

Histograms and Boxplots

Suann Yang
Biology Department, SUNY Geneseo

Focus

This lesson, created for an introductory ecology course, focuses on helping novice R users to import a data file, applying base R plotting functions, and using R Markdown to generate a reproducible report.

Overview

Students will be analyzing a subset of data from Karp et al. (2016) to examine how food safety practices affect the sustainability of lettuce production. This lesson takes students through the first step of visualizing the data using Swirl. This lesson begins with a pre-class homework assignment, followed by the Swirl lesson on boxplots (50-minute class period), and then post-class homework. The lesson sets students up to conduct statistical analyses (subsequent lesson) to follow up on the patterns that they observe in the resulting graphs.

Learning objectives

Students will learn how to:

- A. Visualize central tendency and variability of real data in graphs using R and RStudio.
 - a. Apply `read.csv()` in base R
 - b. Apply `hist()` in base R
 - c. Apply `boxplot()` in base R
- B. Interpret boxplots of real data.
 - a. Define median, interquartile range, whiskers
 - b. Compare and contrast boxplots across treatments
- C. Produce a report of graphical results and interpretation using R Markdown.

Computational thinking practices:

Computational thinking practices from the Weintrop et al. (2016) taxonomy that are addressed with this lesson.

Data Practices	Modeling & Simulation Practices	Computational Problem Solving Practices	Systems Thinking Practices
Visualize data		Programming	Communicating information about a system
		Troubleshooting and debugging	

Lesson sequence:

1. Homework: Students install the *swirl* package and download the .swc file.
2. Mini-lecture to introduce the goal of the analysis; students generate hypotheses and predictions in their groups (Powerpoint slide deck provided with lesson materials).
3. In class activity for students to use swirl and complete the boxplot lesson.
4. Homework: Students finish making their boxplots, interpret, and create report of their work using R Markdown.

Implementation notes:

This lesson was piloted with students who were mostly second-year biology majors in a required course (Ecology) for the major. Prior to this course, most students already had some exposure to using R in introductory biology (with the full gamut of experiences from enjoyment to frustration). At the beginning of the course in which this lesson was implemented, students had a short R refresher in the form of adapting code that was provided to them, followed by an introduction to R Markdown. Note that the pilot implementation's course had two undergraduate teaching assistants, who facilitated the Swirl lesson along with the instructor, plus provided one hour of help sessions each outside of class (heavily attended). Lesson .swc and .csv files were distributed via the Canvas course management software.

References:

Karp DS, Moses R, Gennet S, Jones MS, Joseph S, M'Gonigle LK, Ponisio LC, Snyder WE, Kremen C (2016) Agricultural practices for food safety threaten pest control services for fresh produce. *Journal of Applied Ecology* 53(5): 1402-1412. <https://doi.org/10.1111/1365-2664.12707>

Data for this lesson were modified from:

Karp DS, Moses R, Gennet S, Jones MS, Joseph S, M'Gonigle LK, Ponisio LC, Snyder WE, Kremen C (2016) Data from: Agricultural practices for food safety threaten pest control services for fresh produce. Dryad Digital Repository. <https://doi.org/10.5061/dryad.37cc5>