

Rotman

**Master of
Management
Analytics**

MMA Datathon Fall'21

Team Data Pirates Presents Case Challenge

October 13, 2021 Prepared by Sonali Sachdev, Shulin Gu, Xinghui Li



Rotman School of Management
UNIVERSITY OF TORONTO



Business Problem

Analyzing the Nebraska Mortgage Loan Market

- To Maximize the dollar value while limiting the defaulting loan.
- To Identify the right customer with highest possibility to make repayment.
- Bank collaborated with Team Data Pirate to devise a recommendation to maximize the dollar value while keeping a check on interest rate and state of the economy.

**Here's
where it
changes.**



Co-Creating Solution with Great Lakes Midwest Bank

- Regional Bank in the state of Illinois, Iowa and Wisconsin

As we Know..

- It is imperative for banks to refinance mortgage and take preventive measure to reduce the defaulting loan.
- To target right customer at right time requires solid predictive analytics model and intrinsic recommendations.



Rotman School of Management
UNIVERSITY OF TORONTO

Data Source

Rotman

Data via Fannie Mae(government-sponsored enterprise)

- Dataset of all conventional single-family, fixed interest rate mortgages
 - From 2000 – 2021 Q1
- Analyzed data of more than 270000 customers
 - 43 Variables analyzed
 - 270k rows of transaction to trace the data pattern.
 - Credit Scoring Model

How We Helped the Bank Boost Customer Retentions **Rotman**

Key Challenges

The calculation of this data requires a large number of classification calculations based on the content of the data. Because of the uncertainty of the unpaid time and interest, it is necessary to calculate the actual received payment according to each actual situation.

Our strategy will calculate the actual received and expected receivables, and compare them, find the existing problems to carry out classification discussion.



Monthly Payment Equation:

$$M = P [i(1 + i)^n] / [(1 + i)^n - 1]$$

P = principal loan amount

i = monthly interest rate

n = number of months required to repay the loan

Actual Amount Received (Up to LAST_DTE):

- C, R, L = $M * (LAST_DTE - FRST_DTE)$
- $P = M * (LAST_DTE - FRST_DTE) + LAST_UPB$
- $D = M * (F_DTE - FRST_DTE)$
- T, F, S, N = $NET_LOSS + ORIG_AMT$

Note: Ignore amortization interest when payment is not received

Expected Amount Receivable = $M * ORIG_TRM$



Bank Success through Co-Creation

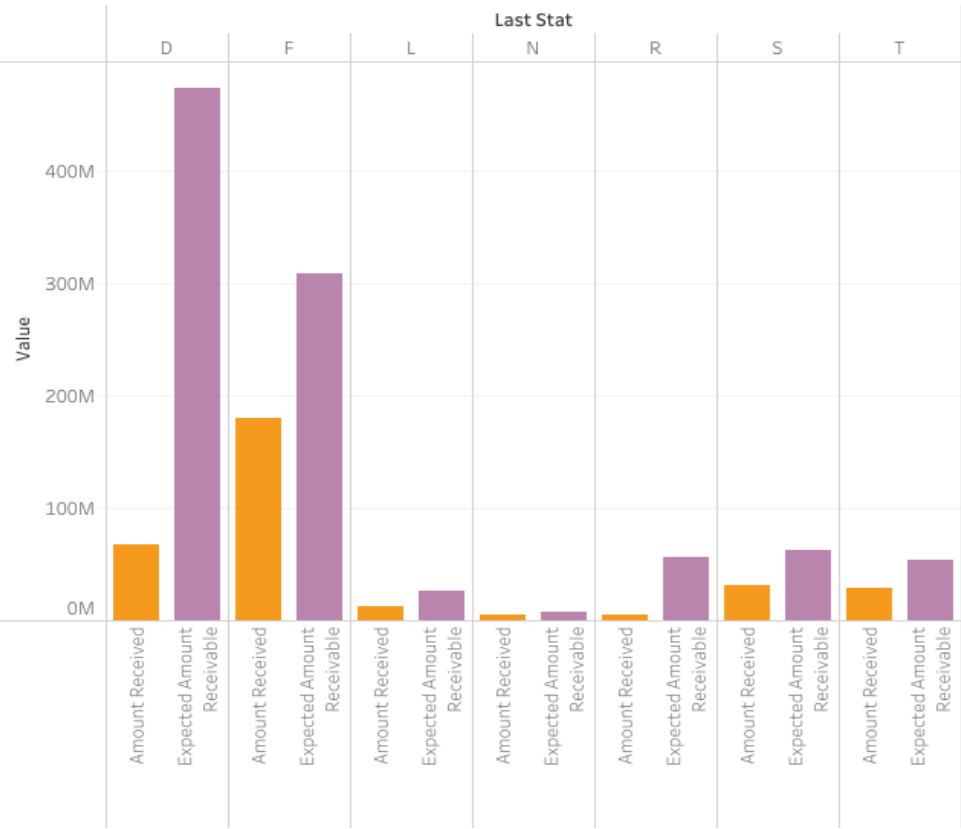
	orig_rt	LAST_RT	orig_amt	ORIG_VAL	LAST_UPB	orig_trm	oltv	ocltv	num_bo
Amount Received	0.056244	0.0545693	0.6670186	0.57980414	0.6344632	0.135483	0.116841	0.143781	0.179006
	dti	CSCORE_B	CSCORE_C	NUM_UNIT	msa	mi_pct	NET_LOSS	MI	MP
Amount Received	0.050492	0.0256393	0.0281798	-0.02498478	0.1468997	0.054203	0.552583	0.056244	0.665029

After the data has been cleaned and new term has been added, a correlation table was made to show the correlation of each term with the amount received. Higher value indicate stronger effect (positive or negative) on the amount received.

From the table, the Original Loan Amount (orig_amt), Last Unpaid Balance (LAST_UPB), Monthly Payment has a higher correlation. Original term (orig_trm), Original Loan to Value Ratio (OLTV) and number of borrowers (num_bo) has a relatively lower correlation. And we will be focusing more on these terms.

Amount Received v.s Expected Amount Receivable

Amount Received v.s Expected Amount Receivable



We compared the the actual amount received and expected amount receivable. We found there are huge different amount D, F & R.

When we go to in-depth analysis, we should pay more attention to these trading categories, and should find problems, why? For example, you can pay attention to what they do for the second mortgage.

With an increase in the number of borrowers, the amount of good loans (C & P) increase. And the amount of bad loan (D,T, L, F...) decrease. We can conclude that 2 or more borrowers will secure the loan better.

Number of Bo v.s Stat

Num Bo	Last Stat								
	C	D	F	L	N	P	R	S	T
Null						0			
1	29,838	1,075	774	59	21	63,806	93	77	181
2	40,419	667	650	55	14	131,109	120	94	92
3	220	11	2			348	1		
4	41					155			
6						2			
7						1			
8						1			

Addition Tables about Credit Score can be found on [github](#).

In the following analysis, we found that for most loans that are not good, they are used to buy new real estate. This may involve the fact that when they use it for investment, they can't get the rent back, which makes them unable to pay the loan.

And, for the loans processed by brokers, their loans are obviously better than those processed by other channels.

Purpose, Orig Chn v.s Stat

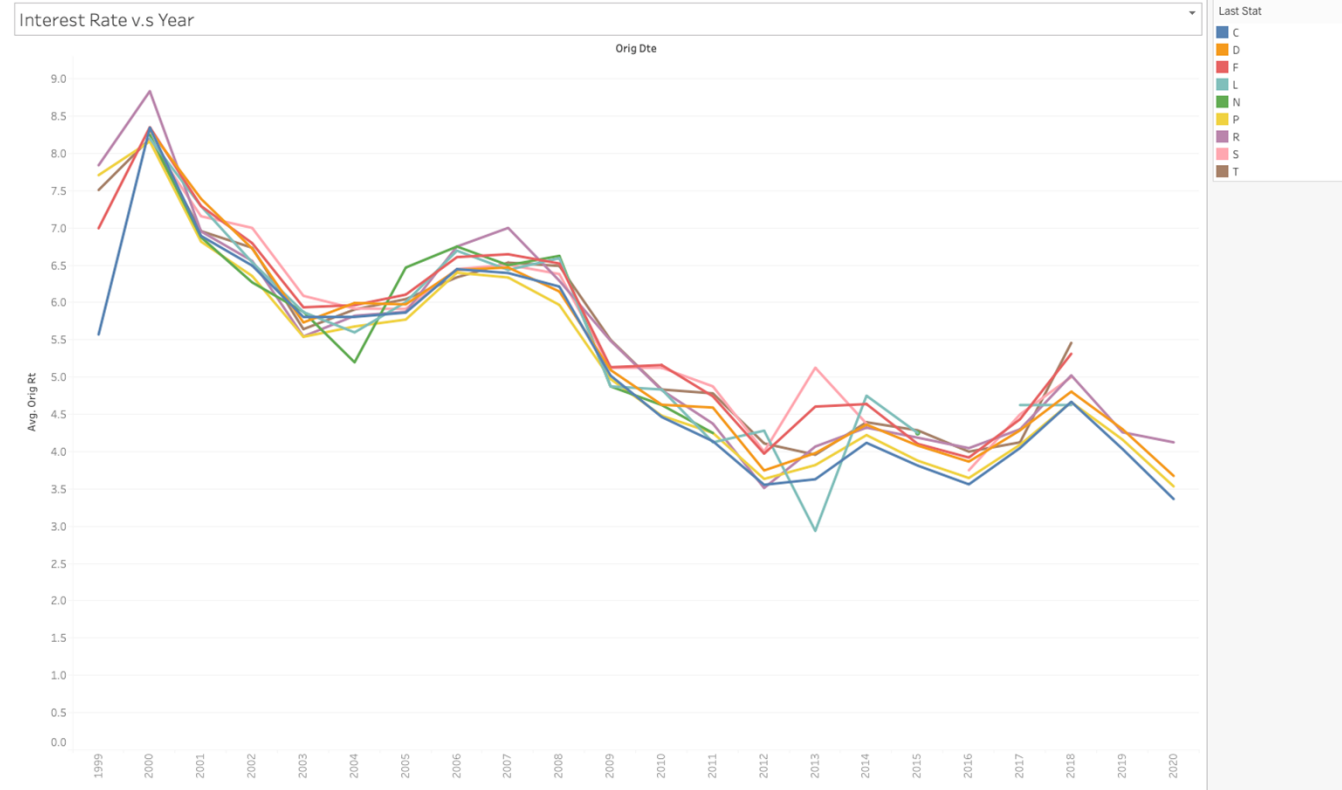
Purpose	Orig Chn	Last Stat								
		C	D	F	L	N	P	R	S	T
C	B	464	13	96	6	4	3,427	6	7	15
	C	5,919	153	224	19	7	22,979	26	19	41
	R	8,894	214	158	16	2	22,221	30	18	37
P	B	874	35	65	7	2	3,282	5	13	12
	C	16,526	588	221	20	6	37,288	44	25	49
	R	15,671	381	201	10	4	35,693	35	29	39
R	B	684	18	79	8	3	3,919	5	9	13
	C	10,367	183	247	14	4	36,716	34	28	36
	R	11,119	168	133	14	3	29,840	28	23	31
U	B						2			
	C			2			68	1		
	R						11			

It is a common sense that the higher the interest rate, the more money the bank will receive.

However, we cannot easily change the interest rate. And a relatively high interest might cause the borrower to pay off ahead of time, or even unable to afford the debt.

From the figure on the right, we can see the interest rate is descending throughout the year.

Thus increase the interest rate will be the last factor to be considered to maximize the profit.



Loan and Borrower Data in Avg

Rotman

From the table on the right shown, it turns out that C,P (good state of loan) have a lower Oltv and Dti value in average. And a higher average credit score.

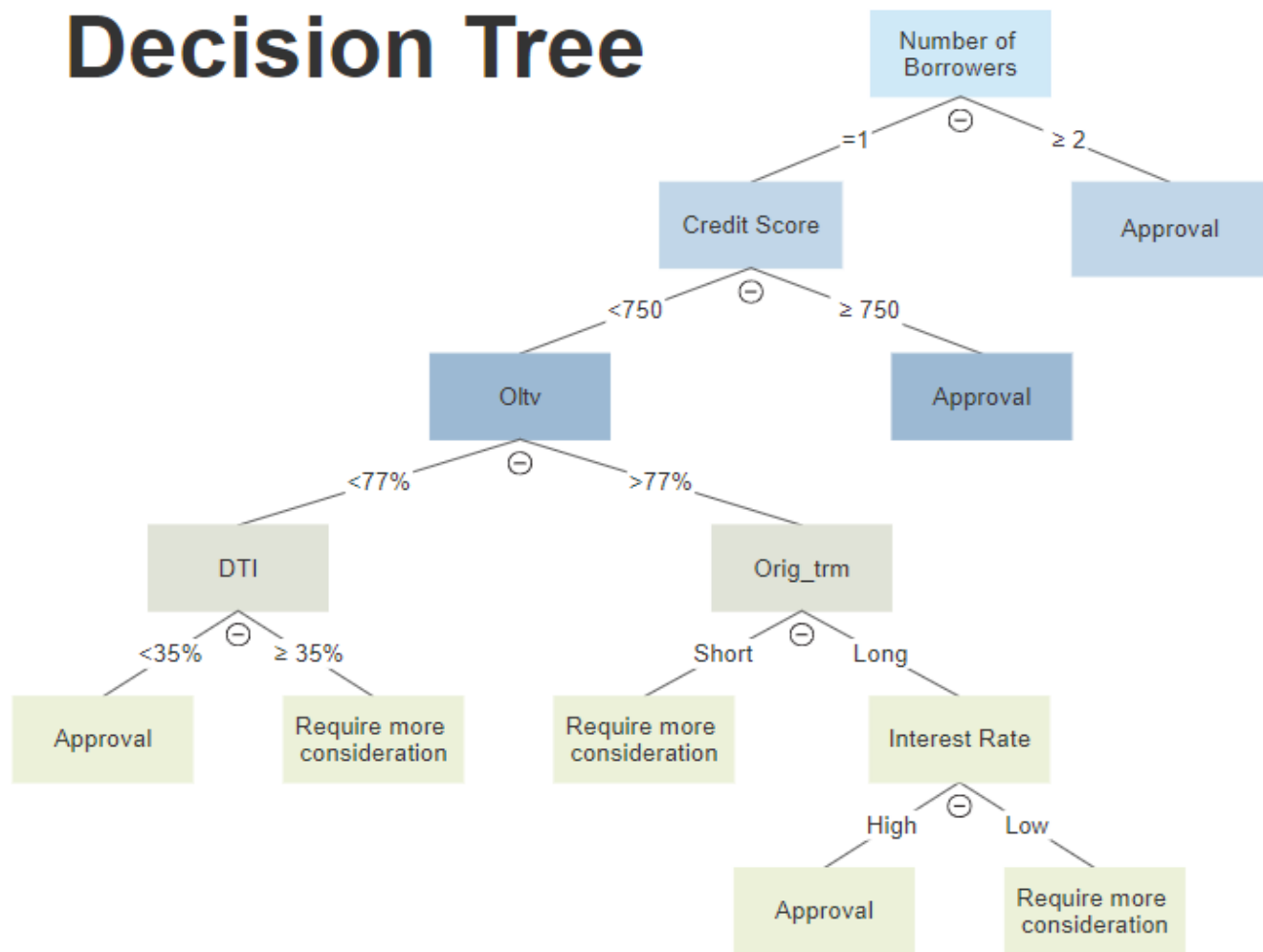
We can conclude for a loan in good status (C,P):

- A lower original amount is lent to the borrower compare to the original value of the property.
- The borrower has a relatively lower monthly income compare to the monthly debt.
- The borrower has higher credit score.

Last Stat	Avg. Orig Amt	Avg. Orig Val	Avg. Oltv	Avg. Dti	Avg. Cscore B	Avg. Cscore C
C	155,400	209,923	76	32	759	764
D	154,500	190,793	83	37	713	716
F	99,585	119,732	83	37	687	686
L	107,982	133,563	81	37	679	682
N	105,743	134,117	78	39	666	702
P	148,933	205,088	75	32	752	758
R	140,449	190,653	77	37	721	737
S	167,111	197,667	86	40	695	695
T	98,282	128,379	77	37	697	698

Decision Tree

Rotman



In order to maximize the dollar value of mortgage and limit the number of defaulting loans, we recommend to focus on the following segment:

Maximizing the dollar value of mortgage:

- Lower Oltv ratio

Limit the number of defaulting loans:

- Number of borrowers
- Higher credit score of borrowers
- Lower DTI ratio

Consideration for future interest rate:

- Interest rate is descending throughout the year.
- Higher interest rate for long term
- Lower interest rate for short term

	Recommended	Need More Consideration
Number of Borrowers	>1	=1
Credit Score	≥750	<750
Oltv	≥77%	<77%
DTI	≤35%	>35%
Orig_Trm	Long	Short
Interest rate	High	Low

Rotman

Thank you

For more information, visit Github Link: <https://github.com/ShulinGu/MMA-DATATHON>

**Here's
where it
changes.**