

Releasing my course materials [DS4B]

[cross-post from [LinkedIn](#)]

I teach data science literacy to different groups, and I've been struggling with a personal dilemma: I believe strongly in open access (data, code, materials), but each hour of lecture materials takes about 15-20 hours of my time to prepare: should I make everything openly available, or try to respect that preparation time?

Which is basically the dilemma that anyone open-sourcing goes through. And the answer is the same: release the materials. Made stronger by the realisation that I teach how to do field data science (data science in places without internet connectivity) and people in places without internet connectivity are unlikely to be dropping into New York for in-person sessions.

Starting this week, the materials are going online in github; I'll be backing this up by restarting the I Can Haz Datascience posts (posts on development data science, with Emily the cat) to cover the topics in them (designing and scoping a DS project, python basics, acquiring data, handling text, geospatial, relationship data etc). Hopefully they'll be of some use to people.