BIOMEDSCI 552: Problem Set 1

1) Create a GitHub account, and provide your username below

2) Go to https://github.com/epimodels/StatsThinking_2025, fork the epimodels/StatsThinking_2025 repository, add your name to Roster.csv, and then create a pull request to incorporate that change into the main branch.

3) In lecture 1, undergraduate psychology students were used as an example of a "sample of convenience". Describe a population for which undergraduate psychology students might be a sample. How does that sample vary from the population (or does it?) in a way that creates bias.

4) Describe a research problem you are interested in. What is the population? Describe a reasonable sample from that population. Do you have a target population? If so, what is it?

5) A researcher comes to you, wanting to do a study in dogs in Banfield pet clinics. They're planning on using a consolidated electronic health record (EHR) from close to 1,000 clinics, looking at the risk factors for the development of parathyroid conditions in elderly dogs, by using clinical diagnoses and the dogs clinical records.

Is this data prospective or retrospective, and why?

Is this data passive or actively collected, and why?

Is this data "Big Data", and why?