

BMS COLLEGE OF ENGINEERING
DEPARTMENT OF ARTIFICIAL INTELLIGENCE & DATA SCIENCE

CREDIT RISK ANALYSIS

Presented By: Pratham S (1BM23AD075)



CREDIT RISK ANALYSIS

- Credit risk analysis evaluates the likelihood that a borrower may default on a loan, helping financial institutions minimize losses and make informed lending decisions.
- In this project, historical loan applicant data is analyzed to predict loan default (loan_status), providing a systematic approach to assess creditworthiness.
- The analysis uses key features like income, employment history, credit score, debt-to-income ratio, and past defaults to quantify risk.
- Machine learning models such as LightGBM, Gradient Boosting, or deep learning models are employed to classify applicants as low-risk or high-risk, improving prediction accuracy.
- Accurate credit risk scoring reduces non-performing loans, optimizes interest rates, and enhances overall portfolio management for lenders.

PROBLEM STATEMENT

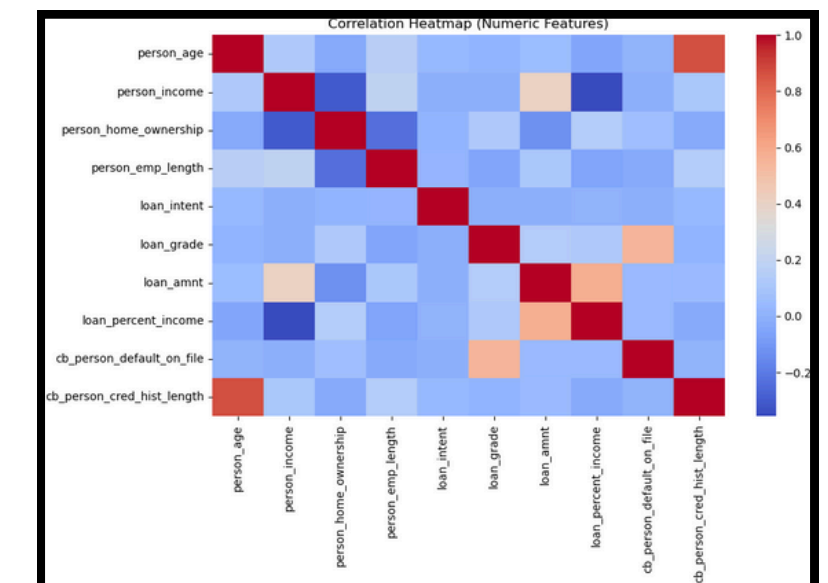
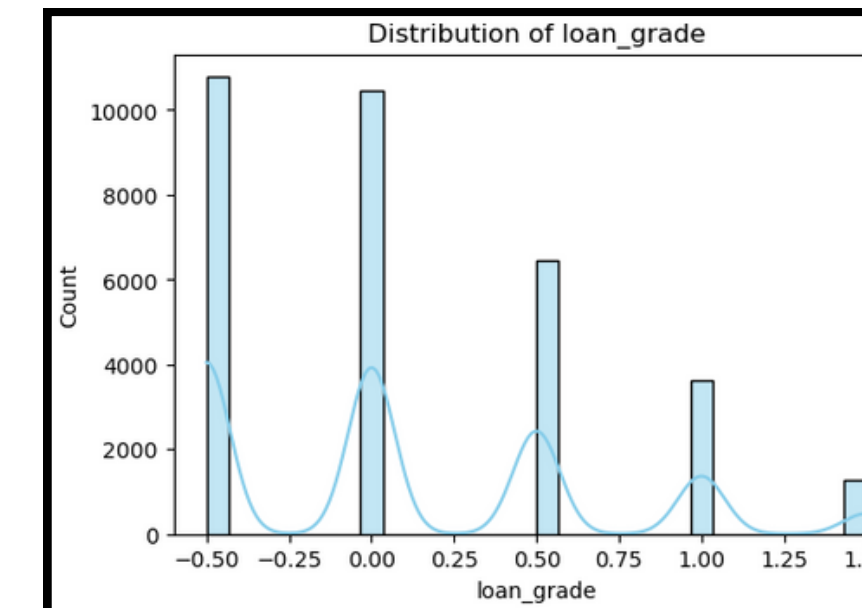
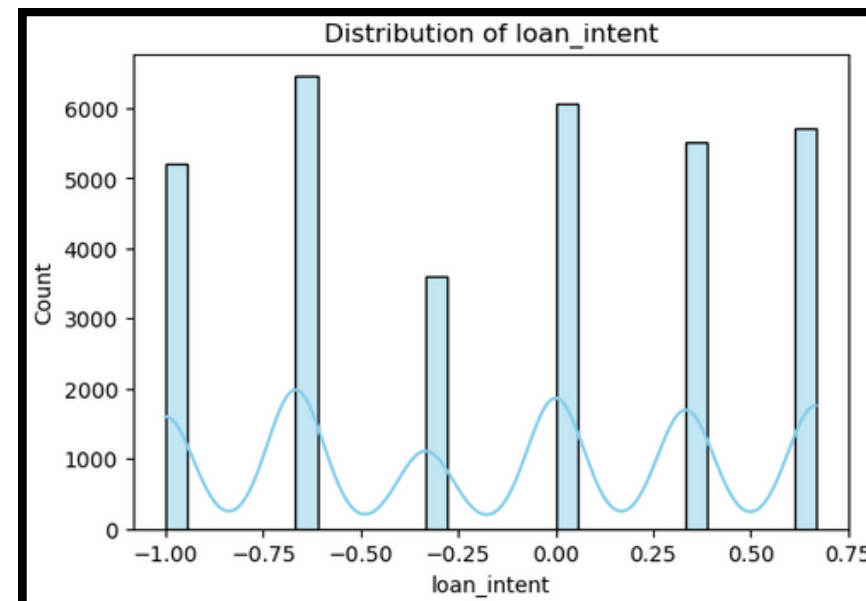
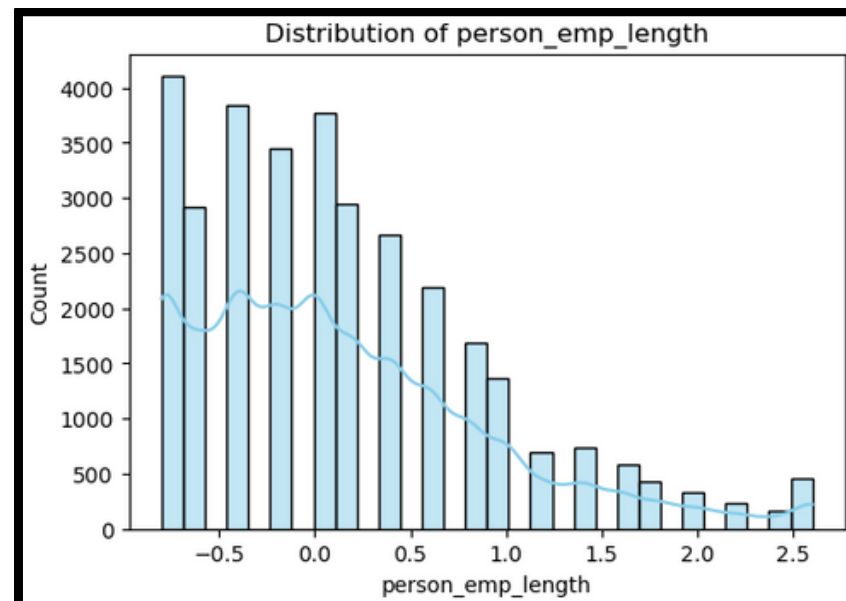
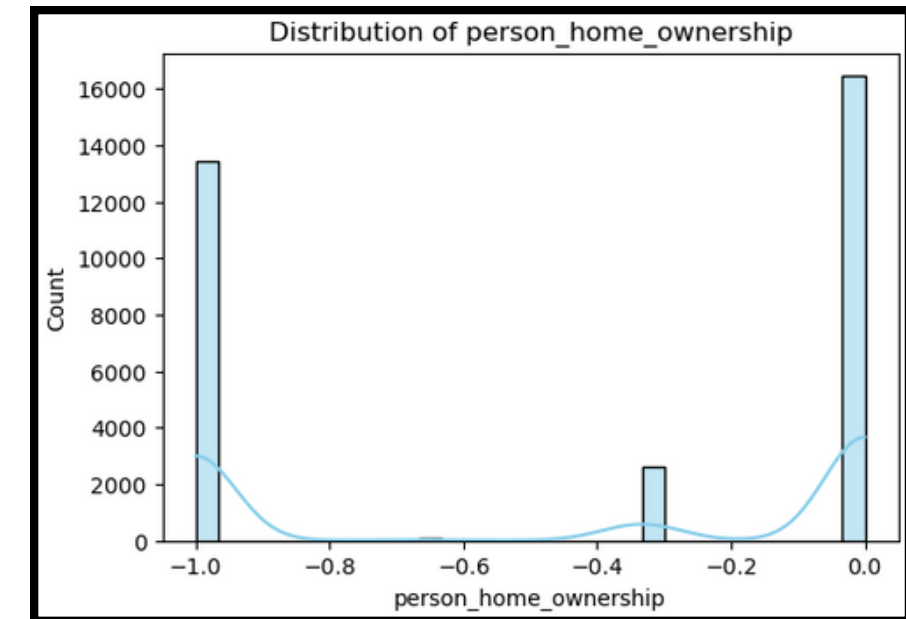
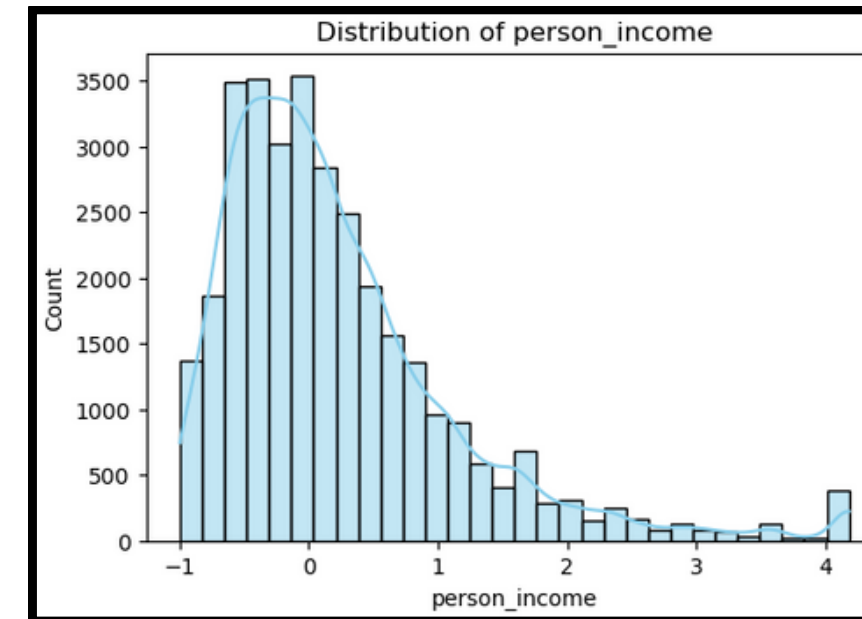
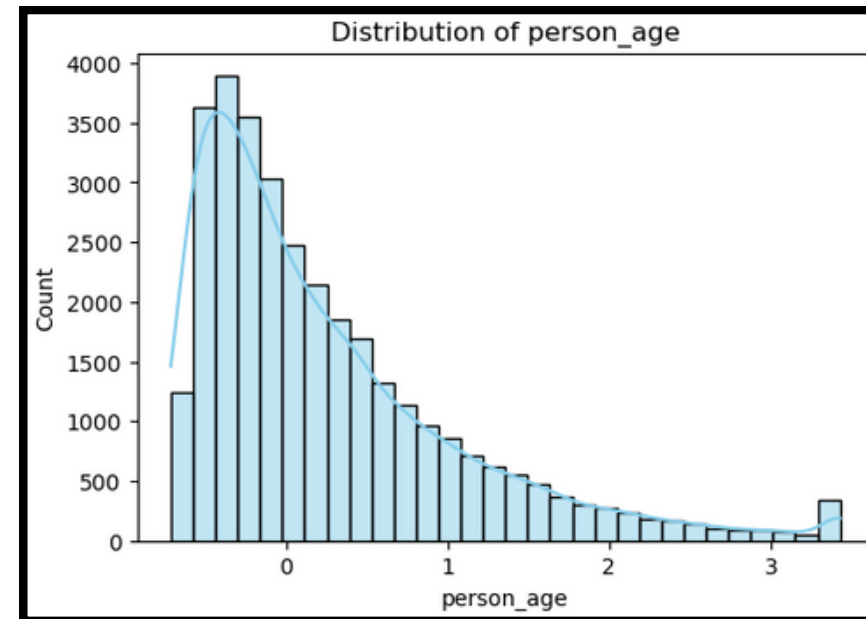
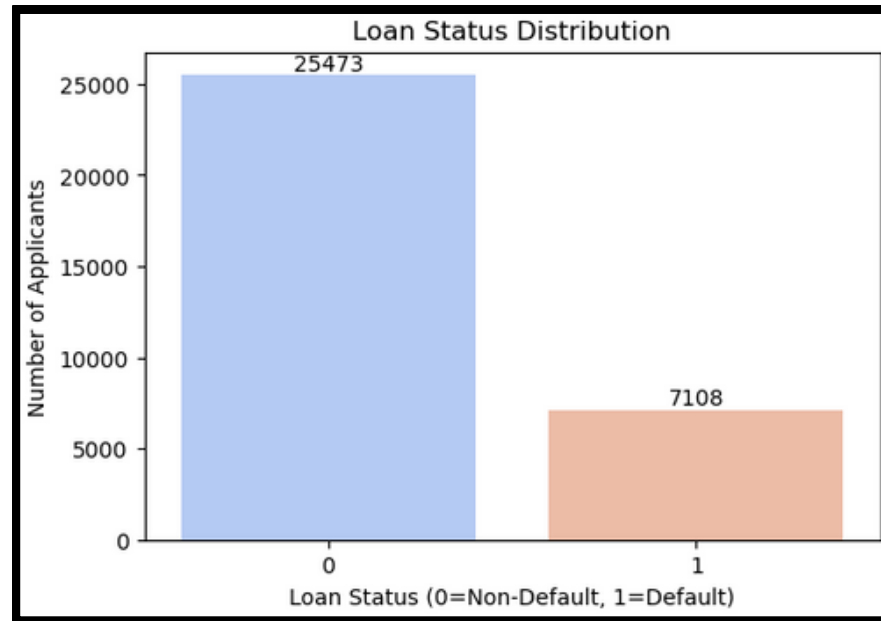
- Financial institutions face losses when borrowers fail to repay loans, making it essential to identify high-risk applicants.
- Traditional credit assessment methods are often manual, time-consuming, and prone to errors.
- There is a need to predict loan defaults accurately using historical applicant data to reduce non-performing loans.
- The challenge involves analyzing multiple factors like income, credit history, employment status, and debt-to-income ratio to evaluate creditworthiness.
- Developing a predictive model will help automate decision-making, improve lending efficiency, and minimize financial risk.

SPECIFIC TOOLS

- Python – Main programming language for data processing, modeling, and backend development.
- Jupyter Notebook – Environment to write, test, and visualize code interactively.
- Pandas & NumPy – For data manipulation, cleaning, and numerical computations.
- Scikit-learn & LightGBM – For building machine learning models and evaluating performance.
- Matplotlib & Seaborn – For data visualization and exploratory data analysis (EDA).
- Joblib – To save trained models for future predictions.
- Gradio– To create a simple frontend interface for model interaction.
- Excel / CSV – For storing and handling input datasets.
- TensorFlow: It is an open-source deep learning framework used to build, train, and deploy neural network models for tasks like credit risk prediction, enabling efficient handling of large datasets and complex model architectures.

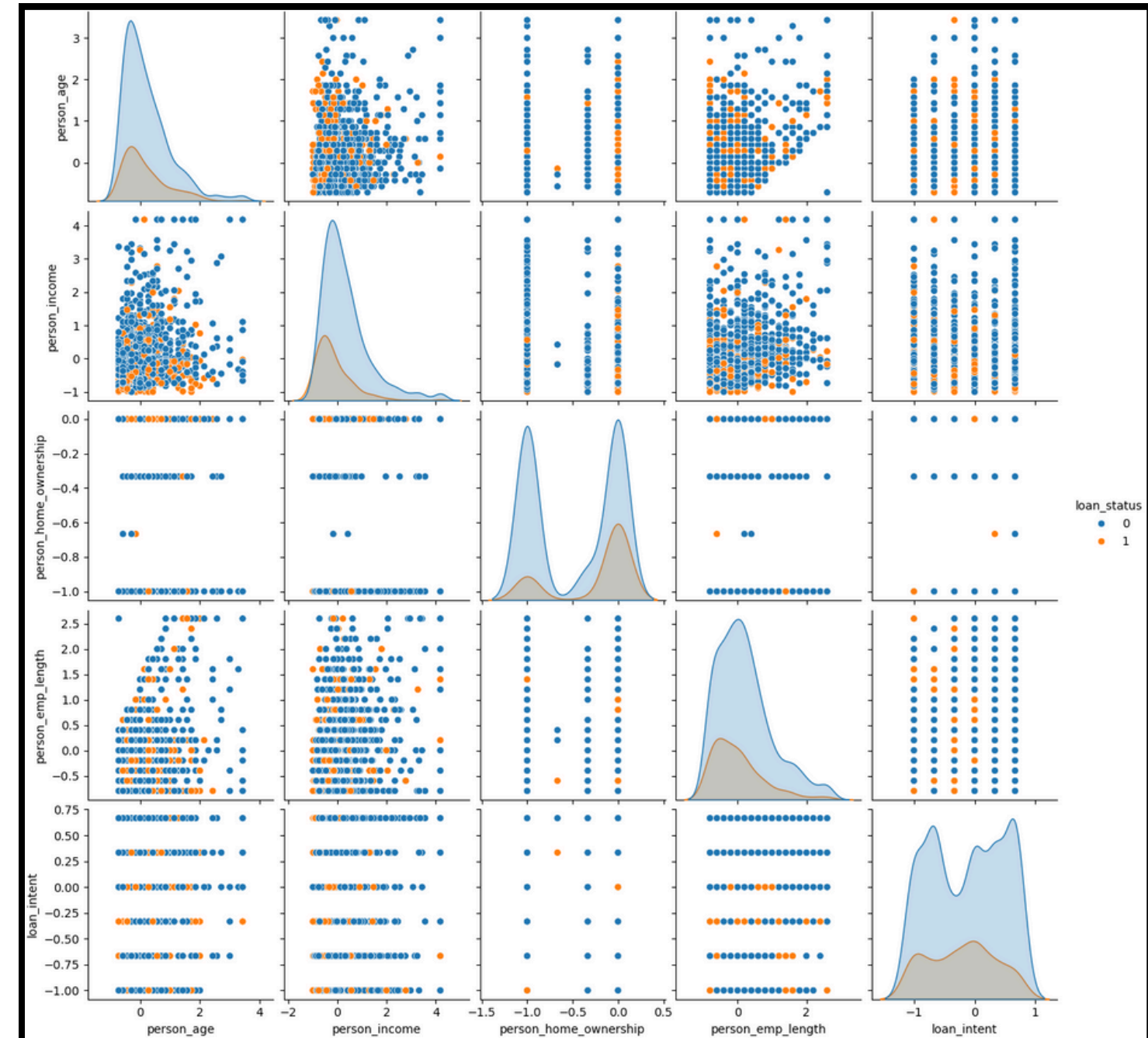
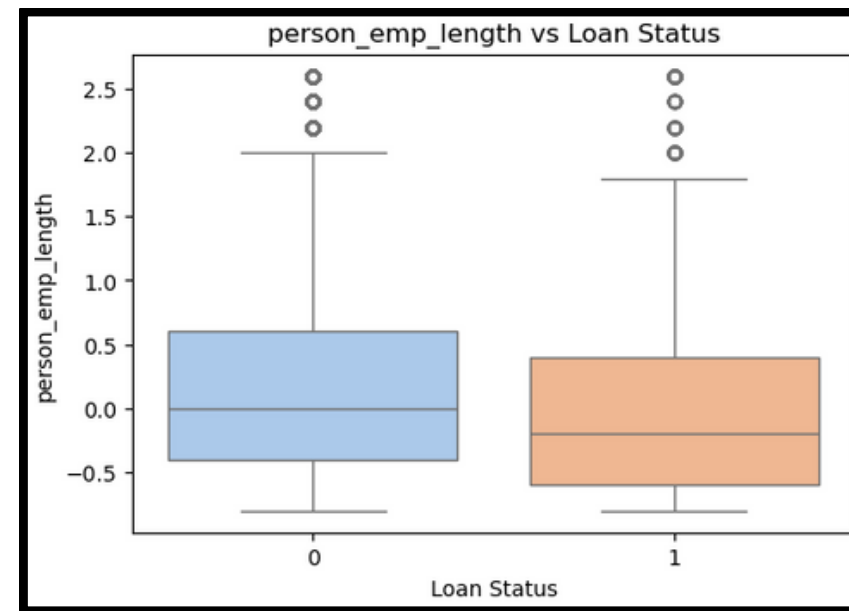
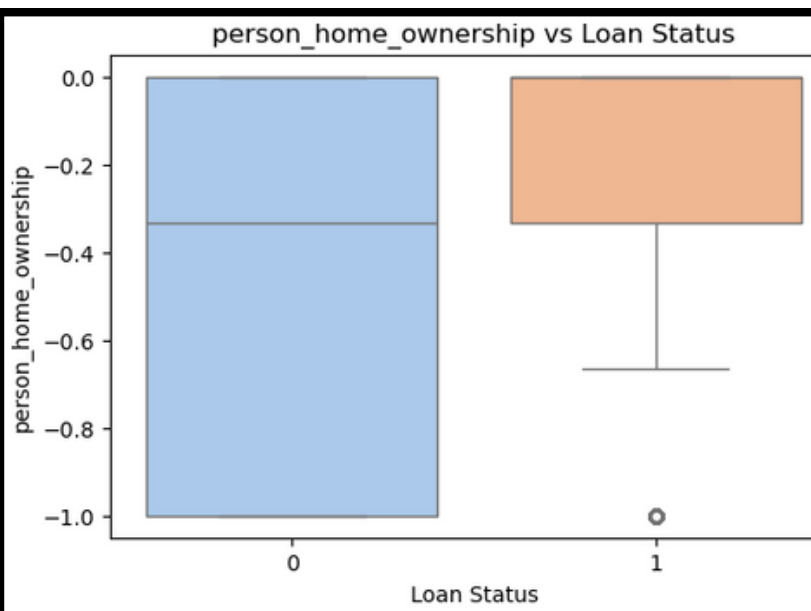
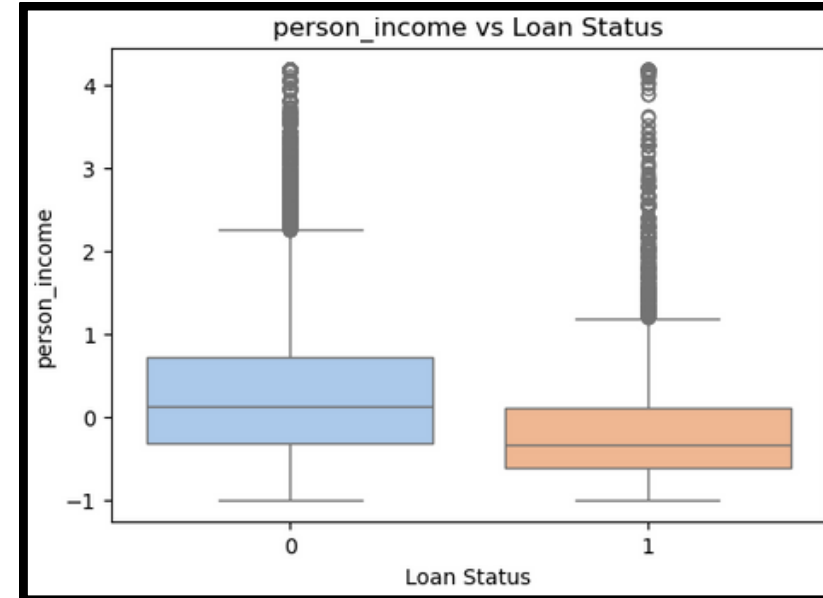
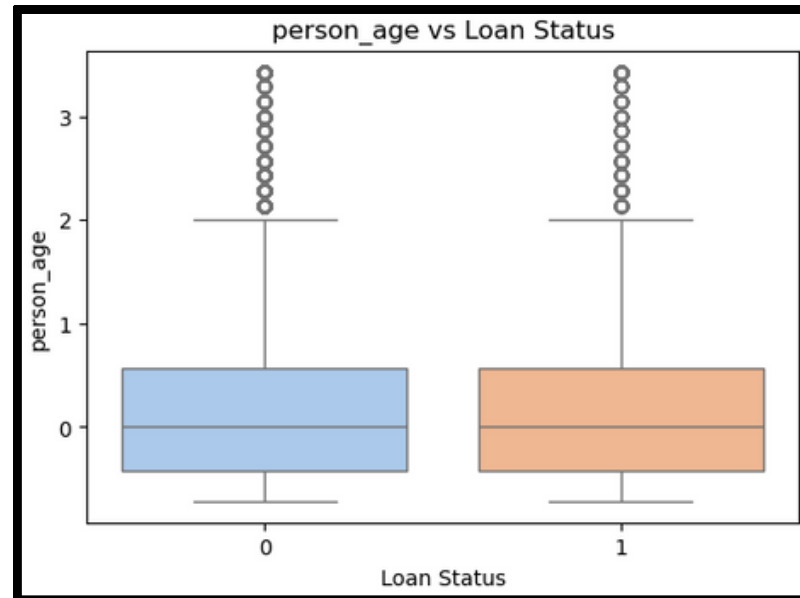
PROJECT DEMONSTRATIONS & VISUALIZATION

- STAGE 1 Exploratory Data Analysis (EDA)



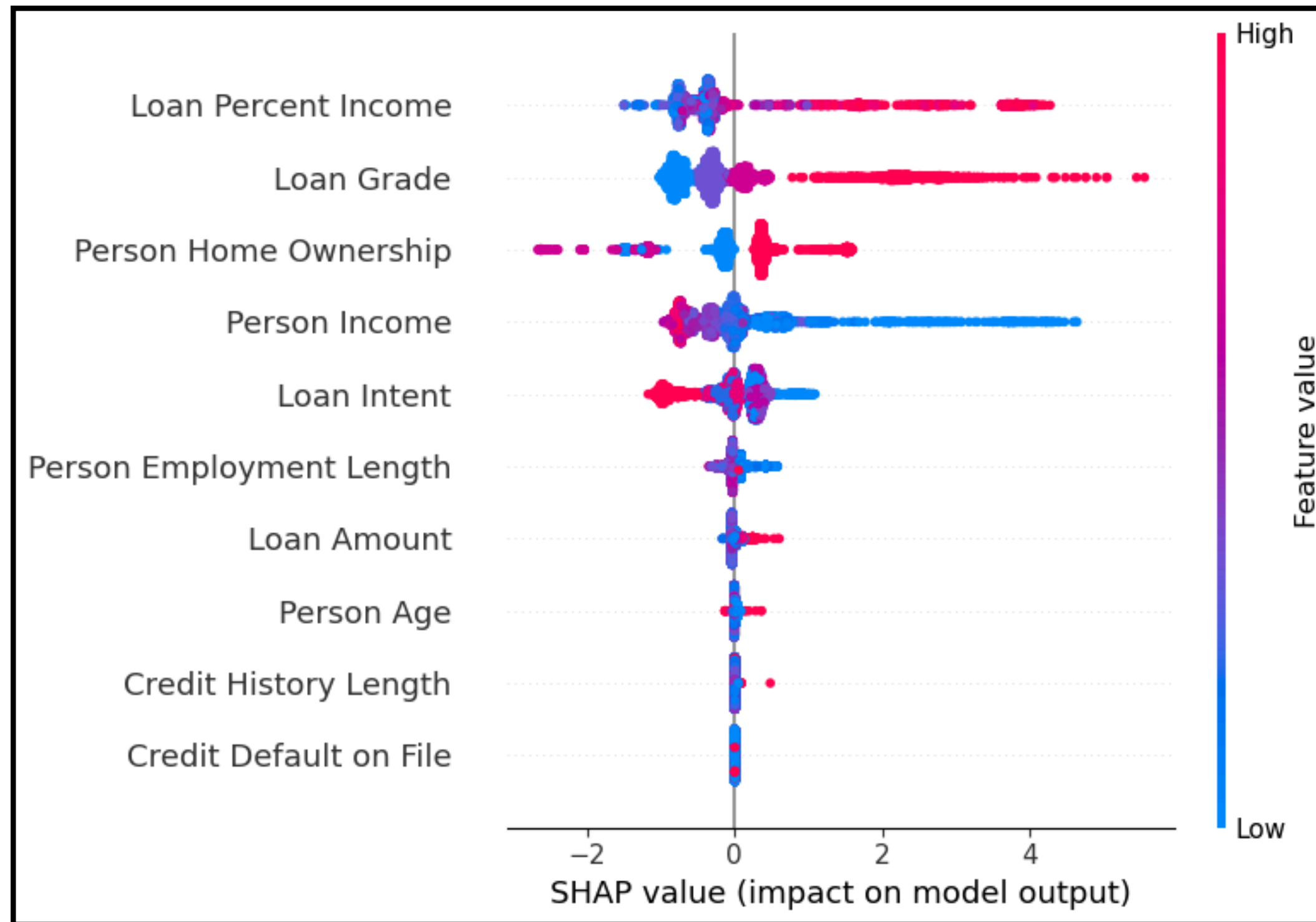
PROJECT DEMONSTRATIONS & VISUALIZATION

- STAGE 1 Exploratory Data Analysis (EDA)



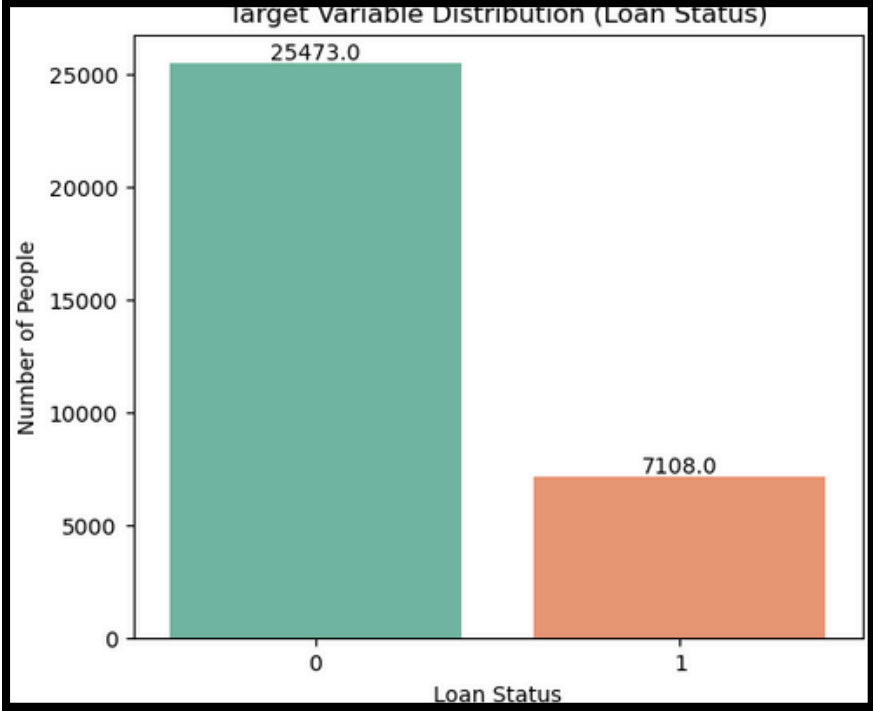
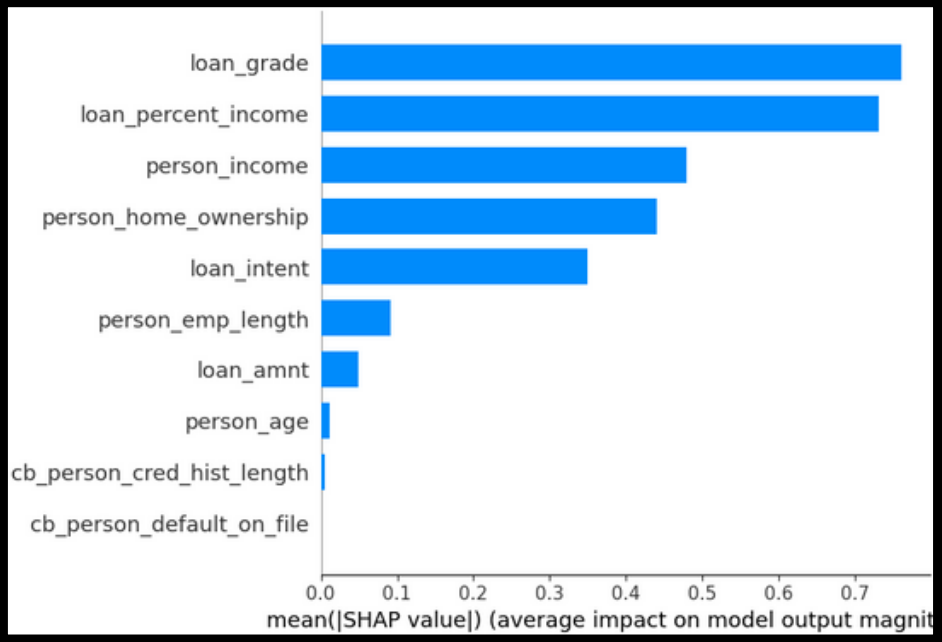
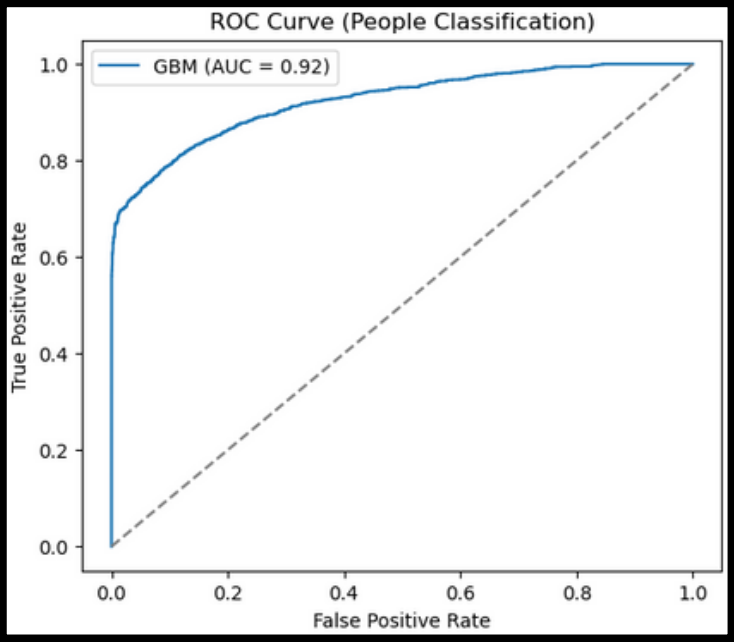
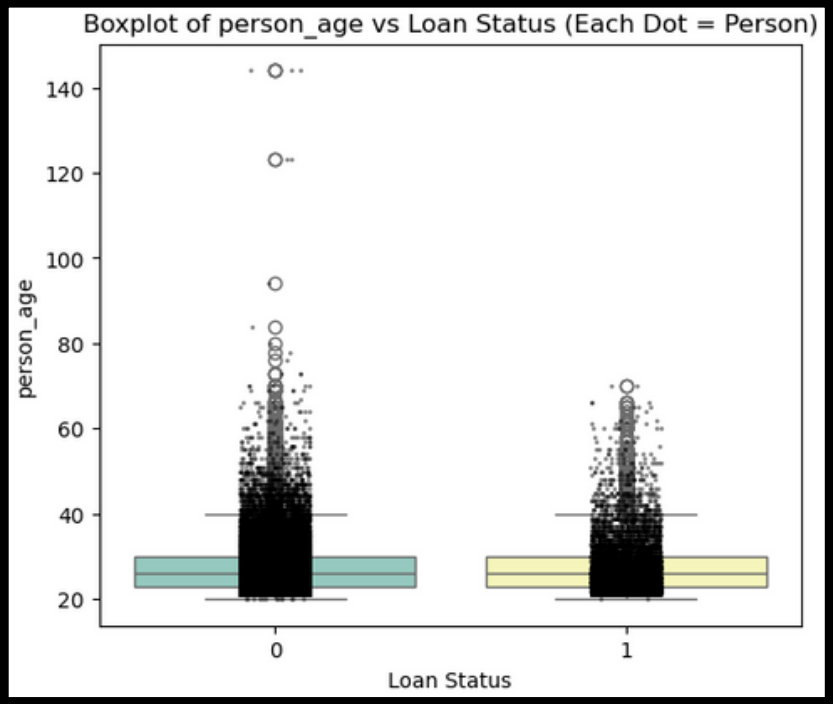
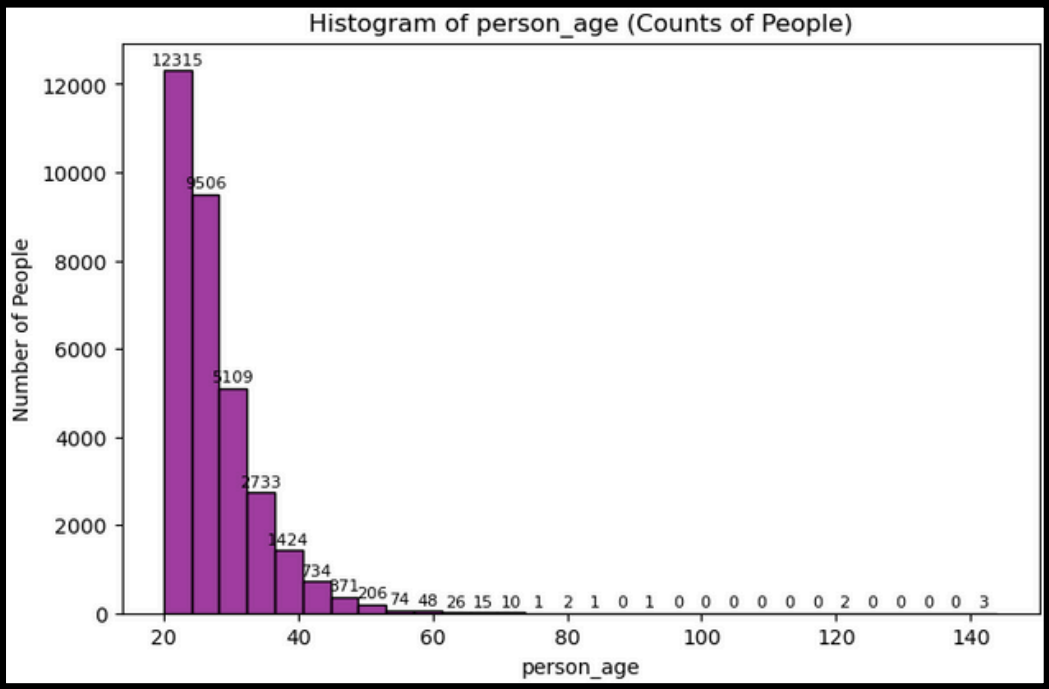
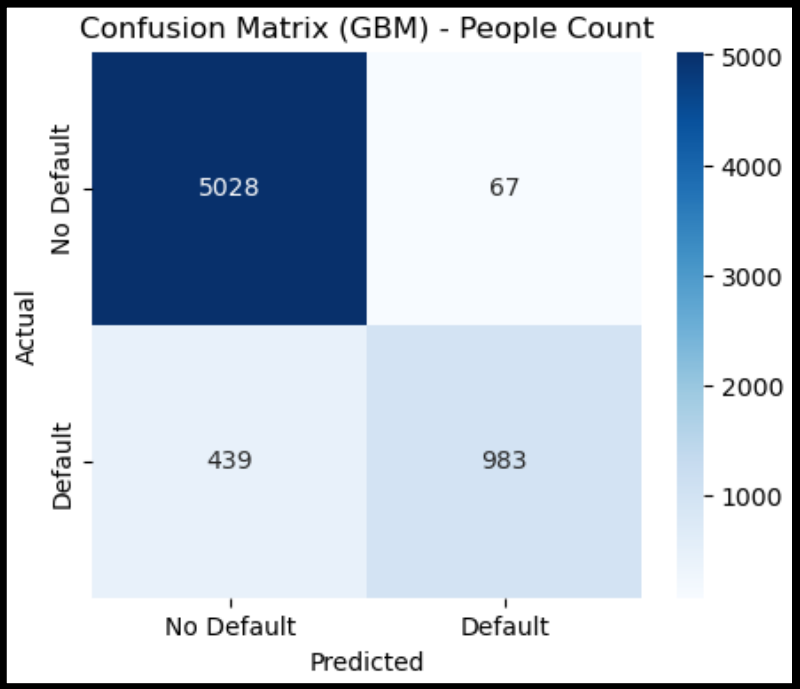
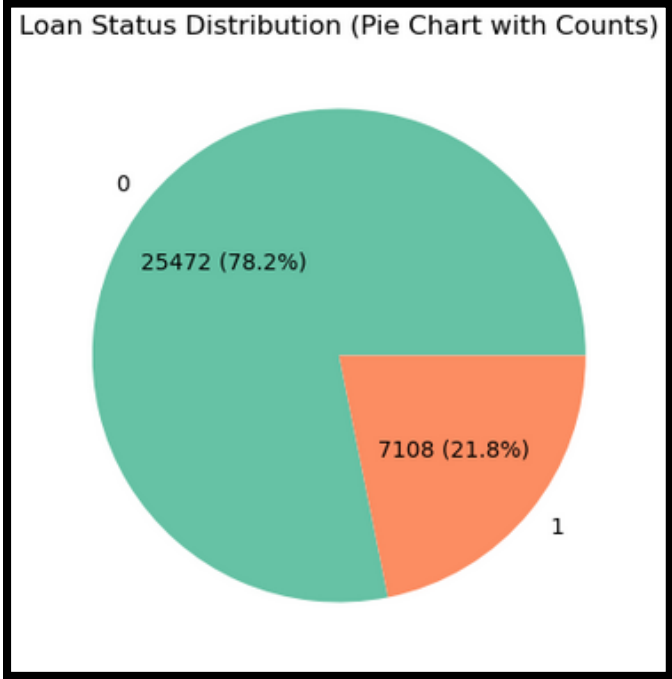
PROJECT DEMONSTRATIONS & VISUALIZATION

- STAGE 2 MODEL



PROJECT DEMONSTRATIONS & VISUALIZATION

- STAGE 3 VISUALIZATION



CONCLUSION

The Credit Risk Analysis project successfully demonstrates how historical loan data can be leveraged to predict the likelihood of borrower default. By applying advanced machine learning and deep learning techniques, the project provides an automated and accurate method for assessing creditworthiness. The models developed, including LightGBM and neural networks, effectively identify high-risk applicants, thereby reducing potential financial losses for lending institutions. Data-driven insights from the analysis help optimize lending decisions, improve portfolio management, and enhance overall risk mitigation strategies. Overall, this project highlights the critical role of predictive analytics in modern banking, ensuring more informed, efficient, and reliable credit decision-making.