

# Class Project: Analysis of PLOS Analyses

*Bonnie Smith*

*September 11, 2017*

**Question and background:** In this project, we perform an analysis of published PLOS papers, identifying the most common techniques and whether they vary by field, as well as trends over time.

The fields represented in PLOS journals are Biology (since October 2003), Computational Biology (since June 2005), Neglected Tropical Disease (since October 2007), Medicine (since October 2004), Genetics (since July 2005), and Pathogens (since September 2005).

**\*\*Data:\*\*** We will begin by forming a list of common statistical methods used by authors publishing in these journals (describe details of how to go about formulating this list, here). We will then use packages (that I will be learning about soon!) to scrape the PLOS website. For each paper, we will record the journal/field, the month/year of the issue it appeared in, and an indicator for each statistical technique from our list as to whether it was used in this paper.

**Analysis:** Once we have the cleaned data, for each statistical technique on our list we will model the proportion of PLOS papers to use that technique, adjusting for time and field.