**Preliminary data analysis**

**GitHub link:** https://github.com/sydwil/ENT6707-G4

**Source:** The data was obtained from the Environmental Data Initiative database

**Any data wrangling undertaken to prepare the data for analysis**

The carnivorous plants have been subset into categories based on their methods of predation (seen in the chart below). Additionally, methods of collection were not consistent across all data (ranging from total number of individuals collected to percentages), so these measures had to be standardized. (we need to standardize them or remove these data ASAP so we can perform the sum. stats)

|  |  |  |  |
| --- | --- | --- | --- |
| **Method of Predation** | **Genera** | **Number of Replicates** | **Number of replicates (w/o % value data)** |
| Passive Trapping | *Pinguicula* | 16 | 10 |
| *Drosera* | 23 | 16 |
| Pitchers | *Nepenthes* | 15 | 1 |
| *Brocchinia* | 1 | 0 |
| *Sarracenia* | 25 | 9 |
| *Triphyophyllum* | 1 | 1 |
| Active Trapping | *Utricularia* | 5 | 5 |
| *Dionaea* | 2 | 0 |

* **Summary statistics for relevant variables, variable types (e.g., categorical vs. continuous)**

**Method of predation:** 3 responses, categorical variable

* **Potential challenges in analysis (e.g., violations of normality; pseudoreplication), and pertinent graphical depictions**
* **If using a published database or data from a publication, please provide a complete citation and explain how your analyses build upon any existing analyses of those data.**

We found through additional research that there was a paper that utilized this data (Ellison & Gotelli 2009) but was not cited on the original Environmental Data Initiative page. Fortunately, this paper compares the prey capture of plants on a genus level, so we may still be able to achieve unique analyses with our subsetting according to their prey capture method. With this prey capture variable, we can compare catch within each group and between the three groups, hopefully demonstrating that the same trap type are more similar to each other. (? I don’t know if we are meant to inject a hypothesis, but this is what I’m thinking?) (also do we *still* think an lm or glm will work for this?)

The first of three graphs using this data is a “star” plot, where each genus has been broken down into proportions of prey. The second graph addresses PIE, or probability of interspecific encounter, which is meant as a means to demonstrate how specialized to certain orders the genus of plants is. The final graph illustrates a similarity index between a few selected samples prey capture. While all analyses are very interesting and may provide useful for further work, we do not intend to use these exact statistical or graphical representations to guide our project. (don’t think I’ll include this in the final draft)

**Citations:**

Ellison, A. M., & Gotelli, N. J. (2009). Energetics and the evolution of carnivorous plants—Darwin's ‘most wonderful plants in the world’. *Journal of Experimental Botany*, *60*(1), 19-42.

Ellison, A. and N. Gotelli. (2023). Prey Capture by Carnivorous Plants Worldwide 1923-2007 ver 18. Environmental Data Initiative. https://doi.org/10.6073/pasta/38b5bf8887b039b3c0b81b4a317f8cbc (Accessed 2024-09-03).