



# CAPSTONE PROJECT: PREDICTING PROBLEM MORTGAGE LOANS

MAGALI SOLIMANO

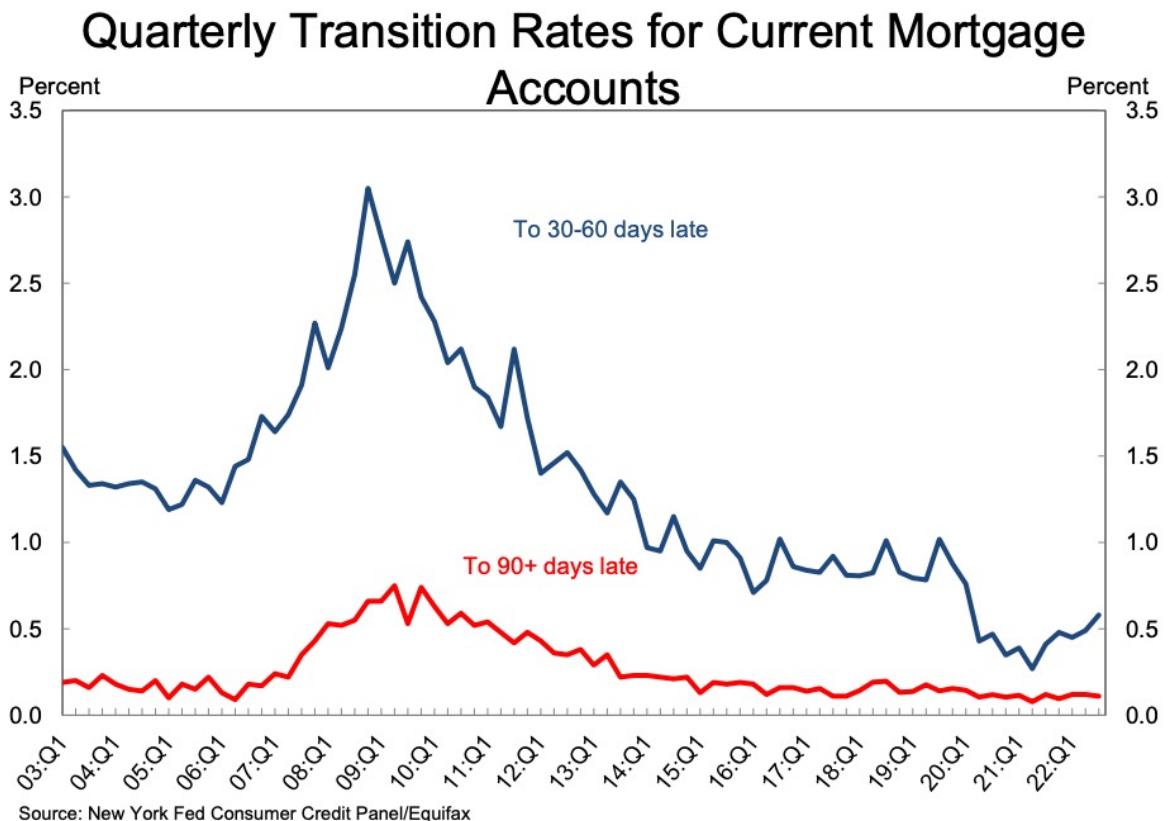
DECEMBER 2022

# TABLE OF CONTENTS

1. Problem Statement
2. Goal
3. Data Understanding
4. Results
5. Recommendations
6. Next Steps

# PROBLEM STATEMENT

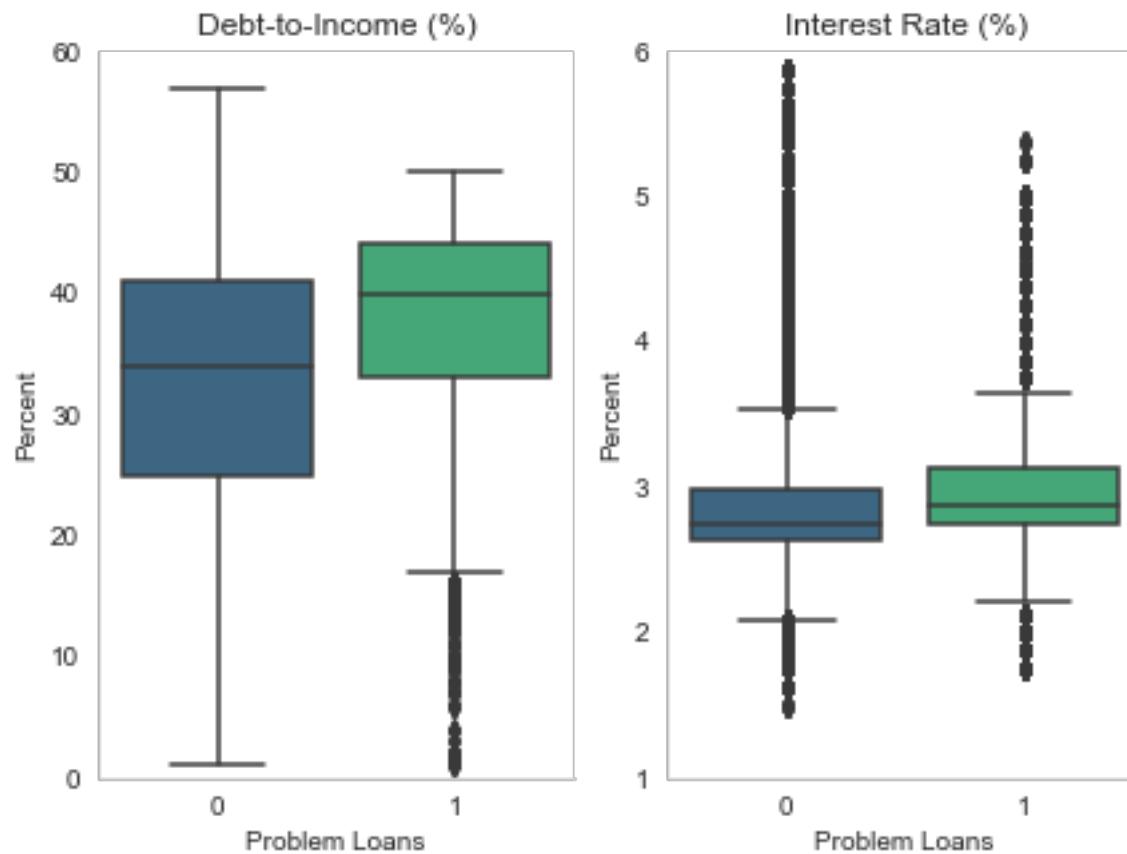
Fannie Mae wants to better predict the likelihood that a borrower will have problems paying its mortgage loan and to identify the features that best identify problem loans.



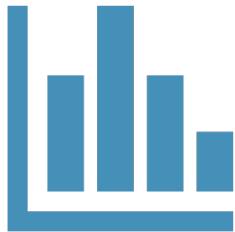
*While mortgage loan delinquencies are currently very low, there are indications of potential increasing stress.*

# GOAL

The goal is to predict the probability of problem loans and to identify their features in order to manage credit risk.



# DATA UNDERSTANDING



## Data

- Fannie Mae Single Family Loan Data
- Loans acquired in Q1 2021
- Over 23 million records



## Methods

- Descriptive data analysis and statistics
- Machine learning modelling

# DATA UNDERSTANDING

## Definitions

### Problem loans

payments that are 90 or more days past due

have been modified, restructured, or received any borrower assistance

have been foreclosed

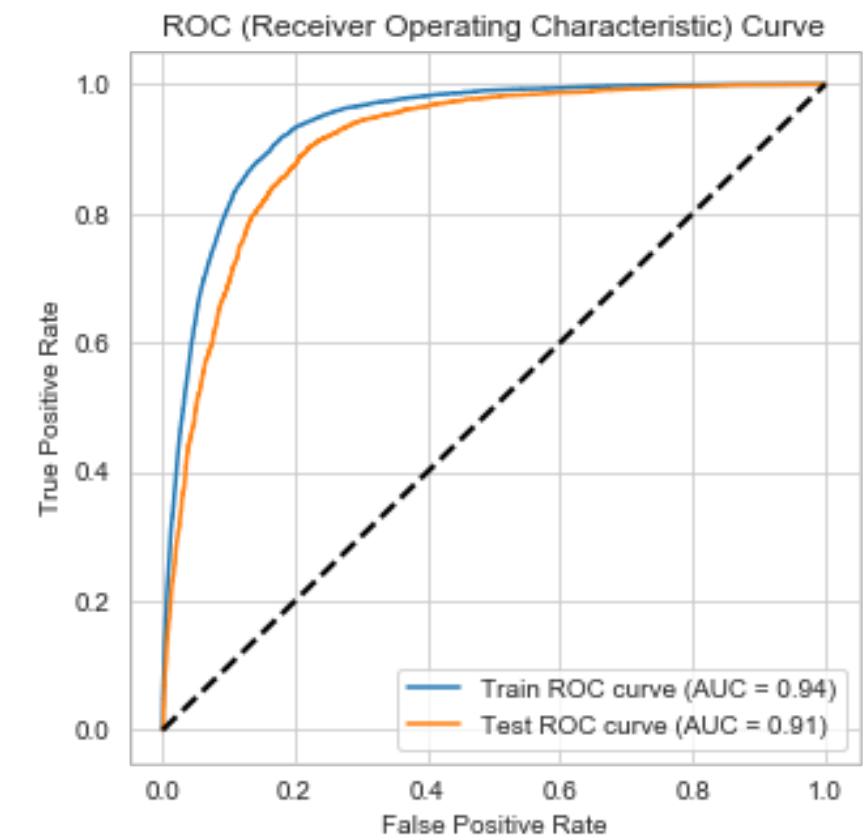
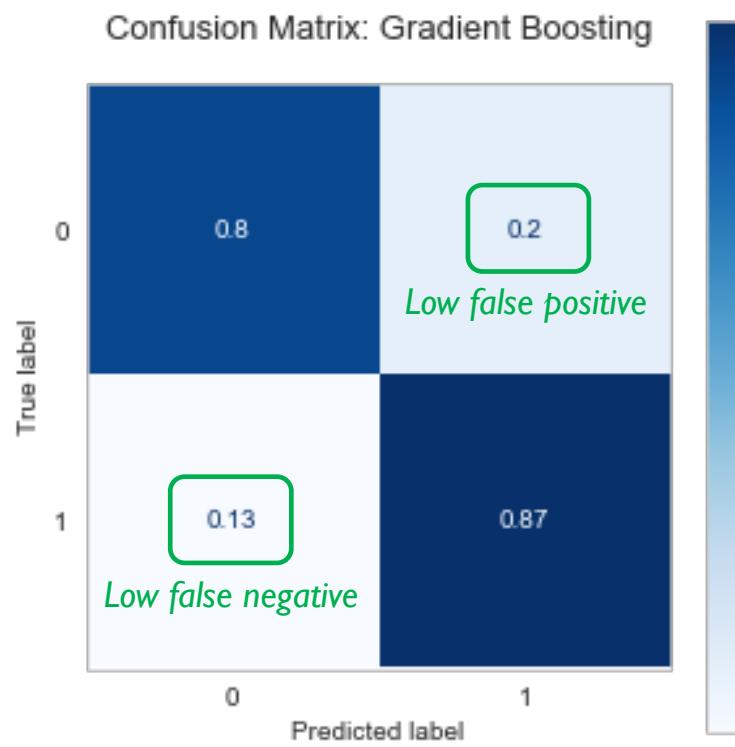
### Binary classification

non-problem, performing loans are classified as '0'

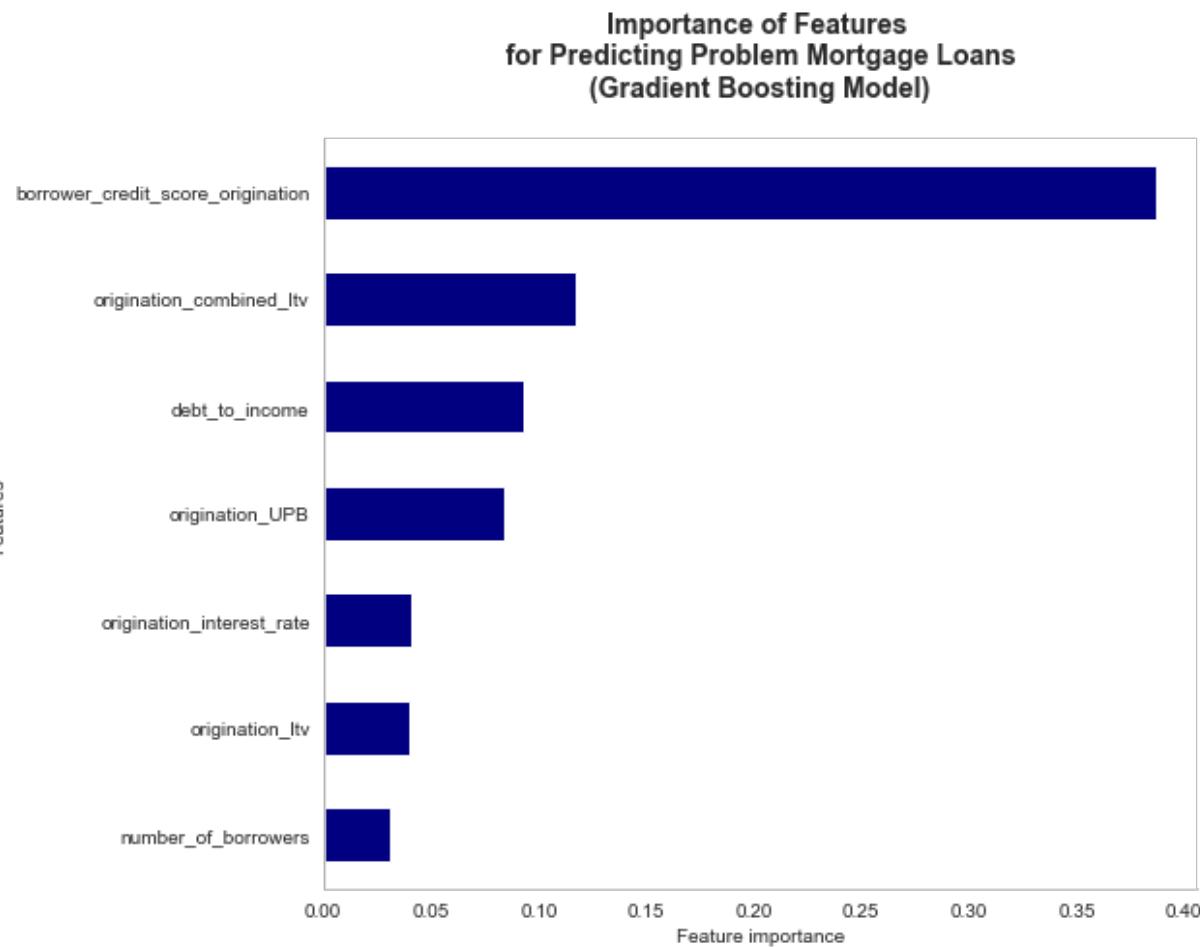
problem loans are classified as '1'

# RESULTS

- Gradient Boosting model selected as best performing model
- True positive rate of 87%



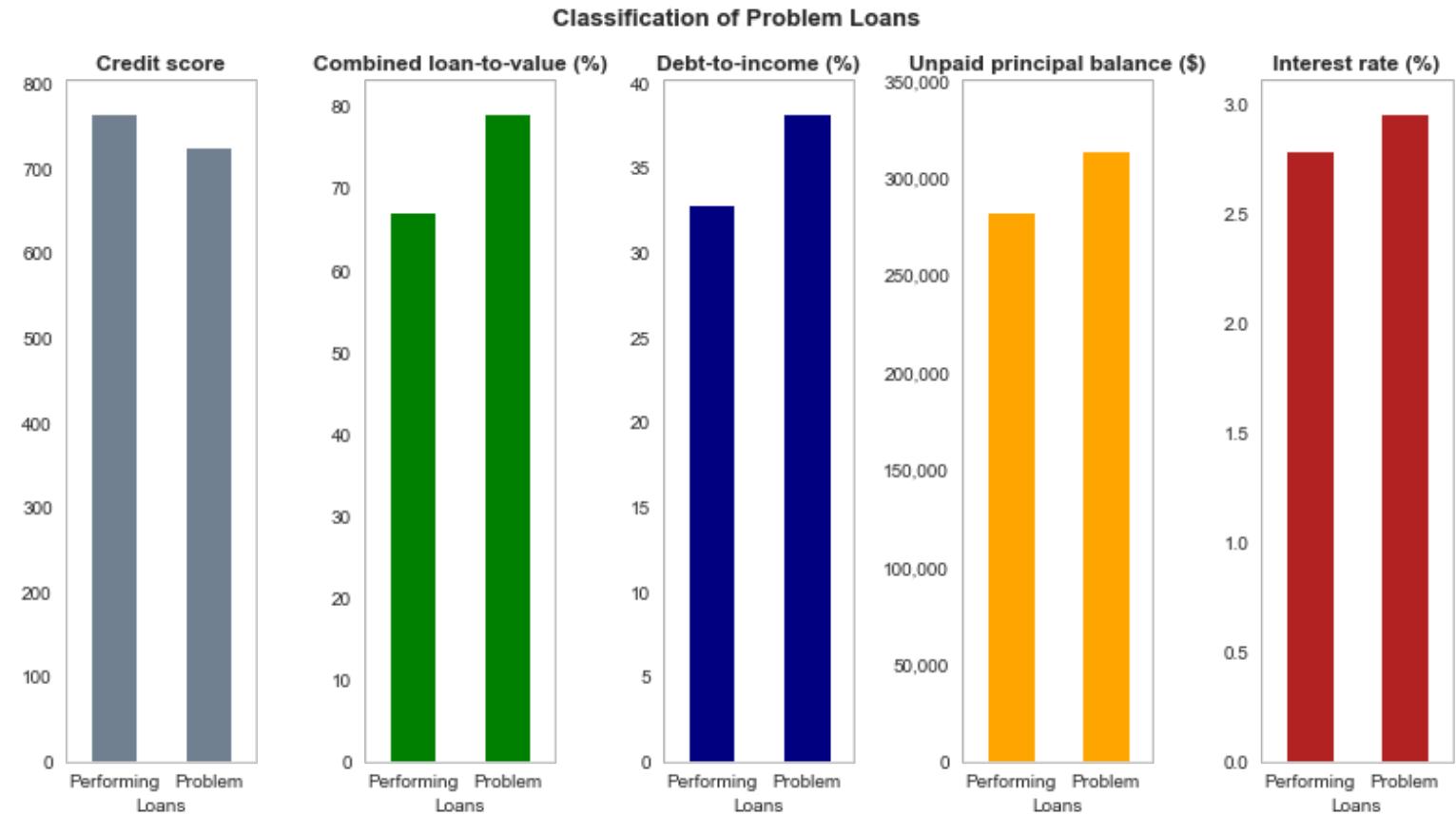
# RESULTS



- Top features represent metrics that are commonly used in credit risk assessments

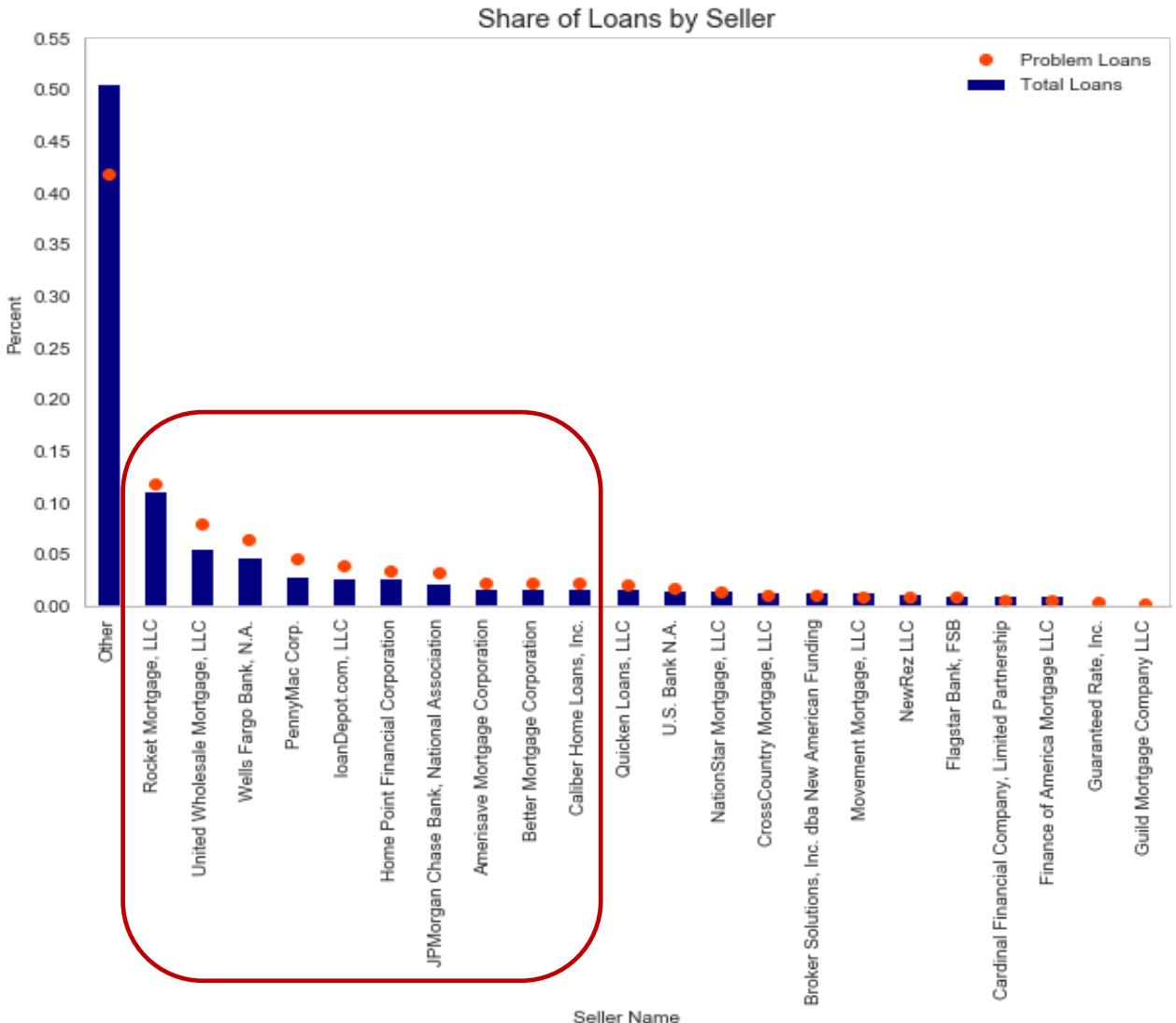
# RECOMMENDATIONS

- Adjust the criteria Fannie Mae applies to the loans they acquire and guarantee
- Focus credit risk management resources on loans that are likely to be problem loans



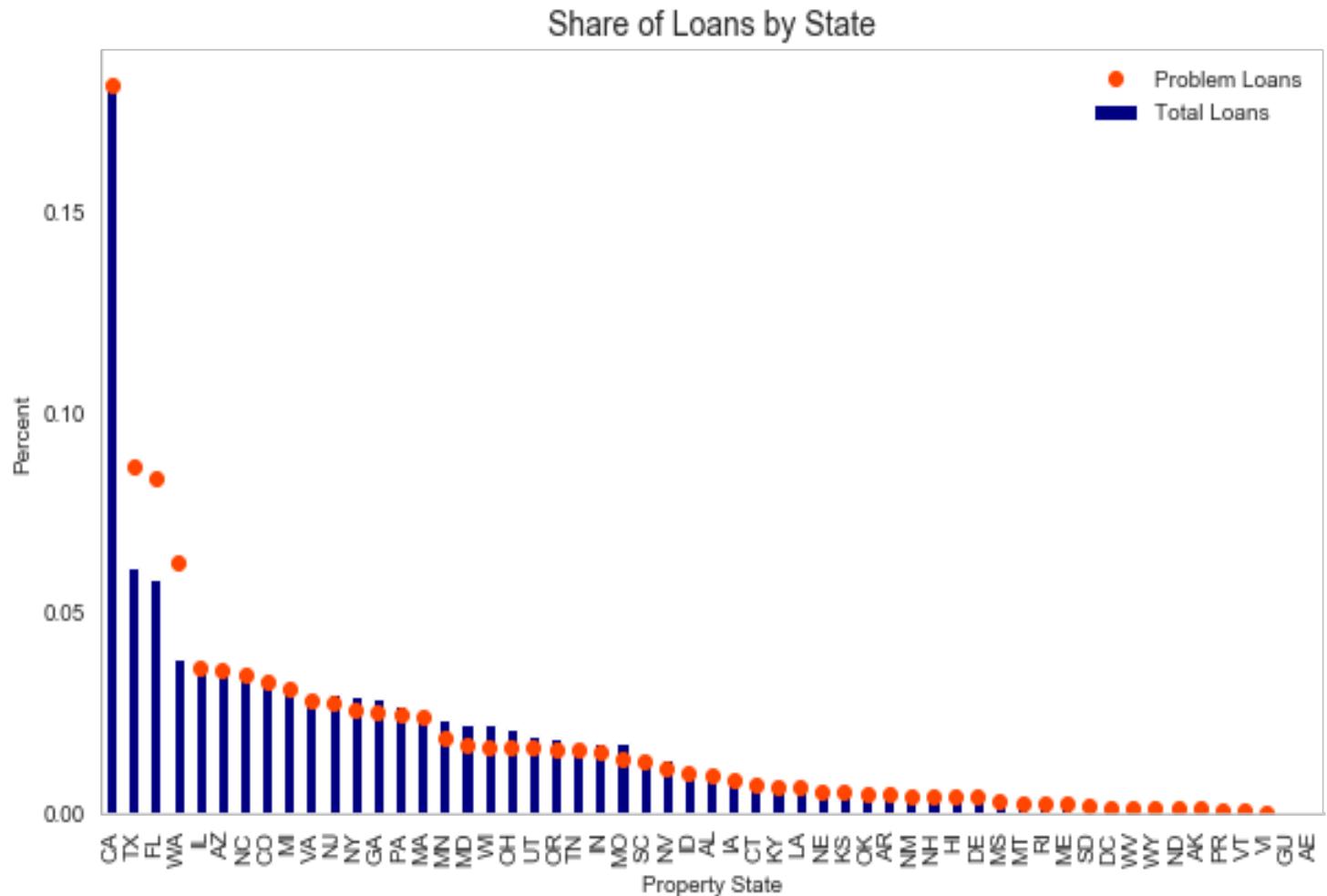
# RECOMMENDATIONS

- Mortgage loans sold by some sellers should be examined more closely for problem loans compared to other sellers



# RECOMMENDATIONS

- Mortgage loans for properties in certain states should also be examined more closely for problem loans compared to other states



## NEXT STEPS: EXTENDED ANALYSIS



Improve model performance



Extend geographical analysis  
to zipcode level



Predict value of losses for  
problem loans

# THANK YOU

***For questions, contact:***

Magali Solimano

[magali.solimano@gmail.com](mailto:magali.solimano@gmail.com)