Behavioral Science Task

The following task is to be completed and submitted by the stated deadline. Please complete this on your own to the best of your ability. We estimate that this task should take between 3 and 6 hours to complete.

PART 1: RESEARCH REVIEW

SCENARIO: For this task, imagine that you are working with a large employer in the United States to help their employees increase their retirement contributions. The company offers a retirement program that requires employees to opt-in to participate. When employees enroll, they are assigned a 3% contribution rate by default, which gets automatically deducted from their monthly paychecks. Additionally, the company also offers a match program up to 5%; currently, however, only 30% of employees are contributing 5% or more, thereby receiving the maximum employer match. The goal of your project is to increase the number of employees receiving the maximum match by contributing at least 5% of their paychecks to retirement.

In your literature review, you find a working paper[1] that seems to be addressing a similar challenge. Your first part of this task is to read the working paper and prepare no more than 5 presentation slides presenting the paper to the team at the company you are working with for possible relevance to this project. Make sure to include the study design, main takeaways, and any limitations you identify.

PART 2: DATA ANALYSIS

Since your team found the working paper to be interesting, you convinced the company to run a small experiment to test this idea. In the experiment, a sample of employees were randomly assigned to receive either a generic informational email ("control") or an email that leverages peer information, highlighting how much others in the company are losing by not contributing enough for retirement to receive the full match ("recommendation"). The hypothesis of your team of experimenters is that treated participants will be more likely to increase their retirement contributions than control participants.

In the email you received you will find attached 2 datasets from this experiment. One dataset represents company data about the participants' experimental conditions and income levels, and the other contains records of actual contribution changes following the experiment.

[1] Perry, E. (2019). Everyone else is making a mistake: Effects of peer error on saving decisions. Available at SSRN 3348672.

Your second task is to analyze this experiment by comparing a control group (informational email) with a treatment group (recommendation email) and write up what you find.

Dataset A contains:

- identifier: a unique identifier for each participant
- condition: control or recommendation (received the peer information email)
- income_level: income categorized as either "low-to-moderate (LMI)" or "non-LMI"
    - LMI is defined as a household income that is below 80% of the Area Median Income of a particular geography

Dataset B contains:

- identifier: a unique identifier
- increased_contribution: whether the individual increased their contribution to the retirement account in the month after receiving the recommendation (tracked through employer-provided records)

You will need to:

1. If necessary, merge, clean, and transform the data.
2. Produce descriptive statistics.
3. Test whether the main hypothesis holds.
4. Explore if other variables play a role in the results.

To complete this you can use any statistical program you wish. Please submit:

1. Your code in a pdf. Please annotate your work so someone can follow it.

2. 3-5 slides sharing your results with the company. These slides should be added to the same presentation as part 1.

    a. This should aim to communicate what you found in simple language with simple graphs/tables/figures. Based on your findings, you can add recommendations for the partner to support their employees' retirement contributions.
    b. You can also include other research questions you would like to explore or that sparked your curiosity as a result of this experiment.
    c. Please also pay attention to the design of your slides – you are presenting them to an external partner. They should convey enough information to be comprehendible without being overly wordy or crowded.

(Note, the data in this task is fictional. It does not reflect a real experiment.)