Statistical Computing with R Laboratory

CS109L Lecture 1

Kevin Shin

March 27, 2015

Outline

Motivation: Why Learn R?

Features:

- Free and open source programming language for statistical computing and graphics
- Massive set of open source packages for statistical modelling, machine learning, visualization, etc
- Cutting edge tools
- Language syntax has high support for data analysis
- Widely used in the statistics and machine learning community.
- Many functional programming features

What will I learn from CS109L?

R application & implementation!

- Strong, concise, and efficient programming practices for data analysis and program implementation.

What will I learn from CS109L?

R application & implementation!

- Strong, concise, and efficient programming practices for data analysis and program implementation.
- How to understand and utilize any new R package.

What will I learn from CS109L?

R application & implementation!

- Strong, concise, and efficient programming practices for data analysis and program implementation.
- How to understand and utilize any new R package.
- A foundational understanding of functional programming that can be applied in future courses (CS240H, CS242).

What will I learn from CS109L?

R application & implementation!

- Strong, concise, and efficient programming practices for data analysis and program implementation.
- How to understand and utilize any new R package.
- A foundational understanding of functional programming that can be applied in future courses (CS240H, CS242).
- A conceptual understanding of data analysis and visualization to be applied for future independent projects.

What will I learn from CS109L?

R application & implementation!

- Strong, concise, and efficient programming practices for data analysis and program implementation.
- How to understand and utilize any new R package.
- A foundational understanding of functional programming that can be applied in future courses (CS240H, CS242).
- A conceptual understanding of data analysis and visualization to be applied for future independent projects.
- "Pain to gain ratio"

Logistics: Course Content

Course Content:

- R Data Structures
- Functional Programming
- R Graphics & Visualizations
- R Workspace Development
- Probabilistic Implementations
- R Machine Learning Applications

Logistics: Lecture Schedule & Office Hours

Weeks 1 - 2:

- Lectures: Tuesdays/Thursdays 2:15 PM - 3:30 PM @ Hewlett 201

Weeks 3 - 9 (No lecture week 10):

- Lectures: Tuesdays 2:15 PM - 3:30 PM @ Hewlett 201

Logistics: Prerequisites & Corequisites

CS109

- Pre/Co-requisite
- A CS109 (or equivalent) background will give a better appreciation from the course. That being said, anybody should be able to benefit from the material that we will cover in CS109L, especially towards the end of the quarter.
- Recommended.

Logistics: Prerequisites & Corequisites

CS109

- Pre/Co-requisite
- A CS109 (or equivalent) background will give a better appreciation from the course. That being said, anybody should be able to benefit from the material that we will cover in CS109L, especially towards the end of the quarter.
- Recommended.

CS106B

- Prerequisite
- A CS106B (or equivalent) background is required for understanding the course and completing assignments as both require prior programming experience.
- Highly recommended

Logistics: Assignments

Assignments:

- Graded on a “1” or “0” rubric

Logistics: Assignments

Assignments:

- Graded on a "1" or "0" rubric
- Two deadlines per assignment for flexibility:
  - "turn in" deadline (optional): If you turn in an assignment, you will receive a grade and have the option to re-submit a final version of the assignment.
  - "redo" deadline (final): Final resubmission deadline after the "turn in" deadline. Assignments will not be accepted past this deadline.

Please refer to `cs109l.stanford.edu` for more detailed information on assignment grading and specific due dates.

Logistics: Course Grading

Course Grading: There are a total of 3 assignments throughout the quarter. To receive credit in the course you accomplish the following:

- Satisfactorily complete *Assignment 0: R Training Bootcamp* by its "redo" deadline.
- Satisfactorily complete at least one of the following by their "redo" deadlines.
  - *Assignment 1a: Liar's Dice*
  - *Assignment 1b: Shiny Development*

Please refer to `cs109l.stanford.edu` for more detailed information on course grading.

Logistics: Final Note

Install R!

- The instructions for installing R are in a handout located on the website.
- Should take < 10 minutes, so please install R by the end of this week!
- Feel free to run through the example code in R to get a better sense of what's going on after lectures.