MarineDNA

Team Meeting 3/13/23

HPC/GPU access:

-Nastassia is working with NOAA admin to get Google Cloud access

-Will schedule a training session with everyone on the team once our request has been approved (should happen this week)

Coding:

-We have a couple short-term goals that we can work on without HPC/GPU access:

1. Identify outlier ASVs from PCA loading scores

-Eric would like to extract a subset of data from each column of a data frame of loading scores

-code is in Eric’s sandbox, ‘outlier\_test.R’ and ‘pca test.ipynb’ for R and Python versions, respectively

-Katie and Nastassia will try to generate equivalent of R in Python

2. Compare PCA ordinations of Aitchison distances generated from raw read counts (DEICODE) with Euclidean distances generated from the random draws of the beta distribution (MarineDNA code)