

In this assignment, the first task is to load the necessary packages for the code which are `vegan`, `BiodiversityR`, and `labdsv`. The dataset, `spaeth`, is then downloaded into a new object. Some columns in the dataset contain non-numeric data, so to exclude these columns from the matrix, each of these columns is downloaded into an object, and then a matrix is created (which is a subset of the original dataset) which does not contain these new objects. This matrix is rarefied to 50 individuals per sample to reduce the impact of high abundance species, and the rarefied matrix is transformed using the `disttransform` package and the “profiles” method to convert the data into proportional abundances. An ANOSIM test was then performed using the `anosim` package to determine if the rank similarity among the data for each year was greater than the rank similarity between the data of each year. In this case, the R statistic was 0.01 which led to a significance of 0.4 which means that there was no significant difference between these two groups. Likewise, an ANOSIM was performed on the data for the creeks with 10,000 permutations, and an R value of 0.4 was received. A significance of >0.001 was also received which means that the difference between the data for one or more creeks was significantly different than the data for one or more other creeks. Similarly, MRPP tests using the `mrpp` package were conducted for both year and creek. Unlike in the ANOSIM test, the similarity within the matrix was not ranked. In the MRPP test for year, an A value of 0.008 was received, and a significance of 0.05 was received, indicating an insignificant difference in data between years. In the MRPP test for creek, on the other hand, an A value of 0.1 and a significance of >0.001 were received meaning that a significant difference exists between data for one or more creeks. A permANOVA test using the `adonis` package was then performed on the matrix, with an analysis of any interaction between year and creek. Pseudo-F values for these values were 2.3 for year, 8.5 for creek, and 0.9 for interaction between creek and year. In the permANOVA, a

significance of 0.02 was received for the year variable, a significance of >0.001 was received for the creek variable, and a significance of 0.7 was received for the interaction between creek and year. These values indicate that observed differences in year and creek data are unlikely to have been generated randomly and that a significant difference exists between data from one or more creeks and years. However, since most of the data are explained by the creek, another permANOVA test was performed. In this permANOVA, data were only randomized within the creek while controlling for the year. In this test, a significance of >0.001 was observed, and the pseudo-F statistic was slightly reduced from 8.51 to 8.46. These values are highly indicative that a significant difference exists between data for one or more creeks.