Predicting H1N1 Vaccination Status:
A Machine Learning Approach

Brooke Smyth
Emine Kesici
Tamiru Berhanu Denka

Outline

Public Health Perspective
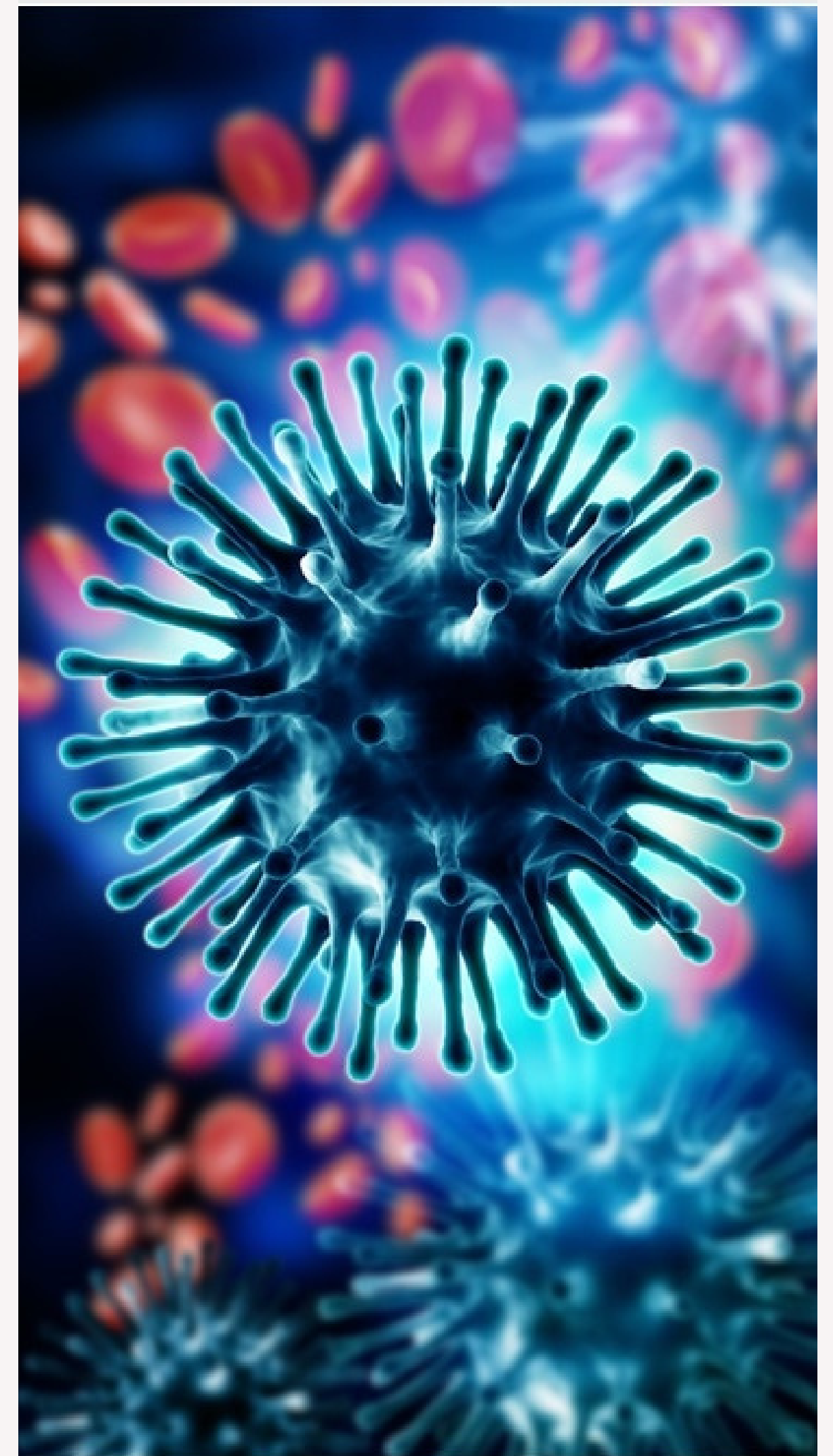
Goals

Data

Methods

Results

Recommendations

Next Steps

Public Health Perspective

- The personal factors that underlie vaccination behavior

- Understanding vaccination patterns from past pandemics can improve future vaccination

- Decrease outbreaks

GOALS

- Build an accurate H1N1 vaccination prediction model

- Find most important demographic, behavioral, and health features affecting vaccination status
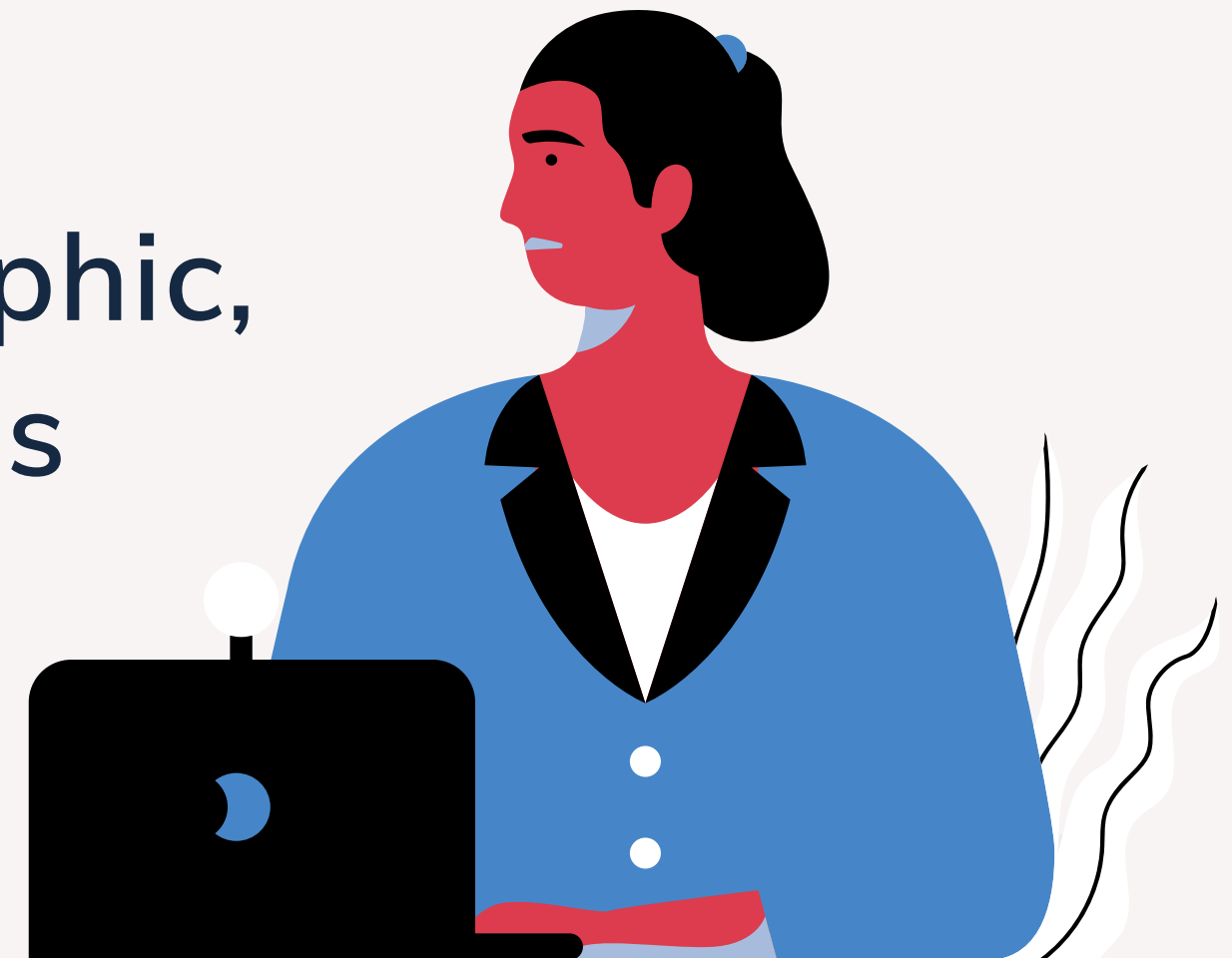
The Data

The National Flu Survey (NHFS, 2009)

26,000 Respondents

79% Did not Get the Vaccine

35 Unique Factors

Modeling Context

FALSE POSITIVE: PREDICTING THAT PEOPLE GOT THE VACCINE WHEN THEY ACTUALLY DID NOT

FALSE NEGATIVE: PREDICTING THAT PEOPLE DID NOT GET THE VACCINE WHEN THEY ACTUALLY DID

Big Problem

Not a Big Problem

Model & Results

False Positive

Confusion Matrix

Gradient Boosting Score

84% Accuracy

Actuals

Not Vaccinated | 0.95 | 0.055

Vaccinated | 0.51 | 0.49

Not Vaccinated | Vaccinated

Predictions
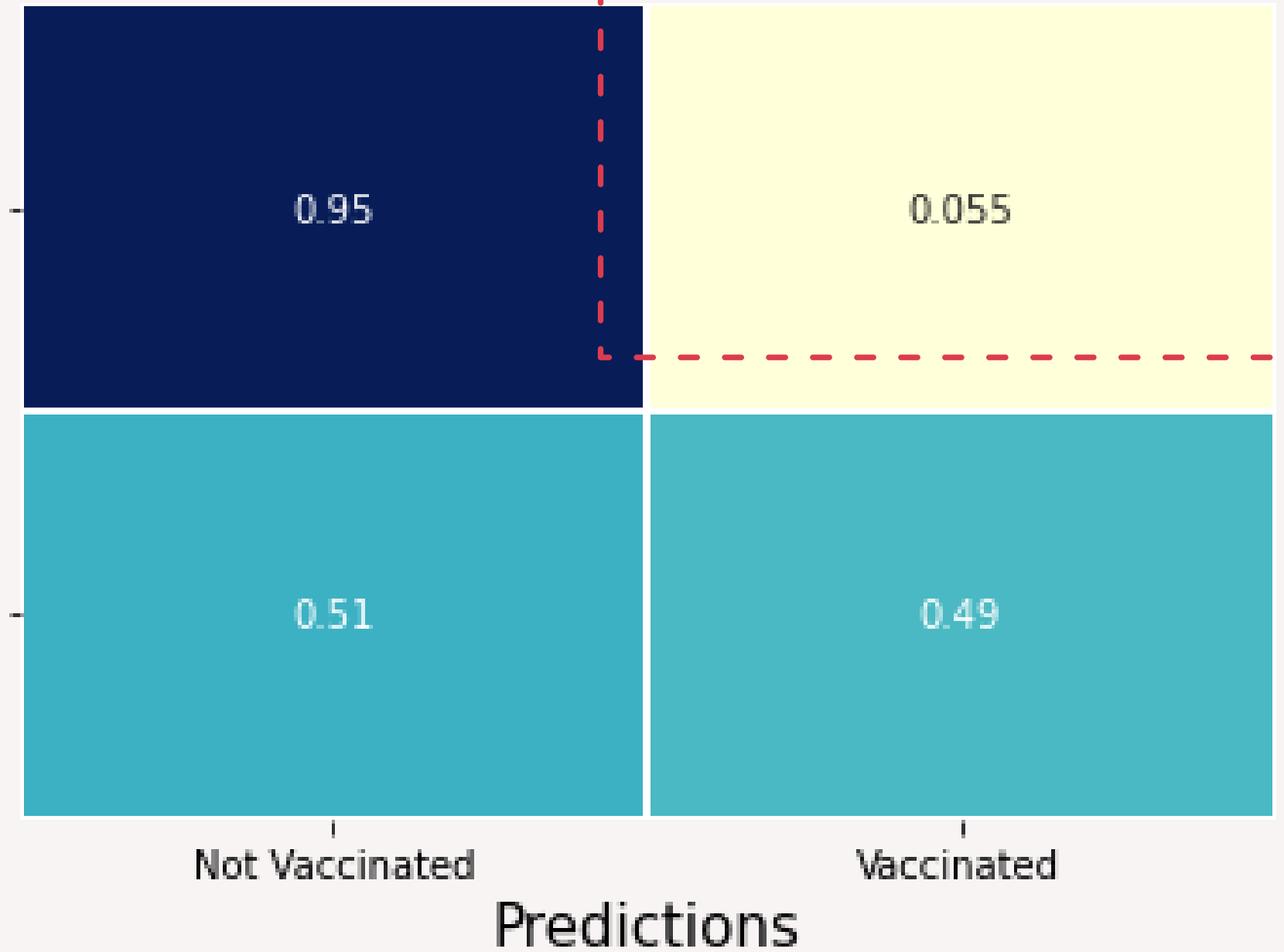
Top 4 Important Features

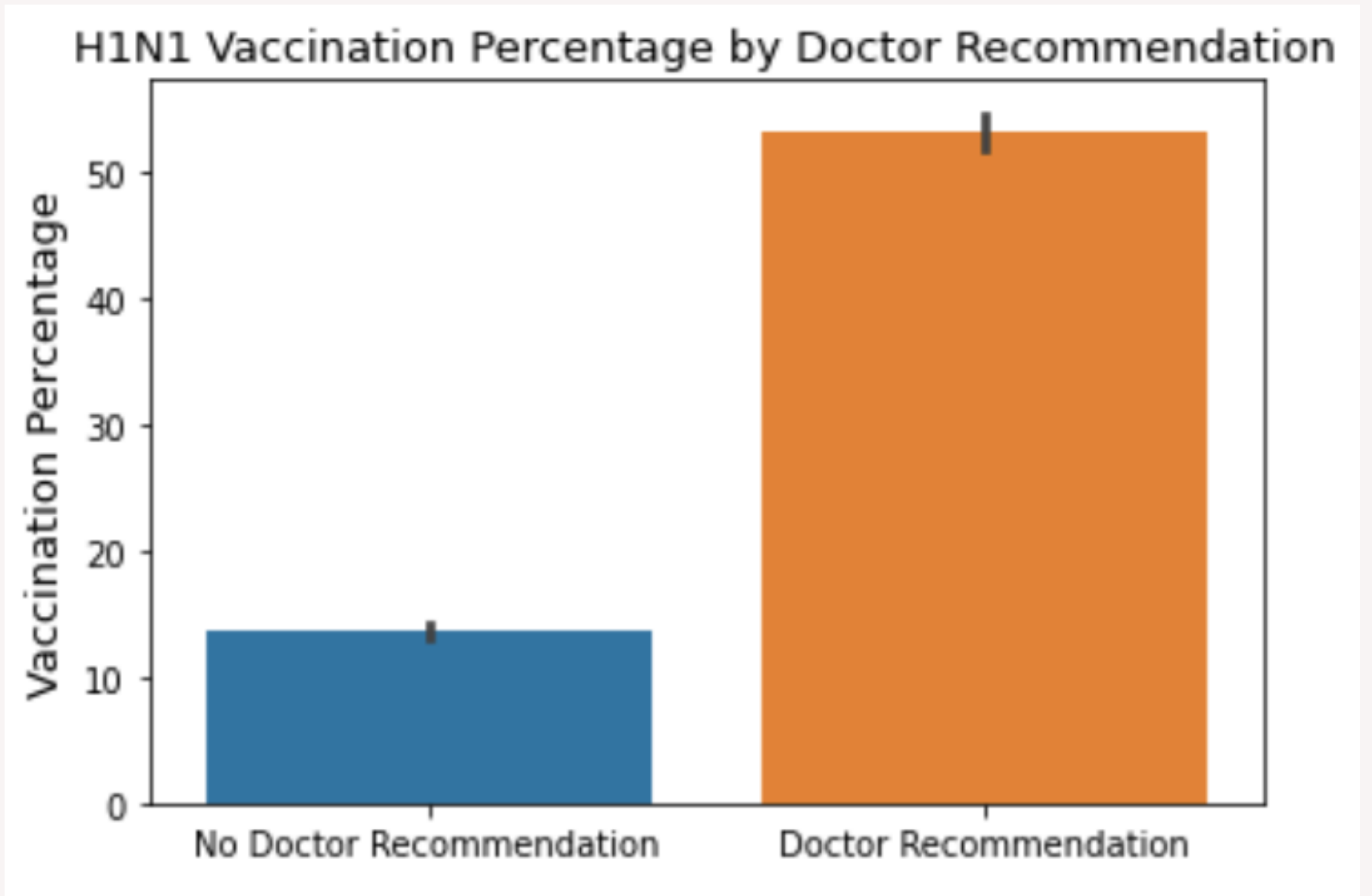**Doctor Recommendation
of H1N1 Vaccine**

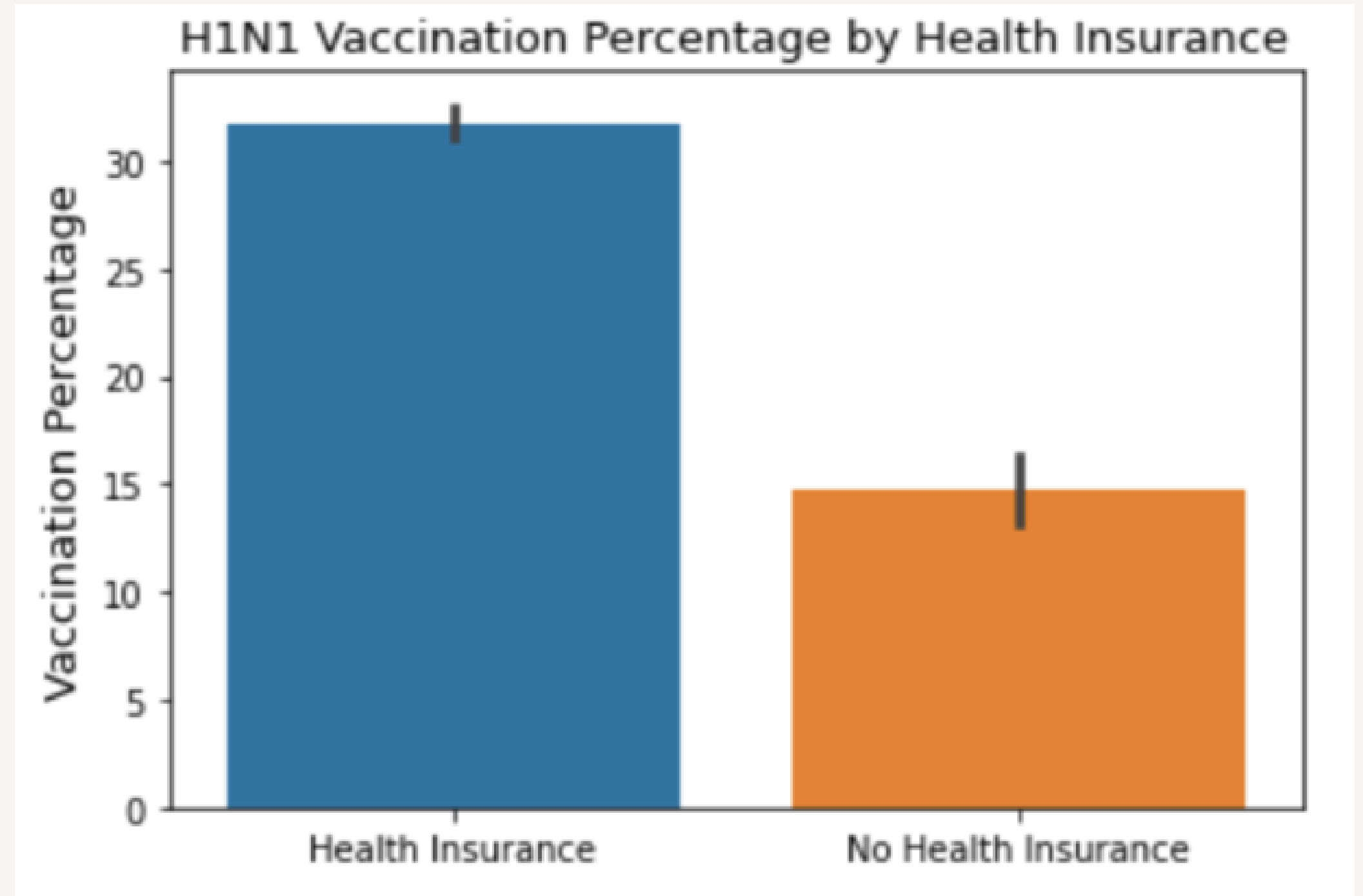**Health Insurance**

**Opinion on H1N1 Vaccine
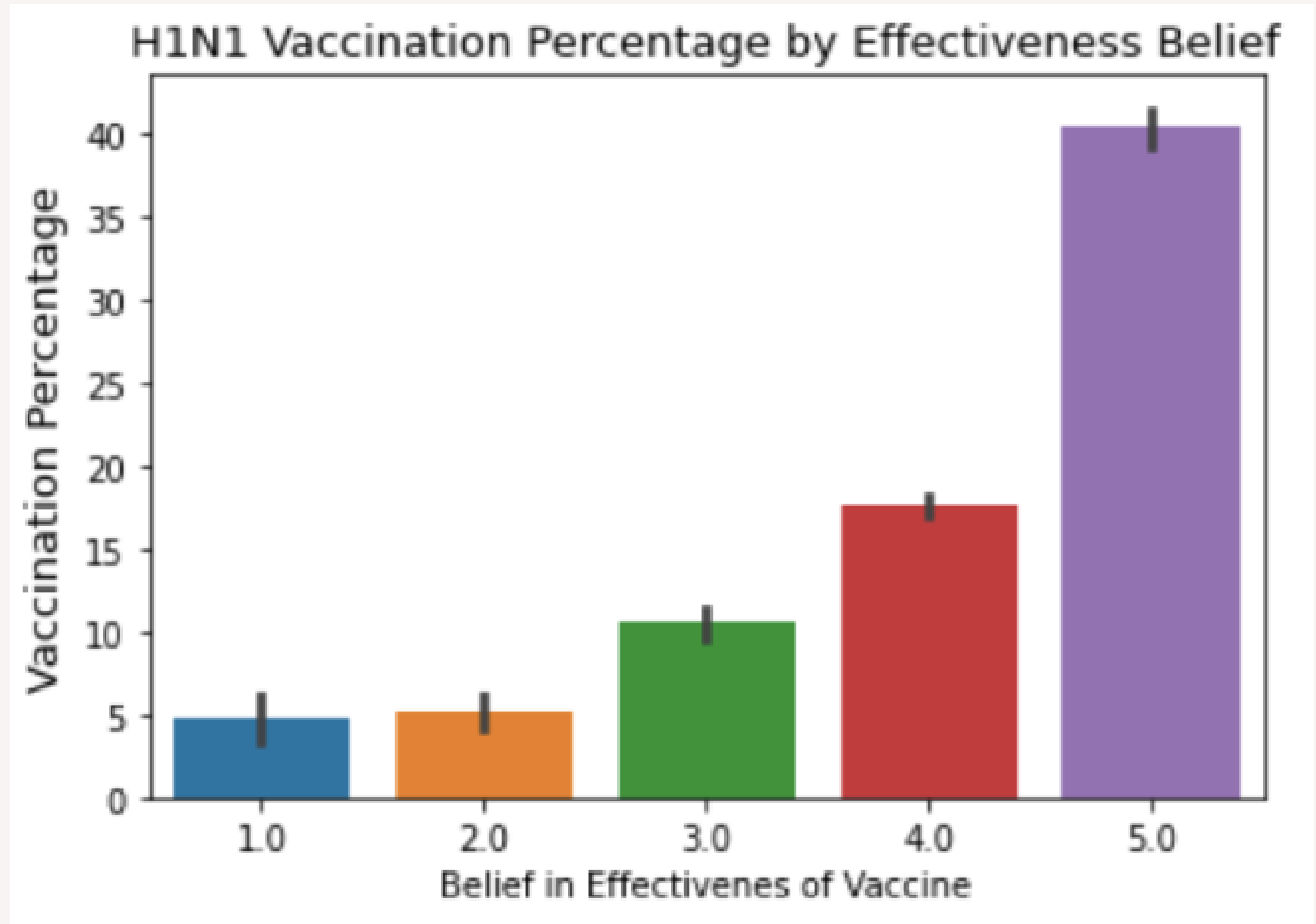Effectiveness**

Opinion on H1N1 Risk

Data Visualizations: Doctor Recommendation
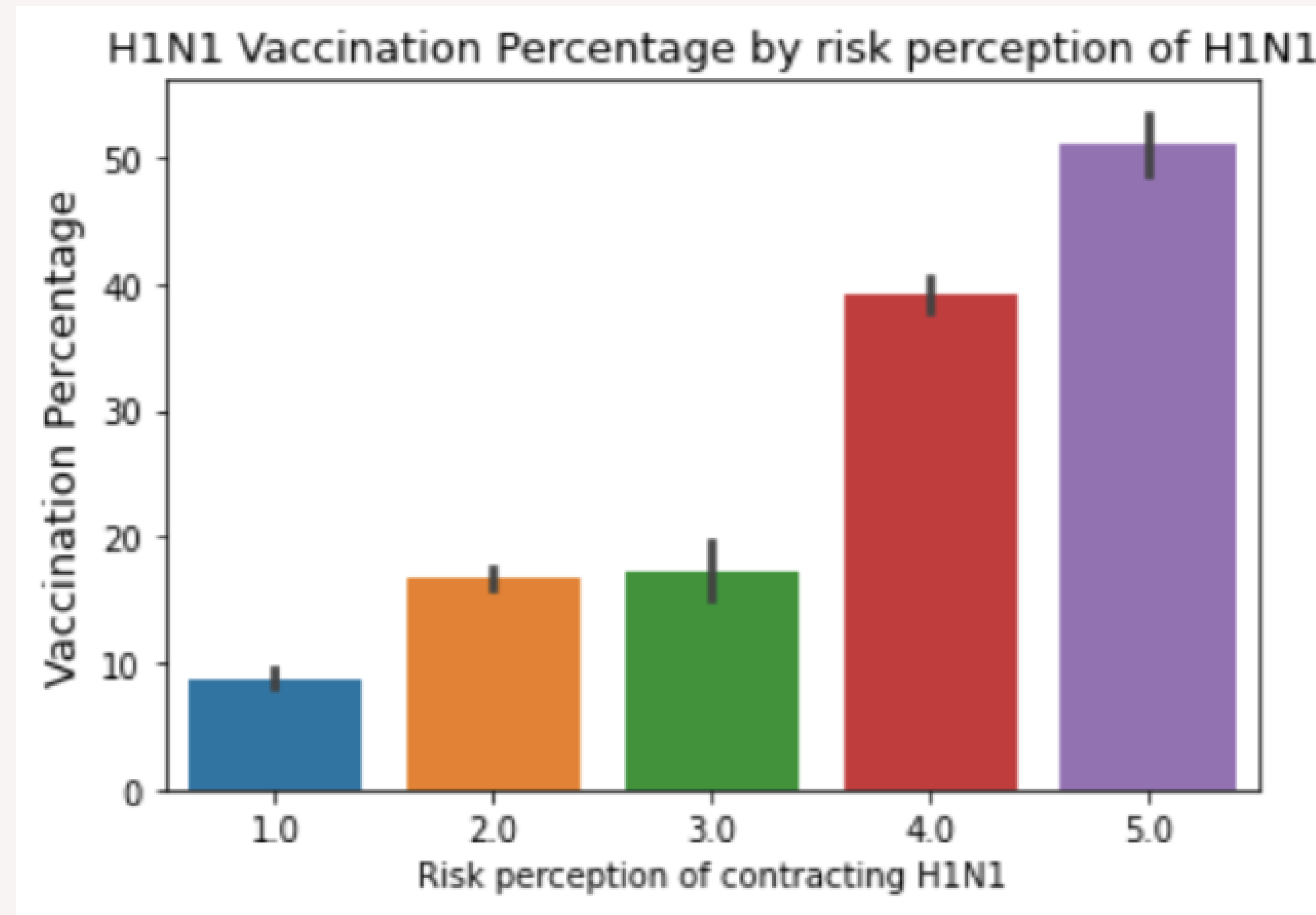

H1N1 Vaccination Percentage by Doctor Recommendation

Data Visualizations: Health Insurance

Data Visualizations: Belief in H1N1 Vaccine Effectiveness

Data Visualizations: H1N1 Risk Perception

Recommendations

Doctor Recommendations

Health Insurance

Educational Outreach

Next Steps

Recent Survey Data
Pre/After Covid-19 Difference

More Feature Engineering
Improve Accuracy

Seasonal Vaccine Prediction
Enhance and Generalize Model

What to do

Thank you!

Research team

Brooke Smyth
@Brooke57

Emine Kesici
@emykes

Tamiru B. Denka
@tamiru3

Appendix

Data Source

https://www.drivendata.org/competitions/66/flu-shot-learning/page/213/

Appendix

Comparison of Multiple Model Results