Week 1 Day 5

1 message

Jonah Sinick <jsinick@gmail.com>   Fri, Feb 19, 2016 at 9:50 AM
To: Ali Bagherpour <ali.bagherp@gmail.com>, Andrew Ho <Kironide@gmail.com>, Chad Groft
<clgroft@gmail.com>, David Bolin <david@bolin.at>, Jacob Pekarek <jpekarek@trinity.edu>, Jaiwithani
<jaiwithani@gmail.com>, James Cook <cookjw@gmail.com>, Linchuan Zhang <email.linch@gmail.com>, Matthew
Gentzel <magw6270@terpmail.umd.edu>, Olivia Schaefer <taygetea@gmail.com>, Sam Eisenstat
<sam.eisenst@gmail.com>, Tom Guo <tomguo4@gmail.com>, Trevor Murphy <trevor.m.murphy@gmail.com>

Today we'll be taking a break from many-part problem pair programming in favor of focusing on fundamentals.
Choose a partner, and pair program or not as you'd prefer.

The book that we'll be looking at *Advanced R* by Hadley Wickham, which I've uploaded to the Google Drive
folder under "books". For now, the most important chapters are

> *1 Introduction*
> *2 Data structures*
> *3 Subsetting*
> *5 Style guide*
> *6 Functions*
> *10 Functional programming*
> *11 Functionals*

Conveniently, the chapters have a lot of exercises at the end. I suggest doing these.

Some of you are looking to learn more about regularization and cross validation. For this, I recommend:

- **Cross Validation:** Chapter 4 of *Applied Predictive Modeling* and Chapter 5 of *Introduction to Statistical
  Learning* (both also uploaded to Google Drive).

- **Regularized Linear Regression:** Section 6.4 of *Applied Predictive Modeling*, Section 6.2 of *Introduction
  to Statistical Learning* and Lecture 7 from Andrew Ng's Coursera course on machine learning.

Try multiple sources and see which one you find it easiest to learn from.