Hey Scott,

I realize we discussed a lot of this stuff on our phone call, but I'll put some of the answers here anyway.

1. No definite idea, but leaning towards something related to machine learning or NLP

2. I have taken (at CUNY):

Knowledge and Visual Analytics: R (ggplot2, Pandas), Python (Seaborn, ggplot), Javascript (D3.js, googlevis)

Coursework centered around visualizing various data sets in different platforms. Here's my final project: http://cuny-is608-vm.cloudapp.net/home.html/SUSAN%20G%20KOMEN%20BREAST%20CANCER%20FOUNDATION

The only major change that Josh made to it was he created that 'smart' text box that autofills. I took the raw filing data from the irs website combined it into a csv, and did a join with information about the companies (such as name as opposed to just tax id). I cleaned that data using Pandas in Python, and then spent a long time figuring out D3 (as I had pretty minimal Javascript experience prior to this). I chose d3, because I knew it would be good to learn--but it wound up taking wayyy longer than I thought it would

Data Acquisition and Management:

MySQL, Neo4j, MongoDB, Hadoop

Cleaned data sets and worked with them doing basic joins and selects, etc.

Computational Mathematics: All the assignments were written in R/LaTeX and centered around relevant concepts in Probability and Statistics, Calculus, and Linear Algebra (all applied)

Advanced programming techniques: This centered around Python and Pandas, and was not a terribly difficult class, but was balanced by the challenges in the other classes (especially Knowledge and Visual).

3. Thesis should be started this summer, and completed by December.

4.Rough estimates:

4:

Python, R, LaTeX

3:

MySQL, PostgreSQL, SQLite, MongoDB, Neo4j, Hadoop, Azure ML, Octave, Matlab, Javascript

2:

C++, Java

5. I have not written scientific research papers, but I was an English major at University of Vermont and was both an editor and writer for the literary magazine, as well as a senior writer for the school's paper. i.e. My writing skills may be better than the average data scientist. I did write some technical papers for some of my Psychology classes (almost a double major, but couldn't fit in the last class I would have needed to declare it as a major).

6. The classes this semester are 'Quantitative Finance', 'Business Analytics', and 'Probability and Statistics'. My calendar is pretty open, as I discussed, and I'm sure our Thursday meetup will be no problem for now.

I am interested in a lot of subjects; I served as an English teacher as a Peace Corps Volunteer for more than five years, working in Lesotho (living in a grass roofed hut) and St. Vincent and the Caribbean (living on the beach), so education, NGO, NPO, etc. is something I'm also very interested in.

My CUNY work is (almost all) here: https://github.com/RodavLasIlad/CUNY

A random forests I wrote for a kaggle competition (well before I started the program, so it's not in great shape) is here: https://github.com/RodavLasIlad/kaggles/blob/master/dataexploration5.R

I'm also attaching two pdfs made with R markdown and a project I made in R's Shiny (interactive visualization) that were all made for Knowledge and Visual Analytics.

Let me know if you have any questions!