How many individuals will receive the seasonal flu vaccine?

Michael Kearns

Overview

1) Business Problem
2) The Data
3) Modeling Process
4) Model Evaluation
5) Conclusions/Next Steps

Business Problem

Goal: *Minimize* vaccine production costs and *reduce* waste.

Solution: *Predict* how many individuals will receive the seasonal flu vaccine this year based on recent data.

The Data

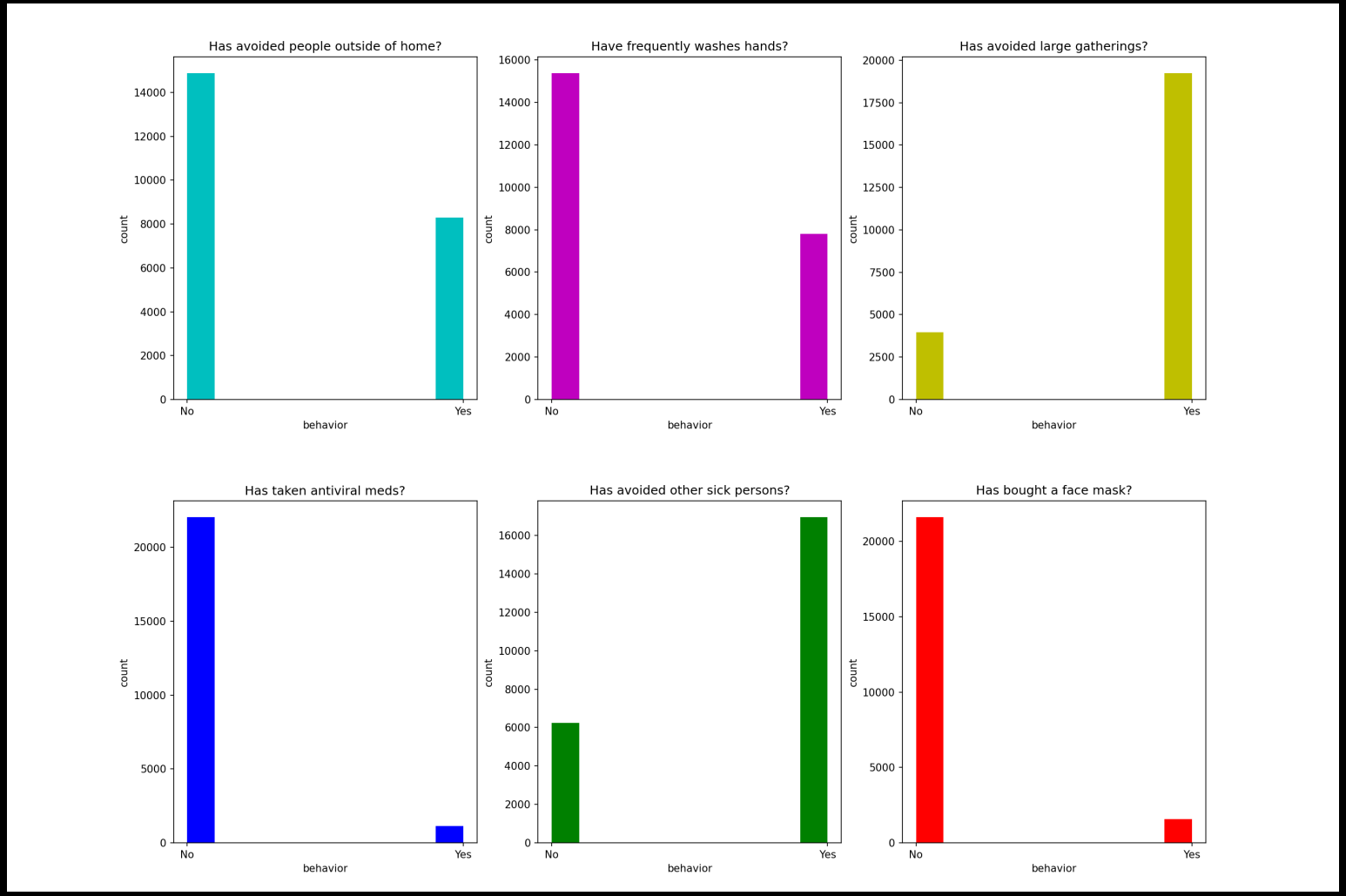
- Data from the National 2009 H1N1 Flu Survey

- Over 26,000 respondents
  - Health Behaviors & Opinions
  - Socioeconomic background
  - Personal background
  - Educational background

- Focus on **Seasonal Flu** data

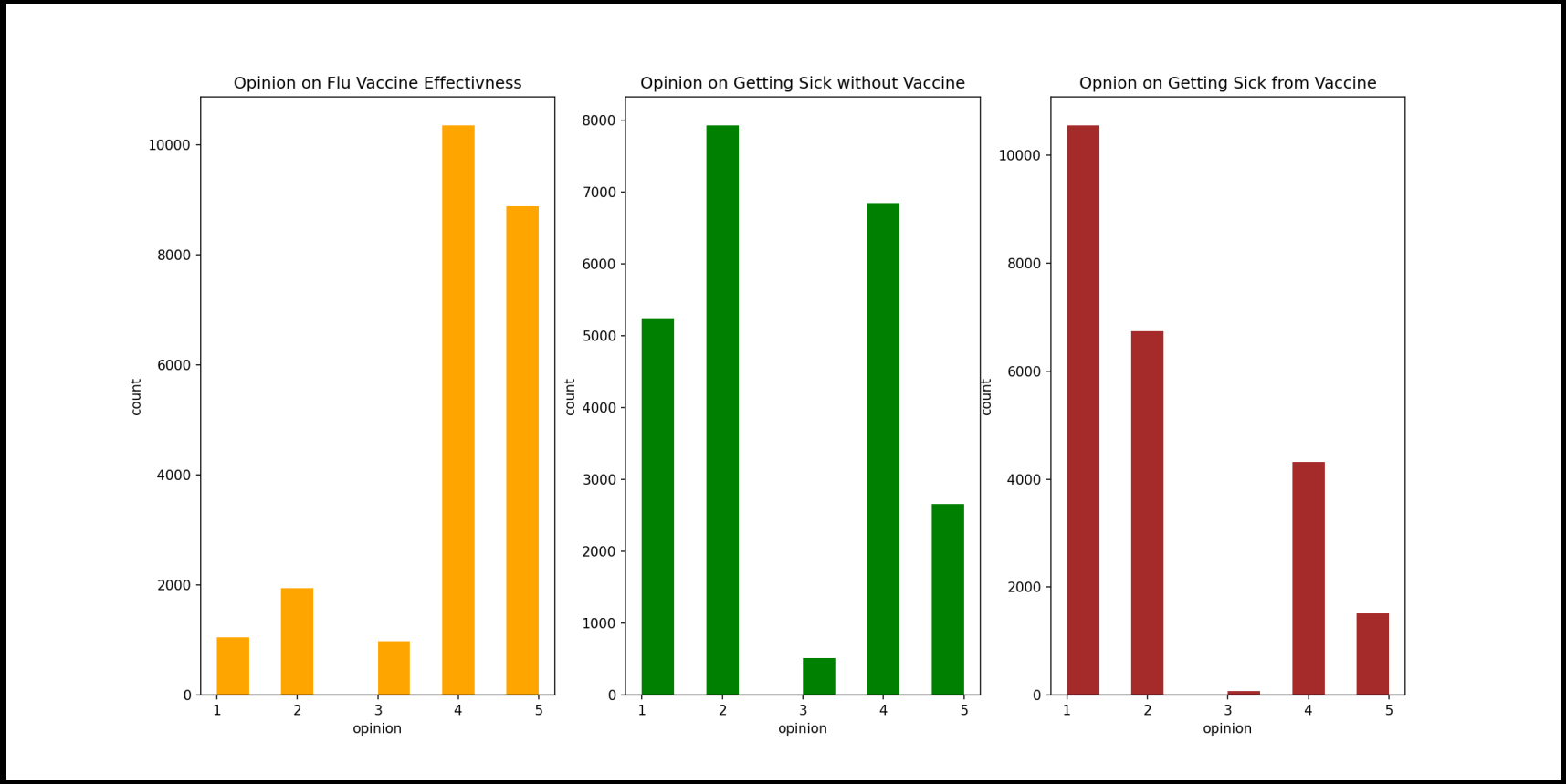
Primary Features

Health Behaviors



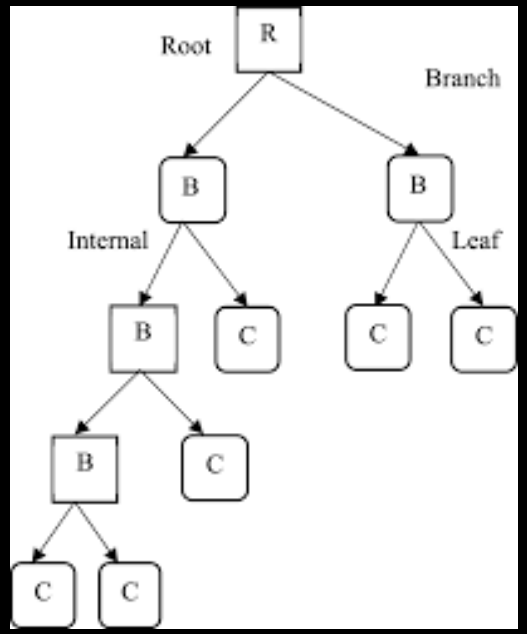
Primary Features

Health Opinions



The Modeling

- Classify an individual on whether or not they will receive the vaccine.
- The Decision Tree Method
- Primary Features:
  - Health Behaviors
  - Health Opinions



Model Evaluation

- Evaluate Model based on **Recall** score.
  - Recall = # True Positive / Total Number of Positives
  - Positive = Received Vaccine

Training Recall: 0.780

Testing Recall: 0.776

Average Cross Validation Scores: 0.780

Conclusions/Recommendations

- This Classification model performs well and can generalize data.
- Apply this model to new data that collects the same information.
- Example # of Vaccines to Produce = # Classified / Recall Score.
  - i.e) 100,000 respondents -> 65,000 classified -> 83,700 actual true cases -> can produce 16,000 fewer vaccines.

Limitations

- Model cannot consider everything…
- Factors beyond individual that can affect their access to vaccines.

The next steps

- Start collecting data on a sample group.
- Apply model to sample group and extrapolate to larger production scale.

Thank you