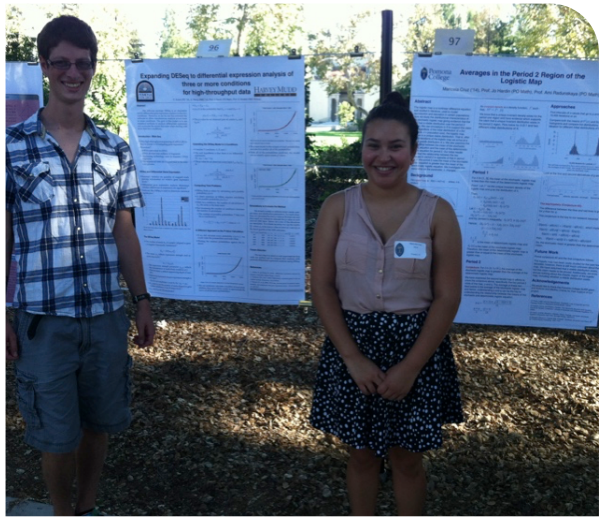
Expectations and Skills for Undergraduate Students Doing Research in Statistics and Data Science

Jo Hardin
Pomona College

jo.hardin@pomona.edu
@jo_hardin47
Github: hardin47

Curriculum Guidelines

- Statistics
- Data science
- GAISE



My background

Journals???

Skills

- Make an argument
- Engage with theory
- Work independently
- Data wrangle

To-do

- I still need to add in examples of my projects as examples for each of the topics I describe below.

Make an Argument

- Theoretical
- Simulation
- Literature

In the Classroom

- Hypothesis test using mean and median to demonstrate two different approaches to the same scientific hypothesis
- Simulation studies (e.g., bootstrap confidence intervals)
- How do we know that? How can we argue that result is better than the other?

Engage with Theory

- Connect theoretical ideas to core principles in statistics

In the Classroom

- Moment generating function: **why** do they uniquely determine a distribution?
- Simulate theoretical results for visualization of the process.

Work Independently

- What did you do?
- Why did you do it?
- What is the next step?
- What do I still not understand?

In the Classroom

- Independent projects (peer assessment!)
- Reflect on assignments (quickly or in detail)

Wrangle Data

- Data Science
- Statistics
- Theoretical

In the Classroom

- Practice, practice, practice
- Learn how to problem solve independently.
- Data wrangling should happen in every class at every level.

What else?

- Becoming a part of the larger community (e.g., quo function in dplyr v 0.7.1, June 22, 2017)
- Using Git & GitHub (http://happygitwithr.com/)
- Bring your love of research to the classroom to generate excitement.

Thank you!

Jo Hardin
Pomona College

jo.hardin@pomona.edu
@jo_hardin47
Github: hardin47