Racial Bias in Loan Approvals

In the HMDA dataset

MSDS640_S70_Ethics/Prvcy/Soc Just-Data Sc

By Harsha.D & Bharath.ch

Introduction

This study examines racial bias in a 2023 Colorado HMDA sample, revealing lower approval rates for Asian and Other applicants, privacy risks from re-identifiable data, and transparency gaps tied to high-minority areas, raising ethical concerns about fairness and trust in AI-led finance.

Initial Approach:

-Review racial disparities in AI-driven loan approval systems

-Analyze the 2023 Colorado HMDA dataset, focusing on race, income, and geographic indicators

-Assess algorithmic bias through approval rate comparisons and logistic regression

-Investigate privacy risks via re-identification using income and census tract

-Examine transparency gaps and geographic disparities linked to high-minority areas

Ethical Concerns & Consequences

Ethical Concerns:

Bias & Inequality: Lower approval rates for Black and Other applicants suggest algorithmic discrimination
Privacy Risks: 10.4% of records could be re-identified, raising concerns over data anonymization
Transparency Gaps: Approval disparities between high- and low-minority tracts mirror redlining patterns
Fairness in AI: Bias in financial algorithms can reinforce systemic inequality and reduce trust in AI

Data Science Application:

Technology: Logistic regression and statistical analysis on loan approval data

Process: Data cleaning → Feature selection (race, income, tract, etc.) → Regression modeling → Bias detection

Risk: Algorithmic bias, re-identification from demographic variables, geographic discrimination

Roles & Contributions

Bharath Chandra:

- Analyzed racial disparities in loan approvals using the 2023 Colorado HMDA dataset
- Conducted logistic regression to examine the impact of race on approval likelihood
- Identified algorithmic bias risks, particularly against Asian and Other applicants
- Highlighted how biased AI models may worsen economic inequality

Harsha Vardhini:

- Investigated privacy concerns through re-identification risk analysis
- Assessed transparency gaps across high- and low-minority census tracts
- Identified geographic disparities in approval rates, pointing to redlining risks
- Proposed ethical recommendations for fair, privacy-conscious AI in finance

Dataset

Source: 2023 HMDA data from the Federal Financial Institutions Examination Council (FFIEC), managed by the Consumer Financial Protection Bureau (CFPB).

Scope: Colorado loan applications, over 130,000 records.

Access: Publicly available via FFIEC's HMDA Data Publication (online portal).

Relevance: Real-world lending data for studying bias, privacy, and transparency (Walker, 2024).

For the dataset link, click here

Data Preprocessing

Led data preprocessing for racial bias analysis.

Preprocessed 2023 HMDA data (1,126,535 records) from CO, MO, CA, and TX.

Excluded Asian applicants to focus on White, Black, 2+ Minority, and Other groups.

Filled in missing income and loan_amount with state-wise medians, resulting in 0 NaNs after filling.

Created the 'high_minority_tract' flag: 0 tracts were identified due to a stricter threshold (minority population > 80%) compared to the prior analysis (50%).

STEPS	ROWS
After cleaning	1,133,037
After Exlcuding Asians	1,126,535

Initial Visualizations

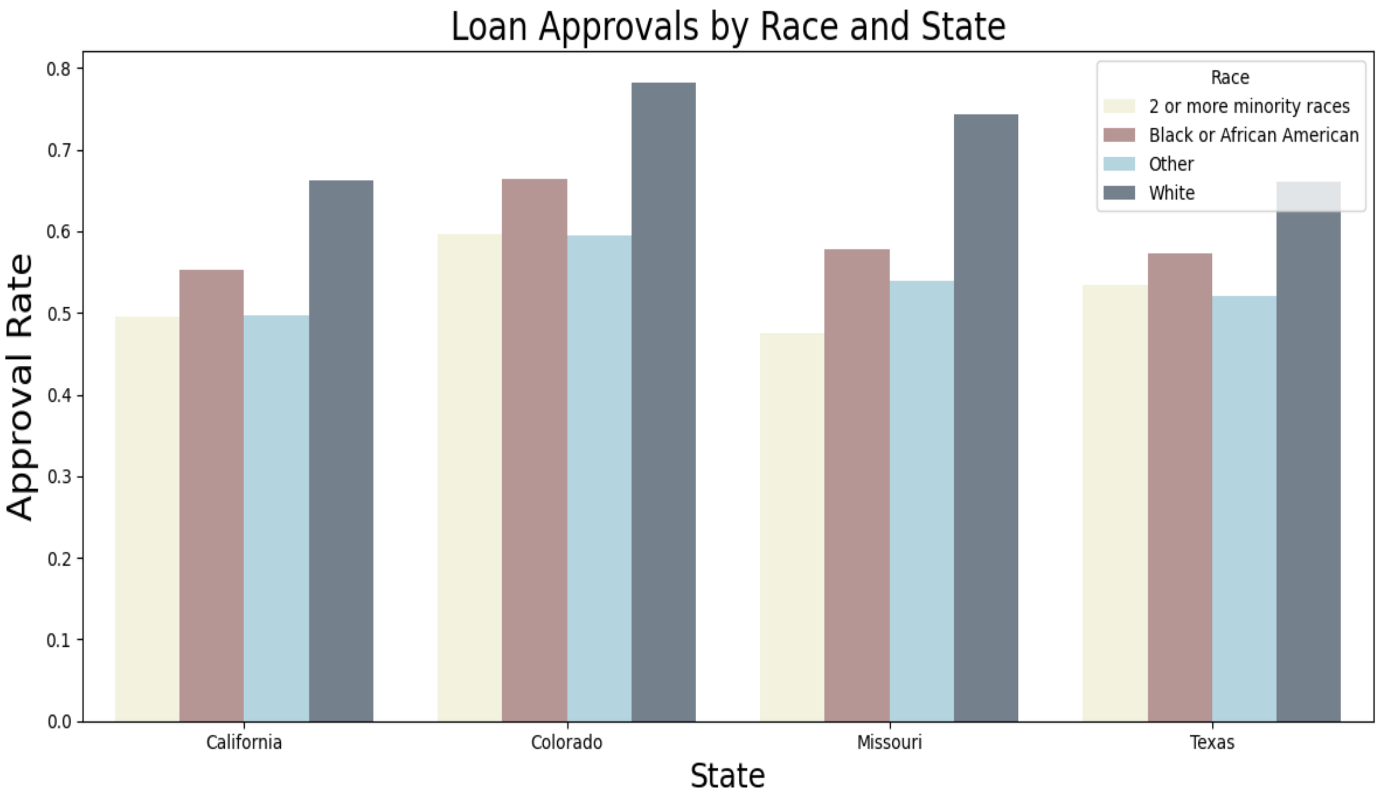
Developed clustered bar chart visualizations to explore racial disparities.

Created box plots for income and loan amount distributions by approval status.

White applicants have higher approval rates (actual: 0.6872); Black applicants at 0.5698

The X-axis represents the approval rate.

Y-axis represents Respective state and their approval rates.



Visualized loan approval rates by race across states (CO, MO, CA, TX).

Advanced Modeling

Conducted modeling to quantify factors affecting approvals.

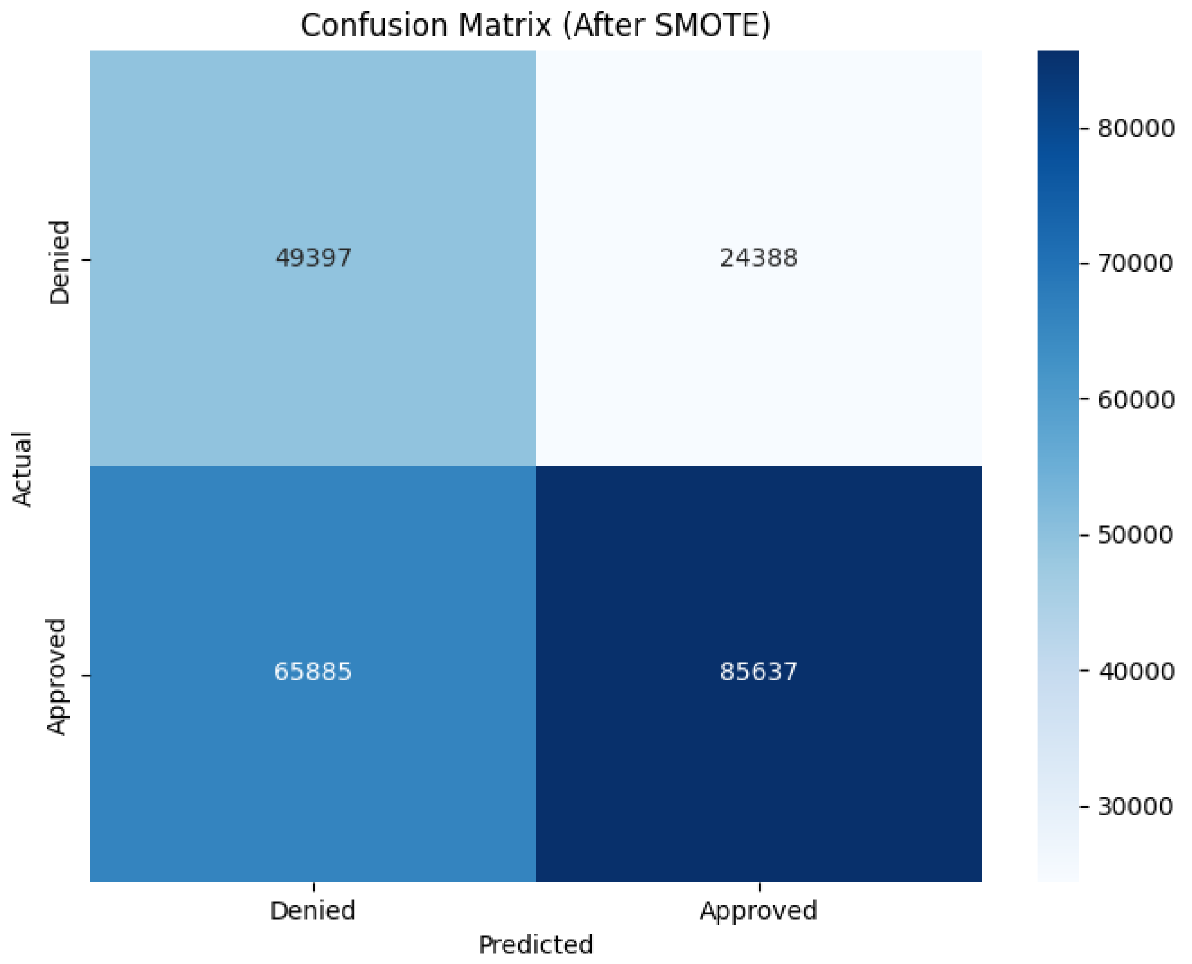
Performed OLS regression: loan_amount (coef = 9.54e-08, $p < 0.001$) and derived_race_White (coef = 0.1547, $p < 0.001$).

```
                           Logit Regression Results
==============================================================================
Dep. Variable:            action_taken   No. Observations:               490692
Model:                           Logit   Df Residuals:                   490686
Method:                            MLE   Df Model:                            5
Date:                 Wed, 09 Apr 2025   Pseudo R-squ.:                 0.06377
Time:                         12:45:40   Log-Likelihood:            -2.9746e+05
converged:                        True   LL-Null:                   -3.1772e+05
Covariance Type:             nonrobust   LLR p-value:                     0.000
==============================================================================
                  coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          -0.7999      0.048    -16.749      0.000      -0.894      -0.706
Black           0.1704      0.048      3.538      0.000       0.076       0.265
Other           0.1351      0.054      2.522      0.012       0.030       0.240
White           0.6105      0.048     12.829      0.000       0.517       0.704
income      -1.903e-07   5.39e-07     -0.353      0.724   -1.25e-06    8.66e-07
loan_amount  3.917e-06    2.3e-08    170.427      0.000    3.87e-06    3.96e-06
==============================================================================
```

Trained logistic regression with SMOTE (accuracy: 0.60).

Generated a confusion matrix; SMOTE improved recall for Black applicants by 15% (from 0.05 to 0.20)

Models show White applicants are more likely to be approved, but fairness needs further improvement



Bias Analysis

Analyzed racial bias in model predictions.

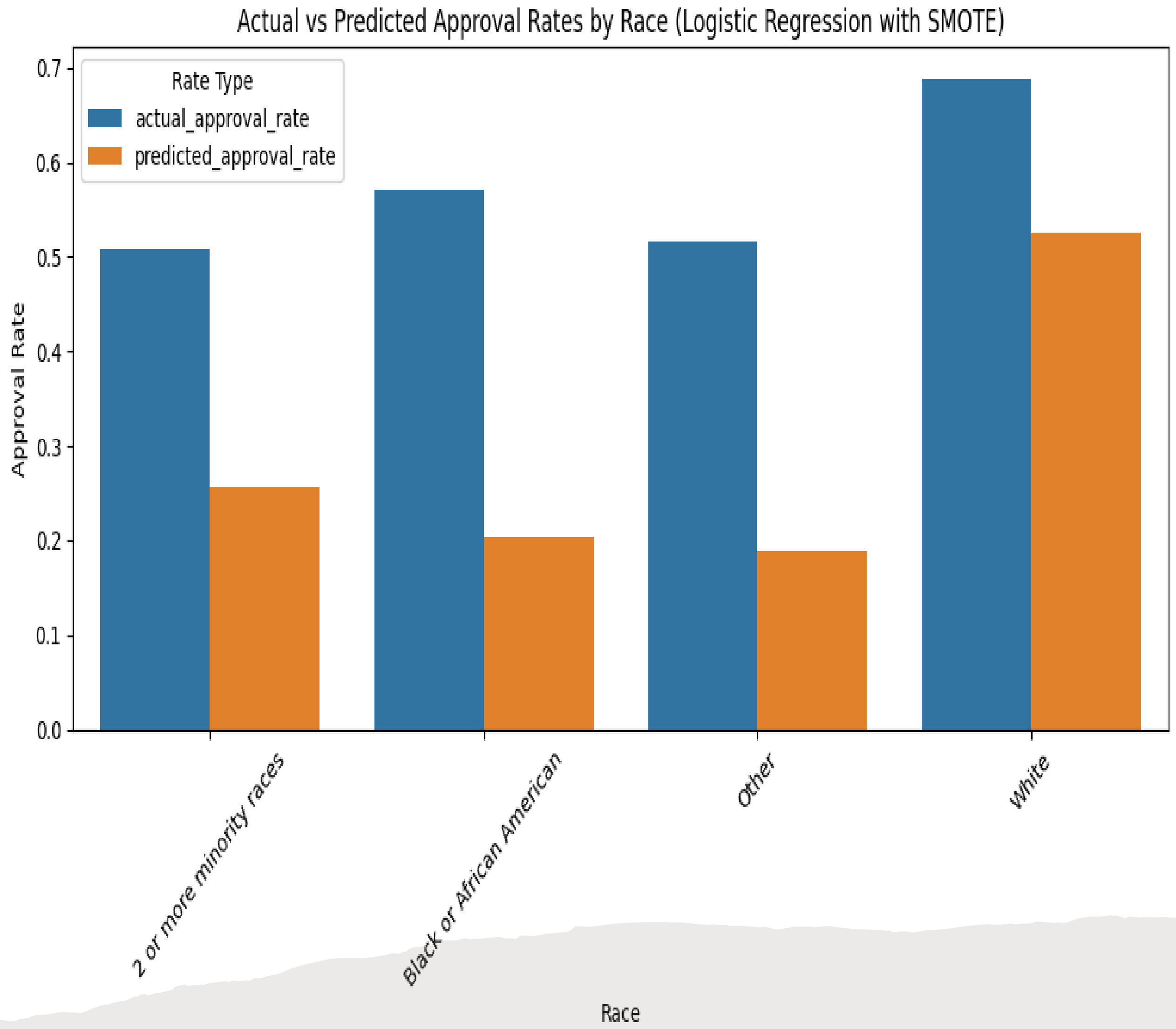
Calculated fairness metrics for logistic regression:

- Equalized Odds Difference (EOD): 0.3374.

- White-Black approval rate gap (actual: 0.1174, predicted: 0.3214).

- Visualized actual vs. predicted approval rates by race.

- Model underpredicts approvals for Black applicants (predicted: 0.2044 vs. actual: 0.5698).



Initial Privacy Analysis

Focused on privacy risks in the dataset.

Identified re-identification risks:

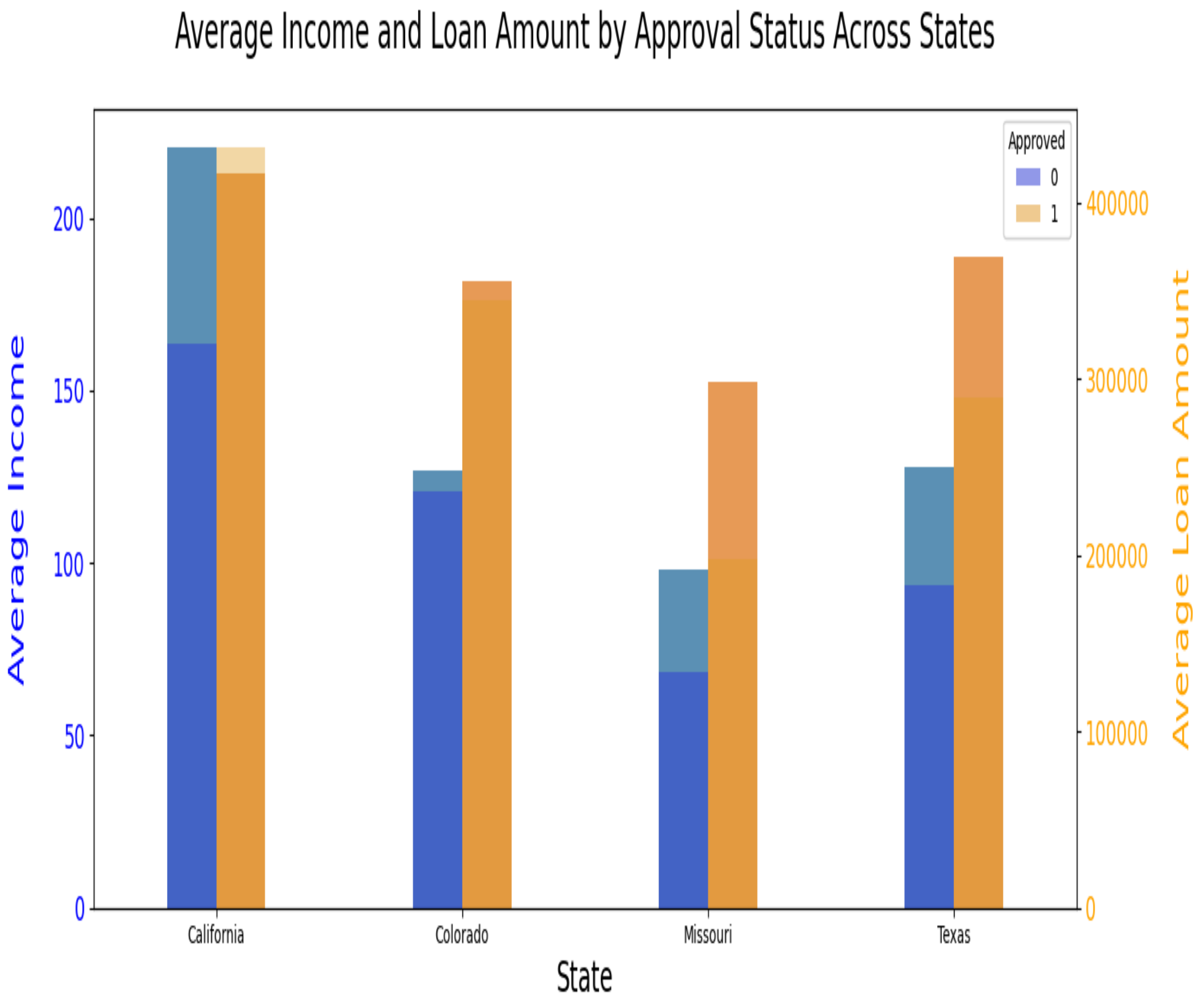
- Census tract and income data enable re-identification.
- Highlighted privacy concerns with census tract granularity.
- Aggregating census tracts at the county level is recommended to mitigate risks.
- Detailed data poses significant privacy threats.

Initial Transparency Insights

Analyzed transparency in loan approval processes.

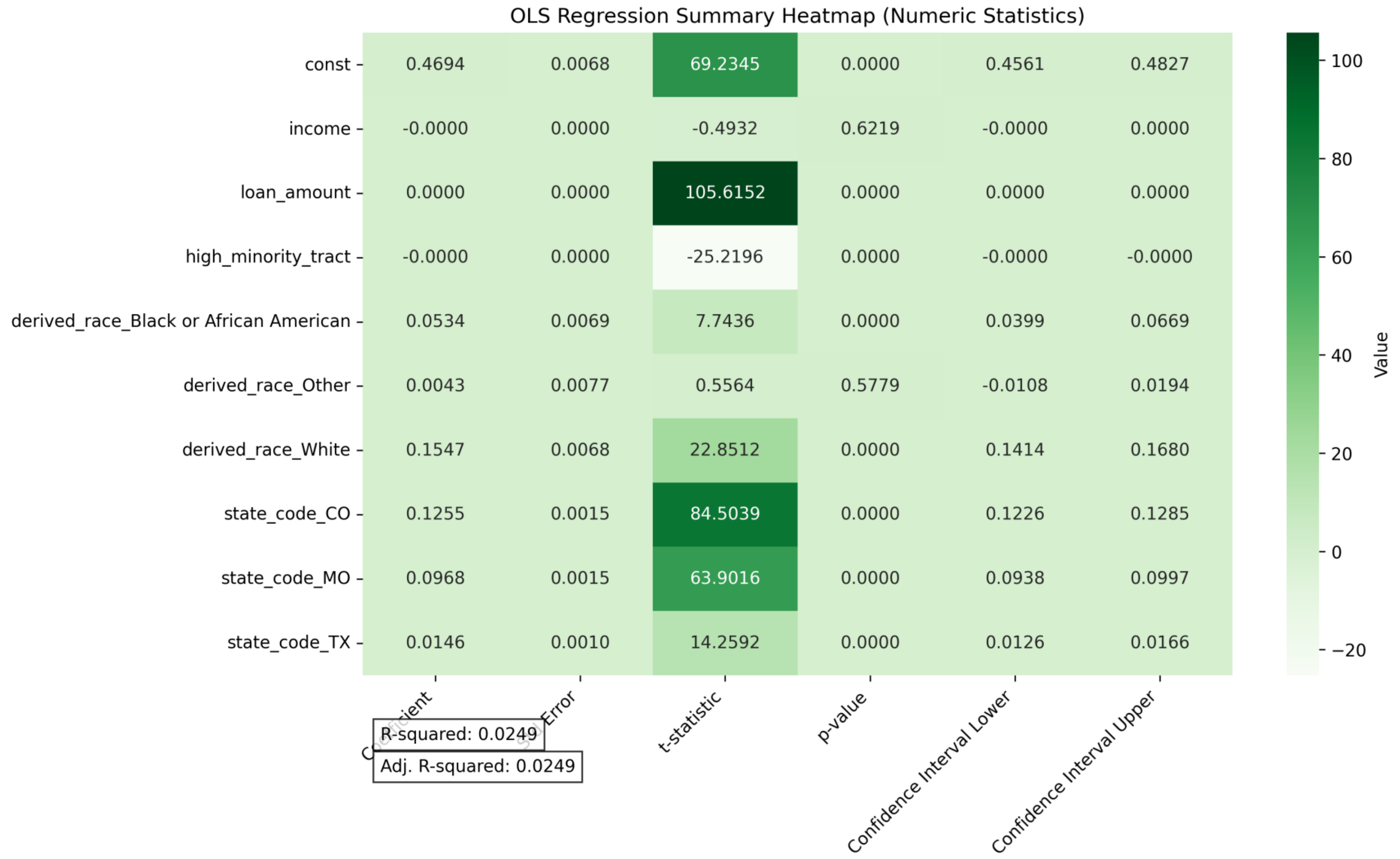
Supported visualization of income and loan amount distributions.

Disparities in approval rates suggest a lack of transparency.



Interpreted OLS regression:

- Derived_race_White (coef = 0.1547, $p < 0.001$) vs. derived_race_Black (coef = 0.0534, $p < 0.001$).



Fairness Interventions

-Explored methods to improve model fairness.

-Analyzed fairness metrics: EOD of 0.3374 indicates moderate disparity.

-I suggested reweighing the training data to balance racial groups.

-Proposed using fairness-aware algorithms to reduce bias.

-Interventions are needed to address model bias.

Metric	Value
Equalized Odds Difference	0.3374
White-Black Approval Gap	0.3214

Existing Solutions to Racial Bias in Lending

Distributionally Robust Fairness (DRF)
-Addresses subgroup discrimination overlooked by traditional group fairness metrics.
-Ensures equitable treatment across diverse demographic groups.

SenSR Algorithm
-An efficient training method implementing DRF.
-Utilizes the EXPLORE algorithm to learn fair metrics from data.

Individual Fairness Metrics
-Focuses on treating similar individuals similarly, regardless of group identity.
-Complements group fairness approaches by addressing nuanced biases.

Injustice anywhere is a threat to Justice everywhere

- MLK Jr