UConn PCS: Data Science, Summer 2022

July 11, 2022

Instructors:
Haim Bar: haim.bar@uconn.edu
HaiYing Wang: haiying.wang@uconn.edu
Jun Yan: jun.yan@uconn.edu

HuskyCT: Login with your NetID and password at `https://lms.uconn.edu`

Lectures: @ Babbidge Library 2118
MWF: 09:00–12:00 ET
TuTh: 14:00–16:00 ET

Office Hours: TBA

Course Description: Data science is a fast developing science of extracting meaningful information from massive data for better decision making. It is interdisciplinary by nature, involving statistics, computing, and domain knowledge. Important principles of data science will be elaborated through interactive simulations, games and examples.

Teaching Notes: Evolving pdf version in HuskyCT.

Course Material: Topics include
1. Introduction to R/RStudio, random number generation, data summaries
2. Exploratory data analysis, visualization
3. Paradoxes explained
4. At the infinity horizon (law of large number; central limit theorem)
5. Hide-and-seek: Estimating the unknowns
6. Sampling design, data collection
7. Correlation and regression
8. Hypothesis testing
9. A live project/capstone

Announcements, homework assignment, and other course information will be posted on HuskyCT.

Computing: Students are required to use `R` and `R Markdown` for homework assignments, exam, and project.

Useful Resources: Pick them up on the fly.
- A quick introduction to R programming language at `https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf`
- For an even shorter introduction see `https://cran.r-project.org/doc/contrib/Torfs+Brauer-Short-R-Intro.pdf`
- Former UConn Stat PhD student Dr. Wenjie Wang's tutorial on R: `https://github.com/wenjie2wang/2018-01-19-siam`
- Introduction to R Markdown by Garrett Grolemund: `https://rmarkdown.rstudio.com/articles_intro.html`
- Hadley Wickham's tidyverse style guide on R: `https://style.tidyverse.org`
- Hadley Wickham's Advanced R: `https://bookdown.org/home/tags/advanced-r/`
- Write cool homework or project report with `bookdown` of Yihui Xie: `https://bookdown.org/yihui/bookdown/`
- Happy git with R of Jenny Bryan: `http://happygitwithr.com/`
- Build your own website with `blogdown`: `https://bookdown.org/yihui/blogdown/`

Grading: Students will grade each other's assignment.
- Within one day after the deadline of an assignment, you will find your randomly assigned classmate on HuskyCT.

- Email your solution to your assigned classmate, exactly as you submitted it online. You have 1 day to provide your feedback.

- You will submit your feedback on HuskyCT. For example, for Homework 1, there will be a follow-up assignment called "Homework 1 - feedback". There, you will enter your numeric assessment using the following scale:
  1 = missing, or totally inadequate
  2 = requires a major revision
  3 = requires a minor revision
  4 = complete and accurate solution
  A template for the feedback is provided on HuskyCT.

- If revisions are necessary, provide the necessary details in the designated box.

Note that your assessment will NOT affect your partner's grade! The purpose of this process is to learn from each other and develop good programming and communication skills, so please be thorough and honest, but also polite.