

H1N1 VACCINE UPTAKE MODEL PREDICTION

"Vaccines are the tugboats of preventive health"

Presented by: Mitchell Chege



Table of Contents



01 Introduction

02 Business Context

03 Exploratory Data Analysis

04 Data Modeling

05 Conclusion

06 Future Improvement Ideas



INTRODUCTION

Vaccines provide immunization for individuals, and enough immunization in a community can further reduce the spread of diseases. However, it was noted in a study that numerous participants expressed reluctance or hesitation towards getting vaccinated.

Vaccine hesitancy is a well-known phenomenon, and the World Health Organization recognizes it as one of the top ten threats to global health.

Having a comprehensive knowledge of the traits linked to individuals' vaccination behavior can aid in the planning and implementation of future public health initiatives.



PROBLEM STATEMENT

Government and NGOs carry out vaccination drives to prevent the occurrence of certain diseases. However due to vaccine hesitancy, the uptake of these vaccines is not as high as expected. This is caused by various factors such as social, economic, and demographic background, opinions on risks of illness and vaccine effectiveness, and behaviors towards mitigating transmission.

It would be beneficial to these stakeholders to identify factors contributing to vaccine hesitancy to allow for proper planning of mitigating grounds that will encourage vaccine uptake.



MAIN OBJECTIVE

The goal of this project is to build a predictive model determining whether people got H1N1 vaccines using information shared about their backgrounds, opinions, and health behaviors.



SPECIFIC OBJECTIVES



To identify the most significant features in determining whether an individual is vaccinated against H1N1.



To build a Decision Tree, Random Forest and SVM model that can accurately predict whether an individual is vaccinated against H1N1.



To assess the performance of the predictive models and identify the best one and potential areas for improvement.



METRIC FOR SUCCESS



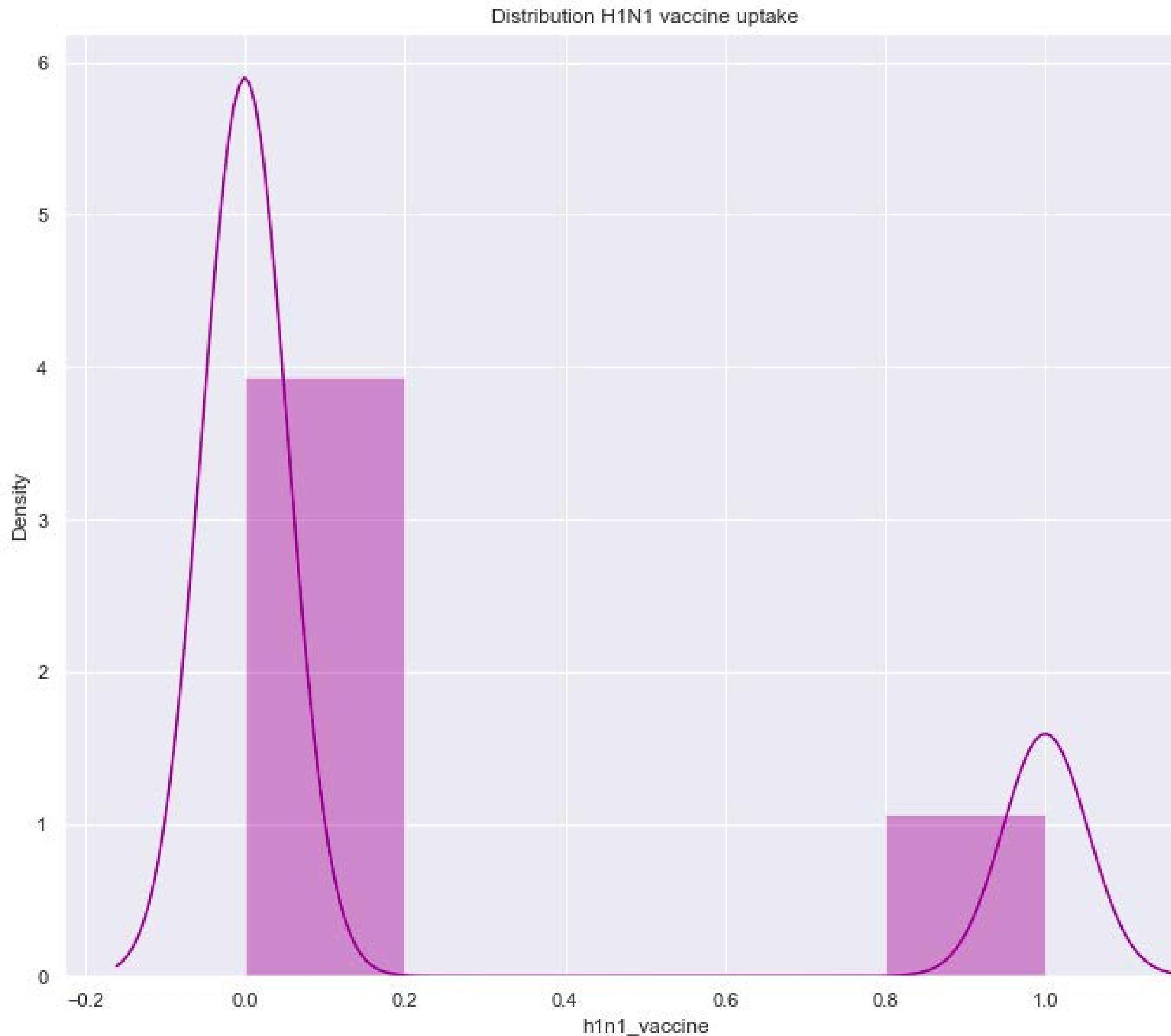
75%

The project will be considered a success if the developed predictive model has a predictive score of 75%

EXPLORATORY DATA ANALYSIS



UNIVARIATE ANALYSIS



The target variable has two unique values, that is, 1 and 0 which are normally distributed.

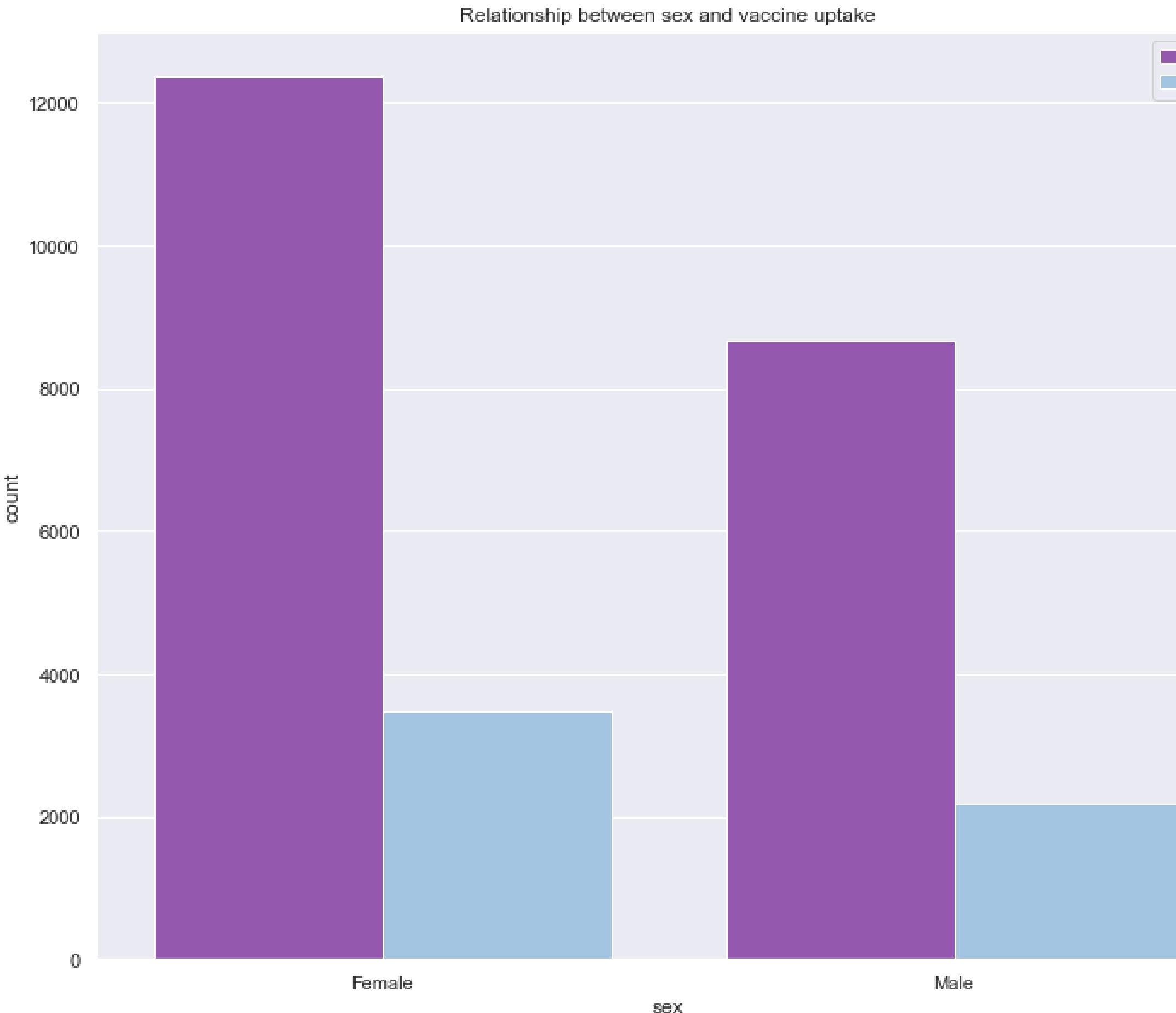
- Skewness - **1.406** (positively skewed)
- Kurtosis - **-0.02** (mesokurtic)



BIVARIATE ANALYSIS



Does the sex of a respondent determine H1N1 vaccine uptake?



Female respondents:

- unvaccinated - **12,500 (47%)**
- vaccinated- **3,500 (13%)**

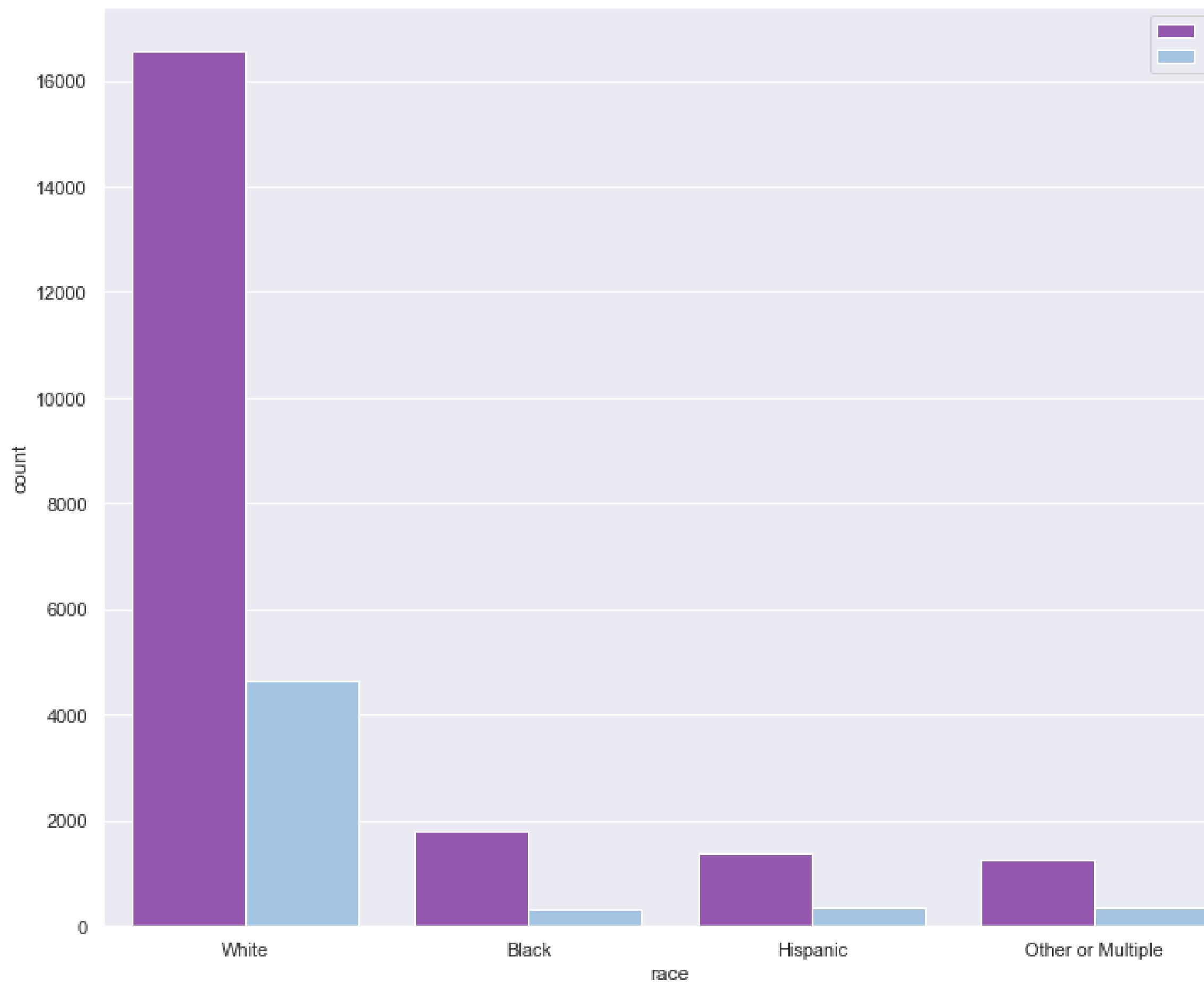
Male respondents:

- unvaccinated - **8,500 (32%)**
- vaccinated- **2,100 (8%)**

Does the race of a respondent determine H1N1 vaccine uptake?



Relationship between race and vaccine uptake



White respondents:

- unvaccinated - **16,500 (62%)**
- vaccinated- **4,500 (17%)**

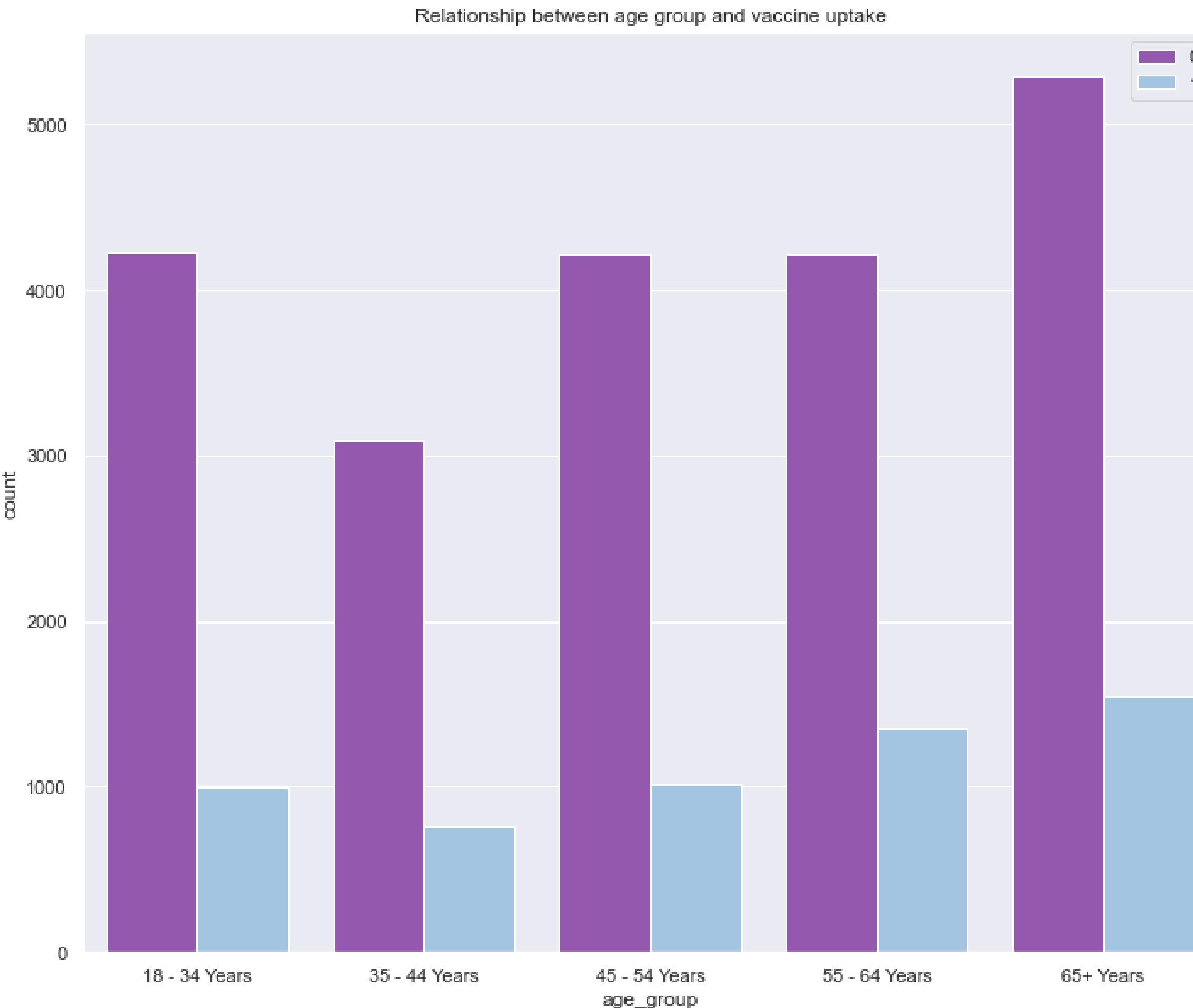
Black respondents:

- unvaccinated - **1,800 (7%)**
- vaccinated- **250 (1%)**

Hispanic respondents:

- unvaccinated - **1,300 (5%)**
- vaccinated- **400 (1.5%)**

Does the age-group of a respondent determine H1N1 vaccine uptake?



The largest number of unvaccinated, 65+ Years:

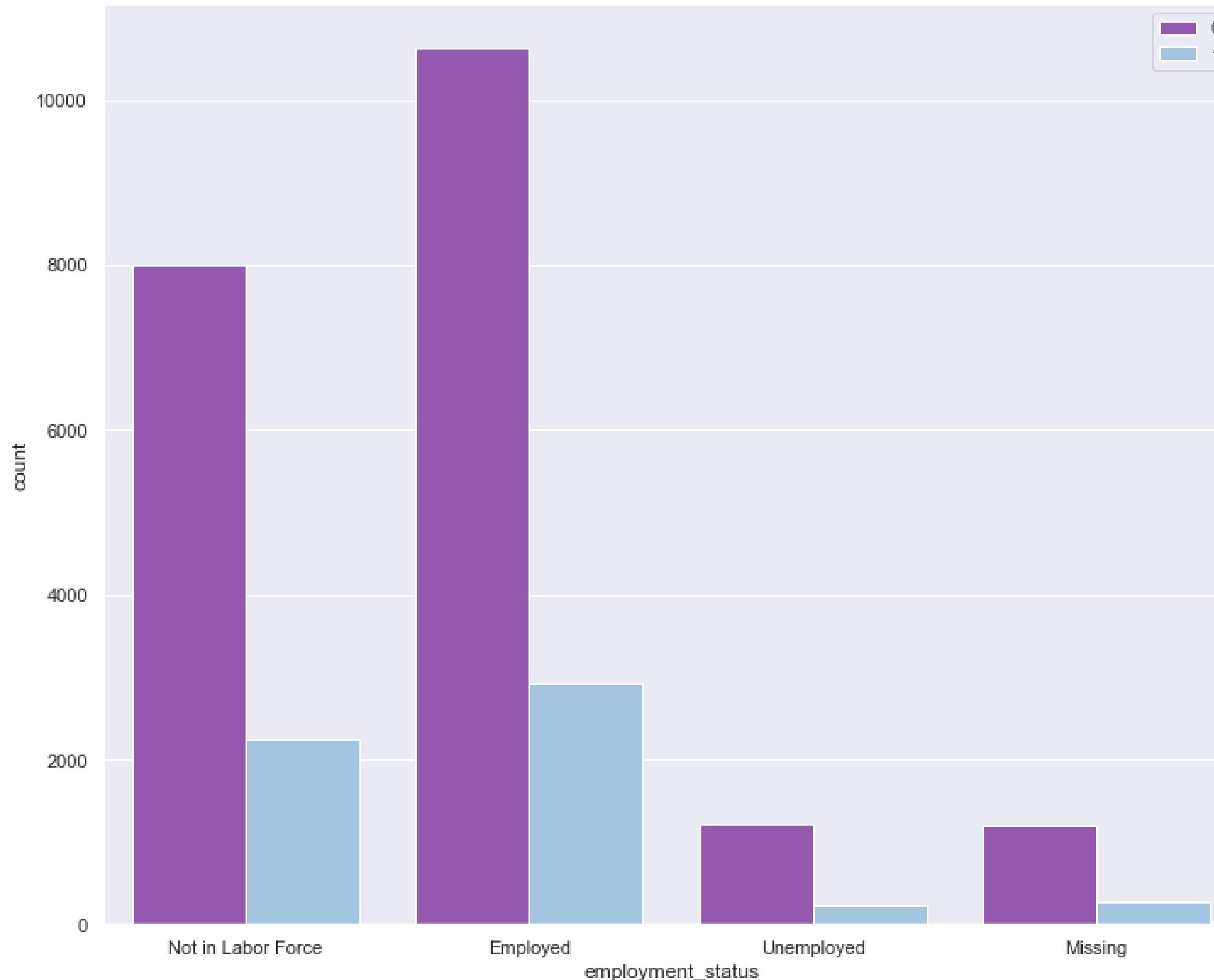
- unvaccinated - **5,300 (20%)**
- vaccinated- **1,600 (6%)**

Age groups 18 - 34 Years, 45 - 54 Years, and 55 - 64 Years, have **4,200 (16%)** unvaccinated individuals.

Does the employment_status of a respondent determine H1N1 vaccine uptake?



Relationship between employment status and vaccine uptake



Employed respondents:

- unvaccinated - **10,600 (40%)**
- vaccinated- **2,900 (11%)**

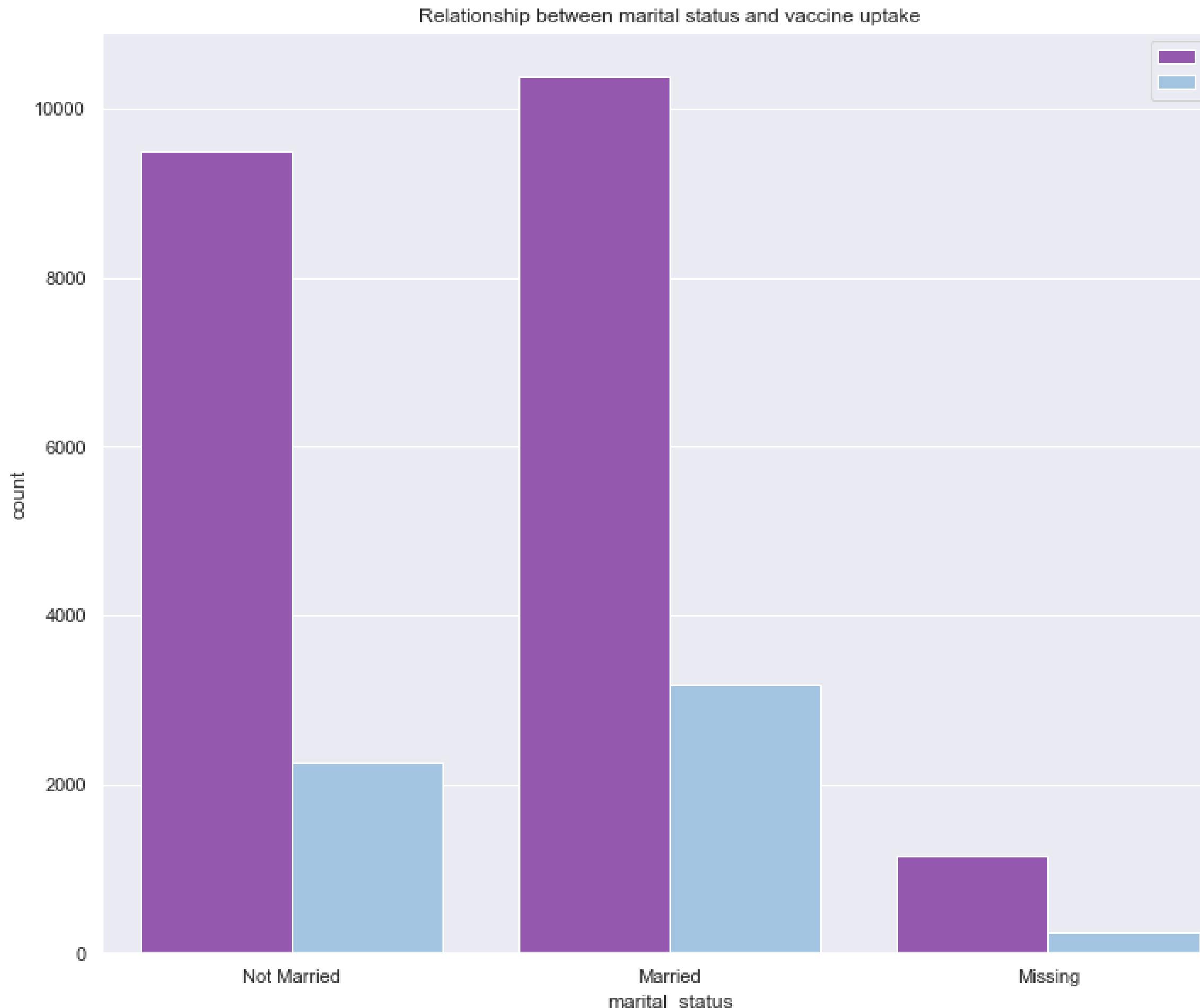
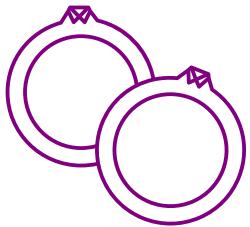
Not in the labor force:

- unvaccinated - **8,000 (30%)**
- vaccinated- **2,300 (8.5%)**

Unemployed individuals:

- unvaccinated - **1,200 (4.5%)**
- vaccinated- **250 (1%)**

Does the marital_status of a respondent determine H1N1 vaccine uptake?



Married respondents:

- unvaccinated - **10,400 (39%)**
- vaccinated- **3,200 (12%)**

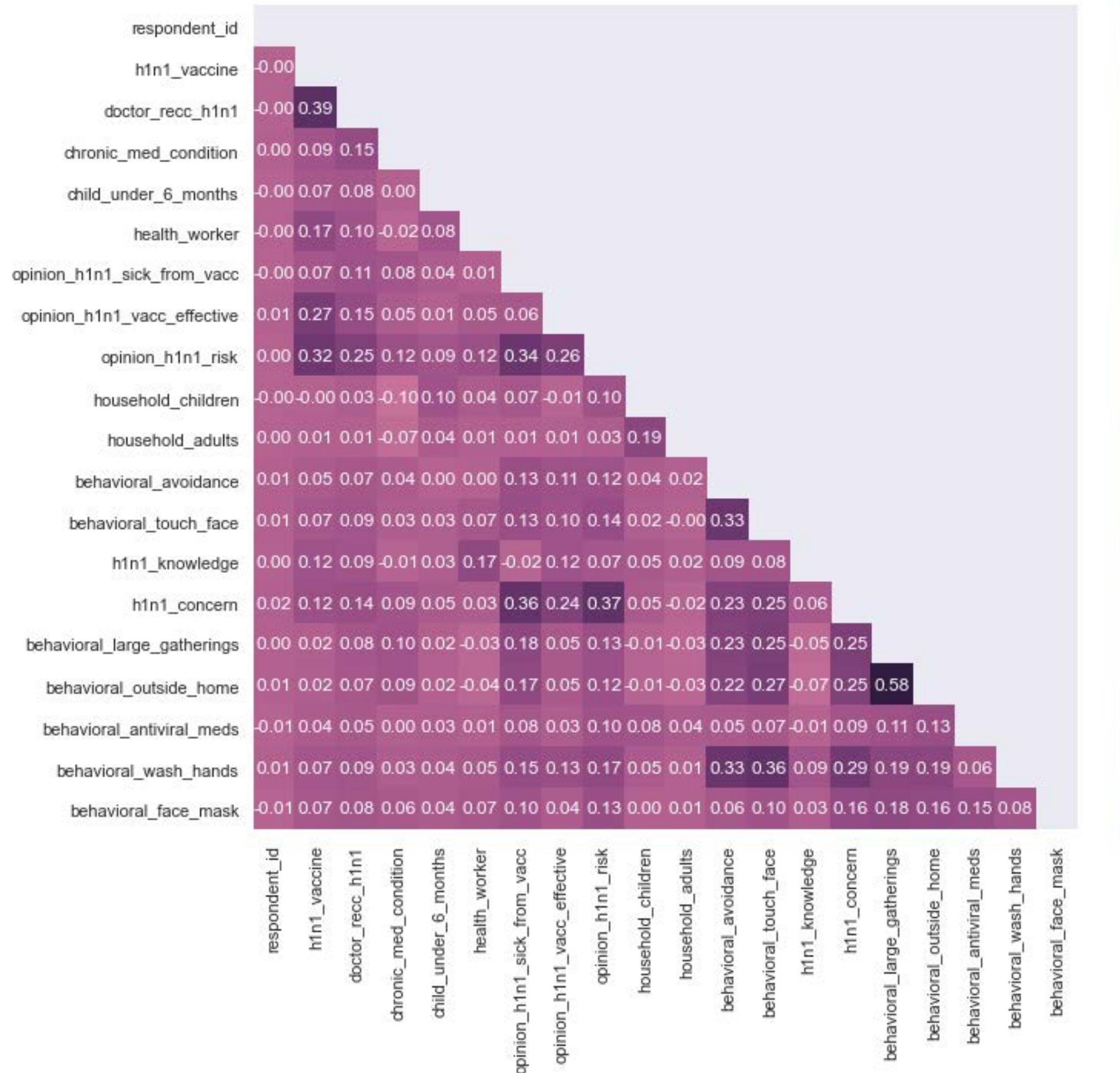
Not married respondents:

- unvaccinated - **9,500 (35.5)**
- vaccinated- **2,300 (8.5%)**

DATA MODELING



Heatmap showing the correlation between target and predictor variables

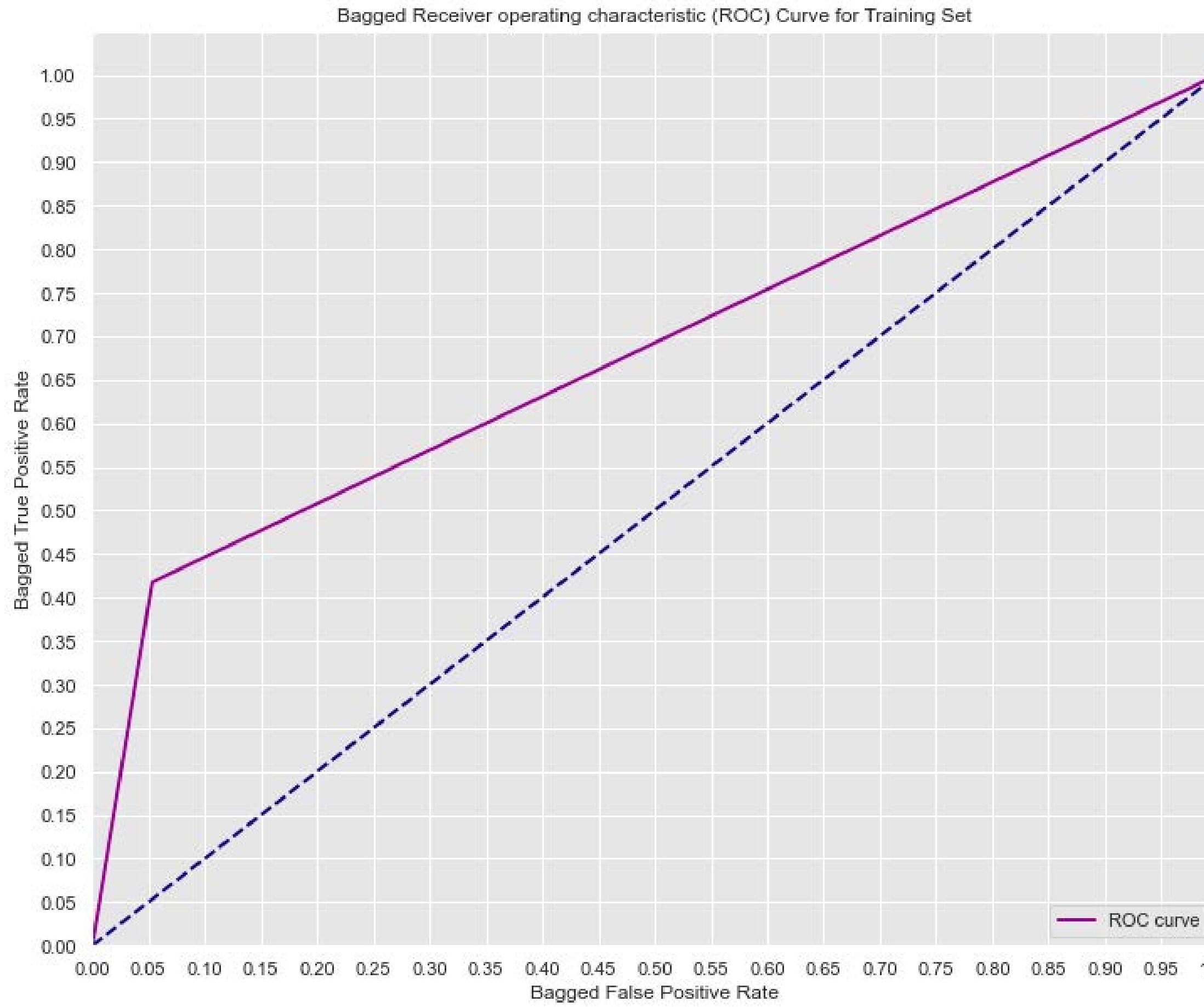


The highest positively correlated variables:

- behavioral_large_gatherings vs behavioral_outside_home at **0.58**.
- h1n1_vaccine vs doctor_recc_h1n1 at **0.39**



1. Decision Tree Classifier



Before tuning hyperparameters and bagging :

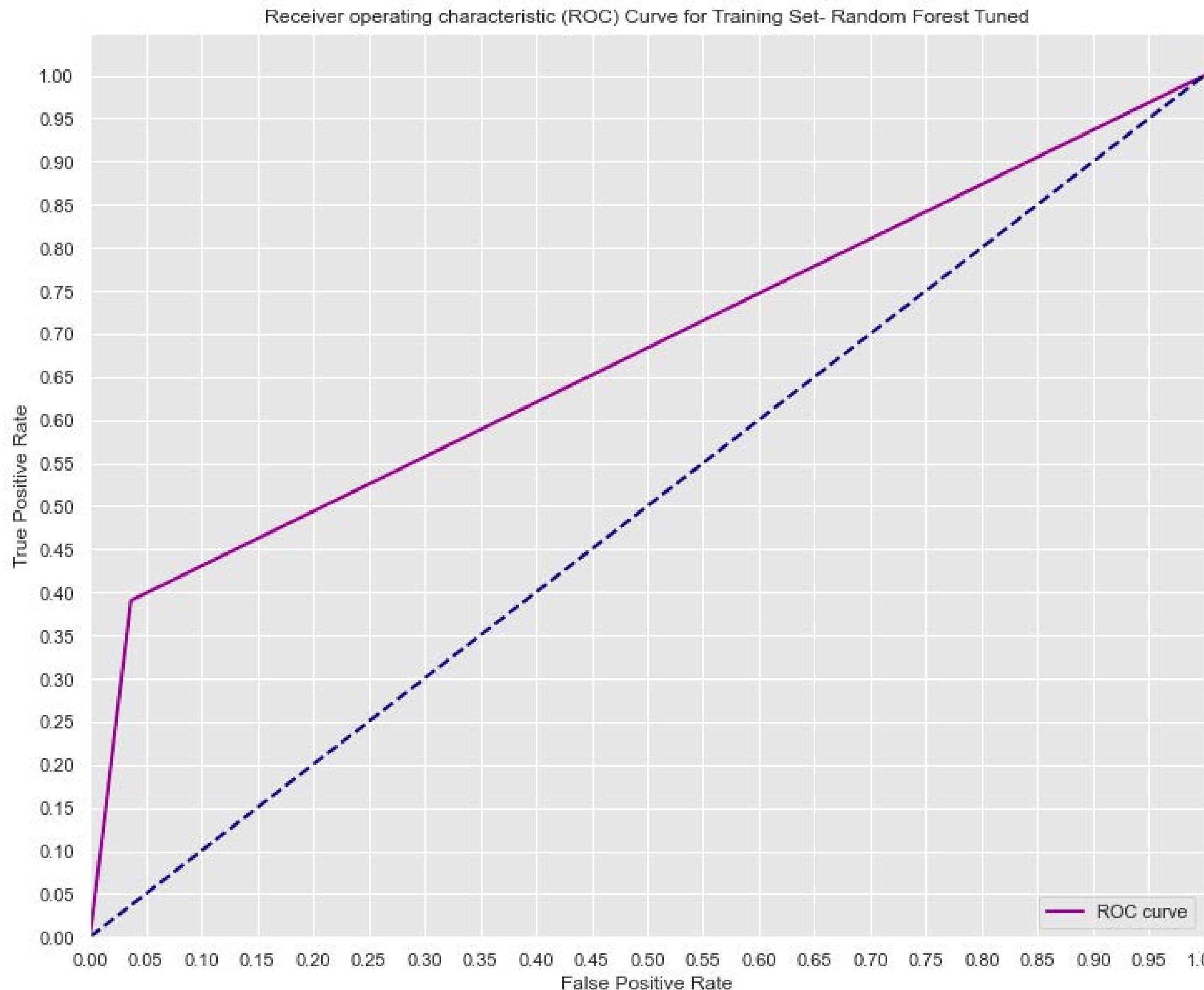
- Accuracy - **76.17%**
- Precision - **44.43%**
- Recall - **43.46%**

After tuning hyperparameters and bagging :

- Accuracy - **83.15%**
- Precision - **66.62%**
- Recall - **43.29%**



2. Random Tree



Before tuning hyperparameters :

- Accuracy - **83.17%**
- Precision - **78%**
- Recall - **65%**

After tuning hyperparameters :

- Accuracy - **83.17%**
- Precision - **78%**
- Recall - **67%**



3. Support Vector Machine (SVM)



Model	Accuracy	Precision	Recall	F1
Decision Tree	83%	76%	68%	71%
Random Forest	83%	78%	67%	70%
SVM	83%	77%	68%	71%

Before tuning hyperparameters:

- Accuracy - 84%
- Precision - 77%
- Recall - 69%

After tuning hyperparameters:

- Accuracy - 83%
- Precision - 77%
- Recall - 68%

CONCLUSION

- 01 The tuned SVM model is the best model for predicting whether or not individuals would get vaccinated.
- 02 doctor_recc_h1n1 is a very important feature in determining the target variable.
- 03 Female respondents who are vaccinated are more than males.
- 04 A higher number of 65+ Years old respondents are vaccinated compared to other age groups.
- 05 A higher number of employed respondents are vaccinated compared to other respondents.
- 06 A higher number of white respondents are vaccinated compared to other races.



Future im

The tuned SVM model is the best model for predicting whether or not individuals would get vaccinated.

doctor_recc_h1n1 is a very important feature in determining the target variable.

Female respondents who are vaccinated are more than males.

A higher number of 65+ Years old respondents are vaccinated compared to other age groups.

A higher number of employed respondents are vaccinated compared to other respondents.

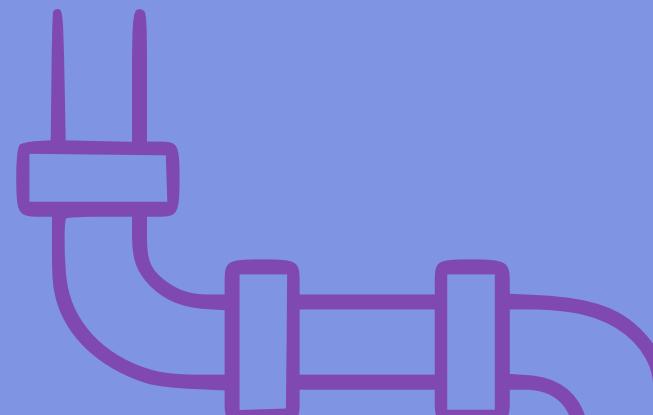
A higher number of white respondents are vaccinated compared to other races.



Future improvement Ideas



Building a
pipeline for this
model.



Tuning the
Random Forest
classifier with a
larger range.



Any Questions?



Presented by: Mitchell Chege

Get In Touch



EMAIL ADDRESS
mitchge@gmail.com