Well, I’m a little late on this, but better late than never, I suppose.

Here are three ideas I have for my Capstone. I’m leaning toward the first.

1. One quickly growing problem in academia is that many students are graduating with very significant debt. If we think of debt as the dependent variable, we can use these [“College Scorecard” data](https://catalog.data.gov/dataset/college-scorecard) to see if there is a correlation between where (in particular, which U.S. states) students go to school and whether students graduate with significant debt. This dataset also includes facts about ethnicities and standardized test scores, which may also be a significant factor in the amount of debt a typical student incurs.

2. Use [data on automobiles](http://archive.ics.uci.edu/ml/datasets/Auto+MPG) to look for predictive patterns in which brands or engine types tend toward higher MPG ratings. As we all are concerning ourselves more with green energy, the challenge of finding a fuel-efficient car becomes more and more pressing.

3. It is undeniable that the U.S. has some underachieving school districts. Many of these districts received federal grants in 2010. When districts applied for help, they were asked to indicate what sort of overhaul they were planning (“turnaround model” vs. “transformation model” etc.). This project would explore patterns in [these data](https://catalog.data.gov/dataset/school-improvement-2010-grants) to see if we can predict what sort of model a typical school district would opt for, given factors like the amount of money received or the district’s location.