

Interactivity online and on-campus: Data Analysis and Statistical Inference

Mine Çetinkaya-Rundel
Duke University

GOAL

Design learning materials for a variety of delivery methods (online, on-campus, hybrid), collect and analyze data to evaluate their effectiveness, and enhance the materials with lessons learned from these analyses

Online

Task: Emulate the brick-and-mortar classroom experience of learning R (a statistical programming language) where the instructor can help the student in real time

Solution: *DataCamp* (web-based interactive platform for learning R) as an alternative to static instructions, fully integrated to the Coursera course via the LTI

Visualizing with box plots

```
my_script.R
1 # The cdc data frame is already loaded into the workspace.
2 # Draw the box plot of the respondents heights.
3 # Print the summary.
4 # Print the summary.
5 # Print the summary.
6 # Print the summary.
7 # Print the summary.
```

Numerical data

With our subsetting tools in hand, we'll now return to the task of the day: making basic summaries of the BRFSS questionnaire. We've already looked at categorical data such as our attention to numerical data. Two common ways to visualize numerical data is with a **box plot**.

To construct a box plot for a single variable, you use the `boxplot()` function. Example: `boxplot(cdc$weight)`.

Instructions

- Draw a box plot of `height` using the `boxplot()` function.
- Use the `summary()` function to also get a numerical summary of the variable.
- Compare the two results.

R Console

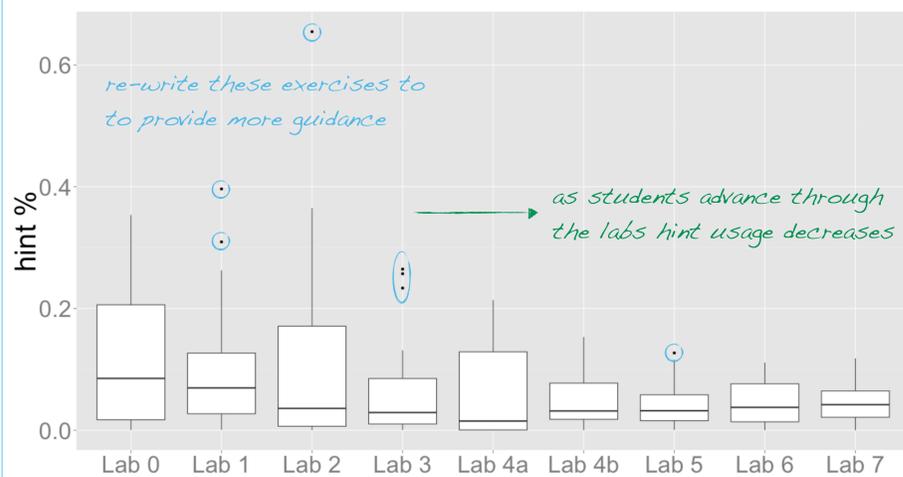
```
> boxplot(cdc$height)
summary(cdc$height)
Confirm that the median and upper and lower quartiles reported in the in the graph. The purpose of a boxplot is to provide a thumbnail sketch of comparing across several categories. So we can, for example, compare height by gender.
```

The notation here is new. The `-` character can be read "versus" or "as a function of".

Target population: Users with little or no background in programming language, especially those who are completely new to R

Data: Click-thru data on viewing hints and solutions as well as performance and time taken to complete labs

[Some] findings: Valuable information on lab components that students most struggle with, data that would be very difficult to effectively collect or glean in brick-and-mortar lab setting:



Next steps: Link up DataCamp and Coursera records to evaluate persistence and performance of students opting to use DataCamp

Student feedback:

"course was a fun, technically cutting-edge (I loved the connection with DataCamp!) way of learning inference and it triggered indeed some curiosity regarding data crunching."

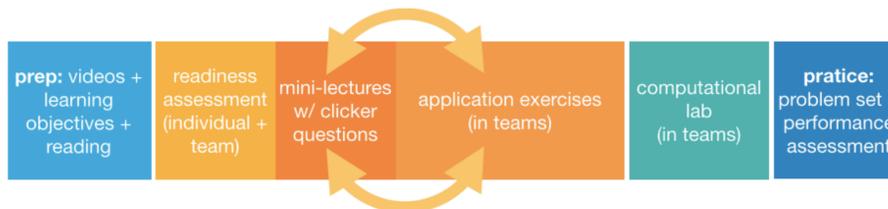
"taught me a lot about R, which was new to me. I like the enthusiasm of the instructor and the user-friendliness of DataCamp."

"lectures, resource materials, books, DataCamp.. everything was so useful and engaging."

On-campus

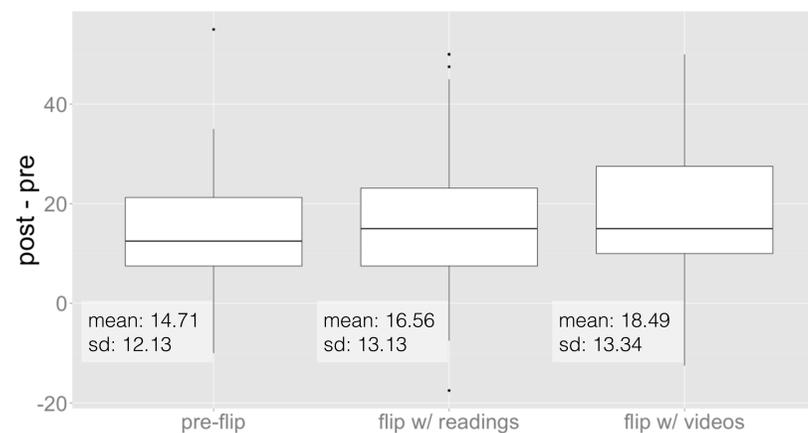
Task: Increase interactivity in the brick-and-mortar classroom by flipping the course, and *do no harm* (to learning and attitude) while doing so

Solution: Use materials developed for Coursera as out-of-class learning materials to prepare students for in-depth hands-on exercises completed in teams in class



Data: Pre-post scores on Comprehensive Assessment of Outcomes in a First Statistics course (CAOS) test

[Some] findings: No statistically significant difference in gains, but no harm done



Next steps: Part of university-wide study on learning and attitude of students in flipped courses, analysis and results forthcoming

Student feedback:

"This class has been extremely helpful and is an excellent example of how a flipped classroom should work. I enjoy the balance of in-class lecture material and applications. The videos outside of class also really address how I learn and have reinforced the material immensely."

"Professor Cetinkaya-Rundel does an amazing job teaching the class. Clickers are a unique way to engage students and to make lectures interesting. The videos, though at times a bit repetitive, are thorough and very helpful. The professor's willingness to gather new data sets and incorporate real-world events into lectures is much appreciated- the associations help make the material much more interesting."

"I just like the idea of the group and the videos because I feel like it helped a lot, especially in explaining difficult concepts. I also like the idea of using the clicker and doing the application question in class because I think it's a better way for me to learn about the information and retain it in my head by putting it into practice."

Other projects

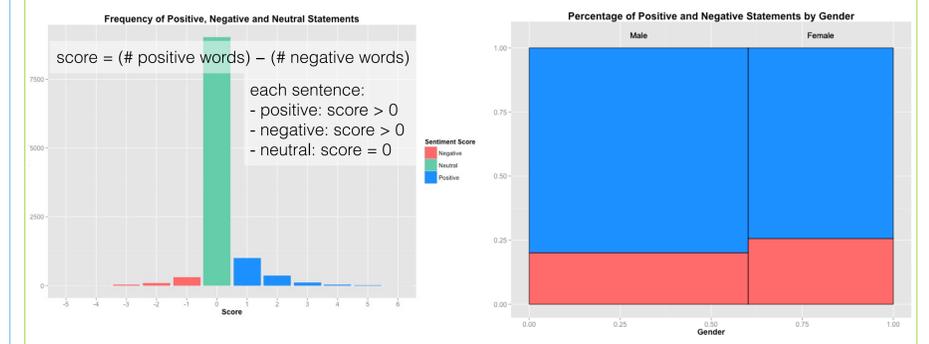
Each project listed below is designed and carried out by one or more members of the project team (Duke students and staff) under the advisement of project leaders (Duke faculty teaching Coursera courses: Dr. Çetinkaya-Rundel & Dr. Canelas)

Identifying characteristics that predict student persistence
by Anthony Weishampel, StatSci

Data: Coursera course records + Duke pre-course survey
Analysis: Define "engaged student" and series of milestones + Model probability of reaching the next milestone + Use variable selection algorithms to identify characteristics associated with increased likelihood of reaching the next milestone

Understanding the mooc student experience through text analysis of interviews
by Heather Shapiro, StatSci & Clara Lee, Chemistry

Data: Pre-study survey for recruitment and personal info + 2x20 30-45 minute semi-structured skype (audio) interviews
Sentiment analysis: Cross reference interview transcripts with an opinion lexicon of positive and negative words
Qualitative analysis: Qualitative coding of transcripts in NVivo + Challenges/barriers, community participation, course components, course feedback, motivation, previous experience — formal analysis and results forthcoming



Engagement, self-regulated learning, and perceptions of motivational strategies
by Kun Li, CIT

Implementation: Attention, Relevance, Confidence, Satisfaction (ARCS) — regularly announce upcoming material & recap past topics, provide accessible materials in many formats, etc.
Data: Interviews
Preliminary observations: In-video questions for attention, own goal most relevant, feedback on assessments would help confidence building, statement of accomplishment helps satisfaction — formal analysis forthcoming
Take-aways: Course emails, pages are minor + instructional video is a big part - applying the ARCS model earlier in video design + Although most students are intrinsically motivated, external rewards seem to have an effect on motivation

Social anxiety and forum posting behavior
by Maria Elena Carvajal, Chemistry

Data: Survey containing questions from Fear of Negative Evaluation (FEN) + Social Avoidance Distress (SAD) + questions on anonymous posting
Stage: IRB approval in progress

find out more

https://stat.duke.edu/~mc301/future_of_moocs/



questions / comments



@minebocek
#CPC15



mine@stat.duke.edu

thank you!



Bass EHD



Online Education Initiatives

