

Utilizing the software package “R” to integrate graphical and statistical analysis in undergraduate laboratory courses



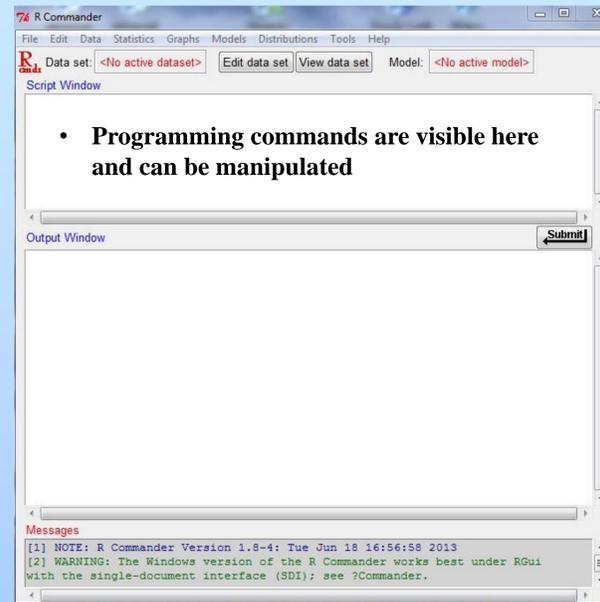
Michael S. Berger
Washington State University Vancouver
msberger@vancouver.wsu.edu



Background

- Student learning goals frequently include graphical and statistical analysis of data
- Students can create graphs and analyze data using point and click software
- Point and click software frequently results in a “black-box” perspective, with limited conceptual understanding of processes involved
- Software that displays coding, which can be modified by the user, allows for a broader understanding of concepts involved in generating graphs and performing statistical analyses
- Proprietary software limits student access
- The software package “R” is open source and can be installed on any computer
- “R” can generate publication quality graphics that can be easily modified to suit students needs and perform a range of statistical analyses

“R Commander”, which functions as a point and click interface



Conclusions

- Steep learning curve for the instructor
- Lots of information available on the Internet
- Graduate teaching assistants will need training
- Detailed instructions that includes specific commands must be provided for students
- Instructions should explain why each command is being used
- Some students did not gain further insight into graphing or statistical analysis, but were able to complete the procedure
- The use of programming forced some students to think about the components involved in graphing or statistical analysis
- Skill set developed by students can be used in future course work

Methods

- An upper division marine ecology course was used to trial the use of “R” in field based laboratory exercises
- Separate labs were used to introduce different graphing and statistical analyses
- Students were instructed on how to format and save data as a tab-delimited text file
- Detailed instructions that included specific programming commands and explained why the commands were being used, were provided to students
- “R Commander”, which is a graphical user interface, was initially used as a learning tool to help students visualize underlying program commands
- As student confidence increased, reliance on “R” Commander” was reduced
- Graphs were exported as an image file or PDF file for use in student papers or presentations

Results

- Students were able to generate scatter plots, bar plots, line, plots, and filled contour plots
- Students were able statistically analyze data using a t-test and ANOVA’s
- Student learning was generally categorized as:
 1. Followed instructions and produced a graph or performed a statistical test; similar to point and click software.
 2. Gained a broader conceptual understanding by going beyond instructions provided. Software coding was manipulated to produce a graph or statistical analysis that suited their specific needs.

Future work

- Integrate “R” into introductory biology labs as a mechanism to increase competency in graphical and statistical analysis
- Teach graduate teaching assistants how to use “R”

Example of instructions provided in a laboratory handout to generate a filled contour plot

Create a contour plot by entering the command:

```
filled.contour(data,col = colorRampPalette(c("yellow", "orange", "red"))(30),  
main = "Salinity",xlab="Distance",ylab="Depth",ylim=rev(range(Depth)))
```

This will produce a contour plot with a scale bar. The number in parenthesis after colorRampPalette(c("yellow", "orange", "red")) determines the number of colors. If colors repeat, increase that number. To change the color scheme, replace “yellow”, “orange”, “red”, with “blue”, “green”, “red” or any color combination! The command [ylim=rev(range(Depth))] reverses the y-axis so zero is on top.

Filled contour plot generated on “R” from student collected data on salinity as a function of depth and distance from a fixed point

