

## Session 4: Got Data?

**Please submit your completed work before 7 AM ET on Thursday, August 5.**

This homework will have you 1) downloading data, in part, to determine if you like creating datasets. If so, you may want to integrate it into your research. If not, you need to think of research with existing datasets. Note, you will need to download codes from Session 3 to complete some of the problems.

1. Read *Bracketology: How can math help?* at <https://ww2.amstat.org/mam/2010/essays/ChartierBracketology.pdf> and in 2-3 paragraphs describe what you found most interesting about the paper. Further, do you enjoy the application part of the paper or the theoretical development or both?
2. Scrape a dataset using python and/or Google Sheets. You may also wish to combine datasets, creating enriched data. In your Word document, explain your dataset to someone who might want to use it. What's there? Further, indicate if you used Google Sheets or python. Finally, include the dataset with your uploaded files to Schoology. (If you download a significant amount of data, just upload a small sample.)
3. The draft of your research proposal is due. In it, you will propose a topic for your research, including what application they intend to rank, which ranking methods you will use or adapt, and what data exists or will be created. Be prepared to give a 2-3 minute description of your topic to the group during our group session on August 2. You may do this entirely verbally or accompanied with slides. You will receive feedback on your draft so it can be further developed for your first individual session. Your final research proposal should be 2-3 pages long. This draft should be at least 1 page in length.

---

**Submitting your work** - Your responses to the first two parts of this homework should be in one document saved as a PDF. The draft of your proposal should be another document, submitted as a PDF. Upload the PDF files and any supplementary materials to Schoology.