

Introduction to visualization in R – 25 pts

PLEASE READ THIS BEFORE CONTINUING

This assignment is designed to help you visualize data in R and practice your R coding skills. You may collaborate with a partner to enhance your learning experience. Please ensure the following:

- Collaboration: If you worked with a partner, include their name.
- Submission: Only one person should submit the assignment.
- Answers: Provide written answers to all questions in the current document or a rendered markdown file.
- Code Submission: Submit your R code in an .R or .Rmd file by sharing a link to your GitHub repository.

Partner 1: Vaibhav Shelar

Partner 2: _____

Data description

Here, we will analyze mycotoxin data collected from wheat heads grown in the greenhouse inoculated with *Fusarium graminearum* (Fg) with or without fungal endophytes to protect against Fg. The experiment was performed by inoculating the heads with Fg (positive control), Fg + 1 of 3 endophytes (Fg + 40, Fg + 70, or Fg + 37), and the non-treated control (NTC). Therefore, the experiment was five treatments with 6-10 technical replicates and three experimental replicates. The authors quantified the parts per million (ppm) of mycotoxins Deoxynivalenol (DON) and the variant 15ADON from the wheat grain, and this is your primary response variable. In other words, we are interested in knowing if the fungal endophytes could reduce the concentration of DON on wheat heads. This was done across two wheat cultivars, a winter wheat variety, Ambassador, and a spring wheat variety, Wheaton. These data are published here:

- 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>
 - The following github repository is associated with this paper: <https://github.com/noelzach/EndophyteBiocontrol>
1. 5 pts. Explain the following concepts about ggplot
 - a. **What three elements do you need to produce a ggplot?**
 - 1) A Dataframe containing data
 - 2) Aesthetics (How the columns of the data frame can be translated into positions, colors, sizes, and shapes of graphical elements)
 - 3) Geom (Geometric object)
 - b. **What is a geom?**

Geom is the basic shape of the elements on the plot.
 - c. **What is a facet?**

A facet is a way to split a single plot into multiple smaller panels. Each display different subset of the data.

d. Explain the concept of layering.

Layering is the process of building a complex plot by adding individual visual elements. It allows to combine different geometric elements like points, bars, and lines with the plot one over another.

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

The x and y variables and map different shapes, colors, and other attributes to the data can be added in the aesthetics function `{aes()}` of the `ggplots2`.

2. 4 pts. Make a boxplot using `ggplot` with `DON` as the y variable, `treatment` as the x variable, and color mapped to the wheat cultivar. Show the code you use to load the libraries you need to read in the data and make the plot. Change the y label to "`DON (ppm)`" and make the x label blank.
3. 2 pts. Now convert this data into a bar chart with standard-error error bars using the `stat_summary()` command.
4. 2 pts. Add points to the foreground of the boxplot and bar chart you made in question 3 that show the distribution of points over the boxplots. Set the `shape = 21` and the outline color black (hint: use `jitter_dodge`).
5. 2 pts. Change the fill color of the points and boxplots to match some colors in the following colorblind pallet.

```
cbbPalette <- c("#000000", "#E69F00", "#56B4E9", "#009E73", "#F0E442",  
               "#0072B2", "#D55E00", "#CC79A7")
```

6. 2 pts. Add a facet to the plots based on cultivar.
7. 2 pts. Add transparency to the points so you can still see the boxplot or bar in the background.
8. 2 pts. Explore one other way to represent the same data <https://ggplot2.tidyverse.org/reference/> . Plot them and show the code here. Which one would you choose to represent your data and why?
9. 4 pts. **Annotate your code and push it to github. Now, find a partner if you don't already have one and have that person "fork" the repository containing the code you just pushed to github. Describe what just happened when you "forked" the repository. Please include the links (URLs) below to answer this question.**

Ans:

When I forked the repository, it asked me to give a name and add description to the new fork repository.

Link: https://github.com/vbs0303/PLPA-6820-Vaibhav_Maryam-Repository.git