The Jitter points make the visualization of the points easier but at the same time is a lot of information which is hard to understand what is happening to the data.