Group 2 (Harper Kates, Himani Patel, Jennifer Sika)

3/6/25

DATA 6550

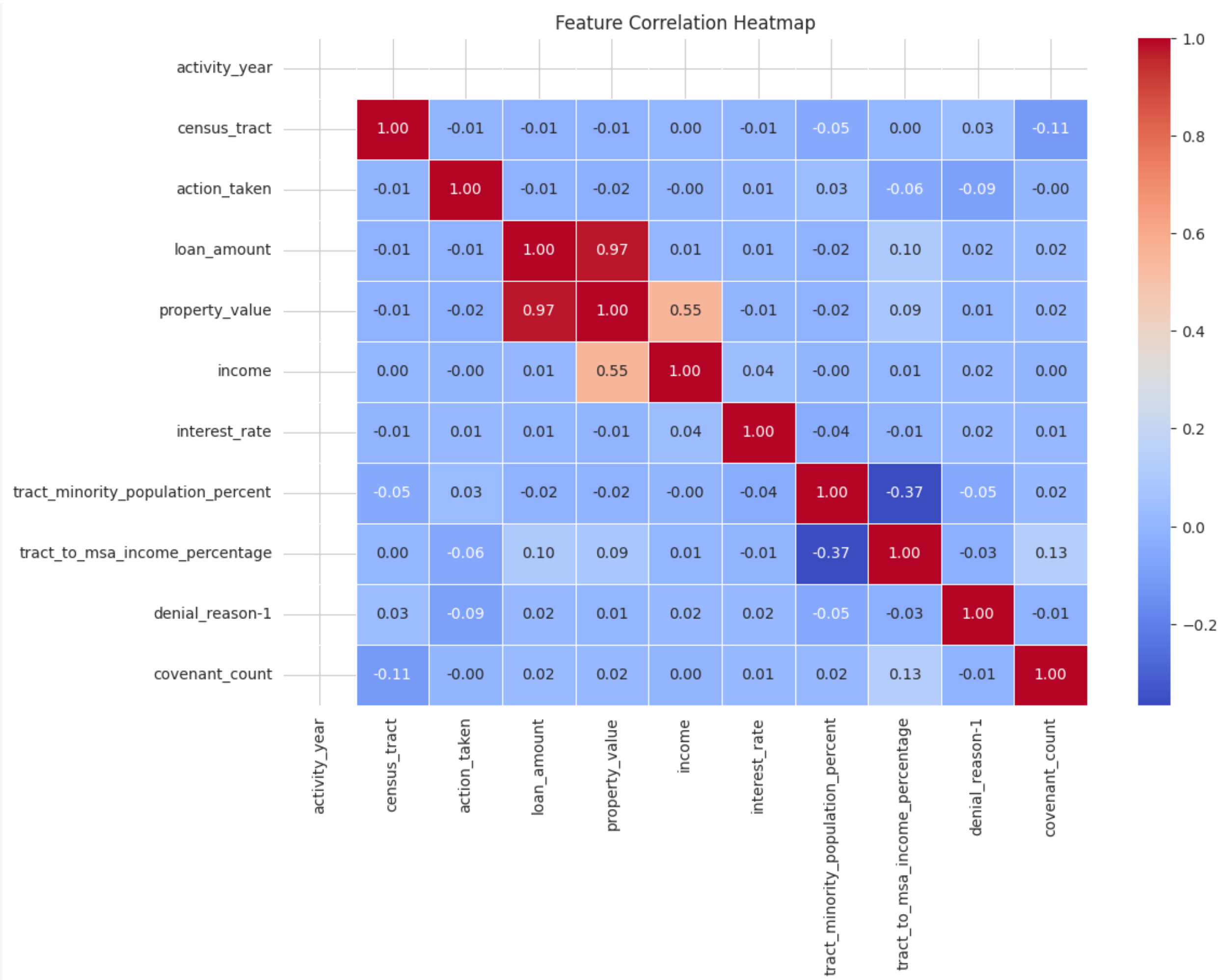
Dr. John Wallin

Module 4 Paper: Analyzing the Effects of Racial Covenants on Mortgage Data

In the 1896 Supreme Court decision in *Plessy v. Ferguson*, segregation was essentially justified under the clause “separate but equal.” In truth, however, this clause is only half correct, as the conditions that minorities had to endure during the segregation era were undoubtedly worse than that of white people. Segregation in the Jim Crow era took many forms, an important one being via the housing market. For example, in the 1926 Supreme Court decision in *Corrigan v. Buckley*, homeowners were given the right to enact racial covenants on their property, which may void the sale of a covenanted house to a racial minority. [1]

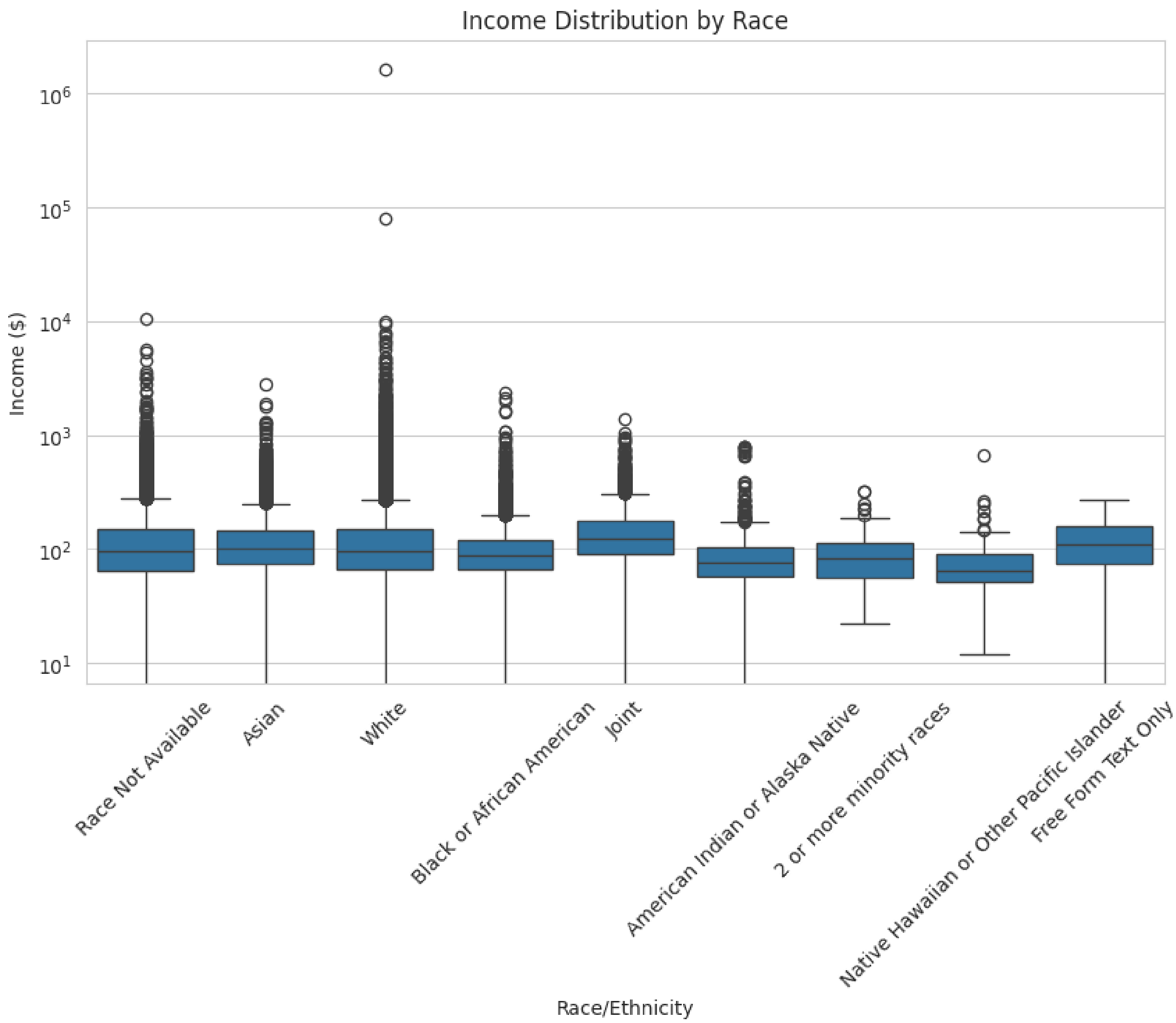
In particular, the city of Minneapolis, which was starting to rapidly grow at the time, had a high concentration of racial covenants, with a peak of 20% of homes being covenanted. [1]. Even after the *Shelley v. Kramer* decision (1948) that ruled racial covenants unenforceable, as well as the Fair Housing Act of 1968, which “explicitly banned housing discrimination by race,” [1], the segregated racial distributions of neighborhoods throughout the city persist today. [1]. Evidence of this de facto segregation can be seen in mortgage data from recent years. In this project, we analyzed mortgage data in Hennepin and Ramsey County (MN) in 2023. The goal of this project is to discover biases in the data that can negatively influence potential loan algorithms, presumably in the context of race.

Correlation Analysis

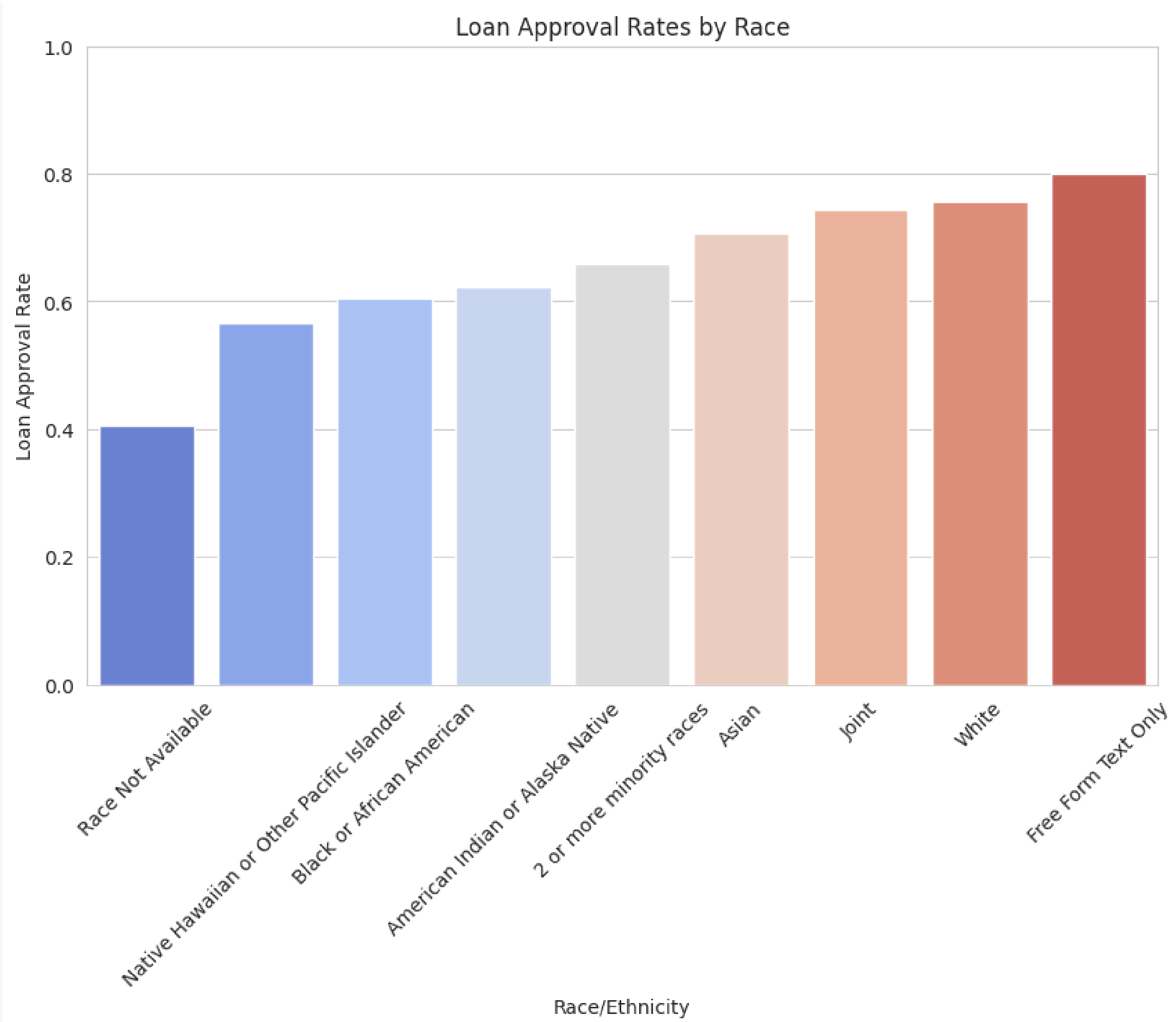


Loan amount and property value have the strongest positive correlation (0.97), which makes sense since larger properties usually require larger loans. Income is moderately correlated with both loan amount (0.55) and property value (0.55), suggesting higher-income applicants tend to purchase more expensive homes. Census tract has almost no correlation with loan **amounts** or property values, indicating that loan size isn't strongly tied to location in this dataset. Which I find to be a bit surprising. Higher minority population percentages in a tract are negatively correlated (-0.37) with median income, suggesting that predominantly minority areas tend to have lower median incomes. Covenant count has a weak negative correlation with census tract (-0.11) and minority population (-0.05), possibly indicating that historically segregated areas still show disparities in housing access. This will influence the focus of Section 4 (Bias / Algorithmic Analysis).

Bias Analysis

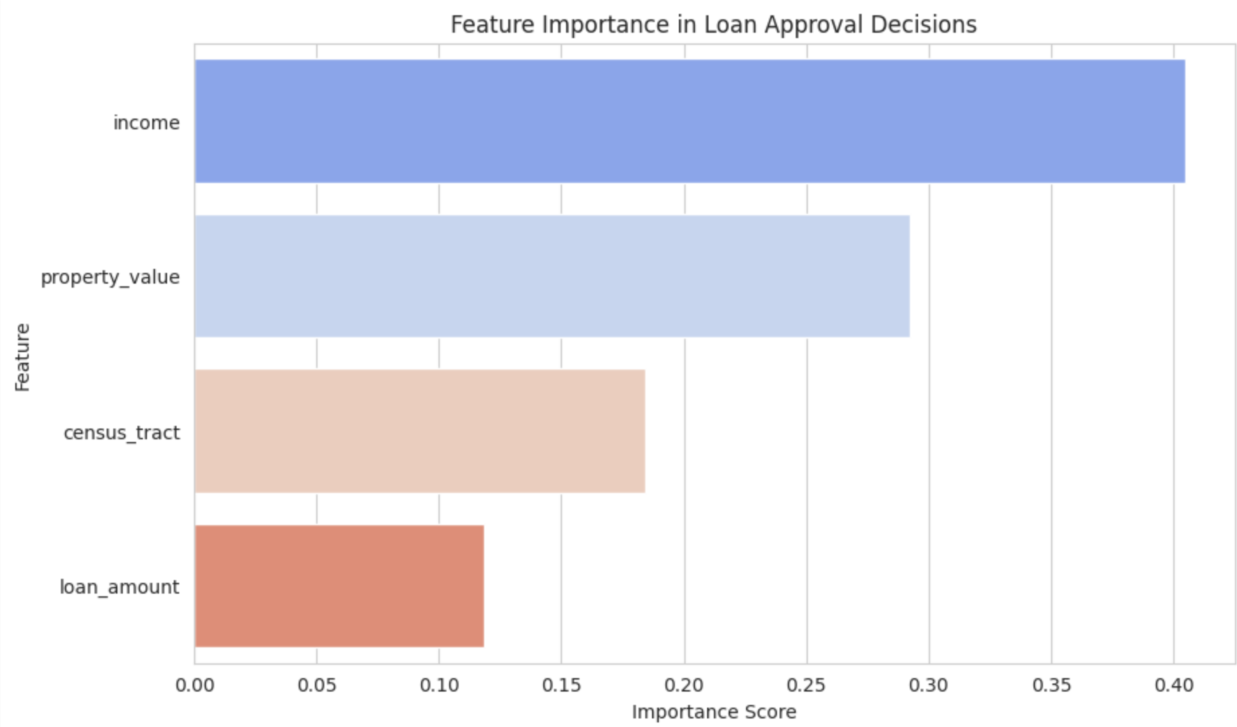


The median income appears relatively similar across most racial groups. Some groups (White, Asian, and Joint applicants) have slightly higher medians compared to others. Some minority groups (Native Hawaiian or Other Pacific Islander, American Indian or Alaska Native) have lower overall income variability compared to others. Assuming that income plays a significant role in loan approvals, minority applicants with lower average incomes may face higher rejection rates. It is also important to note that though the income levels appear similar, there are nuances to consider.



White, Asian, and Joint applicants have the highest loan approval rates.

Black or African American and Native Hawaiian/Pacific Islander applicants have lower approval rates than other groups. The "Race Not Available" category has the lowest approval rate, possibly indicating that applicants who do not disclose race might face negative lending outcomes. Since Black, Native Hawaiian, and Pacific Islander applicants have lower approval rates, this could indicate historical barriers in lending. Disparities in income levels, credit history, or property values might contribute to the gap, but it is worth investigating whether algorithmic bias or discriminatory lending practices play a role.



Income has the most significant role in determining loan approval. Property value is the second most important factor. Since income and property value dominate the decision process, lower-income and lower-property-value applicants may be disproportionately denied loans. As discussed in the previous sections, minority races are often part of the lower income group. This is important to note because the loan approval process hinges on income. This could lead to discriminatory lending decisions.

Census tract also has a moderate influence on loan approval. This is concerning because it might reflect historic discrimination or location-based discrimination. Our analysis proves that applicants from certain neighborhoods face more difficulty getting approved. The chart depicts that census tract does carry a level of importance when determining loan approval. This proves that location influences approvals. This could contribute to persistent discriminatory lending practices against minority or low-income neighborhoods.

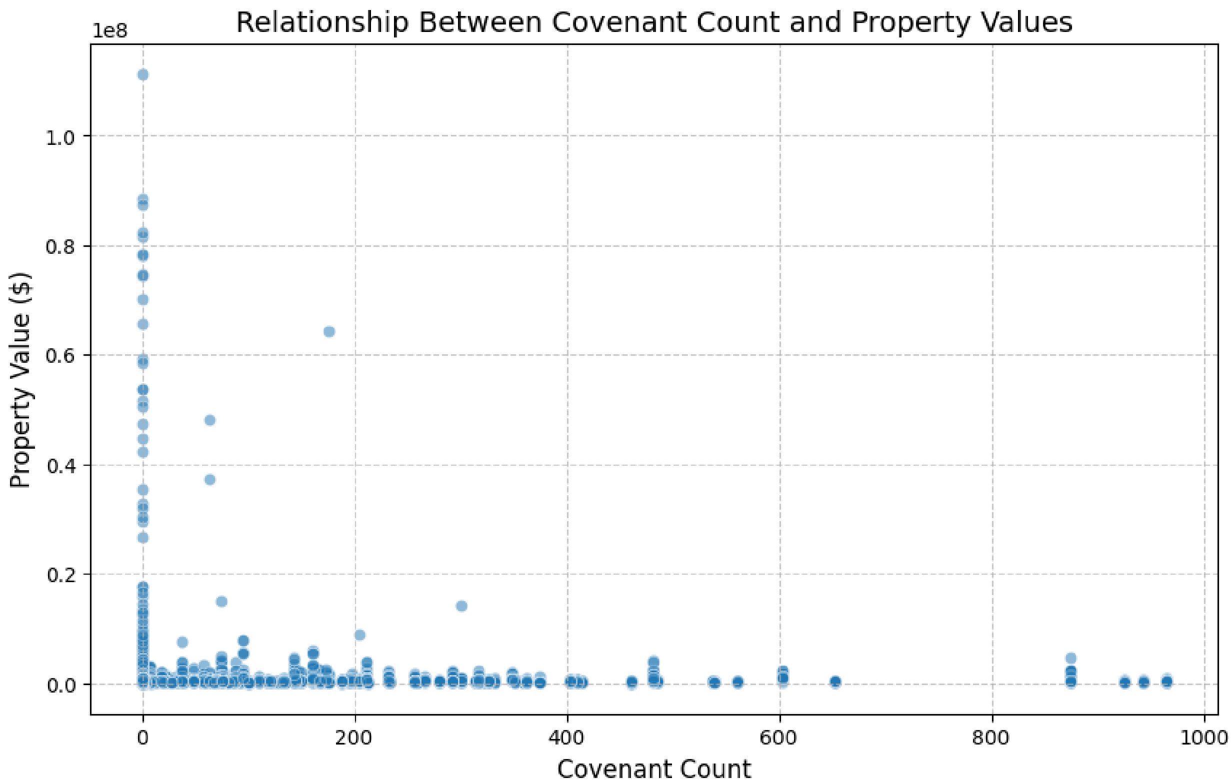
Algorithmic Analysis

Overall Model Performance:

	precision	recall	f1-score	support
False	0.46	0.00	0.01	6731
True	0.67	1.00	0.80	13456
accuracy			0.67	20187
macro avg	0.56	0.50	0.40	20187
weighted avg	0.60	0.67	0.54	20187

When the model predicts a loan denial, it is correct 46% of the time. When the model predicts a loan approval, it is correct 67% of the time. The F1-score is very low for loan denials. This leads me to believe the model is terrible at identifying cases where loans should be denied. Since the model was trained on past loan decisions, the model inherits those biases. If historically, minorities were disproportionately denied loans, the model might reinforce this pattern, especially if it learned from skewed approval criteria. If minority applicants were historically more likely to be denied, and the model is poor at recognizing denials, this could indicate that the model does not correct past discriminatory patterns. Groups with less/poor financial history or lower credit scores may be unfairly assessed due to biased features.

Modern Implications



While the racial covenants have not been put into practice in recent years, the impact of them is still felt. This graph shows that if a property was put under a mortgage covenant, they are more likely to have a lower property value. Most properties with a high covenant count (above 100) have relatively low property values. This leads us to believe that properties with more restrictive covenants may have less market value or be in lower-value areas. This impacts resale and the overall financial investment of the buyer. There are some high-value properties (above $10M) with low covenant counts (between 0 and 50). This might signal that expensive properties may not be as burdened with covenants.

Sources:

1. Sood, Aradhya, William Speagle, and Kevin Ehrman-Solberg. "Long shadow of racial discrimination: Evidence from housing covenants of Minneapolis." *SSRN Electronic Journal* 10 (2019).
2. Portions of this paper were written with AI.