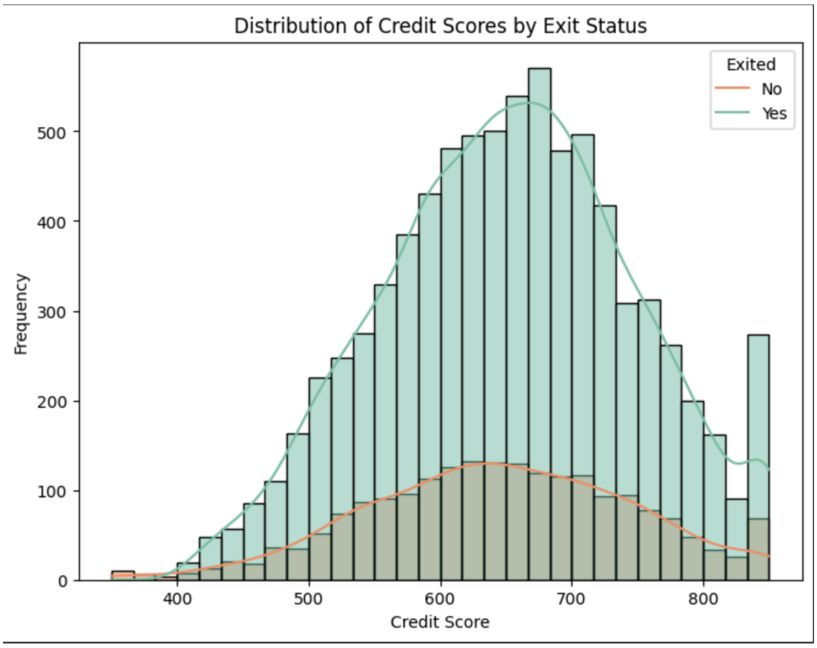
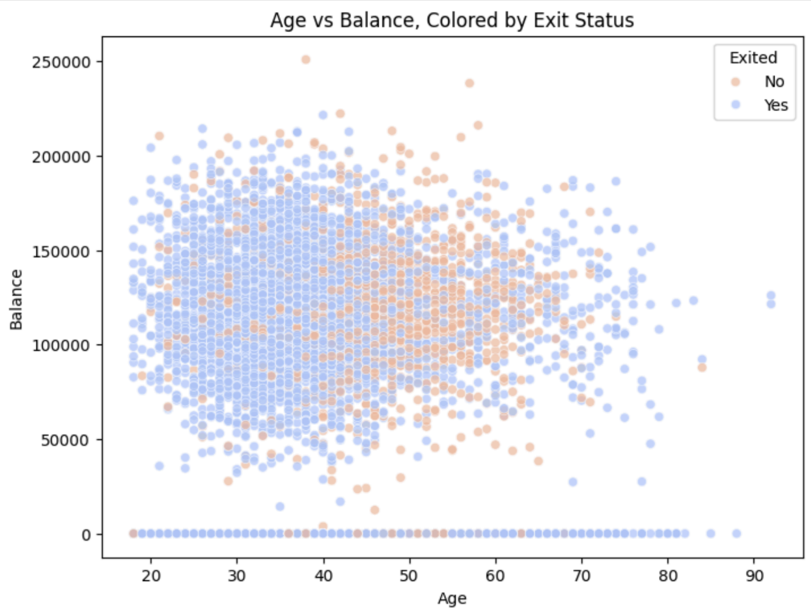
Three visualizations and insights

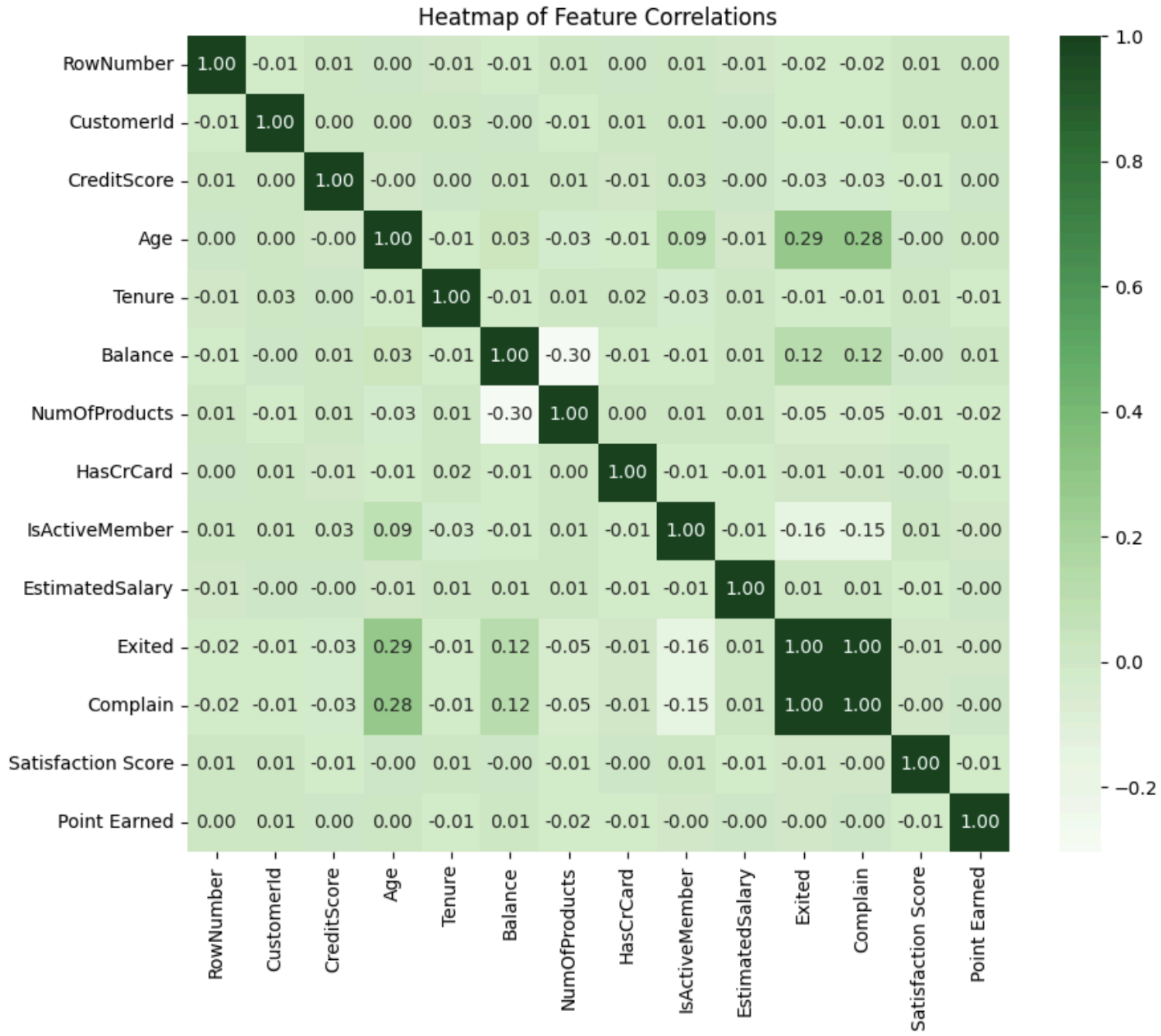
1. Distribution of Credit Scores by Exit Status:
Customers who have exited the bank tend to have lower credit scores, though there is significant overlap with customers who stayed.
The distribution for both groups peaked around similar values but shows different spreads.



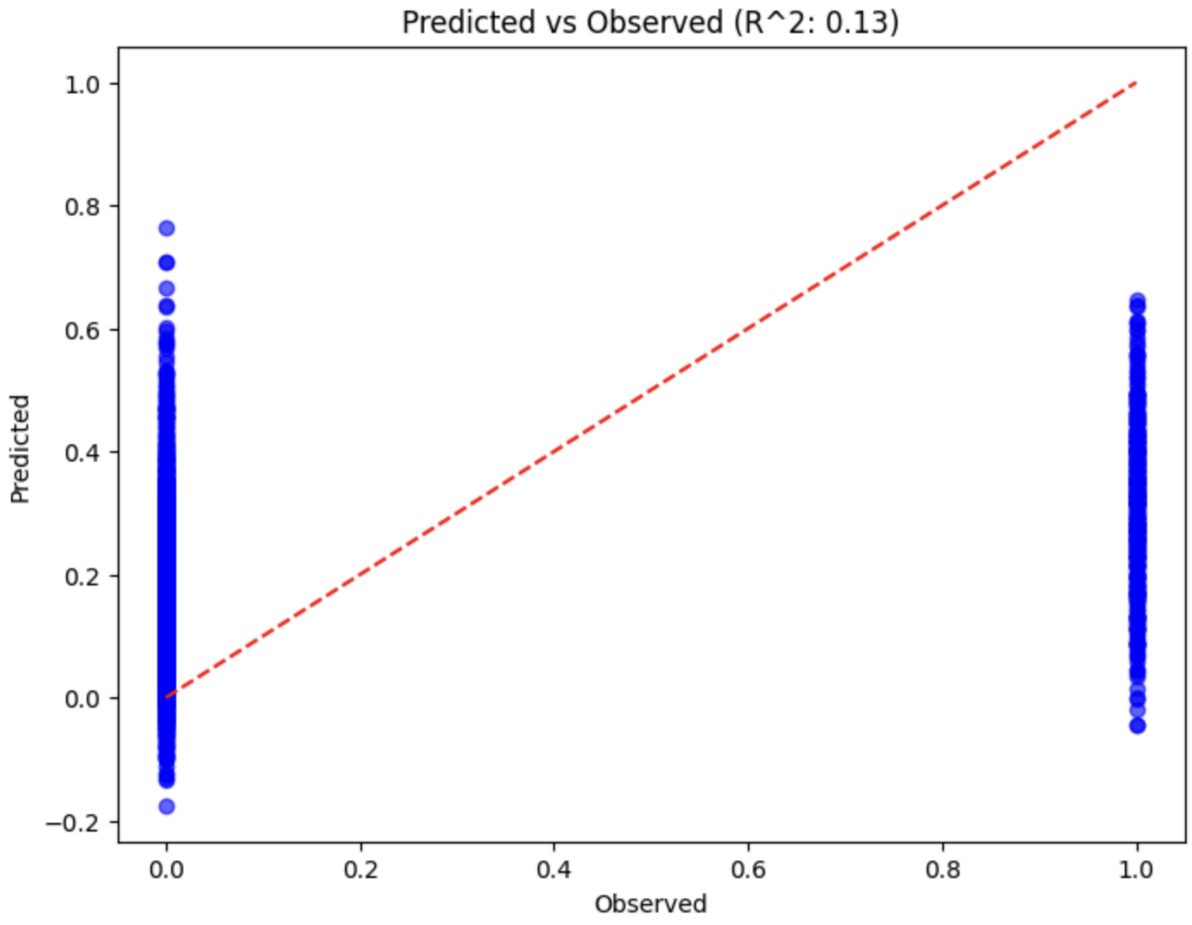
2. Age vs. Balance, Colored by Exit Status:
Exited customers are distributed across various ages and balances but are concentrated more among those with lower balances.
Non-exited customers appear across a wider range of balances and age groups.



3. Heatmap of feature correlations:
   Strong correlations can be found with tenure and satisfaction scores.
   Exited has a noticeable correlation with Balance and Satisfaction Scores.
   Weak correlations among other features suggest additional modeling work is needed to capture their patterns.



Plot of predicted vs. observed data points

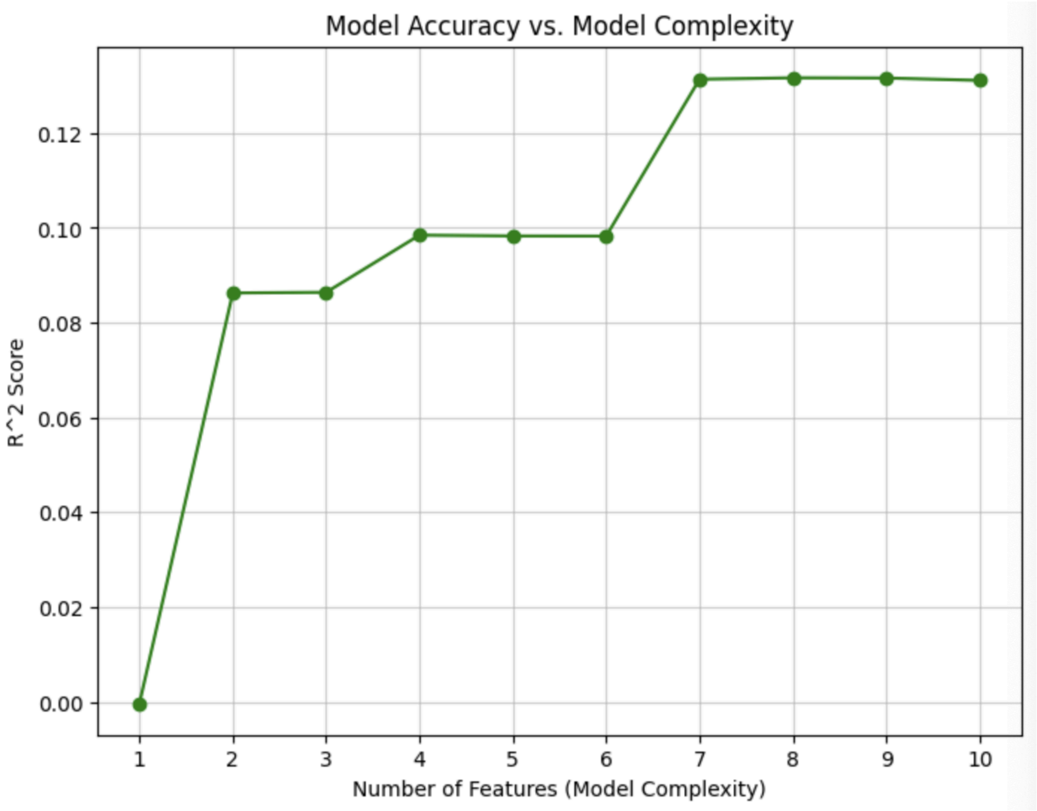


The linear regression model yielded an R^2 value of approximately 0.13.
This indicates that the model explains about 13% of the variance in the response variable (Exited) based on the selected features.
This also suggests that the relationship between the features and the response variable may be non-linear or the model needs more features to make a correct prediction.

Also, the Linear Regression model would not be a good fit for the dataset as the response variable (Exited) is binary, i.e. it has only two outputs (0 or 1) indicating whether the customer has left the bank or not. Linear Regression assumes a continuous response variable, making it unsuitable for binary classification tasks.

Classification models like Logistic regression would be a good fit for the dataset as it outputs probabilities bounded between 0 and 1, which can directly be interpreted.
Other models like Decision Trees or Random Forests can also be used here.

Plot of model accuracy over a range of model complexity



The accuracy vs model complexity curve is shaped this way because of the interplay between bias and variance.

The phases of the line graph would be:

1. Initial phase
   The model underfits the data as it lacks the necessary features to capture the underlying pattern in customer attrition.
   The model exhibits high bias as it oversimplifies the relationship.

2. Growth Phase
   Accuracy improves steadily as more relevant features are added.
   Here the model strikes a balance between bias and variance, improving generalization without overfitting.

3. Plateau Phase
   After a certain point, adding additional features introduces noise or redundancy rather than new predictive power.
   This increases the risk of overfitting, and that would lead to high variance, where the model would perform well on the training data but would perform poorly on the unseen data.

In this dataset, it happens because:

1. Feature Importance: Certain features like Satisfaction Score, Balance, and age) are highly predictive of customer attrition. Once these are included, additional features contribute marginally or add noise.
2. Linear Assumptions: Linear Regression assumes a linear relationship between features and the response variable. If the relationship is non-linear, linear regression cannot leverage additional features effectively.
3. Irrelevant features: Features like HasCrCard or NumOfProducts may have weak correlations with attrition. Including them doesn't improve the predictions but increases the complexity of the model.