1. Concepts you probably need to know
2. Biology
3. Bioinformatics
   1. software
4. Statistical analysis
   1. Setting up your “environment” to perform statistical analysis
      1. R versions
      2. Packages
      3. Use on the server
      4. Conda/Native installation
   2. Data exploration
      1. Sequencing reads
         1. How many raw reads
         2. Quality metrics
         3. How many were reads
      2. Count analysis
         1. Count distributions
            1. Outliers
            2. Summary statistics
            3. Grouped by important study variables?
   3. Statistical analysis
      1. Sequencing stats and alpha diversity (Richness, Shannon’s and Simpson’s diversity)
         1. wilcox.test()
         2. glm()
      2. Ordination cluster testing
         1. https://www.davidzeleny.net/anadat-r/doku.php/en:pcoa\_nmds#:~:text=NMDS%20is%20an%20iterative%20method,properties%20(number%20of%20samples).
         2. PCoA
         3. ANOSIM
         4. ADONIS
      3. Differential abundance testing
         1. fitZig()
         2. ANCOM()
      4. Redundancy analysis
         1. rda()
5. Tips for writing
   1. Example template
6. Challenges
   1. Vocabulary
      1. Overall
         1. ASV
         2. Reads
         3. Scripts
         4. terminal
      2. File names
         1. qza, tsv
   2. manifest creation
      1. first 4 steps difficult
   3. Commands
      1. It’s better to write down your own commands
   4. Setting up R
   5. Navigating server (file download)
   6. Troubleshooting

Contents

[Data exploration 3](#_Toc112668156)

# Data exploration

# Writing tips