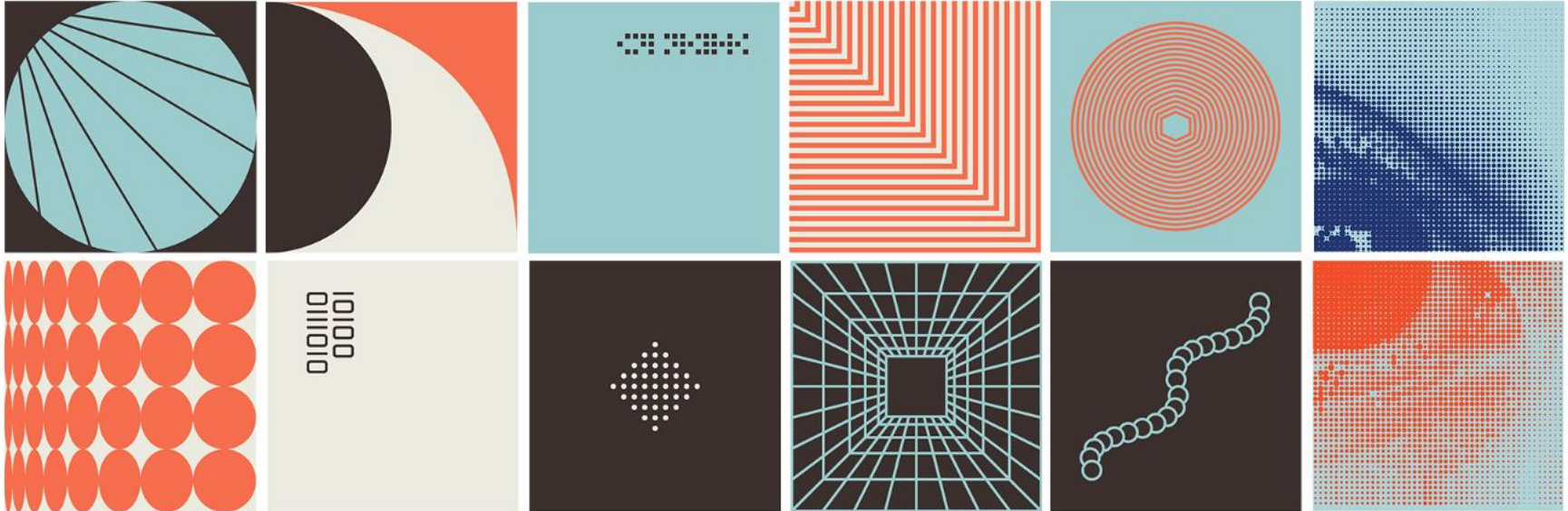


TEAM

Diana Bowser
Marc Abelard
Lisa Thorn
Michaela McCormick

Data Collection: The backbone of racial bias in mortgage lending algorithms



Agenda

Historical Context

How algorithms work and their role in perpetuating bias

Research Plan

Mixed-Methods study proposal

Initial Findings

SDOH dataset

Next Steps

Contributing perspective and nuance

Historical Algorithms

1941

The first computer

The first computer was used to decrypt Nazi communications

1956

Founding of the Fair Isaac Corporation (FICO)

The first computer to use keyboard inputs

1989

FICO reveals first general-purpose lending algorithm

The first laptop

1995

Fannie Mae and Freddie Mac institutionalized FICO

The first mobile device

Research Questions

1	How are mortgage lending algorithms trained?
2	How diverse is the data used to train lending algorithms?
3	How does data collection impact lending for racial and ethnic minorities?
4	What is the national variation in these lending outcomes?

1995 Creditworthiness

Income, Assets and Debt

Racism, sexism, and intersectional forms of bias contribute to the gap in wealth, employment, and financial outcomes

Employment Status

- Internet and the speed of communication has made it easier to work from home
- Tech companies create entire industries of gig workers and lobby to keep their contractor classification

Credit History

Rent, utilities, cell phone bills, internet, subscriptions are all new forms of financial payments. More people are renting for larger portions of their income

Related Policies

1977

Community Reinvestment Act

- Originally intended to help low-income communities and subsequently communities of color
- Due to gentrification these loans end up going to wealthier individuals

2018

EU General Data Protection Regulation (GDPR) & CA Consumer Privacy Act (CCPA)

- Burden on the consumer to opt-out
- Does not limit what can be collected, how long it can be stored, or who has access to the data
- Information is bought and sold for purposes the original collection was never intended to be used for

Agenda

Historical Context

How algorithms work and their role in perpetuating bias

Research Plan

Mixed-Methods study proposal

Initial Findings

SDOH dataset

Next Steps

Contributing perspective and nuance

OWNER-OCCUPIED

89%

Owner-occupied housing units

33%

Owner-occupied housing units

RENTER-OCCUPIED

OWNER-OCCUPIED

6,184

Total households living in the county

180,680

Total households living in the county

RENTER-OCCUPIED

OWNER-OCCUPIED

0.13

Social Vulnerability Index

0.82

Social Vulnerability Index

RENTER-OCCUPIED

OWNER-OCCUPIED

\$172,900

Average home value

OWNER-OCCUPIED

\$55,200

Average household income

OWNER-OCCUPIED

2%

Average unemployment

\$325,700

Average home value

\$56,600

Average household income

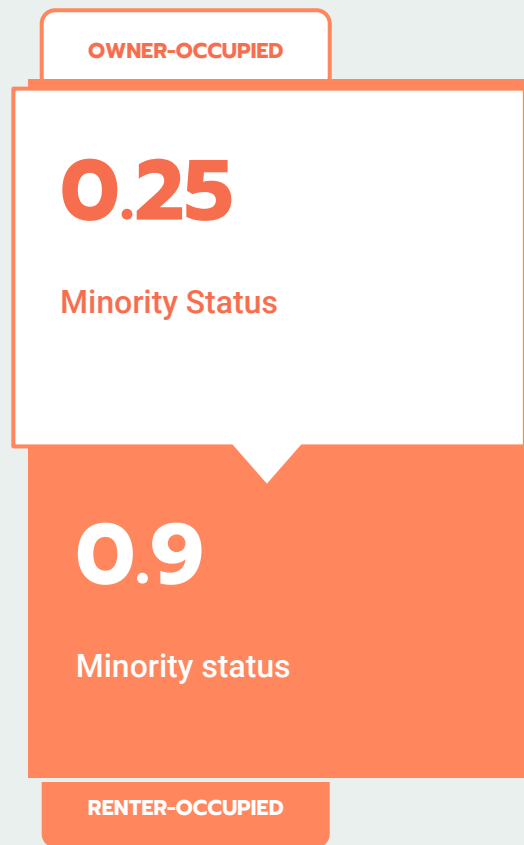
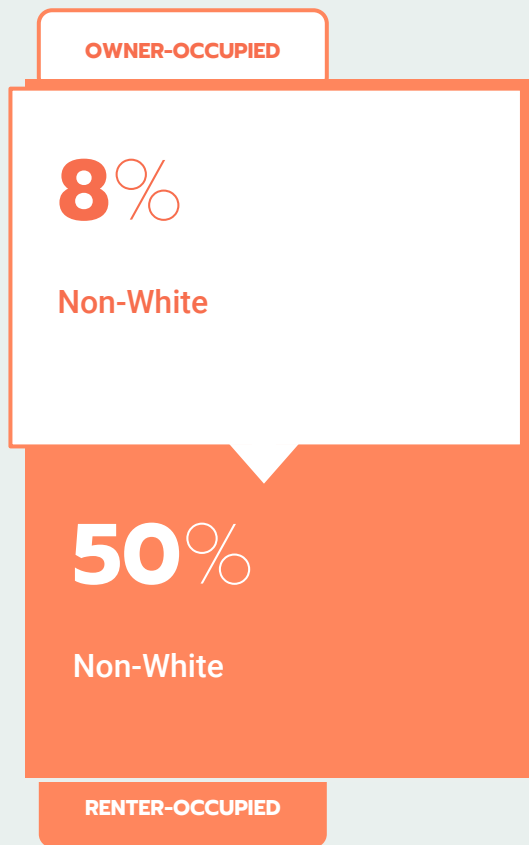
4%

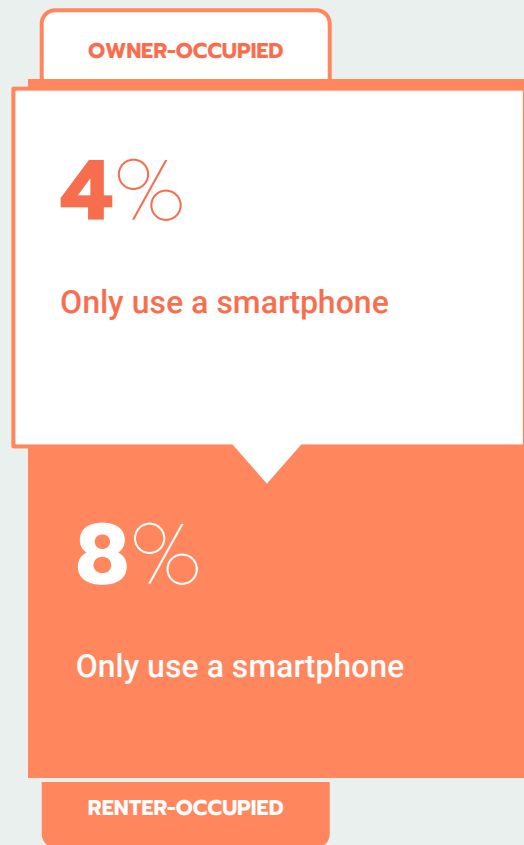
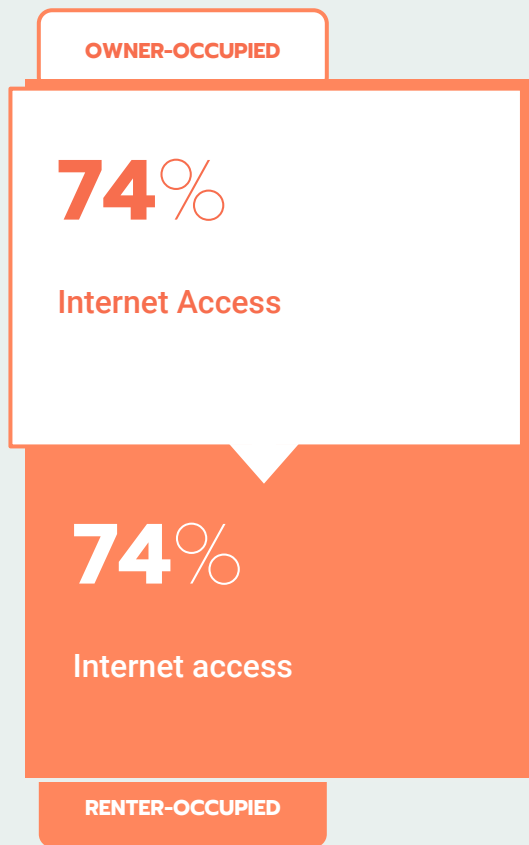
Average unemployment

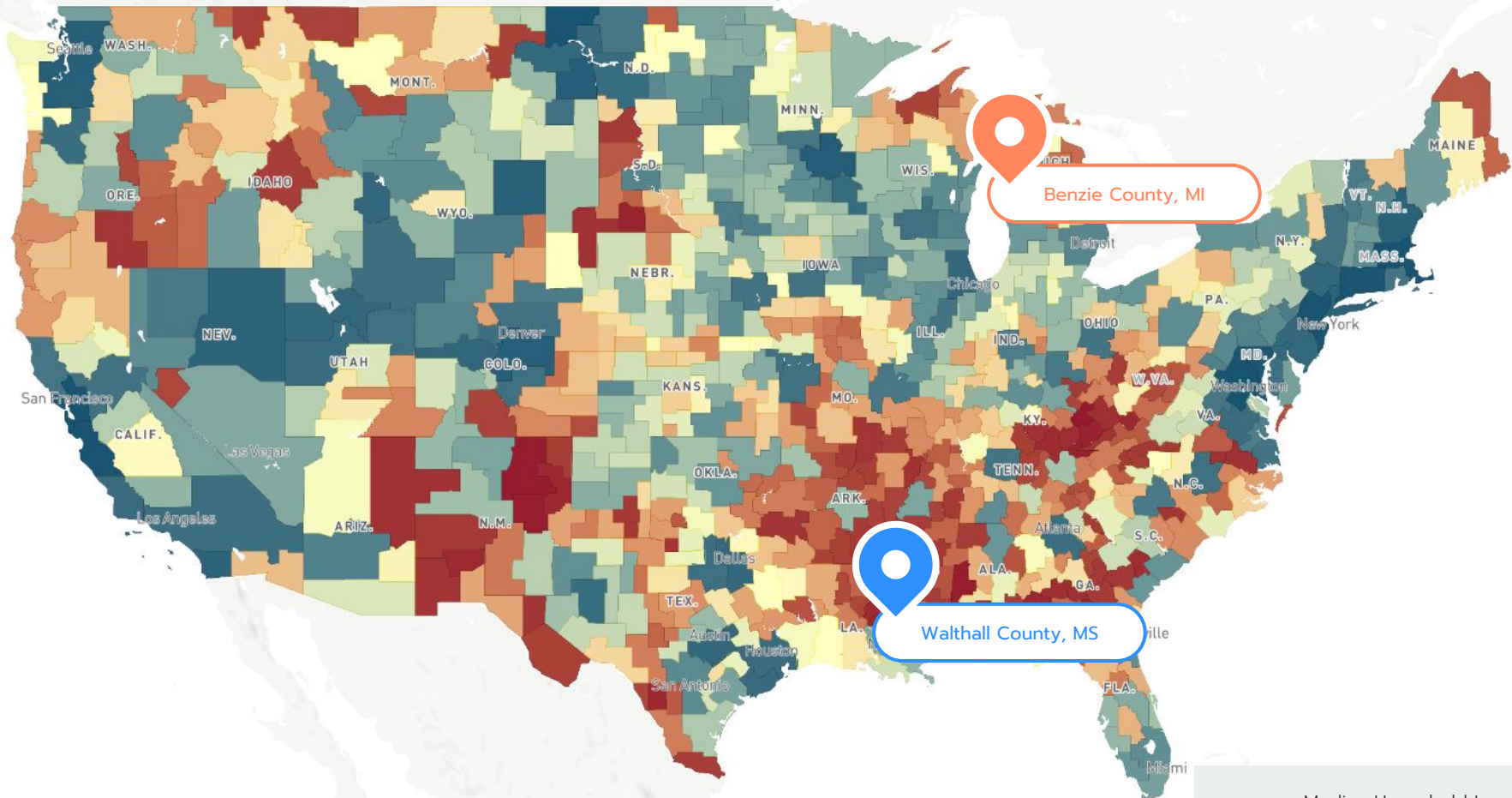
RENTER-OCCUPIED

RENTER-OCCUPIED

RENTER-OCCUPIED







A Tale of Two Counties

AHRQ SDOH Data Set

Benzie County, MI

4%

Non-White

96% White
2% Black
2% Hispanic
<1% NANA, NHPI, AA

\$172,000

Average home value

Income: \$56,600
Education: 62%

82%

Internet access

4% only use a smartphone

90%

Owner-occupied housing units

5% Mobile homes

Walthall County, MS

47%

Non-White

53% White
46% Black
<01% Hispanic, NANA, NHPI, AA

\$90,400

Average home value

Income: \$29,800
Education: 35%

48%

Internet access

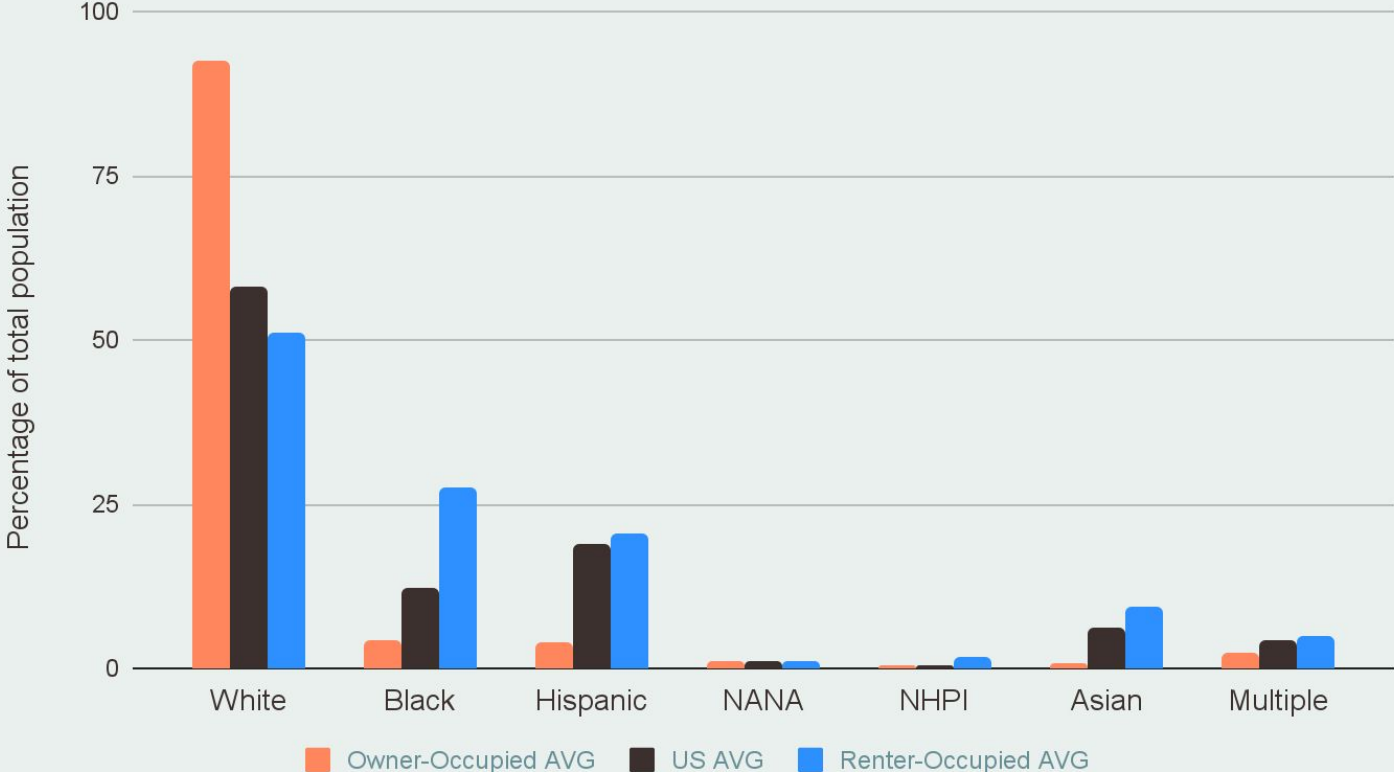
10% only use a smartphone

89%

Owner-occupied housing units

28% Mobile homes

Racial Representation



Agenda

Historical Context

How algorithms work and their role in perpetuating bias

Research Plan

Mixed-Methods study proposal

Initial Findings

SDOH dataset

Next Steps

Contributing perspective and nuance

Research Plan



Literature Review

Further examine the history of mortgage lending for racial and ethnic minorities, machine learning, and automated decision-making systems

Quantitative Sources

Pattern discovery in lending, homeownership, and environment.
Define potential interview criteria and questions

Qualitative Interviews

Develop a proposal that clearly outlines the research project, methodologies, and outcomes

Research Plan

1

Connect lending outcomes to variables that have been used as proxies for race for generations

2

Income, debt, the ability to navigate the process of applying and purchasing a home are all part of a system that has limited people of color from home ownership and future financial mobility

Literature Review

Scholarly

- **Keywords:** Algorithmic Bias, Machine Learning, AI, Mortgage Lending, Housing Discrimination, Racial Bias
- *Defining vocabulary* (ex: AI, ML, data, and algorithms)
- Historical analysis
- How non-race variables have been used as proxies for race in banking, housing, and employment

Popular & Grey

- Tech companies creating lending algorithms
- Platforms using algorithms for banking and/or housing
- Regulation of AI/ML algorithms used in banking, housing, or employment
- The influence of algorithms over racial bias in banking, housing, or employment
- Variables considered in current banking and housing lending algorithms

Quantitative Analysis

Public lending databases

- Capture the mortgage application process to understand demographics, background checks and credit scores for those denied mortgages and whether this varies by type of lender (Home Mortgage Disclosure Act)

Subnational level analysis

- Use county or subnational level data (home value (zillow.com), Social Determinants of Health, County level demographics (income, race/ethnicity, housing, etc.)) to measure variation in background checks, credit scores, and denied loans across the United States

Qualitative Sources

People who have been denied

- Looking for patterns of location, race, age, health to define additional interviews
- Seeking patterns in why they were denied, the application and denial process, overall experience
- Potentially find areas that are doing really well

Mortgage lending and banking experts

- Seeking detailed knowledge of the variables used in banking and housing decision making process

AI and ML development experts

- Seeking detailed knowledge of the variables used in banking and housing algorithms
- How ML algorithms are trained and regulated
- How data influences ML algorithms and non ML algorithms

Limitations

No concrete proof of systemic discrimination

- Information from one database on lending cannot be directly connected to our SDOH database
- Most racism is indirect

Transparency

- Private businesses with private algorithms
- Reverse engineering lending and lifestyle outcomes

Agenda

Historical Context

How algorithms work and their role in perpetuating bias

Research Plan

Mixed-Methods study proposal

Initial Findings

SDOH dataset

Next Steps

Contributing perspective and nuance

Next Steps

Algorithm Editorials

Editorials that reframe algorithms as an old decision-making tool that amplifies their impact and potential biases with the ubiquity of devices and exponential computing power of today's technology.



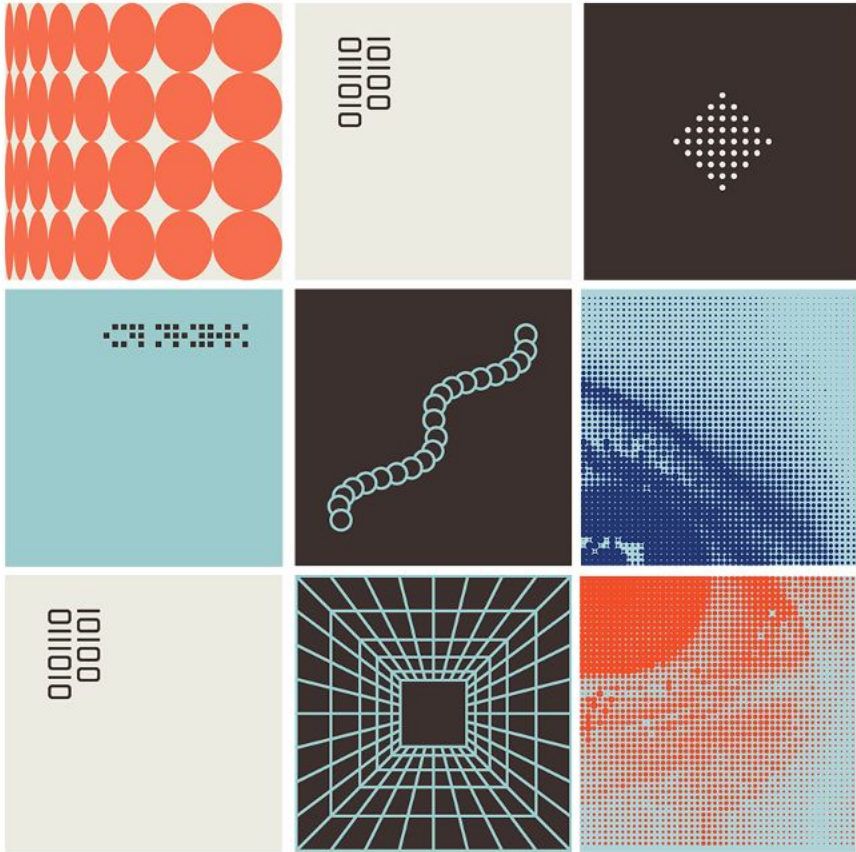
Lending Editorials

Editorials that highlight lending outcomes alongside the experiences of the residents in those areas. Contextualizing the numbers with qualitative stories.



Mixed-Methods Research

- Holistic evaluation of the lending process inclusive of cultural contexts
- How lending resources differ across counties
- Cultural attitudes toward homeownership
- Transparency of who decides what an algorithm includes / excludes
- Deeper Data Dives
 - ◆ Unhoused people
 - ◆ Cost of living, regional markets



Thank You // Q&A