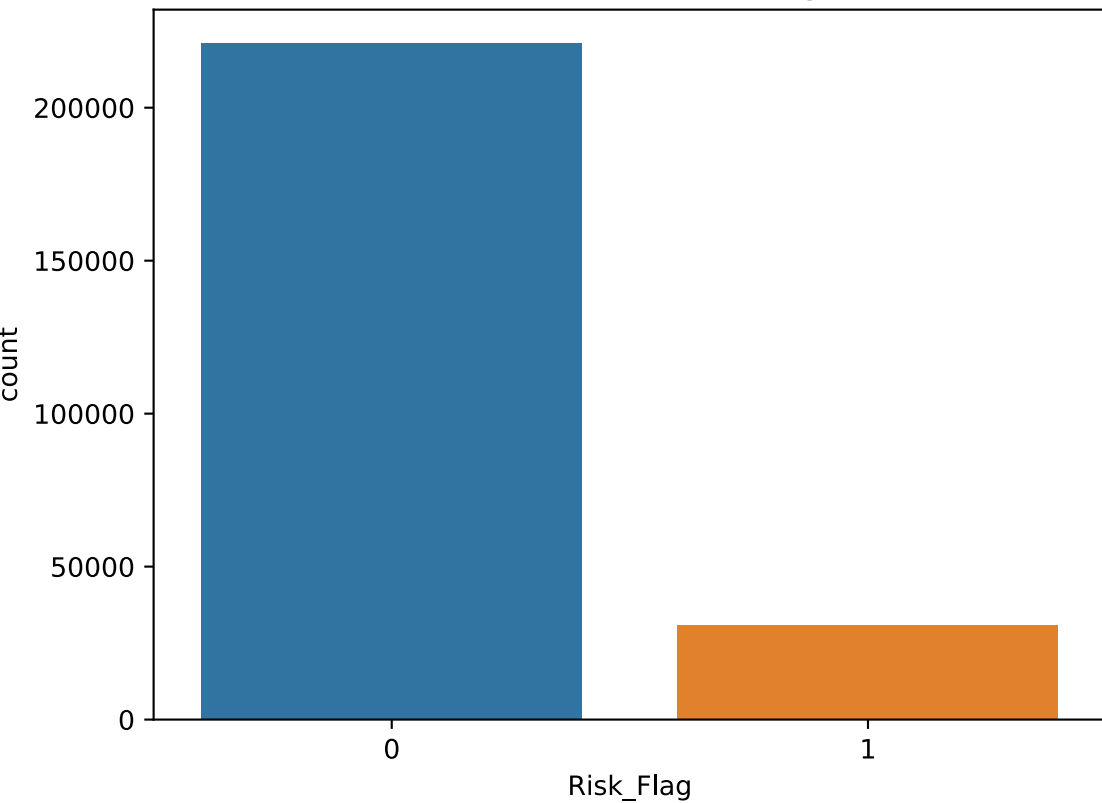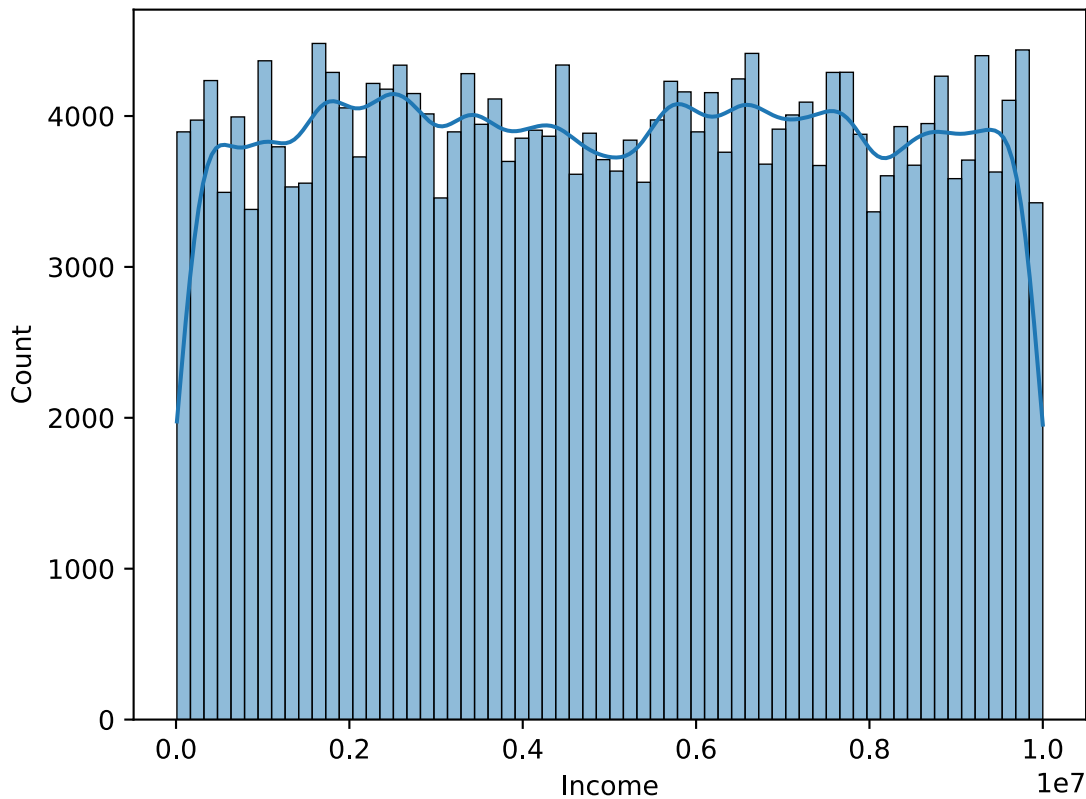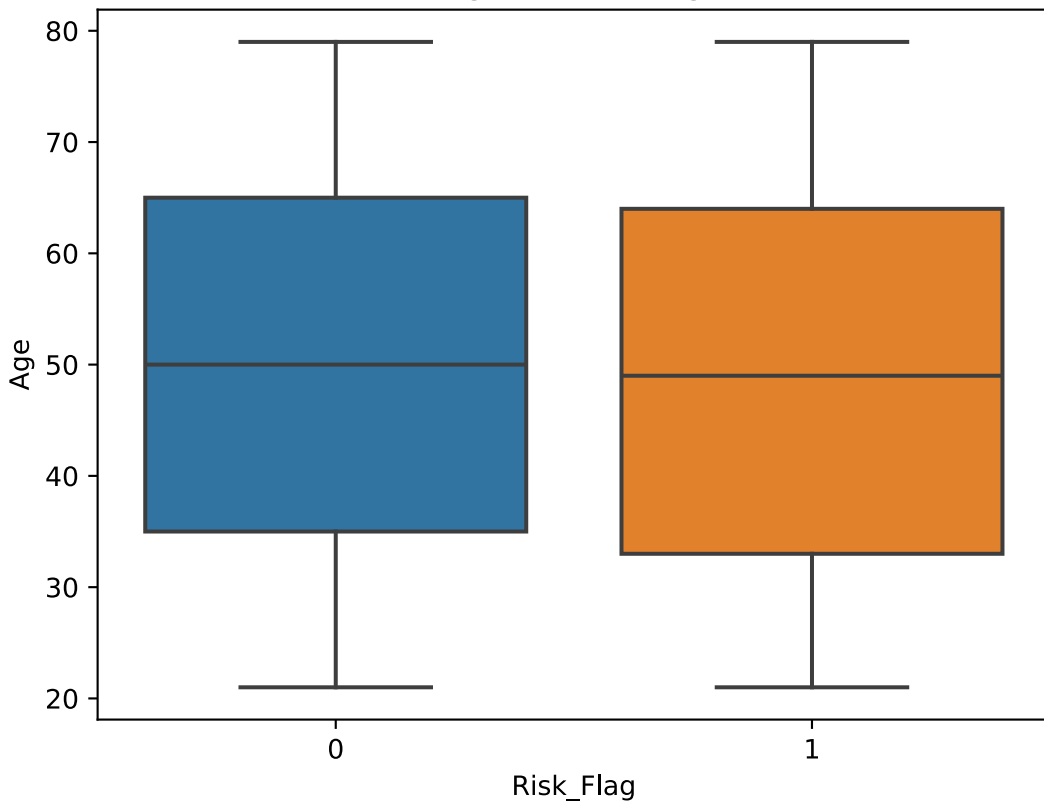Distribution of Risk Flag
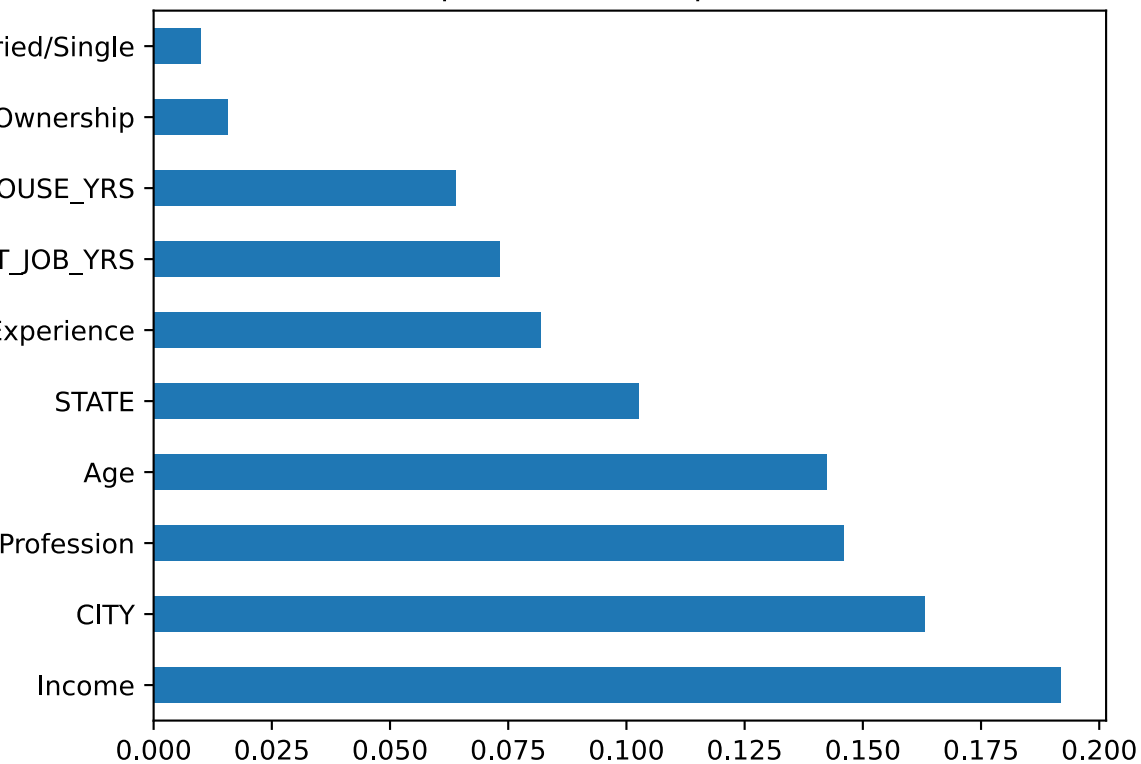
Income Distribution

Age vs Risk Flag

Top 10 Feature Importances

# Accuracy Score

Accuracy: 0.90

## Classification Report

|            | precision | recall | f1-score | support |
|------------|-----------|--------|----------|---------|
| 0          | 0.94      | 0.95   | 0.94     | 44147   |
| 1          | 0.60      | 0.54   | 0.57     | 6253    |
|            |           |        |          |         |
| accuracy   |           |        | 0.90     | 50400   |
| macro avg  | 0.77      | 0.74   | 0.75     | 50400   |
| weighted avg | 0.89    | 0.90   | 0.90     | 50400   |

Confusion Matrix
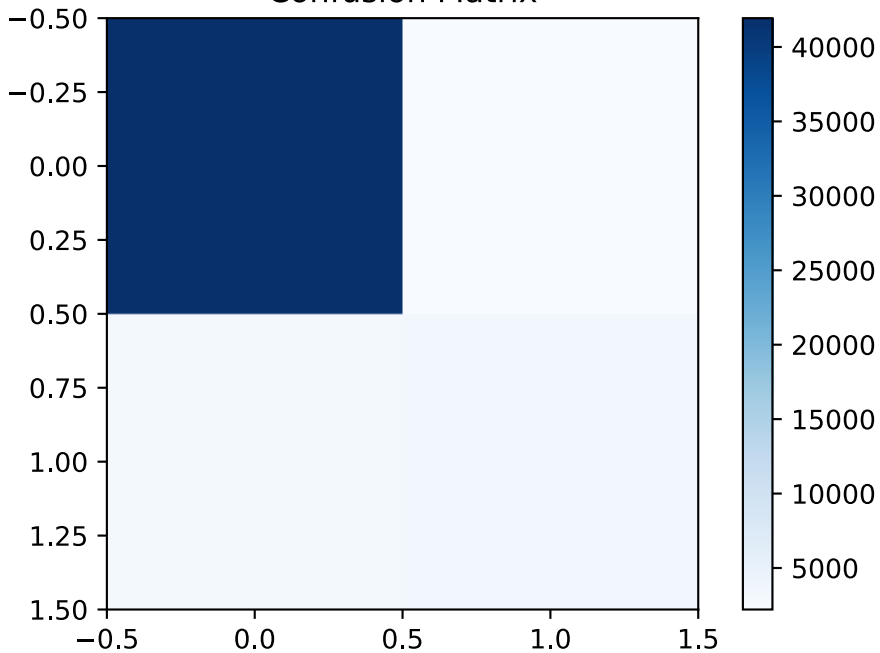
**Confusion Matrix Explained:**
- Model predicted 41936 true negatives, which indicates it is good at identifying the majority class.
- 2211 false positives suggest that the model sometimes incorrectly classifies instances as positive when they actually are negative.
- 2905 false negatives indicates instances that were wrongly classified as negative when they should have been positive.
- Total 3348 True positives shows correctly classified as positive.

**What Next? Reduce False Negative!**

**Why False Negative?**

In this use case if we get risk flag as True, we will have time to verify & financial loss can be saved. But in case of False Negative, we might ignore the risk as it has been predicted non-risky & suffer financial loss.

**How?**

- Adjust class weights to penalize misclassification of the minority class: class_weight='balanced'

- Adjust the decision threshold for classification

- Resample dataset