PRACTICAL 5: DATA ANALYSIS & PRESENTATION OF RESULTS

Overview

Back in Practical 2, you developed the research question, hypothesis, and prediction you want to address in your lab report. **Now you will analyze the data and interpret the results for your lab report.**

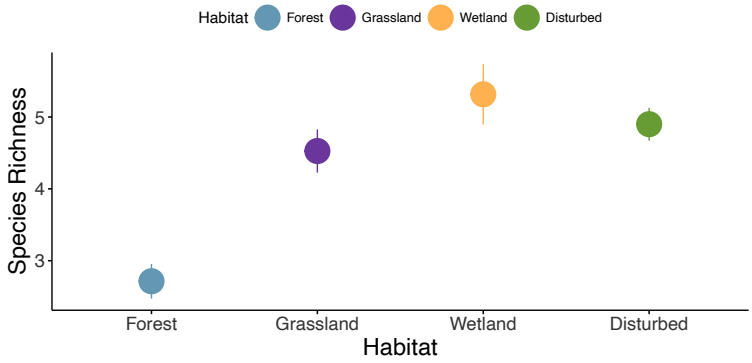
Types of Analyses

There are 5 different script files that can be used, but you only need to use the script(s) suggested for you on the QHP Student Reference posted on Quercus. Below is a brief description of each analysis and an example figure you could create.

Script 1: Continuous Response and Categorical Predictor. In this analysis, you are **comparing a continuous response across habitats following the human disturbance ranking**. Your response can be biotic (e.g., species richness) or abiotic (e.g., pH), as long as it is continuous. **You will also calculate the mean and standard error** for your response variables across the habitats.

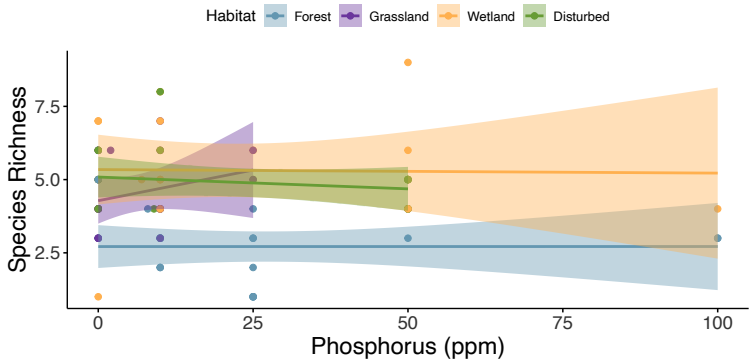
Human disturbance ranking (lowest to highest) = Forest, Grassland, Wetland, and Disturbed

Here is an example figure for Script 1:



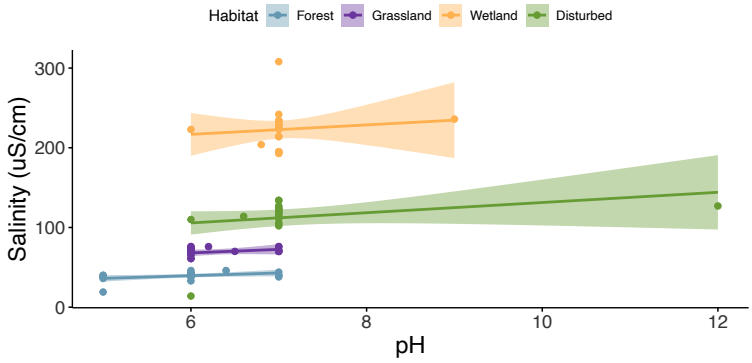
Script 2: Biotic Response and Continuous Predictor. You will use this script if you want to **compare a continuous response against a continuous predictor**. Your **response needs to be a continuous biotic variable**. You will **calculate the slope estimate** for the relationship between the response and predictor for each habitat. **This analysis is similar to a linear regression**, where you are expecting the predictor variable to have a (causal) effect on the response.

Here is an example figure for Script 2:



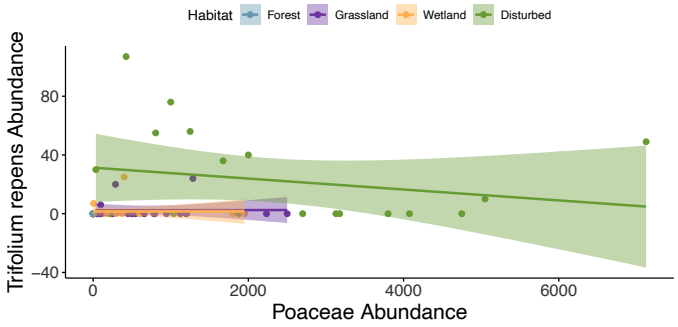
Script 3: Abiotic Y and Abiotic X. You will use this script if you want to **compare the correlation between 2 soil chemistry variables**. You will also **calculate the correlation coefficient for the relationship** between the variables (correlation). **In a correlation, it does not matter which variable is on the x- or y-axis** as we do not make assumptions about which variable is causing the effect on the response. Instead, we are just seeing how the variables are related.

Here is an example figure for Script 3:



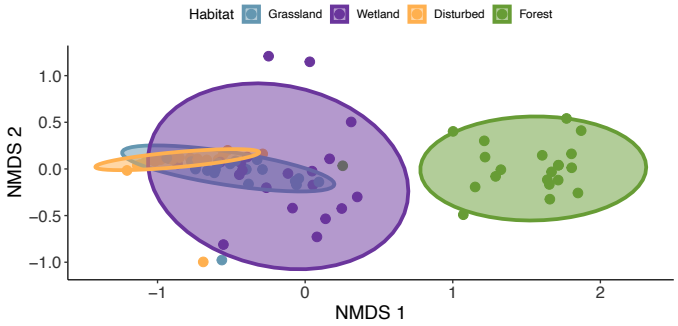
Script 4: Biotic Y and Biotic X. You will use this script if you want to **compare the correlation between 2 biotic variables**. You will also **calculate the correlation coefficient for the relationship** between the variables (correlation). **In a correlation, it does not matter which variable is on the x- or y-axis** as we do not make assumptions about which variable is causing the effect on the response. Instead, we are just seeing how the variables are related.

Here is an example figure for Script 4:



Script 5: Community Composition. You will use this script if you want to **compare community composition across habitats**. Here, there are no summary statistics (e.g., mean and standard error, slope, correlation coefficient), but you will **look for the overlap between ellipses across the habitats**.

Here is an example figure for Script 5:



Instructions

You will use the suggested script(s) to analyze the data, interpret the results, and address your QHP. You should refer to your QHP submission on Quercus, where you will also see the script file(s) suggested for your data analysis. The **provided scripts are a template for you to use and modify**. You will copy, paste, and edit the script(s) until you have completed all the analyses you need to do for your QHP and lab report.

As a reminder, your QHP need to:

1. **Compare biodiversity/abundances/composition between at least 2 habitats.**
2. **Include at least 2 biotic and 1 abiotic variables *or* 1 biotic and 2 abiotic variables.**
3. **Involve at least 2 comparisons/analyses.**
    a. Use 2 analyses to build a story. What is affecting biodiversity or abundances? How do these factors vary across habitats or with other biotic or abiotic variables?
    b. You can repeat analyses! You might have different responses but the same predictor or the same response but different predictors, all using the same analysis.

Please make use of the time in practical and the remaining student hours if you have any questions or problems with the data analysis and interpretation.

Assignment

You will continue analyzing the data until you have all the results you need to answer your QHP, but the practical assignment is just a progress check.

You will submit at least 1 completed figure and a concise (1-2 sentences) interpretation of the results (summary statistics like the mean and standard error or slope are preferred), and **the assignment will be submitted through Quercus by June 11 at 11:59 PM**. A 10% deduction will be applied for each day the assignment is late, weekends inclusive, to a maximum of 5 days. After 5 days, a zero will be recorded. Each day is delineated as a 24-hour period from the deadline.

Your data analysis and results do not need to be completed by this time. The **assignment is designed to make sure you are making progress** for your lab report and able to analyze the data and present the results.

Example interpretations:

- **Script 1** = “Species richness was highest in the wetland (5.32 ± 0.42) and lowest in the forest (2.71 ± 0.24).”
    - **You will report the mean and standard error for the response variable.**
- **Script 2** = “Species richness increased with total nitrogen in the wetland (β = 0.00754) but decreased in the grassland (β = –0.03).”
    - **You will report the slope estimate (β) to 3-5 decimal places.**
- **Scripts 3 and 4** = “The relationship between pH and salinity was positive across all habitats, but strongest in the forest (r = 0.410) and weakest in the wetland (r = 0.117).”
    - **You will report the correlation coefficient (r) to 3-5 decimal places.**
- **Script 5** = “The grassland and disturbed habitats had greatest similarity in composition, while the forest habitat had a unique composition.”
    - There are no formal statistics to report for this analysis, as **you are describing the overlap between ellipses** and how similar or different the habitats are in terms of community composition.

List of Variables

Here are the biotic and abiotic variables that you can use. I have provided the names with the spelling, capitalization, and punctuation you will use in R to perform the analyses:

Variable Name	Description
Biotic	
Species_Richness	Number of different species identified at each flag/quadrat
Species_Evenness	Species richness weighted by relative abundances at each flag/quadrat
Total_Community_Abundance	Sum of the abundances for all plant species at each flag/quadrat
Composition	*Community composition is not necessarily a variable, but it is something that is quantified and plotted using Script 5
Abiotic	
Habitat	Name of the habitat, ranked by human disturbance
Total_Nitrogen	Sum of the nitrite and nitrate measurements (ppm)
Nitrite	Nitrite quantity (ppm)
Nitrate	Nitrate quantity (ppm)
Phosphorus	Phosphorus quantity (ppm)
Potassium	Potassium quantity (ppm)
pH	pH
Salinity	Salinity/conductivity (μS/cm)
Organic_Matter	Ranking of organic matter content, with 1 indicating lowest and 4 indicating highest content

All data were summarized for each flag/quadrat, with separate values for each practical group across all 4 habitats.

Note: R is case sensitive, so make sure you use the correct capitalization and include the underscore in variable names as needed.

Common Plant Species

If you are interested in the abundances of specific plant species, here are the most common plant species:

Common Name	Scientific Name
Grass	Poaceae
Garlic Mustard	Alliaria_petiolata
Black Medic	Medicago_lupulina
White Clover	Trifolium_repens
Canada Thistle	Cirsium_arvense
Crownvetch	Coronilla_varia
Fuller's Teasel	Dipsacus_fullonum

Note: The scientific name is written with the capitalization and punctuation you will need to specify that plant in R.