Assignment 2 - Data visualization

Your name here

9/16/2021

Guidelines

Use R markdown to complete your assignment. Please provide all the code to make your work reproducible. **Extra**: set image size to 6" x 6". (see here)

Data

Load the "Data file of Heat tolerance of chickpea genotypes in thermal zone of Ethiopia.xlsx" dataset and print all the variables it includes.

```
chickpea <-
  readxl::read_excel("../../data/Data file of Heat tolerance of chickpea genotypes in thermal zone of E
```

Points

Plot a 3-dimensional plot (hint: use colors!) that visualizes Biomass and yield as the most relevant features, but also shows a quantification seed weight.

Lines

Take the lotus.csv data and plot a time series of total biomass ("pac"), stem ("tallo"), leaves ("h_t") and roots ("rsum") of only the control group (i.e. trat == "c"). Show points as means and error bars as standar deviation. Make sure that the difference in the dynamics of the genotypes ("planta") are differentiated. Notice the order of the display of the genotypes is alphabetical. Change it to

1. Lotus corniculatus
2. Trébol blanco
3. Trébol frutilla
4. Trébol balansa
5. Lotus tenuis

```r
lotus <- read.csv("../../data/lotus.csv")
```

Histograms

What is the ditribution of frequence of Days to first flowering of the chickpea dataset? And Days to first Pod?

Violins

Do these distributions change depending on the treatment?

Boxplot

What is the Yield distribution between treatments?

Maps

Create a US map with corn yield means at the county level using the USDA data.

Create a US map with corn yield means at the state level using the USDA data, only for years 2012, 2016 and 2018 (one plot each using `facet_*()`). Compare the yields between years.