Biostatistics
Fall 2024

Final Capstone Project Feedback – 35 Points

Student: Kate Gardner Score: 31.25 /35 = 89.3 %

Part 1: Clean up your Repo – 5 pts Score: 4.75

Using Git/GitHub effectively and organizing a project well

Feedback: Project up to date on GitHub and folders well-organized within it for easy understanding. In Final code folder there is an empty file called Final-LR.qmd which is confusing.

Part 2: Finalize statistical analyses- 20 pts Score: 17.5

Remove unneeded code; Follow correct workflow; Reflects feedback; overall challenge

Feedback: Analysis 1: Mean DO between months. Clarify justification to indicate that JULY is when ps would be at peak. Line 72 - why did you run the model as glm and not specify a family? Is effectively a regular anova. Line 96 - HOW DO YOU KNOW from looking at the tables that a little less than half the variance is explained by month? Your biological insights are good. One thing that you didn't address is the fact that none of your observations from a single location are independent within each site becasue the hydrolab was just deployed at each site for an extended period.NICE final plot with annotations!

Analysis 2 - DO and pH - The reason your data look wonky is that they have temporal autocorrelation and indivdiual points within a site are not independent. Line 181 - in ecology, an adj R2 of 0.45 is really high. It means that 49% of variation in pH is due to one single var!

Part 3: Final report – 10 pts Score: 9

Intro, Analysis with biological insight , Challenges; Well-written; Strong use of markdown

Feedback: Generally good. Book would have helped you figure out what to do for glm to get a better fit with ratio data. I'm sorry you didn't try the time series - it would have been a good exercise for these data because they have temporal autocorrelation.

Markdown is fine, good effort on final plot for Analysis 1.