



Andrew Ho &lt;kironide@gmail.com&gt;

---

**Week 2 Day 3 (Morning): self-assessment**

7 messages

**Jonah Sinick** <jsinick@gmail.com>

Wed, Feb 24, 2016 at 9:39 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>

This morning you'll be working on a self-assessment designed to help you determine the extent to which you've been acquiring the skills that you'll need. The assessment will also give you more practice with .Rmd (R Markdown), which you may not have used before. You can open this sort of file in RStudio and turn it into html via "Knit HTML." You may find it convenient to run individual lines of code in the .Rmd file rather than compiling the whole thing every time.

You should work individually to get an unambiguous signal. Feel free to use the web as needed, but please don't copy and paste code verbatim from solutions to past assignments – the intent is for you to assess how much of the material you're retaining.

Two hours are allocated for this. When you're done (or as done as you expect to be), send your solutions to me at [jsinick@gmail.com](mailto:jsinick@gmail.com) for a code review.

If you finish early, you can do one or more of:

- Reading the references on log loss and ROC curves sent on Slack after the presentation.
- Get a head start on what we'll be doing today (dimensionality reduction) by reviewing section 10.2 of ISLR and the lab 10.4.
- Continue working on what we were doing yesterday (examples of fitting logistic regression models) as in the second email in the "Week 2 Day 2 (morning)" thread.

Note that the cross validation function in question #8 can be easily modified to work for logistic regression, by an optional argument of family = "binomial" to cv.glmnet, and replacing lm() with glm() with family = "binomial."

---

 **selfAssessment1.Rmd**  
3K

---

**Andrew** <kironide@gmail.com>  
Reply-To: Kironide@gmail.com  
To: Jonah Sinick <jsinick@gmail.com>

Wed, Feb 24, 2016 at 10:58 AM

[Quoted text hidden]

---

 **selfAssessment1.Rmd**  
6K

---

**Andrew** <kironide@gmail.com>

Wed, Feb 24, 2016 at 11:14 AM

Reply-To: Kironide@gmail.com  
To: Jonah Sinick <jsinick@gmail.com>

[Quoted text hidden]



**selfAssessment1.Rmd**  
6K

**Jonah Sinick** <[jsinick@gmail.com](mailto:jsinick@gmail.com)>

Wed, Feb 24, 2016 at 11:52 AM

To: Ali Bagherpour <ali.bagherp@gmail.com>, Andrew Ho <Kironide@gmail.com>, Chad Groat <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>

A lot of people are still midway through the assessment so we'll have it run until 1 PM. When you submit, note your approximate start time.

I'll send out solutions at that time, but you can hold off on looking at them until later if you'd like.

After finishing, the highest priority thing to do is the lab in section 10.4 of ISLR. We'll break for lunch at 1 PM and reconvene at 2:30 PM to talk about PCA and dimensionality reduction.

**Andrew** <kironide@gmail.com>

Wed, Feb 24, 2016 at 12:34 PM

Reply-To: Kironide@gmail.com  
To: Jonah Sinick <jsinick@gmail.com>

oops, here's the actual HTML-ized version of my responses

[Quoted text hidden]



**self-assessment-answers.html**  
679K

**Andrew** <kironide@gmail.com>

Wed, Feb 24, 2016 at 1:02 PM

Reply-To: Kironide@gmail.com  
To: Jonah Sinick <jsinick@gmail.com>

and my start time = 10:15 approx

[Quoted text hidden]

**Jonah Sinick** <[jsinick@gmail.com](mailto:jsinick@gmail.com)>

Wed, Feb 24, 2016 at 1:09 PM

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>

I've attached solutions to the self-assessment.

Feel free to work more on yours before looking at the solutions if you'd like, but do send me what you have now.

No worries if you found it difficult – if there's interest, we can have a re-test after you've had the chance to study some more.

On Wed, Feb 24, 2016 at 9:39 AM, Jonah Sinick <[jsinick@gmail.com](mailto:jsinick@gmail.com)> wrote:

**2 attachments**

