Unison Home Ownership Investors

Quantitative Financial Analyst – Research and Data

Mini Project 2021

Summary:

The purpose of this mini project is to test candidates on programming, data science, finance and presentation skills. A candidate will be judged by the quality of work across all four dimensions.

Read through the following prompts. Code up your work in Python (Jupyter Notebook preferred) and prepare a PDF (no more than 4 pages) to present your results. Please work independently and refrain from asking questions or sharing your data. We want it to be fair and would like to see how you address an open-ended research project with incomplete information.

PLEASE SUBMIT (1) A PDF OF RESULTS & (2) YOUR CODE IN A ZIPPED FOLDER within a week from the time the project is shared with you.

Prompts:

1. Connect to the MySQL database which has a table of samples of the historical Freddie Mac single-family loan level data. See the Appendix for details.

2. We would like to analyze the breakdown of mortgage statuses (prepay, default, alive) for each origination cohort (group by year of first_pmt_date). Graph the count of mortgages in each status (stacked y-axis) by origination cohort (x-axis). Comment on the change in mortgage behavior through the mortgage origination cycle.

3. Repeat the analysis above but instead of grouping the mortgages by first_pmt_date, group them by age. In addition, briefly explain how you may go about building prepayment and default models for residential mortgages. What data will be required if we want to build such models for Unison home equity agreements (UEA)?

4. Determine the following ***unconditional*** probabilities of mortgage behavior based on the sample data and describe how you arrive at your results. How could prepayment or default behavior of the mortgage affect the Unison equity agreement (UEA) on the same home?

-That a mortgage has defaulted in the first 5 years

-That a mortgage has prepaid in the first 5 years

-That a mortgage remains alive after the first 5 years

5. Select either prepayment or default as a response. Select one or more factors (from the list of mortgage characteristics in the provided table) which you believe would affect the selected response and build a simple model to predict prepayment or default rates. Comment on how your selected response (prepayment or default) might affect the market price of a mortgage and how you could devise an investment strategy around this model's output. Use out-of-sample testing to validate your model. Feel free to highlight if you have additional thoughts on how you could further improve the model.

6. Based on your understanding of what Unison does and how a Unison home equity agreement (UEA) works, why do you think we need to monitor, analyze and understand the mortgage industry in general and our customers' mortgages in particular? Name key metrics that you believe are good indicators of risk and explain how you think we can use those metrics in risk management, investment structuring, etc.

Appendix:

MySQL Database

Host: miniproject-qfa.cef17qxjlavg.us-west-2.rds.amazonaws.com

Port: 3306

Username: newuser

Password: unisonim2.0

Schema: AgencyData

Table: CleanFreddieSample

Data Dictionary:

We have provided a derivative table with a condensed number of fields. The data dictionary provided below highlights those fields. If there is any ambiguity, refer to the latest official data dictionary http://www.freddiemac.com/fmac-resources/research/pdf/user_guide.pdf.

first_pmt_date	Date of first monthly mortgage payment. YYYYMM.
age	Time (in months) from origination to last recorded payment. If the mortgage has prepaid or defaulted, it is the time (in months) to termination.
status	Alive – The mortgage is still outstanding and is currently receiving monthly payments. Prepay – Homeowner has paid the full balance of principal on the mortgage. Mortgage is terminated. Default – Homeowner has failed to pay off the mortgage and is in default. Mortgage is terminated.
first_time_ho_flag	First-time homeowner flag Y – Yes N – No Blank – Not Available
msa_code	Code relating to the metropolitan statistical area
mi_pct	Mortgage insurance premium
credit_score	Credit score of the primary borrower.
num_units	Number of residential units in the building.
occupancy_status	O – Owner-occupied

	S – Second Home I – Investor Property
orig_cltv	Combined loan to value of the mortgage at origination. Loan to value represents the total amount of debt divided by the value of the home.
orig_dti	Origination debt-to-income. Debt to income is the total amount of monthly payments associated with the mortgage divided by the total monthly income of the borrower.
orig_upb	Original unpaid balance. The total amount of principal outstanding on the mortgage.
orig_ir	Interest rate on the mortgage.
prop_type	Property type: SF – Single family CO – Condo PU – PUD MH – Manufactured CP – Co-op 99 - Unavailable
state	State abbreviation
zip	First 3 digits of the zip code
loan_seq_num	Mortgage ID
loan_purpose	Purpose of loan: P – Purchase loan C – Cash out refinance N – Non-cash out refinance
num_borrowers	Number of distinct borrowers