

# **Project Title: Mortgage Insights: An Exploratory Data Analysis (EDA)**

Group:

- Jose Vela
- Ryan Eikanger
- Mark Olson
- Krishna Pandari

## **Description:**

Think of all the people out there paying more interest on their mortgages than they need to. They might not even realize it! And what about those who could save money by refinancing? Their stories are waiting to be told.

The Mortgage Insights EDA project aims to explore mortgage data from various public sources, including the Consumer Financial Protection Bureau (CFPB), Freddie Mac, and Fannie Mae. By analyzing this rich dataset, we seek to answer critical questions related to mortgage rates, customer demographics, and refinancing opportunities. Our goal is to empower homeowners, lenders, and policymakers with actionable insights and save money.

## **Project Objective:**

The objective of this EDA project is to analyze mortgage data from various sources (such as Consumer Financial Protection Bureau (CFPB) HMDA Data, Freddie Mac, Fannie Mae, etc.) to gain insights into mortgage rates by different customer classes, loan types, and demographics.

## **Project Goals, Scope, and Questions to Solve:**

### **1. Interest Rate Disparities:**

- **Question:** Are specific customer segments paying significantly higher interest rates than the current market average?
- **Goal:** Identify disparities and understand the factors contributing to these variations.
  - i. Explore whether certain customer segments (e.g., first-time homebuyers, refinancers, specific income groups) are paying significantly higher interest rates than the current market average.
  - ii. Investigate variations in interest rates by loan type (fixed-rate vs. adjustable-rate) and loan purpose (purchase vs. refinance).

2. Demographic Patterns:
  - **Question:** How do demographics impact mortgage rates?
  - **Goal:** Explore interest rate differences based on age, gender, ethnicity, and income level.
  - Identify any disparities in interest rates across different states or regions.
3. Refinancing Opportunities(Optional):
  - **Question:** Who could benefit from mortgage refinancing to save money on interest payments?
  - **Goal:** Highlight potential candidates/classes for refinancing (and quantify potential savings).
    - i. Individuals paying higher interest rates than the prevailing market rates.
    - ii. Those who could benefit from refinancing to save money on interest payments.

### **Expected Outcomes:**

- Clear visualizations illustrating interest rate distributions.
- Insights into demographic patterns affecting mortgage rates.
- Re-financing opportunities (Optional)

By addressing these questions, we aim to enhance financial literacy, promote informed decision-making, and contribute to a more equitable mortgage market.

### **Identify Data Sets:**

1. Consumer Financial Protection Bureau (CFPB) - Mortgage Data HMDA (Home Mortgage Disclosure Act)
  - a. <https://www.consumerfinance.gov/data-research/hmda/>

If/as needed:

2. CFPB also has the data to find mortgage rates and trends by county.
3. Interest rates historical and forecast are available at
  - a. <https://data.oecd.org/interest/long-term-interest-rates-forecast.htm>
  - b. <https://home.treasury.gov/policy-issues/financing-the-government/interest-rate-statistics>
4. Demographics data by country, zip code is available at <https://www.census.gov/data.html>
5. MN Demographic data and projections <https://mn.gov/admin/demography/data-by-topic/population-data/our-projections/>

### **Data Loading and Initial Exploration**

- Code data loading through API
- Test Data
- Load the dataset into a Python environment (e.g., using Pandas)
- Create data dictionary
- Identify the columns needed for the project
- Handle any missing values or duplicates

### **Descriptive Statistics:**

- Calculate summary statistics (mean, median, standard deviation) for numeric columns like loan amount, interest rate, and loan-to-value ratio.
- Explore the distribution of interest rates, loan amounts, and other relevant features

### **Data Visualization:**

- Create visualizations (histograms, scatter plots, bar charts) to understand the distribution of interest rates, loan amounts, and other variables.
- Plot interest rates against loan amounts to identify patterns.
- Visualize demographic patterns (e.g., ethnicity, race, sex) using bar charts, pie charts, and box plots.

### **Comparative Analysis:**

- Compare interest rates across different customer segments (e.g., first-time homebuyers vs. refinancers).
- Analyze interest rate disparities by loan type (conventional, FHA, VA) and purpose (purchase, refinance).

### **Geospatial Analysis:**

- Explore interest rates by county (using county codes) in MN State.
- Map interest rates geographically to identify regional variations.

### **Correlation Analysis:**

- Investigate relationships between loan amount, interest rate, and loan-to-value ratio.
- Calculate correlation coefficients to identify/understand associations.

**Insights and Interpretation:**

- Summarize key findings from your analysis. Highlight any significant trends, disparities, or opportunities.

**Refinancing Analysis (Optional):**

- Identify potential candidates for refinancing based on interest rates.
- Quantify potential savings for those who could benefit from refinancing.

**Answer Questions Related to Project Objective:**

- Address any specific questions related to your project's goals.
- Interest Rate Disparities:
  - **Question:** Are specific customer segments paying significantly higher interest rates than the current market average?
  - **Goal:** Identify disparities and understand the factors contributing to these variations.
    - i. Explore whether certain customer segments (e.g., first-time homebuyers, refinancers, specific income groups) are paying significantly higher interest rates than the current market average.
      - 1. first-time homebuyers/purchase vs refinancers (loan\_purpose vs interest\_rate)
      - 2. specific income groups
- Geo
  - **Question:** How do demographics impact mortgage rates?
  - **Goal:** Explore interest rate differences based on age, gender, ethnicity, and income level.
    - i. Age
    - ii. Gender
    - iii. Ethnicity
    - iv. Income level
  - Identify any disparities in interest rates across different Counties or regions in MN
- Refinancing Opportunities(Optional):
  - **Question:** Who could benefit from mortgage refinancing to save money on interest payments?

- **Goal:** Highlight potential candidates/classes for refinancing (and quantify potential savings).
  - i. Individuals paying higher interest rates than the prevailing market rates.
  - ii. Those who could benefit from refinancing to save money on interest payments.

### **Documentation and Reporting:**

- Create a Jupyter notebook or report detailing your EDA process. Include code snippets, visualizations, and explanations.

### **Presentation (Group - 5/16 - 20 minutes):**

Include visualizations, code snippets, and explanations to share insights with the groups and stakeholders.

- Original Objective - 1 min: Krishna
  1. Our project aims to explore and analyze mortgage data from various sources (CFPB, Census.gov, MN.gov, Freddie Mac, Fannie Mae, et.c, ) to answer critical questions related to mortgage rates, customer demographics, and refinancing opportunities.
- Key Questions - 1 min: Krishna
  1. Are specific customer segments paying significantly higher interest rates than the market average?
  2. How do demographics impact mortgage rates?
  3. Who could benefit from mortgage refinancing(Optional)?
- Data Sources - 2 minutes: Ryan
- Approach - 5 minutes: Ryan
- Key Findings - 5 minutes:
  1. Interest Rate Disparities:

- First-time buyers/Purchase vs Refinance - first time buyers pay higher rates - Mark
  - Lending institute 549300CN50N3250U7V79 charging an Avg 12.03 interest rate vs Lending institute 549300JGMQJ4R419LR70 at 1.96 - Jose
  - Loan product type - Jose
- 2. Demographic Insights: Jose
  - Age and income impact rates.
  - Age Group A pay more vs Age Group B
  - Ethnicity A pay more vs Ehtinicity B
  - Regional variations exist (Mark)
- 3. Refinancing Opportunities (Optional): Krishna
  - Candidates paying above market rates.
  - Potential savings through refinancing.
- Learnings - 2 minutes: Ryan
  1. Understanding mortgage dynamics is crucial.
  2. Data-driven decisions can lead to cost savings.
- Conclusion - 2 minutes: Ryan/Krishna
  1. Our insights pave the way for informed choices and have the potential to create a fairer mortgage market!
  2. Rate Trends: Keep an eye on rate trends. Economists expect 30-year mortgage rates to decline in 2024, which could create favorable refinancing conditions.
- Questions - 2 minutes: