The goal of the project is for you to use the mathematics of this course in exploring something you're interested in. This is vague and that's deliberate: probability and statistics can show up in many different ways and I want to allow you to pursue your interests. The end result will depend on the exact nature of your project. I expect a minimum of two typed pages explaining your project, the mathematics you used, and your results/conclusions. You will likely want to use some kind of software: I recommend R for statistical analysis and LaTeX for typesetting the document.

Some more details about the final written product:

- Explain clearly the idea behind the project (i.e. what is the idea and why is it worth investigating?);
- Explain clearly where the data you're using comes from (if there is any) with appropriate citations;
- Explain clearly the relevant math and how you used it;
- Address potential problems (e.g. if you made any assumptions that might not be supported by data or if your samples weren't truly random);
- State your conclusions/findings clearly;
- Are there more questions related to your topic that might also be worth answering?

Before all of that, though, you must have a project proposal approved. Your project proposal should be a short (one page) description of your idea and the data/math you'll use to investigate it. The proposal is due (by email) on Friday, April 1. A detailed outline is due on Friday, April 15. The final project is due on Friday, May 6.

The following stats-related sites/blogs may help you generate ideas. All involve math covered in the course and the kind of data you could collect or find.

- Carpe Diem (with a very long url)
- `http://freakonometrics.hypotheses.org/48184`
- `http://www.statisticsblog.com/2014/12/can-pregnant-women-intuit-the-sex-of-their-children/`
- `http://www.wine-economics.org/workingpapers/AAWE_WP16.pdf`
- `http://bayesianbiologist.com/2014/01/22/whats-warren-buffetts-1-billion-basketball-bet-worth/`
- `http://blog.okcupid.com/index.php/page/2/`

Here are a few sources of data:

- `https://books.google.com/ngrams`
- `http://rs.io/100-interesting-data-sets-for-statistics/`
- `http://www.census.gov/`
- `http://www.bjs.gov/`

Finally, some interesting ways to present data:

- `http://benschmidt.org/jobsBroad/`
- `http://freakonometrics.hypotheses.org/14682`