Overview of the Analysis

The purpose of this analysis was to build and evaluate a supervised machine learning model to predict credit risk. Specifically, we used logistic regression to predict whether a loan is "healthy" (low risk of default) or "high risk" (likely to default) based on loan characteristics such as loan size, borrower's income, debt-to-income ratio, and other financial indicators.

Accurately identifying high-risk loans is critical for financial institutions to manage their portfolios and minimize default losses.

Results

- **Accuracy:** 99%
- **Precision (Healthy Loans - 0):** 1.00
- **Recall (Healthy Loans - 0):** 0.99
- **Precision (High-Risk Loans - 1):** 0.84
- **Recall (High-Risk Loans - 1):** 0.94

Summary

The logistic regression model performed **extremely well** overall, achieving a very high accuracy of **99%** on the testing data. It was **near-perfect** in identifying healthy loans, with precision and recall close to 1.00. For high-risk loans, the model achieved a **precision of 84%** and a **recall of 94%**, meaning it was able to correctly identify most high-risk loans while minimizing false positives.

Given the model's strong performance, particularly its ability to catch a high percentage of risky loans, I would recommend using this model to support credit risk decision-making. However, given the slightly lower precision for high-risk loans (84%), some additional model tuning (or more data) could further enhance its reliability before deploying it in a high-stakes production environment.