Executive Summary

1. Objective: This project aims to build a robust and fair machine learning model that predicts whether an auto loan applicant is likely to default (bad loan), using historical lending data. Our primary goals were:
   - Maximize predictive performance on unseen evaluation data.
   - Ensure fairness across gender and race groups.
   - Preserve model interpretability for deployment and auditability.
2. Data and Data Cleaning + Preprocessing:
   - Data: Training dataset ∼ 21,000 records, and testing dataset ∼ 5,400 records.
   - Target Variable: bad_flag, where 1 = bad loan and 0 = good loan.
   - Preprocessing Highlights:
     (a) Removed columns with > 80% missing values.
     (b) Dropped sensitive attributes (Gender, Race and aprv_flag) during training to avoid leakage.
     (c) Handled outliers via quantile capping and scaling via RobustScaler.
     (d) Removed correlated features (corr > 0.7) to reduce redundancy.
3. Models Evaluated:

Table 1: Model Evaluation Summary on Test Set (Bad = 1)

Model	**CV F1 Score**	**Test Precision (Bad)**	**Test Recall (Bad)**
Logistic Regression	0.19	0.11	0.76
Random Forest	0.18	0.11	0.69
Balanced Random Forest	0.16	0.09	0.86
XGBoost	0.18	0.10	0.76
Stacking (XGB + RF)	0.20	0.11	0.75

4. Chosen Model: Stacking Model for final deployment due to good performance on unseen data, good recall for bad loans and interpretability via SHAP. Moreover, the model showed very good fairness among all genders and races (That is why I didn't choose Balanced Random Forest that has better recall).
5. Model Interpretability:
   (a) SHAP (SHaplet Additive exPlanations) (using XGBoost from the base learner of the stacking model) was used to explain predictions.
   (b) Key features driving predicitions included: ltv_1req, fico...
   (c) Visualization showed meaningful feature impact, helping support decision transparency.
6. Fairness Analysis: We reintroduced sensitive attributes (Gender and Race) post-training to analyze fairness on test data.

Table 2: Equal Opportunity Difference (TPR Gap) by Group

Group	**EOD**
Gender	0.0266
Race	0.0776

   Disparity across race was within acceptable range but may need further fairness optimization in production.
7. Sucesses:
   (a) Achieved 75% recall for bad loans on test set using stacking model, capturing high loan effectively.
   (b) Developed modular and reproducible pipelines for data loading, preprocessing, model training, and evaluation.
   (c) Ensured model fairness and interpretability, both essential for financial decision-making.

8. Challenges:
   (a) Severe class imbalance (good loans are way more than bad loans).
   (b) After applying SMOTE, the results didn't change significantly (precision for bad loans was still low).