7. Worksheet: Diversity Synthesis

Student Name; Z620: Quantitative Biodiversity, Indiana University

04 January, 2023

OVERVIEW

In this worksheet, you will conduct exercises that reinforce fundamental concepts of biodiversity. Specifically, you will construct a a site-by-species matrix by sampling confectionery taxa. With this primary data structure, you will then answer questions and generate figures using tools from previous weeks, along with wrangling techniques that we learned about in class.

Directions:

1. In the Markdown version of this document in your cloned repo, change "Student Name" on line 3 (above) to your name.
2. Complete as much of the worksheet as possible during class.
3. Refer to previous handout to help with developing of questions and writing of code.
4. Answer questions in the worksheet. Space for your answer is provided in this document and indicated by the ">" character. If you need a second paragraph be sure to start the first line with ">". You should notice that the answer is highlighted in green by RStudio (color may vary if you changed the editor theme).
5. Before you leave the classroom, **push** this file to your GitHub repo.
6. For the assignment portion of the worksheet, follow the directions at the bottom of this file.
7. When you are done, **Knit** the text and code into a PDF file.
8. After Knitting, submit the completed exercise by creating a **pull request** via GitHub. Your pull request should include this file `7.DiversitySynthesis_Worskheet.Rmd` and the PDF output of `Knitr` (`DiversitySynthesis_Worskheet.pdf`).

CONSTRUCT A SITE-BY-SPECIES MATRIX

We will construct a site-by-species matrix using confectionery taxa (i.e, gummies). The instructors have created distinct **sources communities** that vary in the composition of gummy taxa. It might be fun to consider them as distinct geographical regions experiencing different environmental regimes, or different experimental units under different treatments. Each student will sample a source community and then use a taxonomic key to identify gummies and their abundances. While students will measure **alpha diversity** for each sub-sample, students will also collate data with their peers sampled the same source community and create a site-by-species matrix that will be shared via GitHub. Students in in the same group will explore their sampling efforts and their effects on species richness using **coverage** and **rarefaction** concepts, as well as the **data wrangling** tools they have learned during the class. What about the differences among different source communities? Now all class can combine the data using GitHub and explore **beta diversity** with the tools they have learned. It might be interesting to explore differences within (different sites) and among communities (different sources).

SAMPLING PROTOCOL

1. Instructors will assign you to sample confectionery taxa from one of the three designated source community bucket (A, B and C).

2. After randomly sampling one scoop (imagine as an equal biomass) from the source community, each student will count the total number of individuals (N), identify the taxa using the species key and quantify their abundances.

3. Work with other students in your group to assemble data into site-by-species matrix on the white board. One person needs to create a .csv or .txt file and share your group's site-by-species matrix with the class using GitHub. Make sure that you include a sample identifier (student name) and what community you sampled from.

GROUP BRAINSTORM

In smaller groups, take 15 minutes to brainstorm questions, code, statistical tests, and "fantasy figures" using the site-by-species matrix the class generated.

5. Using this data, explore how well did your sampling effort go. You can use rarefaction and coverage tools you have learned earlier.

6. Investigate alpha diversity based on the methods you have learned in the rest of the handout and accompanying worksheet. For example, you can measure richness, Shannon diversity and Simpson index. You can also convert them to effective number of species using the Hill numbers concept.

7. Measure beta diversity using orientation and multivariate statistics methods. For example, you can create a PCoA plot, based on Bray-Curtis dissimilarity, of sites and communities using different shape and color codes. Use Permanova to test if there are differences between communities.

1) Sampling coverage and rarefaction curves

Question 1: Using this data, explore how well did your sampling effort go. Compare your sampling efforts with other groups. Do you think that your samples cover actual diversity? You can use rarefaction and coverage tools you have learned earlier.

Use the space below to generate a rarefaction curve/sample coverage based on the data we collected in class for each community. Make sure to annotate your code using # symbols so others (including instructors) understand what you have done and why you have done it.

Answer 1:

2) Alpha diversity

Question 2: Compare alpha diversity measures within sites and among communities. For example plot richness, Shannon diversity and Simpson index. You can convert them to effective number of species using Hill numbers concept. This would make comparisons easier. How is the variation among the samples in your group and between the communities of other groups? Generate a hypothesis around the diversity metrics you chose and test your hypothesis. Interpret your findings.

Answer 2 - Analysis:

Answer 2 - Plot: With your analysis, create one (and only one, although it can have multiple panels) *publication-quality* figure. Make sure to annotate your code using # symbols so others (including instructors) understand what you have done and why you have done it.

Answer 2 - Interpret results: Write an informative yet succinct (~5 sentences) caption that creates a "stand-alone" figure. Take a peek at figures and figure captions in a paper published in your favorite journal for inspiration.

3) Beta diversity

Question 3: Measure beta diversity using orientation and multivariate statistics methods. For example, you can create a PCoA plot, based on Bray-Curtis dissimilarity, of sites and communities using different shape

and color codes. Use Permanova to test if there are differences between communities. Generate a hypothesis around your chosen analysis and test your hypothesis. Interpret your findings.

Answer 3 - Analysis: Use the space below for code that is being used to analyze your data and test your hypotheses on your chosen beta diversity tool. With your analysis, create one (and only one, although it can have multiple panels) *publication-quality* figure. Make sure to annotate your code using # symbols so others (including instructors) understand what you have done and why you have done it.

Answer 3 - Plot:

Answer 3 - Interpret results: Write an informative yet succinct (~5 sentences) caption that creates a "stand-alone" figure. Take a peek at figures and figure captions in a paper published in your favorite journal for inspiration.

SUBMITTING YOUR ASSIGNMENT

Use Knitr to create a PDF of your completed 7.DiversitySynthesis_Worksheet.Rmd document, push it to GitHub, and create a pull request. Please make sure your updated repo includes both the pdf and RMarkdown files.

Unless otherwise noted, this assignment is due on **Wednesday, February 15th, 2023 at 12:00 PM (noon)**.