Math 422 **Final Project** **March 25, 2018**

The goal of the final project is for you to use statistics to investigate something interesting. This is deliberately vague: probability and statistics can show up in many different ways and in many different places and I want you to pursue your interests. The end result will depend on the exact nature of your project. I expect a minimum of three 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 description of your idea and the data you'll use to investigate it. This may be a plan for collecting the data or a description of an existing data set you have access to. The proposal is due (by email) on Friday, April 6. A detailed outline/preliminary analysis is due on Wednesday, April 25. The final project is due on Thursday, May 10.

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.
- Where people live.
- Can women intuit the sex of their babies before they're born?
- Warren Buffet's big basketball bet.

Here are a few sources of data:

- `http://www.kaggle.com/datasets`
- `http://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`