Introduction

- In statistics courses for social-science undergraduate (and most graduate) students, teaching statistical computing isn’t an end in itself.
- Statistical software should support the fundamental goals of the course.
- Because (what I take to be) the fundamental course goals differ by the level of the course, use of statistical software should differ as well.
- I’ll focus on two typical courses: An introductory statistics course, and a second course that emphasises applied regression analysis.
The emphasis (in my opinion) should be on statistical reasoning and basic statistical concepts.

Teaching data analysis is a secondary goal.

Students completing the course should be able to read critically research reports using basic statistical methods; use basic statistical methods in their own work; and pursue further coursework in social statistics.

Examining Data
- Categorical vs. numerical data
- Distribution and variation
- Association
- Visualizing data with statistical graphs

Producing Data
- Observational vs. experimental data (random selection vs. random assignment; causal inference)
Basics of Statistical Inference
- Sampling variation and sampling distribution
- Confidence intervals
- Hypothesis tests
- Possibly likelihood, Bayesian inference

Simplicity and transparency of use
- “Low threshold/high ceiling” (borrowed from Logo)
- Mesh with course goals, integration into course content
- Ability to use on students’ own computers and after course ends
- Encourage good habits (e.g., reproducible research)
Course Goals

- The emphasis (again in my opinion) should be on practical data analysis
- Teach statistical modeling as part of the work-flow of data analysis
- Encourage sound data-analysis practices (e.g., examination of data, model criticism, reproducible research)
- Provide basis for further study, including self-study

Content, based, e.g., on approximately the first half of Fox (2008)

- Examination of data, statistical graphics, transforming data
- Linear regression, linear models
- Generalized linear models
- Model diagnostics
- Visualizing/interpreting fitted models
- Possibly other topics: e.g., missing data, model selection, ...
A Second, Intermediate-Level Statistics Course
Role/Properties of Appropriate Statistical Software

- Flexibility
- Extensiveness
- Extensibility
- Mesh with course goals, integration into course content
- Ability to use on students' own computers and after course ends
- Encourage good habits (e.g., reproducible research)
- These characteristics apply as well to more advanced courses (but may be weighted differently)

R Commander Design Goals
Many present at the origin of the project, circa 2003 — see Fox (2005)

- Familiar menu/dialog-box interface
- Simple work-flow, based on an “active,” modifiable data set
- Where appropriate, an active statistical model on which the user can compute; models are associated with data sets
- Keep menus and dialog boxes simple, with uniform (tabbed) structure
- Make it difficult to do unreasonable things (e.g., take the mean of a categorical variable)
- Simple to install on all platforms — implemented as an R package
- Generate visible, editable scripts of R commands
- Generate dynamic, editable documents with executable R code (in Markdown or LaTeX) to encourage reproducible research
- Cover the content of a basic-statistics course and beyond, including statistical graphics
- Extensibility (via plug-in packages)
The R Commander is more appropriate for the basic statistics course than for the intermediate course (though it would support both).

Students in an intermediate course are better served by learning to write R commands.

Using R within the RStudio interactive development environment facilitates the production of dynamic documents written in Markdown or \LaTeX\.

References

