**Potential Dataset 1**

Name: Patient-level dataset to study the effect of COVID-19 in people with Multiple Sclerosis

Link: <https://physionet.org/content/patient-level-data-covid-ms/1.0.1/>

Description: This dataset was collected to study the effects of COVID-19 in people with Multiple Sclerosis due to their heightened susceptibility to infections and concerns about how the disease interacts with their central nervous system. The data was collected via direct entry into a created platform, done by clinicians, the person with MS being studied, or a representative either in a hospital setting or at home. Although not specified, collection presumably began toward the start of the pandemic and was terminated on February 3rd, 2022. It contains information about smoking habits, COVID-19 symptoms, treatment, and whether patients recovered from 1141 people with MS.

Preliminary Investigation: This dataset holds great potential for investigating relationships between factors such as whether or not recovery was reached among smoking and non-smoking MS patients. I would assume non-smokers have a greater probability of achieving a full recovery due to healthier lungs. One relationship I’m very interested in examining is whether or not patients made a full recovery depending on the type of MS they have (progressive or relapsing). Considering relapsing is typically observed in younger ages, a factor that plays a role in recovery, I’d predict they have a higher probability of achieving a full recovery.

**Potential Dataset 2**

Name: Effect of obesity on arithmetic processing in preteens with high and low math skills. An event-related potentials study

Link: <https://openneuro.org/datasets/ds004019/versions/1.0.0>

Description: Researchers from the University of Arkansas for Medical Sciences (UAMS) generated this dataset to examine the effects obesity may have on cognitive processing in preteens. Participants were 62 obese and non-obese (according to BMI scale) preteens with varying levels of math skills. Basic arithmetic tasks were displayed using E-prime software version 2 and the “Geodesic Net Amps 300 system running Netstation 4.5.2 software using the 128-channel Geodesic Hydrocell Sensor Net™” to obtain EEG signals at the Arkansas Children Nutrition Center (ACNC). The period in which the data was collected is not obviously stated, but the research article was published in 2022.

Preliminary Investigation: The clear relationship to investigate with this dataset is how obesity (based on the BMI scale) affects ability to perform well with arithmetic tasks (researchers used a binary system where 1 indicates a question was answered correctly and 0 indicates it was answered incorrectly, which I think will make generating probabilities straightforward). Based on what I learned from taking NRSCI-389, I predict obese preteens will have a lower probability of scoring correctly on arithmetic tasks because their dopaminergic systems can often be overstimulated, which makes focusing more challenging.