Capstone Project 2
Model Metrics

Machine Learning Algorithm: XGBoost

Sampling Method: Random Undersampling

Model Parameters:

- Number of trees: 100
- Learning rate: 0.3
- Max tree depth: 3
- L2 regularization constant: 10
- Subsampling: None

Model target: Heart Disease

Model features:

- Good Health
- Health Insurance
- Hypertension
- High Cholesterol
- Asthma Status
- Arthritis
- Race
- Age Category
- BMI
- Education Level
- Income Level
- Smoker Status
- Heavy Drinker
- Physical Activity
- Seatbelt
- Flu Shot
- Pneumonia Vaccine
- HIV
- Bronchitis
- Depression
- Kidney Disease
- Diabetes
- Sex
- Marital Status
- Sodium Intake
- Sample Weights

Cross-validation performance metrics:

	XGBoost, Random Undersampling
precision_pos_mean	0.2217
precision_pos_std	0.0034
precision_neg_mean	0.975
precision_neg_std	0.001
recall_pos_mean	0.8075
recall_pos_std	0.0102
recall_neg_mean	0.7253
recall_neg_std	0.0087
f1_pos_mean	0.3479
f1_pos_std	0.0033
f1_neg_mean	0.8318
f1_neg_std	0.0053
roc_auc_mean	0.8405
roc_auc_std	0.0033