

Class 22: Course wrap-up

June 20, 2018



General

Announcements

- Homework 4 and extra credit Homework 5 due **tonight by 11:59pm**
 - Homework 4 must be submitted before you can turn in Homework 5
- **Final project due dates**
 - **Annotations first draft:** 12:00pm noon on Thursday, June 21st (tomorrow)
 - **Peer reviews:** 6:00pm on Thursday, June 21st (tomorrow)
 - **Annotations and final draft:** 9:00am on Friday, June 22nd
 - **Comparative discussion of simulations:** 10:30am on Friday, June 22nd
- **Final interviews scheduled during final exam period:** Friday, June 22nd between 10:30am and 1:15pm

Fill out evaluation forms

Overview of Final Interview

Overview of final interview

- Arrive at Research Hall 249 for your designated time slot (schedule available on Slack)
- Upon arrival, wait outside until you are asked to enter
 - I may still be meeting with another student
 - I need a couple minutes to reset things after each student leaves
- You do not need to bring anything with you

Overview of final interview

- Final interview is structured to be 10 minutes in length, with 2 extra minutes for a buffer
 - Will be enforced using a timer
- This is loosely modeled after "code interviews" that some employers include in their interview process
- Your objective will be to demonstrate that you can complete tasks similar to what we've practiced during the semester in order to answer simple questions
- When you enter the room, I will hand you a sheet with questions that I want you to answer
- You will be provided with a computer (my laptop) logged into RStudio Server, and printed versions of the RStudio cheat sheets that you can consult at any time while answering the questions
- You may also use R's documentation (running `?geom_histogram` for example)
- You must stay on the RStudio Server page the entire time, meaning you cannot perform internet searches

Overview of final interview

- Questions will ask you to manipulate data in some way (transformation, filtering, aggregation) and also to visualize it
- **Grading**
 - You begin the interview with a 100%, to keep it you need to answer the questions within the allotted time
 - I will provide hints if you get stuck at different parts of a question, however hints lower your score
 - The more hints I need to give you (or the larger those hints are), the more points that I deduct
 - Not finishing all the questions in the allotted time will also deduct points

Overview of final interview

- Be familiar with the primary functions we learned for `ggplot2` and `dplyr`!
 - For `dplyr`, this includes `filter()`, `mutate()`, `group_by()` and `summarize()`, it is important you remember how to use these two commands
 - For `ggplot2`, you should definitely be able to make histograms (including how to convert it into a PMF, how to adjust bin size, etc.) scatter plots, and you should also know how to use the `color = variable` and `fill = variable` inputs for `aes()` and how to use `facet_wrap()` and `facet_grid()`
- Spend time looking at the cheat sheets so that you can efficiently use it to look up information
- Review slides from the first two weeks of classes, and chapters 3 and 5 in the *R for Data Science* textbook