OPTIMIZING H1N1 VACCINATION EFFORTS THROUGH PREDICTIVE MODELING

A PRESENTATION BY LISA MWIKALI

INTRODUCTION

The H1N1 virus remains a global health concern. Predictive modeling offers insights into vaccination strategies, enhancing preparedness and response. This project aims to predict the likelihood of individuals receiving the H1N1 vaccine to enhance public health strategies

BUSINESS PROBLEM

Optimizing public health efforts related to H1N1 vaccination by accurately predicting vaccination uptake and identifying key influencing factors.

PROJECT GOALS

Primary Goal: To predict individuals' likelihood of receiving the H1N1 vaccine to improve vaccination campaigns.

Specific Objectives:

- Identify significant predictors of vaccination uptake.
- Explore interaction effects between predictors.
- Develop a predictive model for classification.

STAKEHOLDERS

Key Stakeholders:

- **Public Health Authorities**: To enhance vaccination campaigns and outreach initiatives.
- **Policymakers**: To make informed decisions about vaccination policies and resource allocation.
- **Healthcare Providers**: To identify high-risk communities for targeted interventions.

DATA OVERVIEW

Training Set Features: Includes demographics, health behavior, and vaccination status.

- **Demographics**: Age group, education level, marital status, income, and employment status.
- **Health Behavior**: Doctor's recommendation, perceived risk of contracting H1N1, health insurance status, and previous vaccination history.
- **Vaccination Status**: Whether the individual received the H1N1 and seasonal flu vaccines.

Test Set Features: Similar to the training set, with demographics and health behavior data used to predict vaccination status.

Labels: H1N1 and seasonal vaccine status.

METHODOLOGY

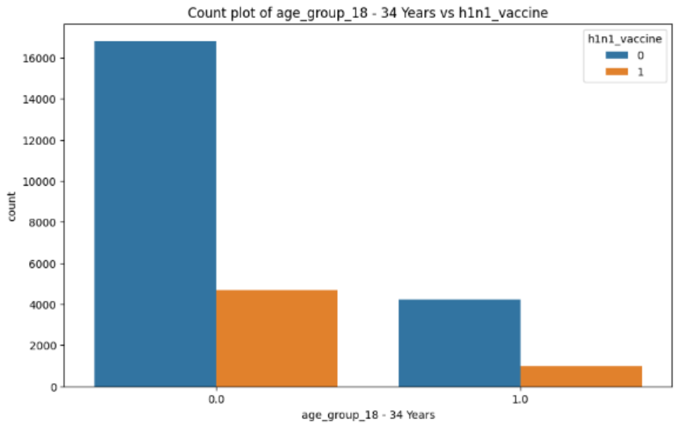
Steps:

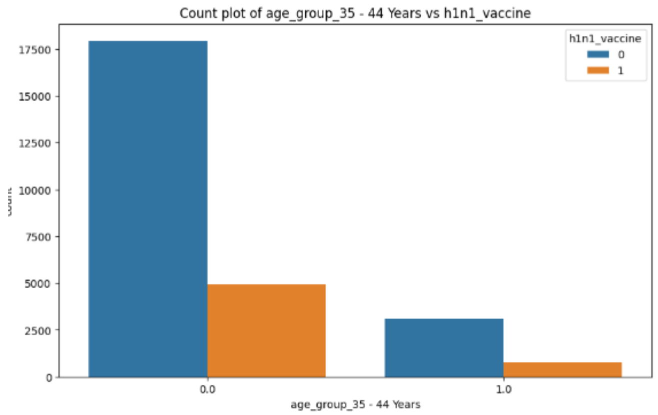
- **Data Loading and Inspection**: Cleaning and understanding the dataset.
- **Data Cleaning and Preprocessing:**
- Handling missing values and normalizing data.
- Mapping coded variables to meaningful labels for better interpretability.
- **Model Training**: Using various classification models to predict vaccination status.
- **Evaluation**: Comparing model performance using metrics like accuracy, precision, recall, and ROC-AUC.

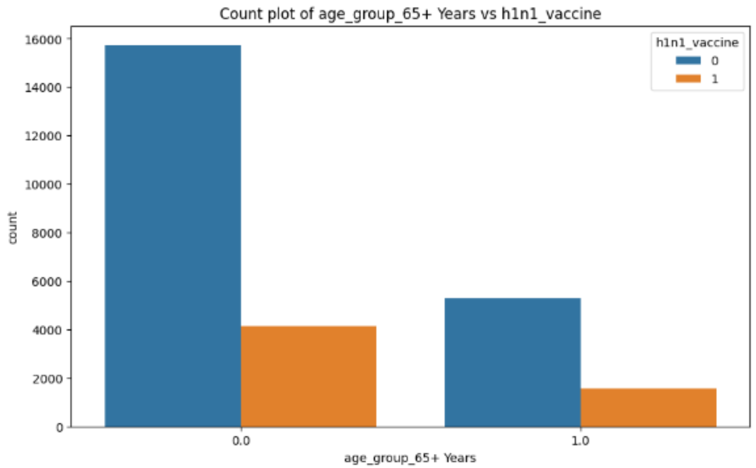
SAMPLE DATA VISUALIZATIONS

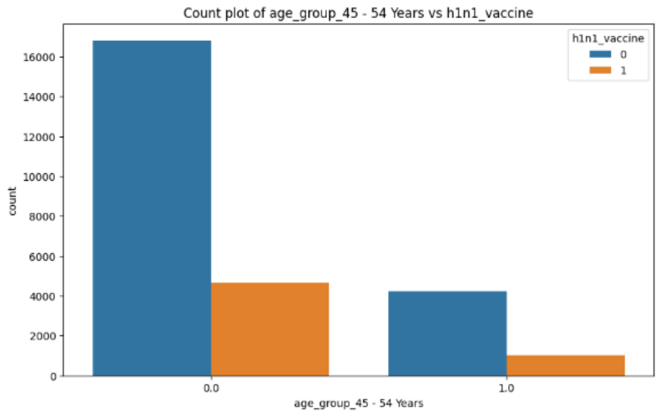
- **Demographics Distribution**: Visual representation of age groups, education levels, and income brackets.
- **Health Behavior Insights**: Charts showing the correlation between doctor's recommendations and vaccination uptake.

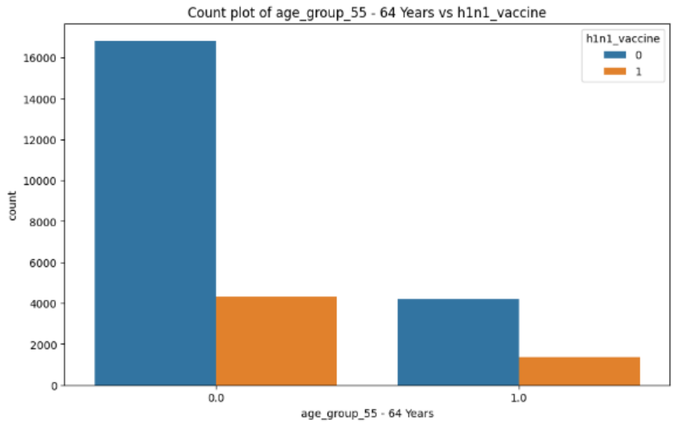
VISUAL REPRESENTATION OF AGE GROUPS VS H1N1 CONCERN



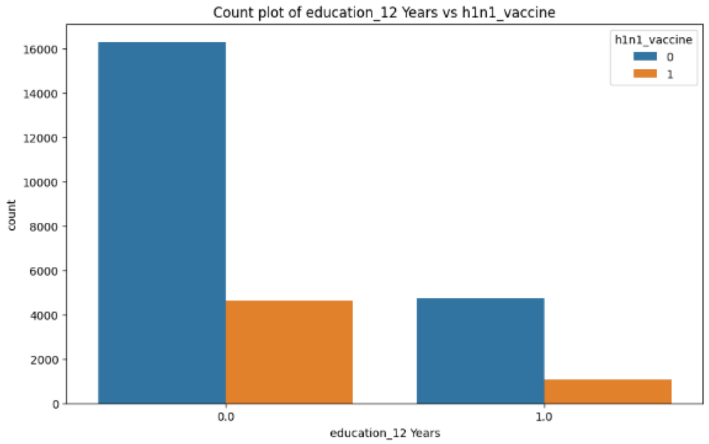


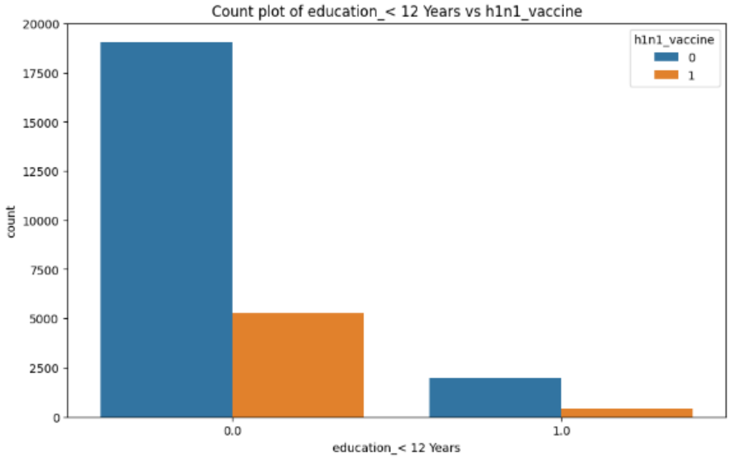


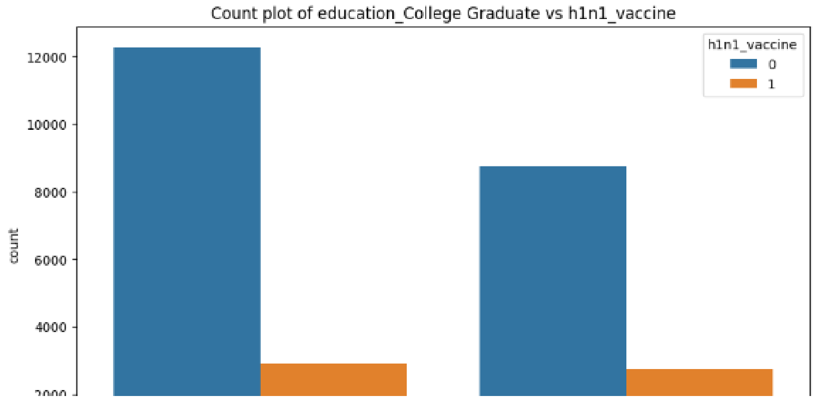


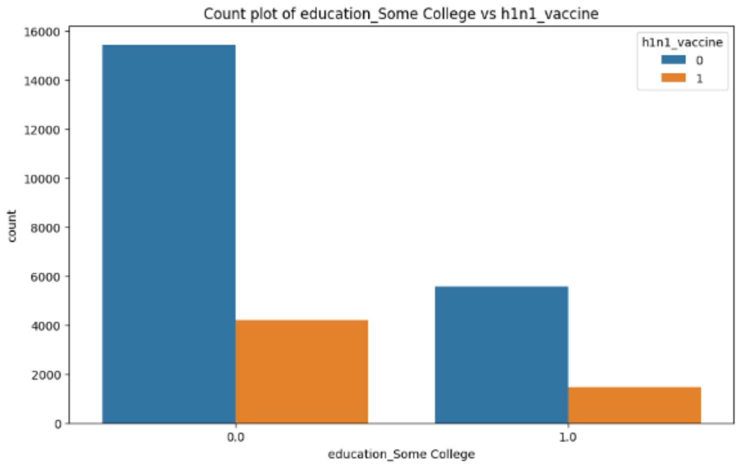


VISUAL REPRESENTATION OF EDUCATION LEVEL VS H1N1 CONCERN

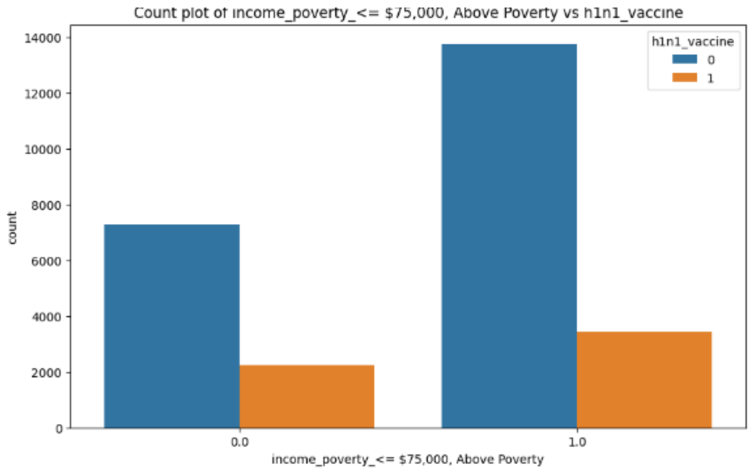


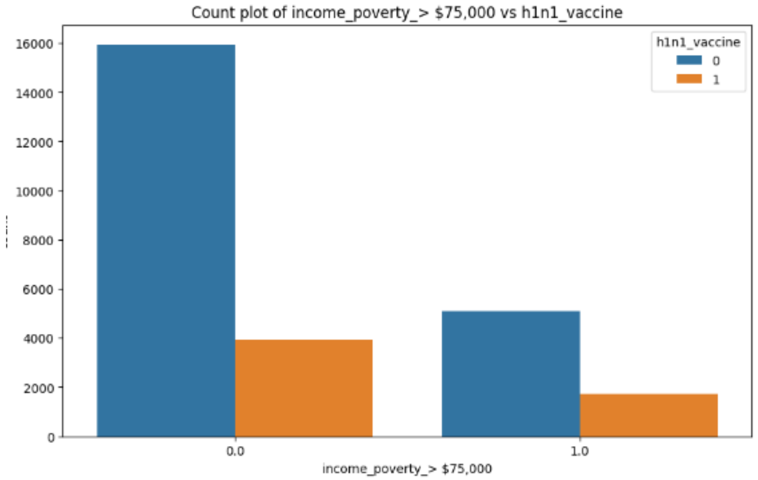


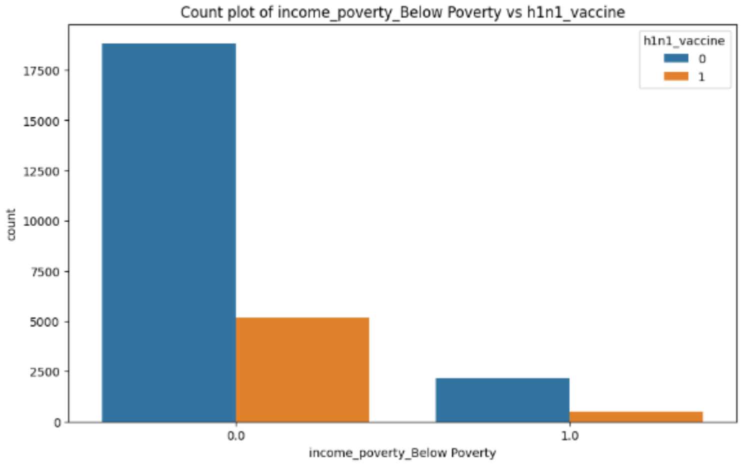




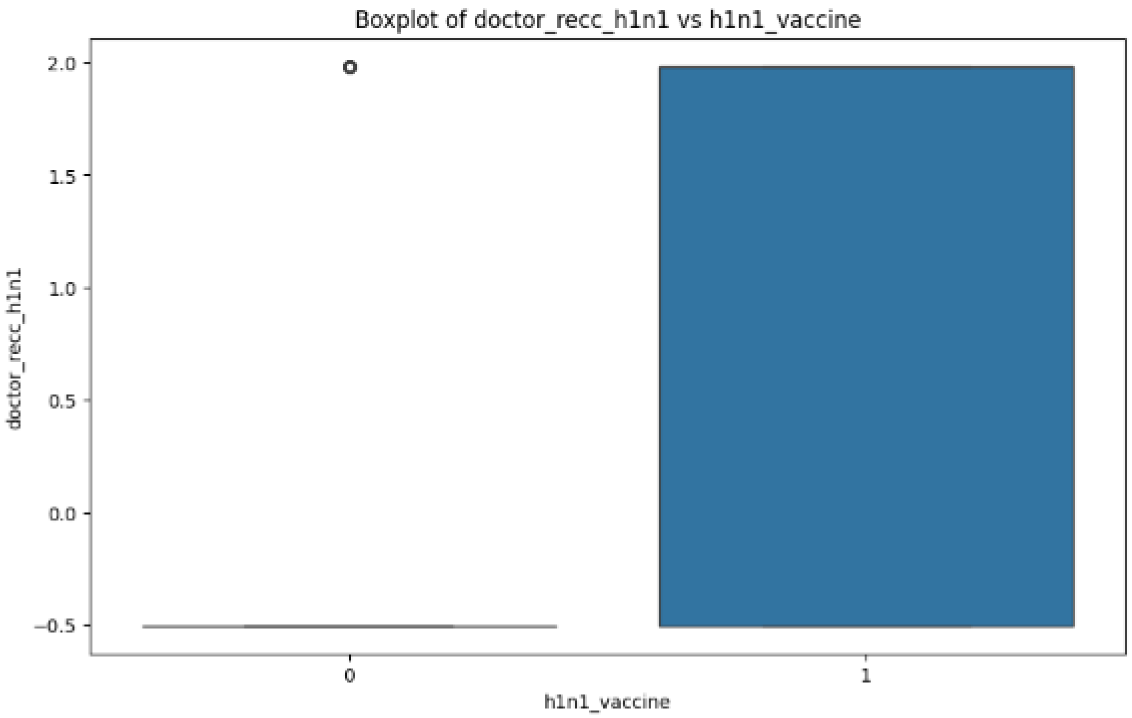
VISUAL REPRESENTATION OF INCOME LEVEL VS H1N1 CONCERN







BOX PLOT SHOWING DOCTOR RECOMMENDATION NOTE VS H1N1 VACCINE



SIGNIFICANT PREDICTORS

- **Demographics**: Age, education level, employment status.
- **Health Factors**: Doctor recommendations, perceived risk, and vaccine effectiveness.
- **Socioeconomic Factors**: Income, marital status.

KEY INSIGHTS FROM DATA EXPLORATION

- **Demographic Trends**: Younger individuals and those with higher education levels are more likely to get vaccinated.
- **Health Influences**: Doctor's recommendation significantly increases the likelihood of vaccination.
- **Socioeconomic Factors**: Income and employment status influence vaccine uptake, with higher-income individuals more likely to get vaccinated.

BASELINE MODEL USING RANDOM FOREST RESULTS

```
Validation Accuracy: 0.8367652564582553
Validation ROC AUC: 0.8259515585474287
Classification Report:
              precision    recall  f1-score   support

           0       0.85      0.96      0.90      4212
           1       0.72      0.38      0.49      1130

    accuracy                           0.84      5342
   macro avg       0.79      0.67      0.70      5342
weighted avg       0.82      0.84      0.82      5342
```

PREDICTIVE MODELING

Evaluation Metrics:

- Accuracy
- Precision
- Recall
- F1 Score

PREDICTIVE MODEL PERFORMANCE OF OTHER MODELS

```
Validation Accuracy: 0.8398
Precision: 0.6986
Recall: 0.4265
F1-score: 0.5297
ROC-AUC: 0.8314
------------------------------------------
Model: Decision Tree
Cross-Validation Accuracy: 0.7530 +/- 0.0032
Validation Accuracy: 0.7551
Precision: 0.4253
Recall: 0.4487
F1-score: 0.4367
ROC-AUC: 0.6430
------------------------------------------
Model: Random Forest
Cross-Validation Accuracy: 0.8341 +/- 0.0027
Validation Accuracy: 0.8373
Precision: 0.7208
Recall: 0.3770
F1-score: 0.4951
ROC-AUC: 0.8319
------------------------------------------
Model: Support Vector Machine
Cross-Validation Accuracy: 0.8362 +/- 0.0045
Validation Accuracy: 0.8398
Precision: 0.6986
Recall: 0.4265
F1-score: 0.5297
ROC-AUC: 0.8052
```

```
------------------------------------------
Model: Gradient Boosting
Cross-Validation Accuracy: 0.8380 +/- 0.0032
Validation Accuracy: 0.8411
Precision: 0.6993
Recall: 0.4363
F1-score: 0.5373
ROC-AUC: 0.8390
------------------------------------------
```

Best performing model: Gradient Boosting with an accuracy of 84.11%.

RECOMMENDATIONS

- **Targeted Educational Campaigns**: Focus on low-education groups to increase vaccine awareness.
- **Improved Access**: Enhance healthcare services and outreach programs for low-income individuals.
- **Cultural Tailoring**: Address cultural beliefs and preferences in vaccination campaigns.
- **Policy Adjustments**: Allocate resources to areas with low vaccination uptake.
- **Healthcare Provider Training**: Strengthen provider recommendations through education and training.
- **Clear Communication**: Address misconceptions with accurate information about H1N1 risks and vaccine effectiveness.

CONCLUSION

Our data-driven approach provides valuable insights into vaccination behaviors, enabling stakeholders to design more effective public health strategies.