Learning Data Analysis Skills in Intro Biology Labs with R

Linda Forrester, Rachel Schwartz, Bryan Dewsbury
Dept. of Biological Sciences

Goal

"While colleges should include courses in programming, data visualization and statistics, more students develop digital fluency more quickly and easily when digital tools are integrated throughout the curriculum." 1

Objectives

- Students will be able to
  - Apply fundamental quantitative reasoning skills to analyze data and solve problems
  - Develop professional-quality data analysis skills
  - Visualize and analyze large data sets
  - Conduct data analysis using software R

Rationale

- Writing code to require students to
  - consider the goal of their analysis
  - specify x and y variables
  - consider how changes to code impacts their analysis
  - think and communicate in a new language
- Written code is easily repeatable and modifiable for subsequent labs.
- Our approach allows students to reinforce and scaffold skills.
- R is a data analysis program that is used in many fields.

Description of labs

Lab 1 – Narragansett Bay Plankton Diversity
- Students count plankton samples from two sites in Narragansett Bay
- Students enter counts into a shared spreadsheet to collate data
- Students calculate mean and standard deviation of diversity for different sites
- Students learn the basics of how to use the program R
- Students create boxplots of diversity at different sites for different tidal cycles using explicit code and directions

Lab 2 – Mechanisms of Evolution – effects on snail populations
- Students count changes in snail populations as affected by birds and physical processes
- Students enter counts into a shared spreadsheet to collate data
- Students calculate mean and standard deviation of diversity for different sites
- Students create line graphs of changes in snail populations
- Students add standard deviation error bars to plots
- Students use a t-Test to compare effects of birds and physical processes on snail populations

Lab 3 – Photosynthetic pigments of Narragansett algae (seaweed)
- Students measured pigments in local algae using a spectrophotometer
- Students enter counts into a shared spreadsheet to collate data
- Students create line graphs representing photosynthetic pigment data

Lab 4 – Seedling growth
- Students measured growth of seedlings under two conditions
- Students enter counts into a shared spreadsheet to collate data
- Students calculate mean and standard deviation of diversity for different sites
- Students create line graphs of changes in snail populations

Approach

- Students collect data in lab covering different biological concepts.
- Students analyze their collected data using R.
- Students utilize server-based access to RStudio, which eliminates difficulties with installation on personal computers and allows access anywhere.
- Each lab builds on previous R code.
- When students learn new skills, they are given explicit code.
- When students build skills based on previous labs, they are given suggestions but not complete code.
- When students practice skills, they are expected to write their own code independently based on prior labs.

Rationale for labs

- Labs with R

Future directions

- Students are ready to continue learning applications of data analysis programs and want more instruction to improve their skills.

Discussion

- Students are already reporting progress with writing code, making graphs, and working with data.
- Students with little computer coding experience can develop data analysis skills early in their academic career with this approach. Using coding instruction embedded in lab to analyze data they have collected, students find coding skills to be relevant and useful.
- Students are ready to continue learning applications of data analysis programs and want more instruction to improve their skills.

THANKS TO:

- Dean Kirby and CELS for funding to purchase lab laptops.
- Office of Student Learning Outcomes Assessment & Accreditation for funding to support grad student time for writing R instruction manuals for each assignment.
- Chi Shen, University Computing Systems for maintaining URI R server (CELSRS.uri.edu)

References

2. Black Panther, 2018