Dataset Information – Dataset A

You are provided a dataset with 30 columns. The first column respondent_id is a unique and random identifier. The remaining 29 features are described below.

For all binary variables: 0 = No; 1 = Yes.

1. flu_concern - Level of concern about the flu.
   - 0 = Not at all concerned; 1 = Not very concerned; 2 = Somewhat concerned; 3 = Very concerned.
2. flu_knowledge - Level of knowledge about flu.
   - 0 = No knowledge; 1 = A little knowledge; 2 = A lot of knowledge.
3. behavioral_antiviral_meds - Has taken antiviral medications. (binary)
4. behavioral_avoidance - Has avoided close contact with others with flu-like symptoms. (binary)
5. behavioral_face_mask - Has bought a face mask. (binary)
6. behavioral_wash_hands - Has frequently washed hands or used hand sanitizer. (binary)
7. behavioral_large_gatherings - Has reduced time at large gatherings. (binary)
8. behavioral_outside_home - Has reduced contact with people outside of own household. (binary)
9. behavioral_touch_face - Has avoided touching eyes, nose, or mouth. (binary)
10. doctor_recc_seasonal - Seasonal flu vaccine was recommended by doctor. (binary)
11. chronic_med_condition - Has any of the following chronic medical conditions: asthma or an other lung condition, diabetes, a heart condition, a kidney condition, sickle cell anemia or other anemia, a neurological or neuromuscular condition, a liver condition, or a weakened immune system caused by a chronic illness or by medicines taken for a chronic illness. (binary)
12. child_under_6_months - Has regular close contact with a child under the age of six months. (binary)
13. health_worker - Is a healthcare worker. (binary)
14. health_insurance - Has health insurance. (binary)
15. opinion_seas_vacc_effective - Respondent's opinion about seasonal flu vaccine effectiveness.
    - 1 = Not at all effective; 2 = Not very effective; 3 = Don't know; 4 = Somewhat effective; 5 = Very effective.
16. opinion_seas_risk - Respondent's opinion about risk of getting sick with seasonal flu without vaccine.

- 1 = Very Low; 2 = Somewhat low; 3 = Don't know; 4 = Somewhat high; 5 = Very high.
17. opinion_seas_sick_from_vacc - Respondent's worry of getting sick from taking seasonal flu vaccine.
    - 1 = Not at all worried; 2 = Not very worried; 3 = Don't know; 4 = Somewhat worried; 5 = Very worried.
18. age_group - Age group of respondent.
19. education - Self-reported education level.
20. race - Race of respondent.
21. sex - Sex of respondent.
22. income_poverty - Household annual income of respondent with respect to 2008 Census poverty thresholds.
23. marital_status - Marital status of respondent.
24. rent_or_own - Housing situation of respondent.
25. employment_status - Employment status of respondent.
26. census_msa - Respondent's residence within metropolitan statistical areas (MSA) as defined by the U.S. Census.
27. household_adults - Number of *other* adults in household, top-coded to 3.
28. household_children - Number of children in household, top-coded to 3.
29. employment_sector - Type of industry respondent is employed in.

Target variable: seasonal_vaccine