PREDICTING H1N1 VACCINE UPTAKE

The data for this project is sourced from the National 2009 H1N1 Flu Survey conducted in the United States following the Influenza outbreak of 2009. The datasets can be found on the Driven Data Website.

The dataset primarily consists of categorical variables with binary and numerical values. Additionally, certain columns contain coded data.

Business understanding

Vaccine uptake is a crucial aspect of public health, directly impacting disease prevention and population well-being. By analyzing the factors influencing vaccine acceptance, public health authorities can develop targeted strategies to improve vaccination rates and enhance overall community immunity. Understanding and predicting vaccine uptake enables proactive measures, such as tailored communication campaigns and accessible vaccination programs, to address barriers and increase vaccination coverage. This analysis aids in identifying groups at risk of low vaccine uptake and informs resource allocation for effective public health interventions, ultimately contributing to better health outcomes and disease control.
Objectives

1. Develop a Robust Predictive Model capable of estimating the likelihood of H1N1 vaccine uptake for individuals.

2. Feature importance: Identify Key Predictors of H1N1 Vaccine Acceptance

3. Generate insights and recommendations: Analyze the outcomes of the analysis and models to derive actionable insights and recommendations for enhancing vaccine uptake strategies.Model Success Criteria
For evaluating model performance, I'll use the following metrics:
Model success criteria

The Baseline model initially employed two simple linear regression models to showcase the correlation between price and two highly correlated variables: sqft_living and bathrooms.

Subsequently, we delved into crafting multiple linear regression models and iteratively refining them through feature selection.

Our process included feature selection and the development of a polynomial regression model, followed by validation using cross-validation techniques.
Data understanding

The data for this project originates from the National 2009 H1N1 Flu Survey conducted in the United States during the 2009 Influenza outbreak. You can access the datasets on the Driven Data website. Primarily, the dataset comprises categorical variables with binary and numerical values. Additionally, certain columns contain coded data.

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

h1n1_concern - Level of concern about the H1N1 flu.

0 = Not at all concerned; 1 = Not very concerned; 2 = Somewhat concerned; 3 = Very

concerned.
h1n1_knowledge - Level of knowledge about H1N1 flu.

0 = No knowledge; 1 = A little knowledge; 2 = A lot of knowledge.
behavioral_antiviral_meds - Has taken antiviral medications. (binary)

behavioral_avoidance - Has avoided close contact with others with flu-like
symptoms. (binary)

behavioral_face_mask - Has bought a face mask. (binary)

behavioral_wash_hands - Has frequently washed hands or used hand sanitizer.
(binary)

behavioral_large_gatherings - Has reduced time at large gatherings. (binary)

behavioral_outside_home - Has reduced contact with people outside of own household.
(binary)

behavioral_touch_face - Has avoided touching eyes, nose, or mouth. (binary)

doctor_recc_h1n1 - H1N1 flu vaccine was recommended by doctor. (binary)

doctor_recc_seasonal - Seasonal flu vaccine was recommended by doctor. (binary)

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)

child_under_6_months - Has regular close contact with a child under the age of six
months. (binary)

health_worker - Is a healthcare worker. (binary)

health_insurance - Has health insurance. (binary)

opinion_h1n1_vacc_effective - Respondent's opinion about H1N1 vaccine
effectiveness.

1 = Not at all effective; 2 = Not very effective; 3 = Don't know; 4 = Somewhat
effective; 5 = Very effective.
opinion_h1n1_risk - Respondent's opinion about risk of getting sick with H1N1 flu
without vaccine.

1 = Very Low; 2 = Somewhat low; 3 = Don't know; 4 = Somewhat high; 5 = Very high.
opinion_h1n1_sick_from_vacc - Respondent's worry of getting sick from taking H1N1
vaccine.

1 = Not at all worried; 2 = Not very worried; 3 = Don't know; 4 = Somewhat worried;
5 = Very worried.
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.
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.
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.
age_group - Age group of respondent.

education - Self-reported education level.

race - Race of respondent.

sex - Sex of respondent.

income_poverty - Household annual income of respondent with respect to 2008 Census poverty thresholds.

marital_status - Marital status of respondent.

rent_or_own - Housing situation of respondent.

employment_status - Employment status of respondent.

hhs_geo_region - Respondent's residence using a 10-region geographic classification defined by the U.S. Dept. of Health and Human Services. Values are represented as short random character strings.

census_msa - Respondent's residence within metropolitan statistical areas (MSA) as defined by the U.S. Census.

household_adults - Number of other adults in household, top-coded to 3.

household_children - Number of children in household, top-coded to 3.

employment_industry - Type of industry respondent is employed in. Values are represented as short random character strings.

employment_occupation - Type of occupation of respondent. Values are represented as short random character strings.
Recommendations

1. Promote Doctor Recommendations: Encourage healthcare providers to actively recommend the H1N1 vaccine to their patients.

2. Public health messaging should emphasize the health risks associated with contracting H1N1 flu without vaccination.

3. Address Perceived Risks: public health messaging should address and clarify any misconceptions or concerns regarding the perceived risks associated with H1N1 vaccinations.

4. Strengthened awareness and communication: Strengthened communication about the effectiveness of the H1N1 vaccine to enhance public confidence and trust in its efficacy would positively influence uptake rates.

5. Target Health Workers: Promote vaccination among healthcare workers and emphasize the importance of their role as advocates for vaccination within their communities.

6. Behavioural interventions: Promoting preventive behaviors such as frequent hand hygiene practices and avoiding large gatherings in high-risk situations.

7. Implementation of targeted interventions to address underlying concerns related to vaccine safety and side effects, aiming to increase confidence in vaccination among hesitant individuals.
Next steps

1. Training: Effective training of healthcare providers to effectively communicate the importance and benefits of H1N1 vaccination to their patients.

2. Further research: Further research can be conducted to understand the specific reasons behind negative associations with certain predictors, allowing for tailored interventions to address barriers to vaccine acceptance and uptake.

3. Monitoring and evaluation: Monitor and evaluate the impact of implemented strategies on vaccine acceptance and uptake rates, adjusting approaches as needed for continuous and sustainable improvement.