**1. Childhood Obesity Prevention Study**

**Data Introduction**

You have data from 300 children enrolled in a community-based obesity prevention program. Each child is weighed at baseline (age 8) and then annually for 4 years (ages 9, 10, 11, 12). The main outcome is **BMI percentile**. Key measured factors include:

* **Physical activity** (minutes per day, self-reported)
* **Dietary habits** (measured using a food frequency questionnaire, e.g., daily sugar-sweetened beverage intake)
* **Socioeconomic status** (SES), categorized as low, medium, or high
* **Gender** (male/female)

**Hypothesis of Interest**

Whether the **BMI percentile** trajectory over time differs based on participation in the program and how these differences are moderated by **SES** or **gender**.

**2. Blood Pressure in a Multi-Center Clinical Trial**

**Data Introduction**

A pharmaceutical company sponsors a new **antihypertensive medication** trial. The study recruits 500 patients from 10 different clinics. Each patient’s **systolic blood pressure (SBP)** is measured at months 0, 1, 3, 6, and 12. Each patient is randomized to **Drug A** vs. **Placebo**.

Additional information:

* **Age** (continuous)
* **Clinic ID** (to capture site-level differences)
* **Medication adherence** (percentage of prescribed doses taken)
* **Baseline comorbidities** (e.g., diabetes: yes/no)

**Hypothesis of Interest**

Patients on **Drug A** will have a **steeper reduction** in systolic blood pressure over 12 months compared to placebo, controlling for clinic-level differences.

**3. Environmental Exposure and Respiratory Symptoms**

**Data Introduction**

You have daily measurements of **air pollution (PM2.5)** from 20 different neighborhoods. In each neighborhood, you have **daily counts of emergency room visits** for **asthma-related** complaints over a one-year period (365 days).

Additional contextual variables include:

* **Neighborhood average SES** (continuous, standardized)
* **Season** (categorical, e.g., winter, spring, summer, fall)
* **Day of week** (to account for systematic variation)

**Hypothesis of Interest**

Higher **PM2.5 levels** are associated with **increased asthma-related emergency visits**, possibly differing by **season**.

**4. Mental Health Assessment in a Telehealth Intervention**

**Data Introduction**

A mental health support program enrolls 200 individuals who exhibit symptoms of moderate depression. They receive **weekly telehealth counseling** sessions for 10 weeks. Each week, participants complete a **depression symptom score**(e.g., PHQ-9). They also report:

* **Stress level** on a 1–10 scale
* **Medication use** (yes/no)
* **Number of hours of sleep** in the past 24 hours

**Hypothesis of Interest**

Whether **depression symptoms** improve more quickly among those with **consistent medication use** and/or higher **sleep**quality, and whether stress level modifies the rate of change in symptoms.

**5. Vaccination Coverage in School Districts Over Multiple Years**

**Data Introduction**

You have an annual survey of **vaccination rates** for measles in 50 school districts over 5 years. For each district-year, you record:

* **Measles vaccination coverage** (as a percentage of eligible children)
* **Population density** (continuous)
* **Median household income** (continuous)
* **Geographic region** (e.g., rural vs. urban)

Additionally, a statewide policy to **increase vaccination awareness** was implemented at the start of year 3.

**Hypothesis of Interest**

Vaccination coverage will **increase** following the statewide policy, and the **effectiveness** of the policy depends on **population density** and **income** levels.

**6. HIV Viral Load Monitoring in an Infectious Disease Clinic**

**Data Introduction**

In an observational study, 100 HIV-positive patients are monitored for 2 years, with **viral load** measurements taken every 6 months (baseline, 6m, 12m, 18m, 24m). Data collected includes:

* **Medication regimen** (two categories: standard vs. new regimen)
* **Patient age** at baseline
* **CD4 count** at baseline
* **Adherence** (continuous measure of proportion of doses taken)
* **Time since diagnosis** (years)

**Hypothesis of Interest**

Patients on the **new regimen** have a **faster decline** in viral load compared to those on the standard regimen, and the effect is possibly influenced by **adherence** levels.