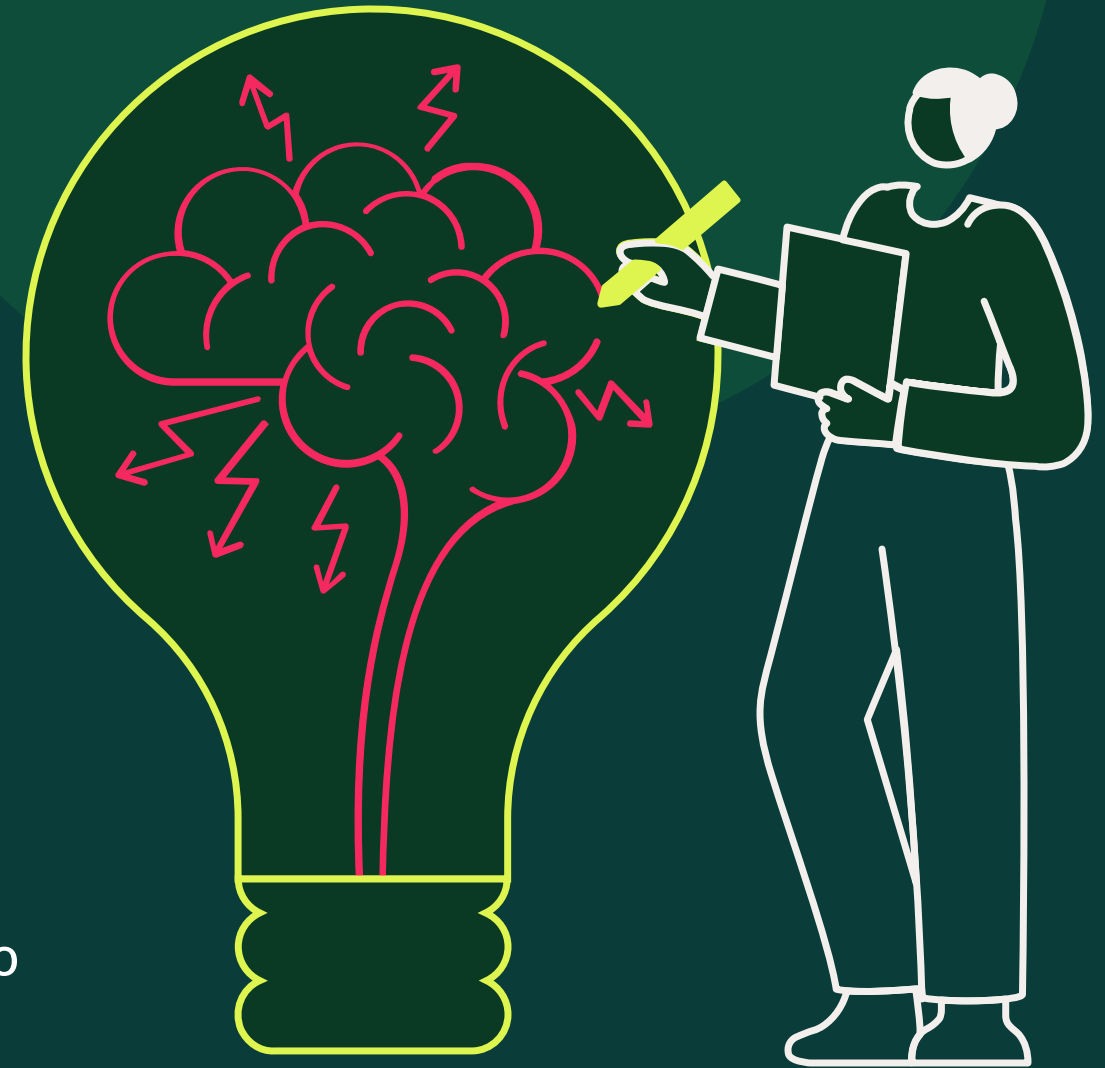


# Business Framing Document

Davya Vuyyuru, Devni Shah, Chandan N A, Rishi Rao



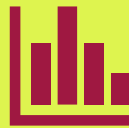
# Agenda



Business – Description,  
Objectives and  
Background



Success Criteria



Data



Risks and  
Contingencies



Stretch Goals and  
Plan

# Business Problem And Objectives



**Prepayment vs. Default Risks:** Freddie Mac must balance the competing risks of early loan payoffs and defaults, which directly affect financial stability.



**Economic Sensitivity:** Borrower behaviors vary widely with economic changes, requiring adaptable forecasting models.

## OBJECTIVES

- **Monthly Transition Rate Tracking:** Predict monthly mortgage status changes and minimize forecast errors.
- **Error Minimization:** Compare predicted vs. actual rates using root mean square error.
- **Period Analysis:** Evaluate performance across pre-COVID, COVID, and post-COVID phases.

# Business Background

- **Stakeholders:** Freddie Mac provides liquidity to mortgage lenders, allowing them to issue loans to borrowers, while investors rely on Freddie Mac's mortgage-backed securities for returns, and regulatory bodies oversee these processes to ensure stability and affordability in the housing market.
- **Resources Available:**
  1. Freddie Mac Loan-level Dataset  
<https://www.freddiemac.com/research/datasets/sf-loanlevel-dataset>
  2. Freddie Mac House Price Index Data  
<http://www.freddiemac.com/research/indices/house-price-index.page>

# Problem History

- **California's Housing Volatility:** Sharp price and foreclosure shifts from tech changes, regulations, and economic cycles demand adaptable models.
- **Prepayment vs. Default Risks:** High property values drive prepayments; downturns increase defaults, impacting Freddie Mac's stability.
- **Crises Reveal Vulnerabilities:** Events like 2008 and COVID-19 exposed risks of sharp price drops and increased foreclosures in California, showing the need for resilient, tailored models.

# Success Criteria

## Minimizing Forecasting Errors

The primary success metric is reducing the monthly average RMSE between the predicted and actual transition rates of mortgages. Smaller errors indicate a more reliable model.

## Pre-COVID, COVID, and Post-COVID Period Analysis

Success also involves demonstrating how the model performs across these periods, ensuring the model adapts to changing conditions.

# Data Requirements

## 1. Freddie Mac Loan-Level Dataset

Contains detailed loan-level data for mortgages purchased by Freddie Mac.

Key Variables shown on the right.

## 2. Freddie Mac House Price Index

Provides house price indices at national, state, and MSA levels.

Key Variables: Year, Month, Geo\_Name, Index\_NSA, Index\_SA

## 3. Federal Reserve Economic Data

Provides the data for the underlying economic drivers throughout the years.

Key Variables: GDP, unemployment rates, inflation rates, etc.

Sl No	Column Name
1	CREDIT SCORE
2	FIRST PAYMENT DATE
3	FIRST TIME HOMEBUYER FLAG
4	MATURITY DATE
5	METROPOLITAN STATISTICAL AREA (MSA) OR METROPOLITAN DIVISION
6	MORTGAGE INSURANCE PERCENTAGE (MI %)
7	NUMBER OF UNITS
8	OCCUPANCY STATUS
9	ORIGINAL COMBINED LOAN-TO-VALUE (CLTV)
10	ORIGINAL DEBT-TO-INCOME (DTI) RATIO
11	ORIGINAL UPB
12	ORIGINAL LOAN-TO-VALUE (LTV)
13	ORIGINAL INTEREST RATE
14	CHANNEL
15	PREPAYMENT PENALTY MORTGAGE (PPM) FLAG
16	AMORTIZATION TYPE
17	PROPERTY STATE
18	PROPERTY TYPE
19	POSTAL CODE
20	LOAN SEQUENCE NUMBER
21	LOAN PURPOSE
22	ORIGINAL LOAN TERM
23	NUMBER OF BORROWERS
24	SELLER NAME
25	SERVICER NAME
26	SUPER CONFORMING FLAG
27	PRE-RELIEF REFINANCE LOAN SEQUENCE NUMBER
28	PROGRAM INDICATOR
29	RELIEF REFINANCE INDICATOR
30	PROPERTY VALUATION METHOD
31	INTEREST ONLY INDICATOR (I/O INDICATOR)
32	MI CANCELLATION INDICATOR

# Assumptions

1. Historical loan-level data is complete and accurately represents loan performance and status transitions.
2. Borrower behavior remains relatively consistent within each defined period.
3. External economic factors influencing borrower decisions are adequately captured in the dataset.
4. Data quality is sufficient for analysis, with minimal missing values and inconsistencies.

# Limitations

1. Changes in borrower behavior and economic conditions during COVID-19 may not be fully reflected in the data.
2. The impact of forbearance programs may not be fully reflected in the dataset.
3. Unforeseen external factors like regulatory changes that occurred during the specified periods may not be captured in the dataset.

# Risks and Contingencies

<b>Data Quality Issues</b>	<p>Missing or incomplete loan performance data may lead to incorrect predictions.</p> <p><b>Mitigation:</b> Implement data cleaning and validation techniques to handle gaps in the dataset.</p>
<b>Model Overfitting</b>	<p>If the model fits the training data too well, it may not generalize to the testing period (COVID and post-COVID).</p> <p><b>Mitigation:</b> Use cross-validation and regularization techniques and find new variables that can improve prediction for these time periods.</p>
<b>Effects of Covid-19</b>	<p>The effect of COVID-19 on the economy may not be significantly represented in the data, potentially leading to inadequate model predictions for the COVID period.</p> <p><b>Mitigation:</b> Analyze the data to identify any gaps related to the pandemic's impact and consider integrating external data sources to better capture these effects.</p>
<b>Regulatory Changes</b>	<p>Policy changes during the pandemic, both short and long term, would have affected borrower behavior.</p> <p><b>Mitigation:</b> Stay informed about regulatory developments during this period and make changes to the model accordingly.</p>



# Stretch Goals

## Alternative Data Sources

Explore using Google Trends or financial news sentiment data to improve predictions, as this could capture public sentiment about economic conditions or housing markets.

## Advanced Machine Learning Models

Investigate deep learning approaches to improve accuracy.

## Real-Time Forecasting

Implementing a model that updates predictions in real-time as new performance data becomes available could provide immediate insights to stakeholders.

# Long Term Plan



**Thank you**

