ACCESS TO HOME OWNERSHIP

EXAMINING MORTGAGE APPROVAL & DENIAL BASED ON RACE

Project 5

Barb Allen
Christopher Duro
Vadim Serenbrinskiy
Peter Wentzel

EFFECT OF RACE ON HOME LOAN APPROVAL

Project 5

Barb Allen
Christopher Duro
Vadim Serenbrinskiy
Peter Wentzel

AGENDA

01
BACKGROUND
Context of the Study and Problem Statement

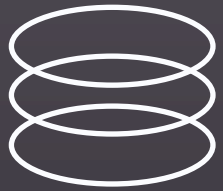
02
DATA PROCESSING
Gathering, Cleaning, Feature Engineering

03
EXPLORATORY DATA ANALYSIS
Insights from Data

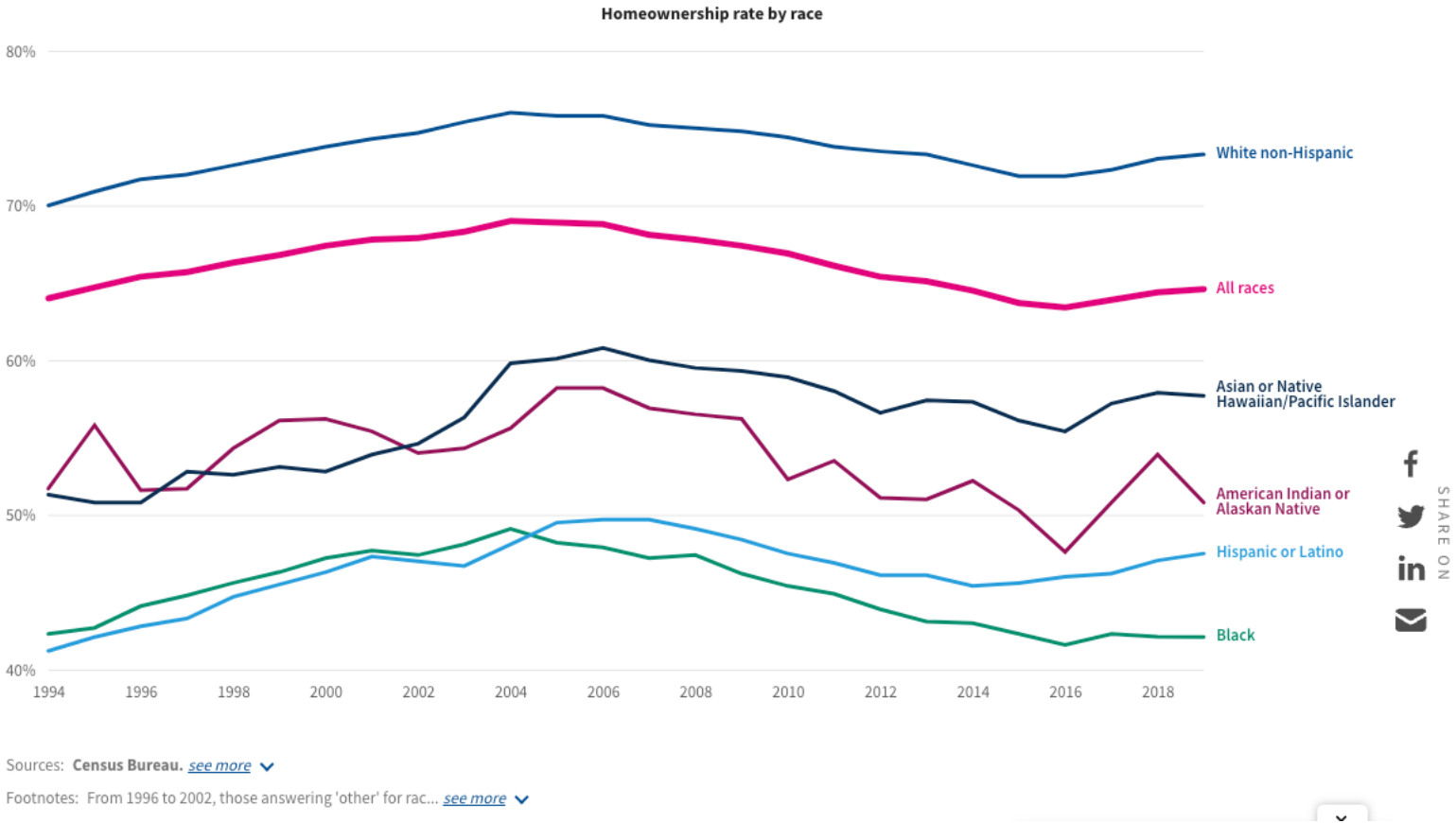
04
MODELING
Logistic, KNN, Random Forest

05
RECOMMENDATIONS
Moving Forward

01



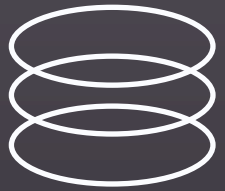
BACKGROUND


Homeownership rate by race
80%
70%
60%
50%
40%
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
White non-Hispanic
All races
Asian or Native Hawaiian/Pacific Islander
American Indian or Alaskan Native
Hispanic or Latino
Black
SHARE ON
Sources: Census Bureau. see more
Footnotes: From 1996 to 2002, those answering 'other' for rac... see more

PROBLEM STATEMENT

There are large disparities in home ownership between white, non-Hispanic Americans and minority American populations.

FOCUS OF STUDY



What is the relationship between home loan approval and race in major metropolitan areas in the four largest states in the US?

DATA SET

DATA SOURCE

Consumer Financial Protection Bureau 2014-2017

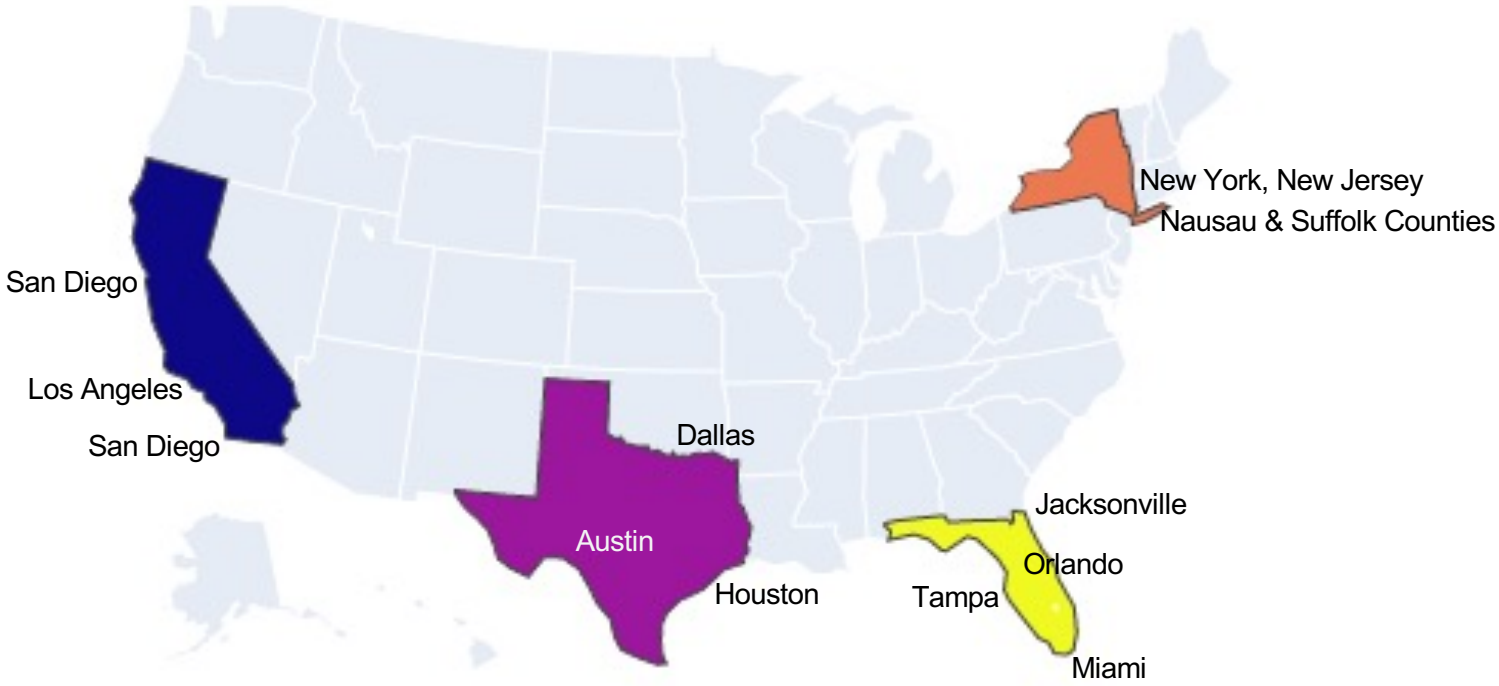
GEOGRAPHIC AREA

California, Texas, New York, Florida

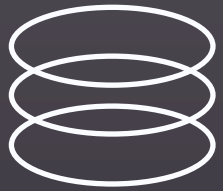
MERGED DATAFRAME SIZE

4,715,850 rows
52 features

12 METROPOLITAN STATISTICAL AREAS/ METROPOLITAN DIVISIONS (MSAMD)



02



DATA PROCESSING

DATA PROCESSING

GZIP COMPRESSION

CSV reduced from 612.9MB to 99.8MB

COMPRESSED DATA

Numeric key:value pairs

DATA REDUCED

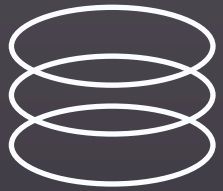
California, Texas, New York, Florida

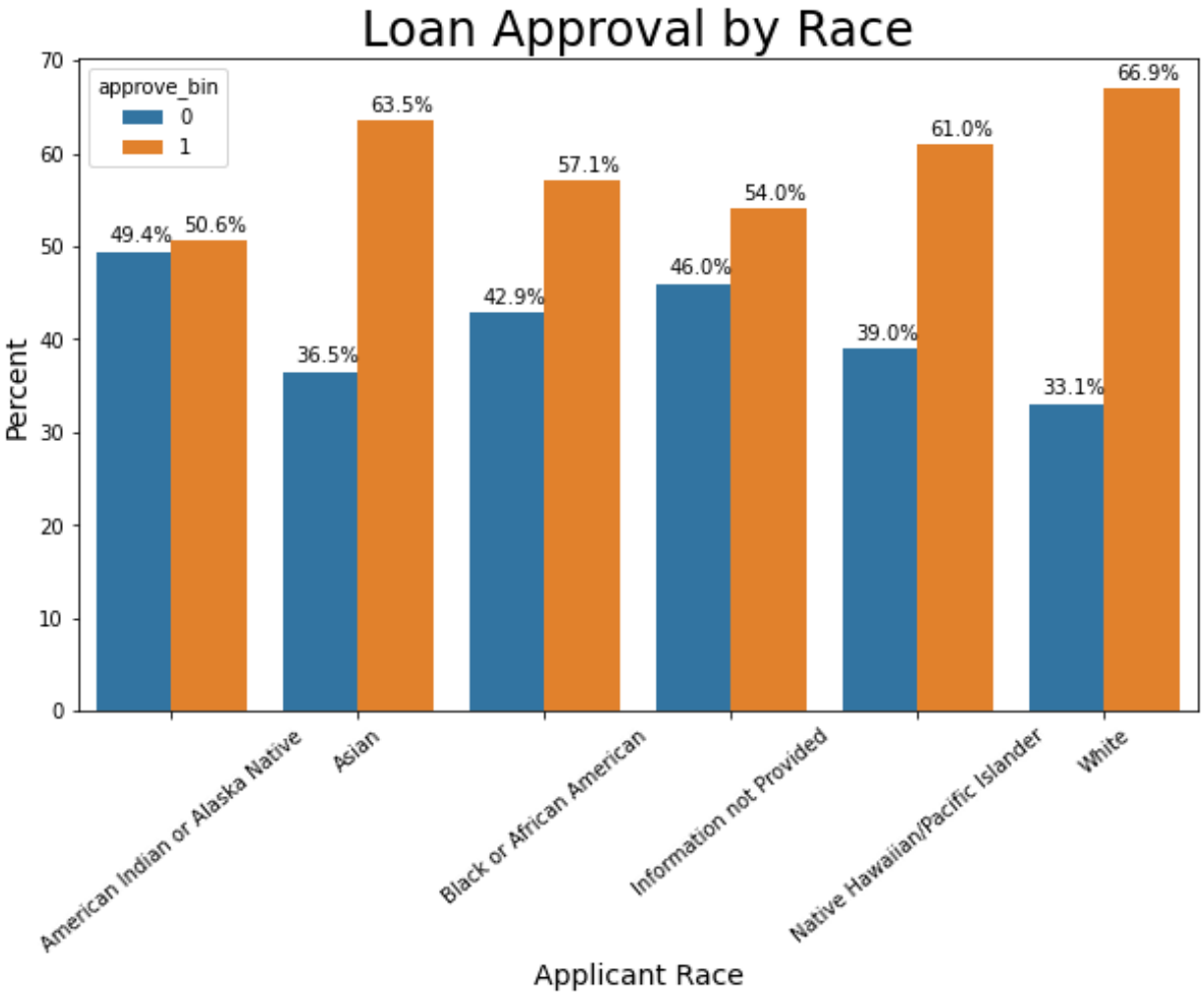
REMOVED INCOME OUTLIERS

Applicants below 20th percentile & over the 90th percentile

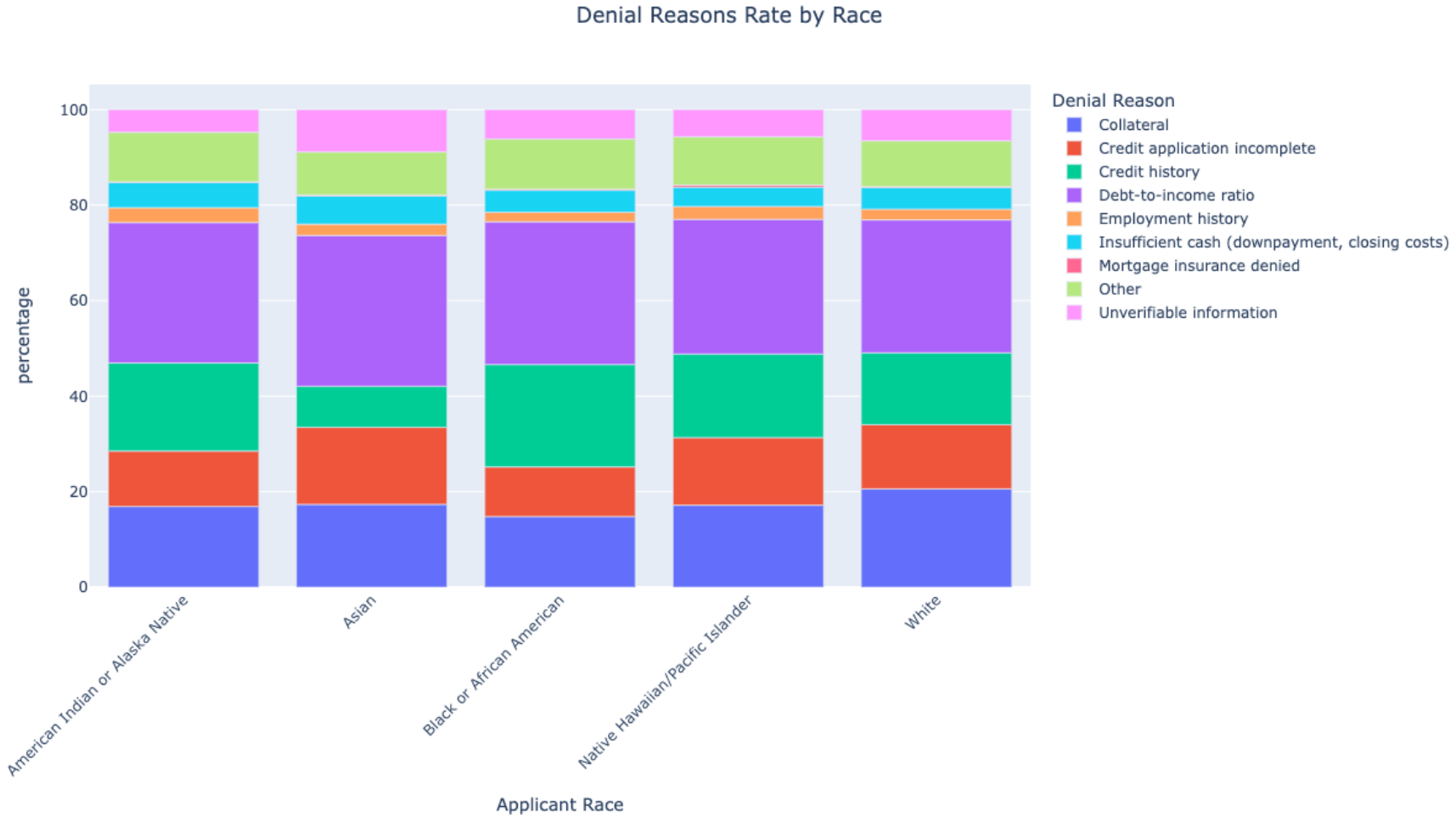
03

EXPLORATORY DATA ANALYSIS

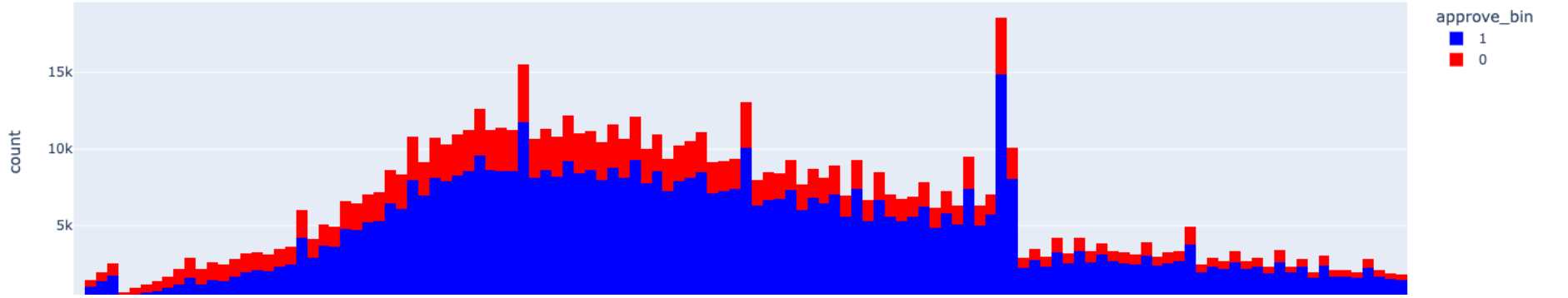



Loan Approval by Race
approve_bin
0
1
Percent
70
60
50
40
30
20
10
0
49.4%
50.6%
36.5%
63.5%
42.9%
57.1%
46.0%
54.0%
39.0%
61.0%
33.1%
66.9%
American Indian or Alaska Native
Asian
Black or African American
Information not Provided
Native Hawaiian/Pacific Islander
White
Applicant Race

Loan Approval Rate by Race
Approved
Denied
Approval Rate (%)
0
10
20
30
40
50
60
70
80
73.5%
26.5%
78.2%
21.8%
71.6%
28.4%
77.3%
22.7%
79.2%
20.8%
American Indian or Alaska Native
Asian
Black or African American
Native Hawaiian/Pacific Islander
White
Applicant Race

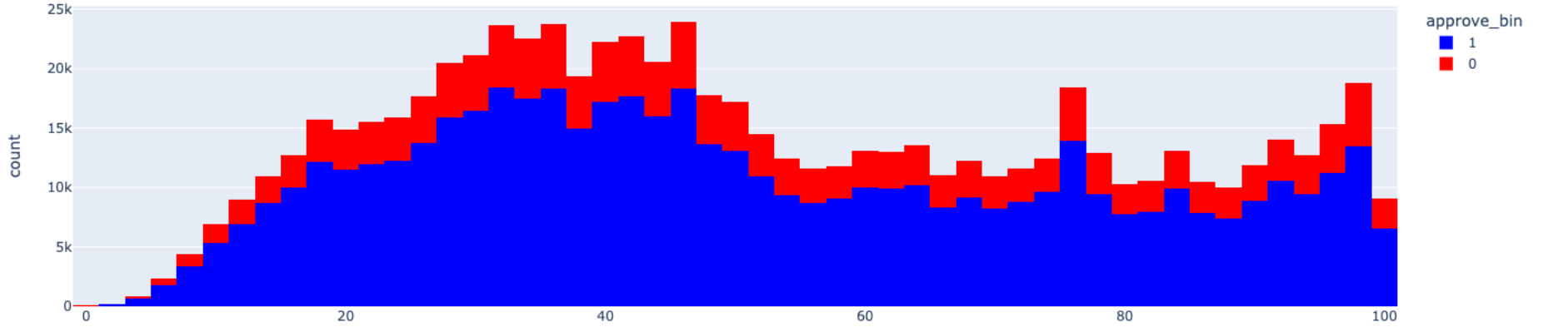

Denial Reasons Rate by Race
percentage
100
80
60
40
20
0
American Indian or Alaska Native
Asian
Black or African American
Native Hawaiian/Pacific Islander
White
Applicant Race
Denial Reason
Collateral
Credit application incomplete
Credit history
Debt-to-income ratio
Employment history
Insufficient cash (downpayment, closing costs)
Mortgage insurance denied
Other
Unverifiable information

Minority Applicant Loan Amount by Approved/Denied Count

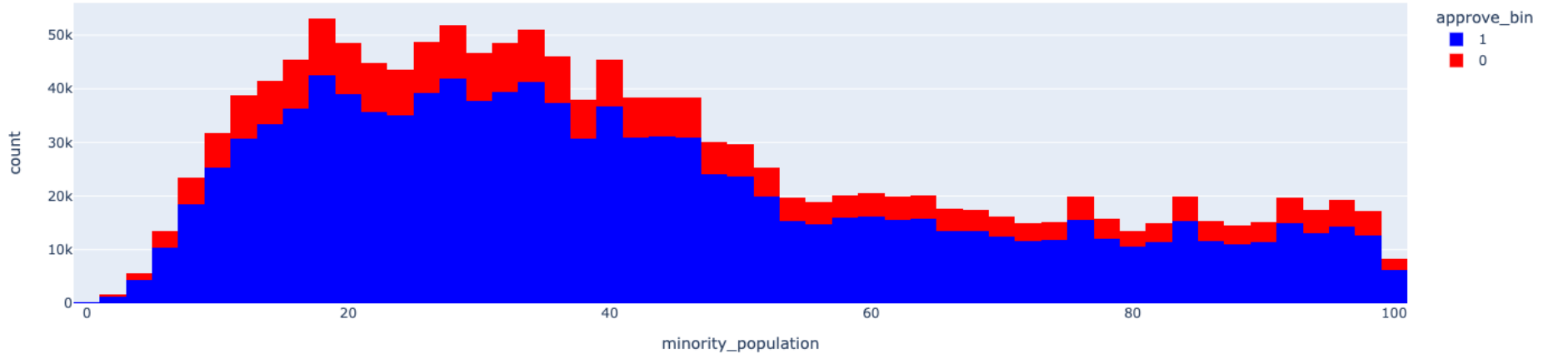


White Applicant Loan Amount by Approved/Denied Count

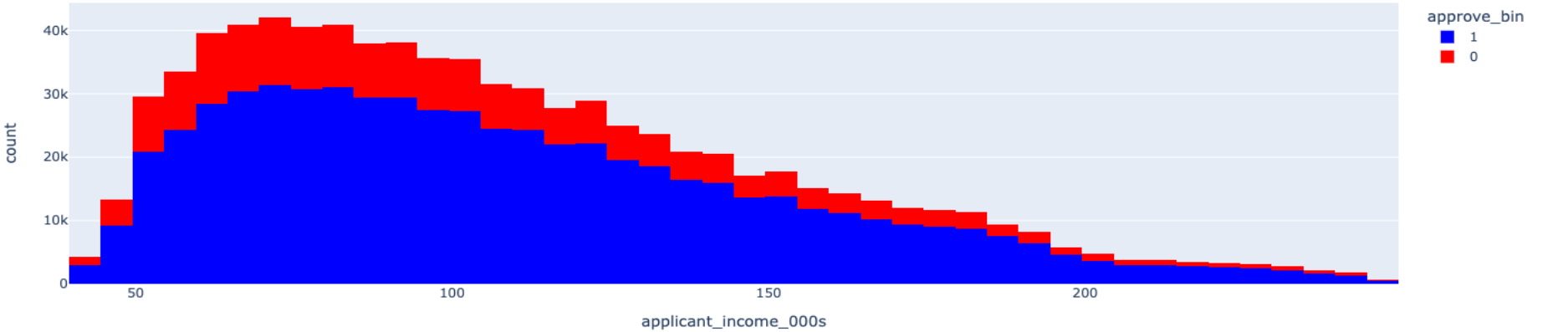
Minority Applicant Population Tract Percentage by Approved/Denied Count



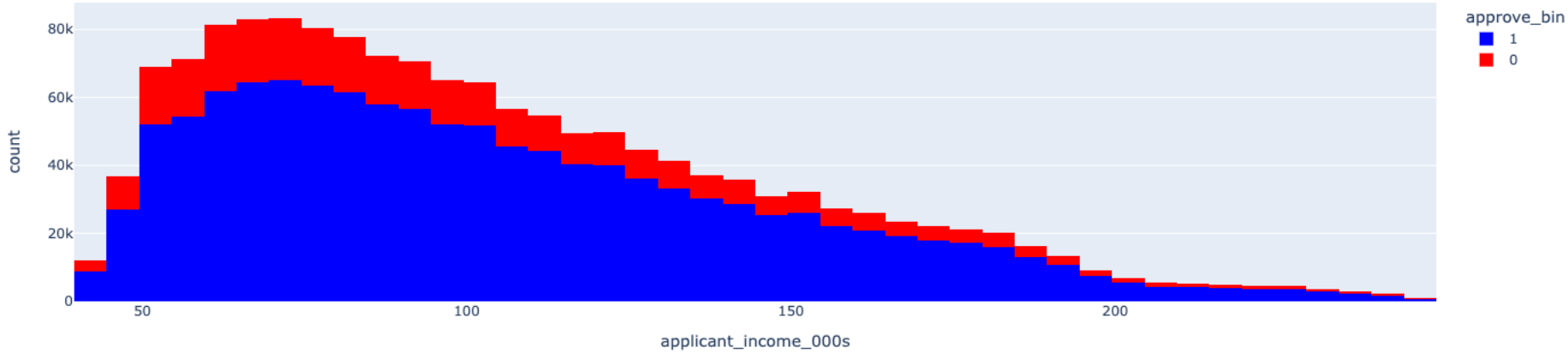
White Applicant Minority Population Tract Percentage by Approved/Denied Count

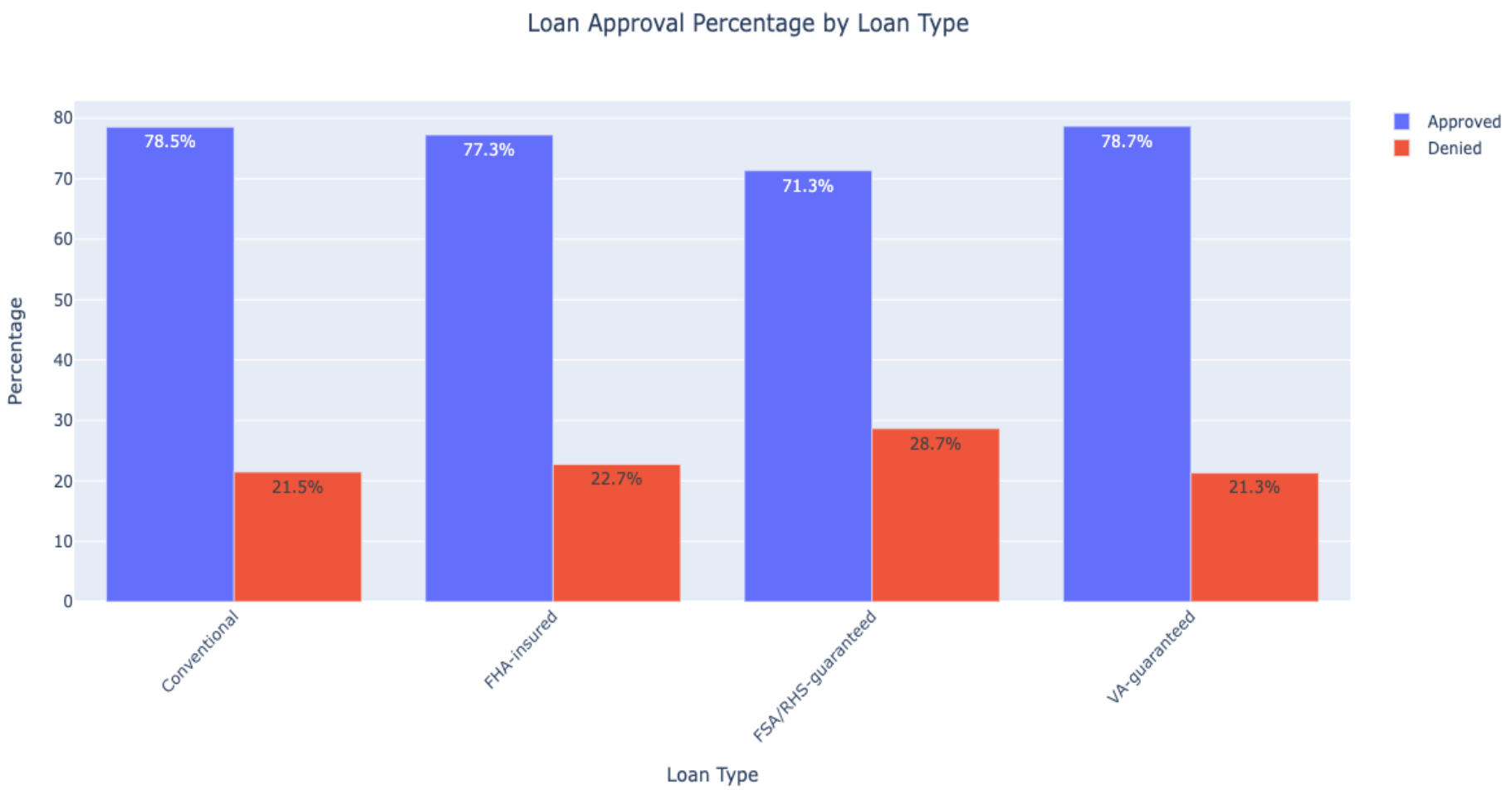


Minority Applicant Income < $300,000 by Approved/Denied Count

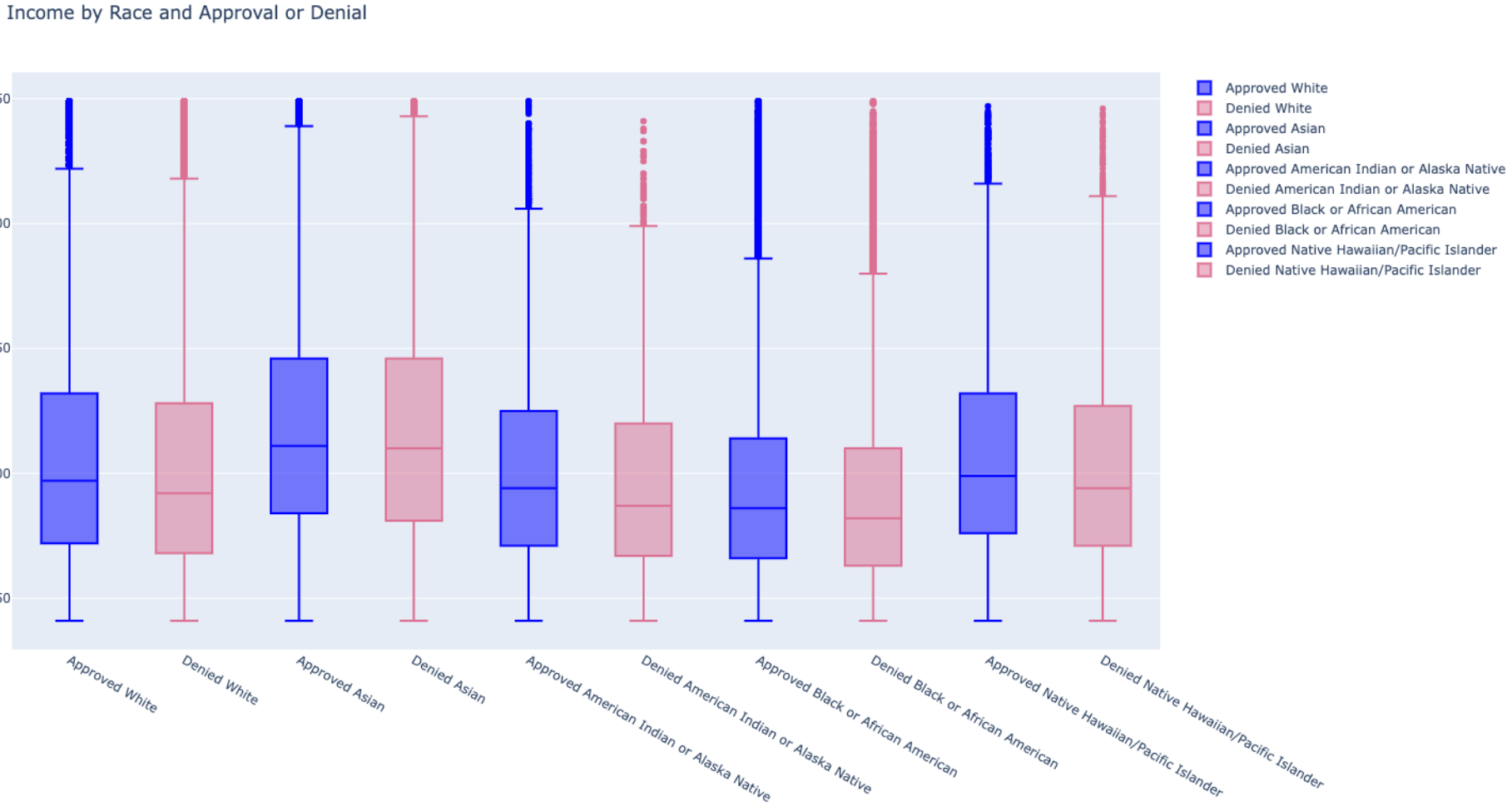


White Applicant Income < $300,000 by Approved/Denied Count

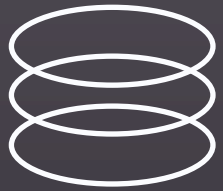



Loan Approval Percentage by Loan Type
Approved
Denied
Percentage
0
10
20
30
40
50
60
70
80
78.5%
21.5%
77.3%
22.7%
71.3%
28.7%
78.7%
21.3%
Conventional
FHA-insured
FSA/RHS-guaranteed
VA-guaranteed
Loan Type

Minority Applicants vs Minority Population
Tampa, St. Petersburg, Clearwater - FL
San Diego, Carlsbad - CA
Orlando, Kissimmee, Sanford - FL
Oakland, Hayward, Berkeley - CA
New York, Jersey City, White Plains - NY, NJ
Nassau County, Suffolk County - NY
Miami, Miami Beach, Kendall - FL
Los Angeles, Long Beach, Glendale - CA
Jacksonville - FL
Houston, The Woodlands, Sugar Land - TX
Dallas, Plano, Irving - TX
Austin, Round Rock - TX
MSA/MD
0
10
20
30
40
50
60
70
80
Minority Percentage
Minority Applicants
Minority Population


Income by Race and Approval or Denial
250
200
150
100
50
Approved White
Denied White
Approved Asian
Denied Asian
Approved American Indian or Alaska Native
Denied American Indian or Alaska Native
Approved Black or African American
Denied Black or African American
Approved Native Hawaiian/Pacific Islander
Denied Native Hawaiian/Pacific Islander
Approved White
Denied White
Approved Asian
Denied Asian
Approved American Indian or Alaska Native
Denied American Indian or Alaska Native
Approved Black or African American
Denied Black or African American
Approved Native Hawaiian/Pacific Islander
Denied Native Hawaiian/Pacific Islander

04



MODEL PROCESSING

BASELINE SCORE

77.54%

DROPPING DATA LEAKAGE ITEMS

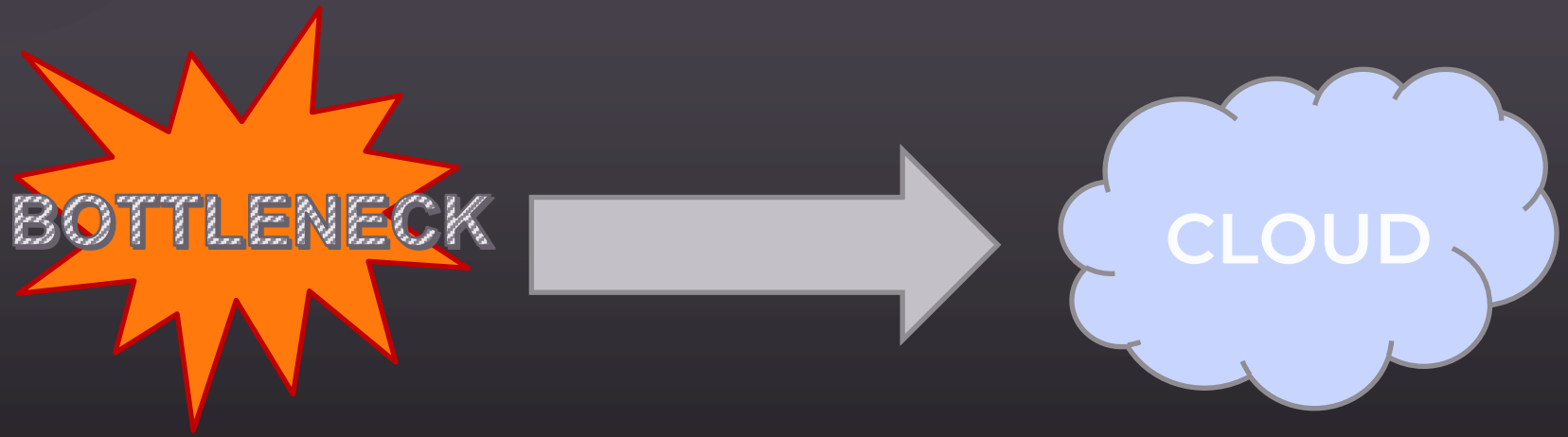
Denial Reason

Action Taken

Agency

Name Analog Filtering

PIPELINE TRANSFORM


BOTTLENECK
CLOUD

FEATURE SELECTION

	Coefficient	Value	abs_Val
35	msamd_name_Nassau County, Suffolk County - NY	0.671963	0.671963
16	agency_name_Federal Reserve System	0.591762	0.591762
37	msamd_name_Oakland, Hayward, Berkeley - CA	0.569985	0.569985
42	state_name_Florida	-0.390041	0.390041

- CONVERGENCE
- PCA: 15 Components (85%)
- REFIT: {Logistic, KNN, Random Forest}

05

RECOMMENDATIONS FOR FUTURE WORK

RECOMMENDATIONS

CONTINUE EXPLORATION	EXPAND DATA	MODELING & FEATURE ENGINEERING
Collect HMDA data from more metro areas	Use Census information	Exploring new modeling techniques to provide better insight and modeling accuracy

RESOURCES

- https://apnews.com/article/lifestyle-technology-business-race-and-ethnicity-racial-injustice-b920d945a6a13db1e1aee44d91475205
- https://www.consumerfinance.gov/