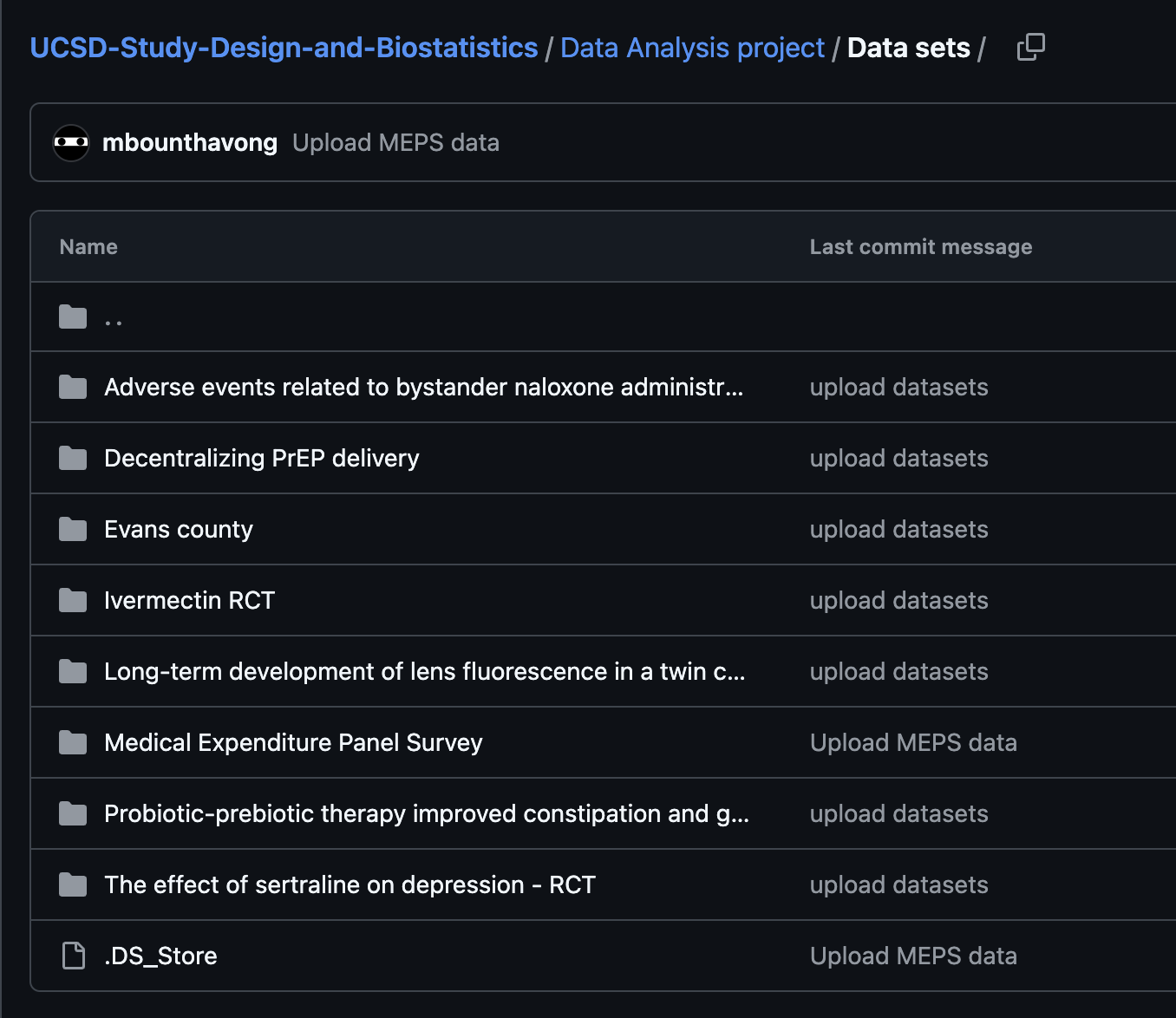
**SPPS 208 – Study Design & Biostatistics II**

**Data Analysis Project**

**Research questions for each dataset**

Location of the data are in the course [GitHub site](https://github.com/mbounthavong/UCSD-Study-Design-and-Biostatistics). Each data set has a README file or its equivalent. Please read through these as they contain important information about the data. Once your team has selected the data set for the Data Analysis Project, please send an email with the name of the data set to [mbounthavong@health.ucsd.edu](mailto:mbounthavong@health.ucsd.edu). Afterwards, each team will set up a one-hour meeting with Mark Bounthavong to review their report, analysis, presentation, or any other issues between 03 February 2025 and 06 March 2025.

The final report due date is **Friday, 07 March 2025**. Presentations will be on **Friday, 07 March 2025**.

**Adverse events related to bystander naloxone administration in cases of suspected opioid overdose in British Columbia**

This was an observational study to examine whether an association exists between the number of ampoules of naloxone administered and the odds that the recipient will experience withdrawal symptoms. Subjects in British Columbia received Take-Home Naloxone kits. Data collection through the administration form was anonymous, and information from paper forms was entered manually into a database at the BC Centre for Disease Control (CDC). This study utilized the Overdose Response Information form data from program inception on 31 August 2012 to 31 December 2018. The primary exposure variable was the number of naloxone ampules administered. The primary dependent variable was withdrawal symptoms reported (“None or mild” and “Moderate or severe”).

For this data analysis project your team will need to answer the following research question:

*Was there an association between naloxone ampule use and patients who experienced opioid withdrawal symptoms? Describe this difference.*

**Decentralizing PrEP delivery data**

This was a dissemination and implementation study to investigate the feasibility of delivering pre-exposure prophylaxis (PrEP) to men who have sex with men (MSM) through two means: (1) primary care physicians and (2) sexual health clinic nurses in Toronto, Canada. In the patient-initiated continuing medical education (PICME) strategy, participants saw their family doctors and gave them the card, which also contained a link to a Continuing Medical Education module. In the nurse-led strategy, participants visited one of two participating clinics to obtain PrEP. Patients were provided a baseline questionnaire and a 6-month follow-up questionnaire.

For this data analysis project your team will need to answer the following research question:

*Was there a difference in the number of patients who completed PrEP therapy among those who intended to start via their nurse, family doctor, or other (hint: use the* PrEPIntent *variable)? Describe this difference.*

**Ivermectin RCT dataset**

This was a randomized, double-blinded, placebo-controlled trial comparing the efficacy of ivermectin versus placebo in the treatment of mild-moderate COVID-19 in India. All adult patients (18 years old and older) were enrolled if they had a PCR-confirmed diagnosis of COVID-19 or a positive rapid antigen test and were admitted for mild-moderate COVID-19. Patients were excluded if they had a known allergy or adverse drug reaction with ivermectin, unwilling or unable to provide consent to participate in the study, had a prior history of ivermectin during the course of current illness, and were pregnant and lactating. The primary outcome was having a negative PCR for COVID-19 at day 6.

For this data analysis project your team will need to answer the following question:

*Was there an association between patients who were randomized to ivermectin or placebo with mortality (“*final outcome*” variable)? Describe this difference.*

**Long-term development of lens fluorescence in a twin cohort dataset**

This is a prospective cohort study that followed twins over a 21-year period to evaluate changes to their lens fluorescence. The study cohort was from Denmark as part of the Copenhagen Twin Cohort Eye Study. Previous studies have reported that lens fluorescence is elevated among patients with diabetes and long-term complications related to diabetes. The authors sought to understand the correlations between glycemic control, smoking, and age on lens fluorescence.

For this data analysis project your team will need to answer the following question:

*Was there an association between smokers and non-smokers at baseline with lens fluorescence change over a 21-year period? Describe this relationship.*

**Probiotic-prebiotic therapy improved constipation and gut motility in Parkinson’s disease dataset**

This study was a randomized, placebo-controlled, trial evaluating the impact of probiotics on constipation and gut motility among patients with Parkinson’s disease compared to placebo 8 weeks after initiation. The study was conducted among patients with Parkinson’s disease in Malaysia between October 2018 and February 2019. The main end point was improved constipation and gut motility.

For this data analysis project your team will need to answer the following question:

*Was there a difference in gut transit time change from baseline to 8 weeks between patients who received probiotics and did not receive probiotics? Describe this relationship.*

**Effect of sertraline on depression dataset**

This study was a randomized, placebo-controlled trial to evaluate the impact of sertraline on depression compared to placebo among patients with active infection of meningitis in Uganda. Patients were evaluated for depression using the Center for Epidemiologic Studies Depression Scale (CES-D) scale at one and three months.

For this data analysis project your team will need to answer the following question:

*Was there an association between the proportion of patients with severe depression at 12 weeks according to the CES-D and treatment assignment (sertraline versus placebo)? Describe the relationship. (hint: Your team needs to define “severe” depression.)*

**Medical Panel Expenditure Survey (MEPS) dataset**

The United States MEPS dataset is a based on a cross-sectional design that contains information about a household’s annual healthcare expenditure. The Agency for Healthcare Research and Quality are the data stewards of the MEPS data. Total healthcare expenditure includes all costs associated with a household’s health-related consumption.

For this data analysis project your team will need to answer the following question:

*Was there an association between poverty level (*povcat*) and healthcare expenditure (*totexp*) among adult household respondents in the United States in 2022? Describe the relationship.*