IOWA

Tippie College of Business

US Bank Mortgage Churn Analysis: Final Presentation



Raymund Boltron



Chayse Hawbaker



Matthew Hodges



Ramya Pandurangan



Xianxi Yang

FIN:9390 Applied Financial Analysis: Industry Project

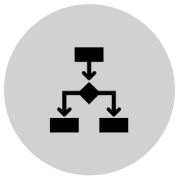
Presentation Agenda



REVIEW FROM THE MIDTERM



TALE OF TWO BORROWERS



METHODOLOGIES

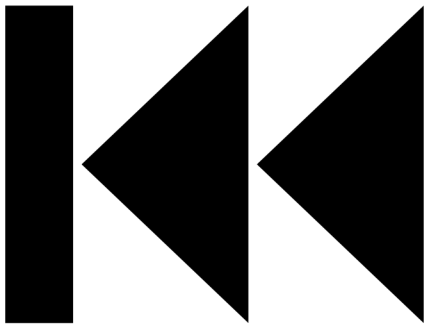


MODELS



SUMMARY & RECOMMENDATIONS

Review from the Midterm



The Problem

The Data

Refinance Retention

The Problem

How likely are borrowers to be retained after their initial payoff from US Bank?

What factors can US Bank target in order to retain customers while minimizing retention costs?



Data

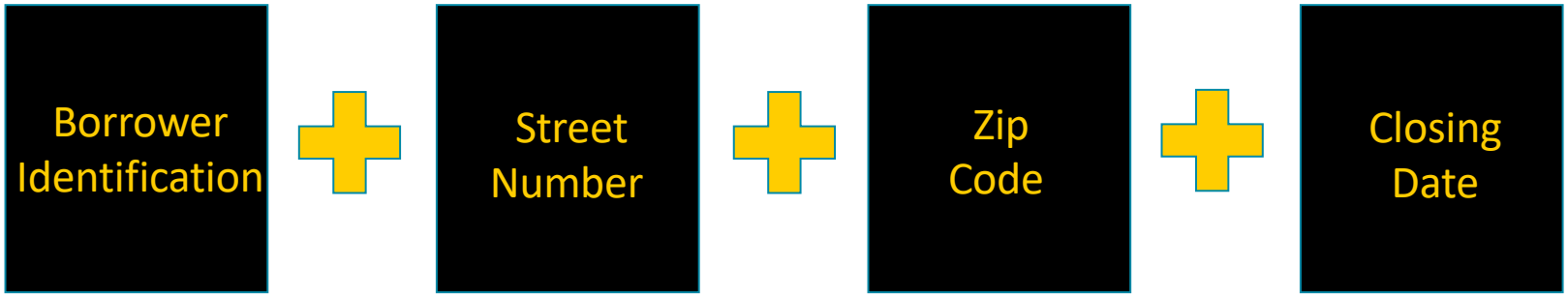
"Payoffs" Dataset

Time-series data for loans within the servicing portfolio

"All" Dataset

A record of all loans ever serviced by US Bank

Unique Retention Identification



Refinance Calculations

Retention rate by borrower and property

13.35%

56,427 out of 422,713

Retention rate by unpaid balance

14.01%

12.15 billion out of 86.79 billion

A Tale of Two Borrowers



Given that borrowers and loans have different characteristics, can we analyze those features in such a way as to predict who will be retained by US Bank?

Two Borrowers: Tom and Nick



Two-loan Tom



No-way Nick

Loan Characteristics



Two-loan Tom



No-way Nick

Unpaid balance
Annual rate
Original loan amount
Remaining balance %
Loan age

Loan Characteristics

UPB: $50,000
Annual rate: 5.4%
Original amt: $150,000
Remaining: 33.33%
Loan age: 3650 days



Two-loan Tom

UPB: $150,000
Annual rate: 3.4%
Original amt: $300,000
Remaining: 50.00%
Loan age: 4150 days



No-way Nick

Unpaid balance
Annual rate
Original loan amount
Remaining balance %
Loan age

Delinquency Status



Two-loan Tom



No-way Nick

Current

30 days past due

60 days past due

90 days past due

120 days past due

Delinquency Status



Two-loan Tom



No-way Nick

Current
30 days past due
60 days past due
90 days past due
120 days past due

Delinquency Status



Two-loan Tom



No-way Nick

Current

30 days past due

60 days past due

90 days past due

120 days past due

Origination Channel



Two-loan Tom



No-way Nick

Retail
Wholesale
Bulk transfer
Correspondent
Consumer direct

Origination Channel



Two-loan Tom



No-way Nick

Retail
Wholesale
Bulk transfer
Correspondent
Consumer direct

Origination Channel



Two-loan Tom



No-way Nick

Retail
Wholesale
Bulk transfer
Correspondent
Consumer direct

Economic Data (time sensitive)



Two-loan Tom



No-way Nick

Housing starts

MBA 30Y rate

Home price index

Home vacancy rate

New home sales

Economic Data (time sensitive)

January 2020
Housing starts: 1589
MBA 30Y rate: 3.71
HPI: 446.22
Home vacancy: 1.4
New home sales: 756



Two-loan Tom

April 2021
Housing starts: 1514
MBA 30Y rate: 3.18
HPI: 484.26
Home vacancy: 0.9
New home sales: 796



No-way Nick

Housing starts
MBA 30Y rate
Home price index
Home vacancy rate
New home sales

State-based Characteristics



Two-loan Tom



No-way Nick

State - home address
Number of USB loc.
Unemployment rate
Education rate
Poverty rate

State-based Characteristics

State: Wyoming
# of USB locations: 15
Unemployment rate: 6.1%
Education rate: 26%
Poverty rate: 15%



Two-loan Tom

State: California
# of USB locations: 619
Unemployment rate: 5.1%
Education rate: 35%
Poverty rate: 12%



No-way Nick

State - home address
Number of USB loc.
Unemployment rate
Education rate
Poverty rate

Two Borrowers: Significant Features



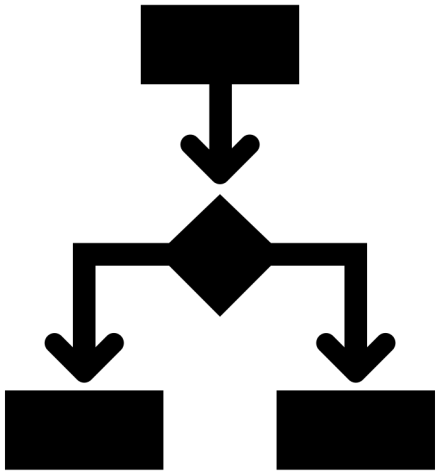
Two-loan Tom



No-way Nick

Loan Characteristics
Delinquency Status
Origination Channel
Economic Data
State-based Features

Methodology



Data Preparation

Feature Selection

Three Predictor Models

Scaling and Derived Variable

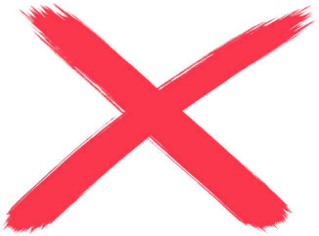


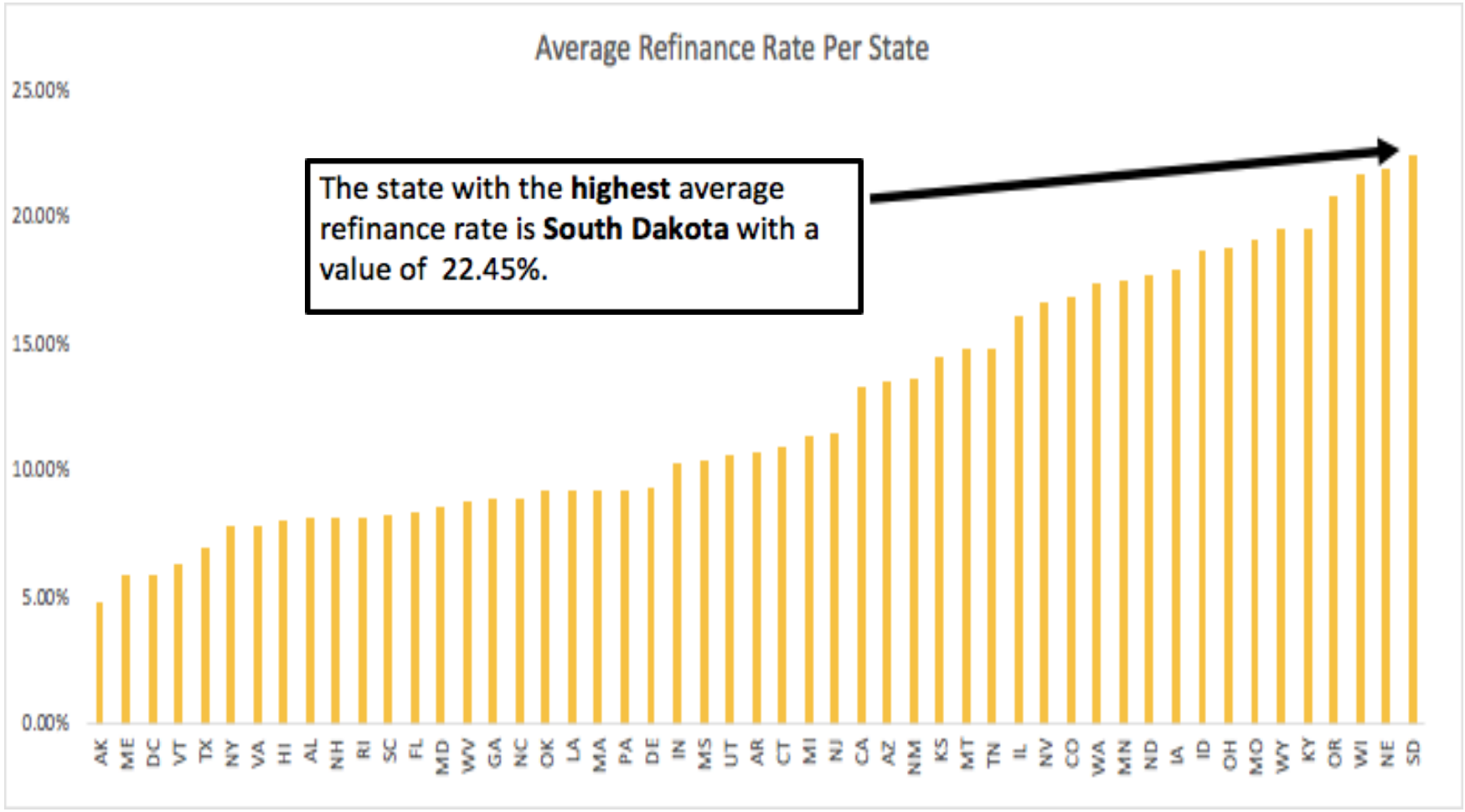
- **Loan Age In Years** = Date - Origination Closing Date
  - Scaled by years
- **Unpaid Mortgage Balance** - scaled by 100,000
- **State ratio** of origination channels
- ****Potential Savings** = (Interest Rate-Current Interest Rate)*Unpaid Balance

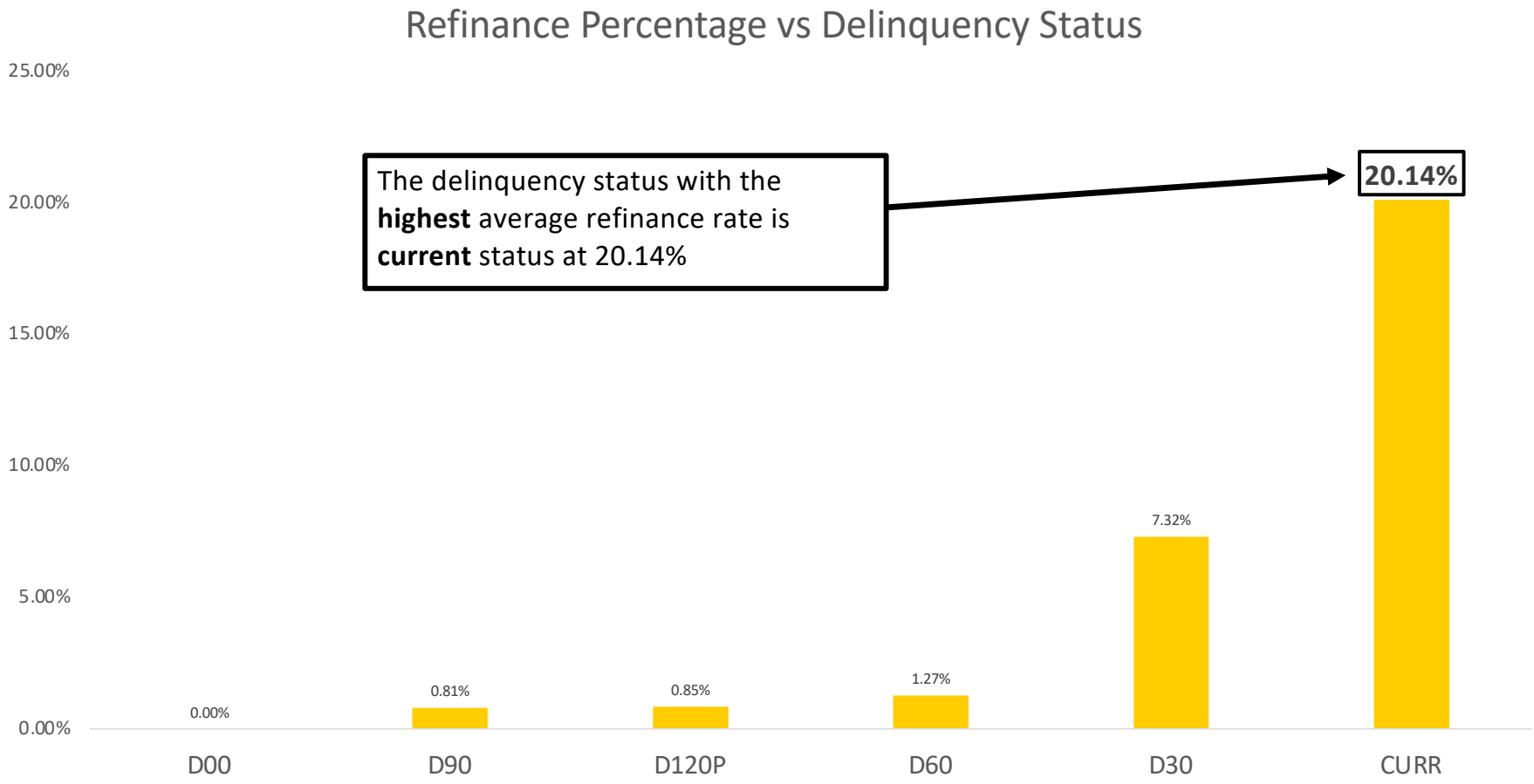
Categorical and Baseline Variable Selection Methodology Variable Removal

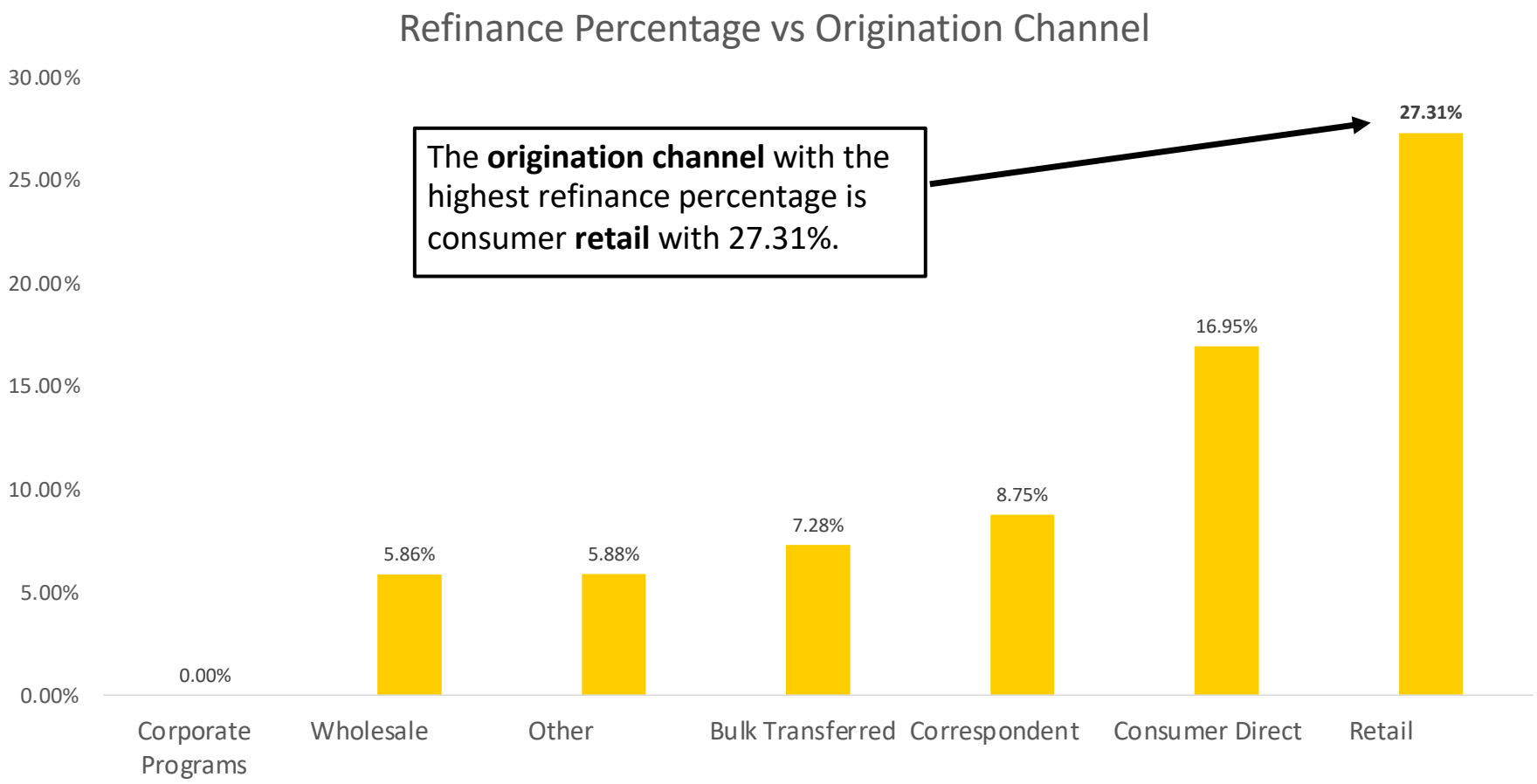
- 5 categorical variables to consider
- **Included**
  - Origination Channel
  - State
  - Delinquency Status
- **Removed**
  - Date
  - Investor Sub






Average Refinance Rate Per State
25.00%
20.00%
15.00%
10.00%
5.00%
0.00%
The state with the highest average refinance rate is South Dakota with a value of 22.45%.
AK ME DC VT TX NY VA HI AL NH RI SC FL MD WV GA NC OK LA MA PA DE IN MS UT AR CT MI NJ CA AZ NM KS MT TN IL NV CO WA MN ND IA ID OH MO WY KY OR WI NE SD


Refinance Percentage vs Delinquency Status
25.00%
20.00%
15.00%
10.00%
5.00%
0.00%
The delinquency status with the **highest** average refinance rate is **current** status at 20.14%
20.14%
0.00%
0.81%
0.85%
1.27%
7.32%
D00
D90
D120P
D60
D30
CURR


Refinance Percentage vs Origination Channel
30.00%
25.00%
20.00%
15.00%
10.00%
5.00%
0.00%
The origination channel with the highest refinance percentage is consumer retail with 27.31%.
0.00%
5.86%
5.88%
7.28%
8.75%
16.95%
27.31%
Corporate Programs
Wholesale
Other
Bulk Transferred
Correspondent
Consumer Direct
Retail

Methodology

How do we select features?



RECURSIVE FEATURE ELIMINATION



MODELING METHODS



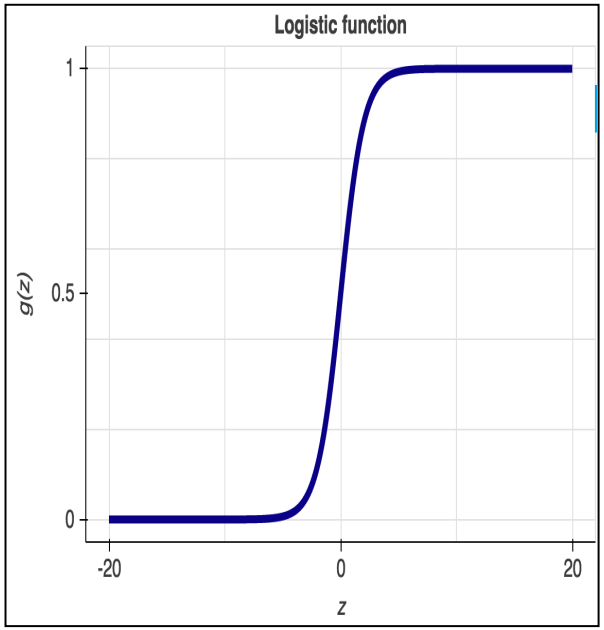
DROP COLUMNS WITH P-VALUE>5%



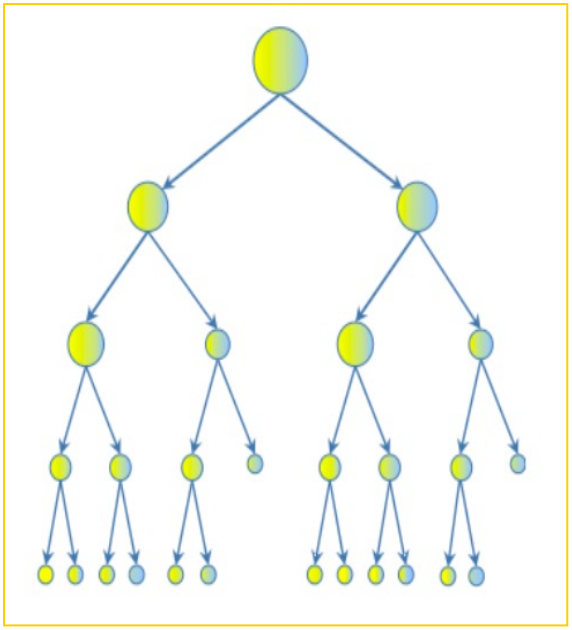
RERUN LOGISTIC REGRESSION UNTIL SIGNIFICANT

Predictors Brief Explanation

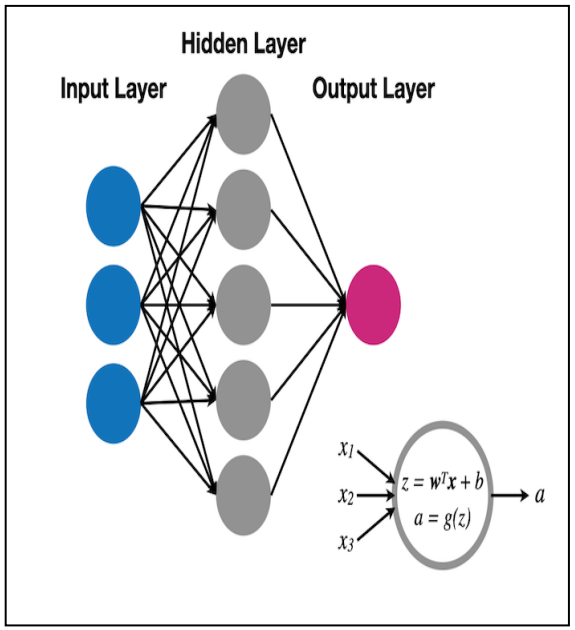
Logistic Regression



Random Forest



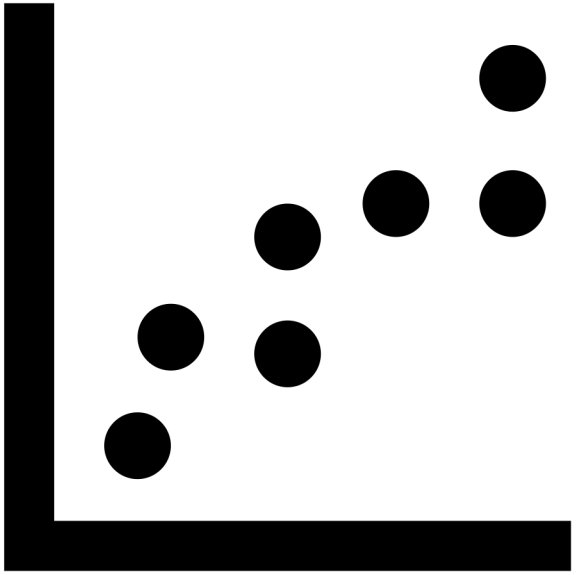
Neural Network



Criteria for Predictor Selection

- **Accuracy** → $\frac{\text{True Positive + True Negative}}{\text{Total}}$
- **Precision** → $\frac{\text{True Positive}}{\text{True Positive + False Positive}}$
- **F1-Score** → $F_1\text{−score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision + Recall}} = \frac{2\text{TP}}{2\text{TP + FP + FN}}$

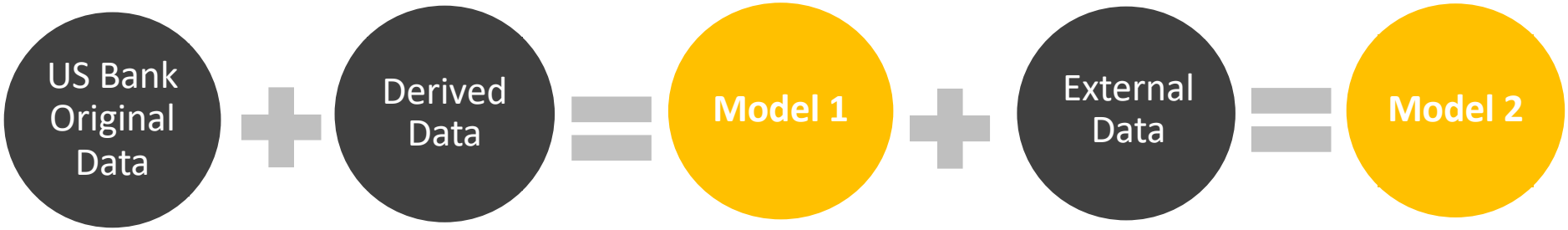
Models



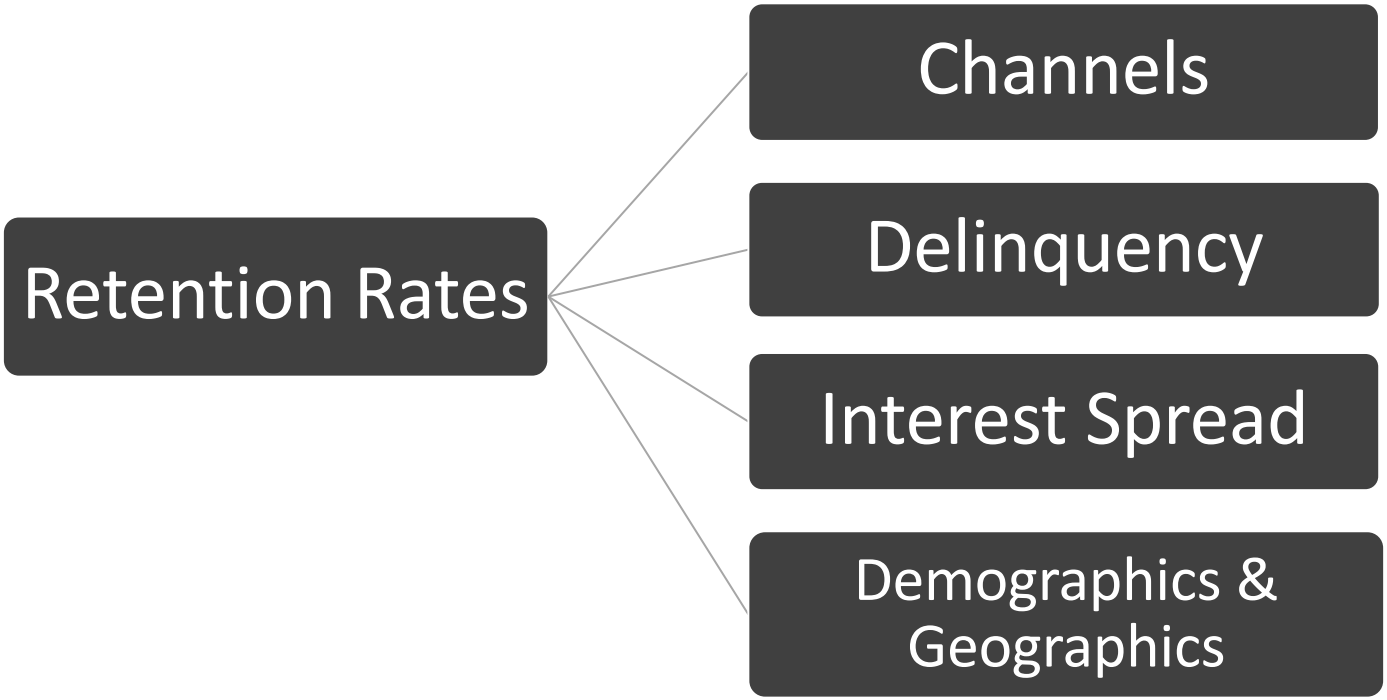
Data-derived findings provide insights for the profiles of mortgage customers

What differentiates various models?

Model 1 & Model 2



State-level Models



Model 1: US Bank Original/Derived Data

73 input features

- Original variables US Bank provided
- Dummy Variables for categorical features
- Derived Variables from US Bank data

1 output feature

- Refinance

Groups of Significance



Loan-related Data:

- Remaining Balance Percentage
- Annual Interest Rate



Loan-related Data:

- Loan Age
- Origination Channels

(Bulk-Transferred, Consumer Direct, Correspondent, Wholesale)

Demographic Data

- States

(AL, AR, AZ, FL, GA, IN, MA, MD, MI, NC, PA, SC, TX, UT, VA)

Delinquency Status

- D30, 120

Model 1 Predictors Results

Logistic



Accuracy
72%



Precision
27%



F1-Score
37%

Random Forest



Accuracy
65%



Precision
22%



F1-Score
33%

Neural Networks



Accuracy
73%



Precision
28%



F1-Score
37%

*Precision and F1-score are test result for Refinance = 1

Model 2: US Bank Data & External Data

95 independent features

- US Bank Original/Derived Data
- External data
  - Economic, demographic, & geographic variables

1 dependent feature

- Refinance

Groups of Significance

Loan-related Data:

- Remaining Balance Percentage

Economic Data:

- Home Price Index
- Home Vacancy Rate

Demographic & Geographic Data

- Number of Branches
- Sates (KY, OR)

Loan-related Data:

- Loan Age
- Origination Channels

(Bulk-Transferred, Consumer Direct, Correspondent, Wholesale)

Economic Data:

- House start
- New home sales
- MB30 contract rate

Demographic & Geographic Data

- Poverty rate
- States (AL, AR, AZ, IN, MD, SC, UT, CA, LA, NV)
- Education rate

Delinquency Status

- D30,60,120

Model 2 Predictors Results

Logistic



Accuracy
61%



Precision
21%



F1-Score
31%

Random Forest



Accuracy
65%



Precision
23%



F1-Score
33%

Neural Networks



Accuracy
55%



Precision
19%



F1-Score
29%

*Precision and F1-score are test result for Refinance = 1

State-leveled Models: Linear Regression

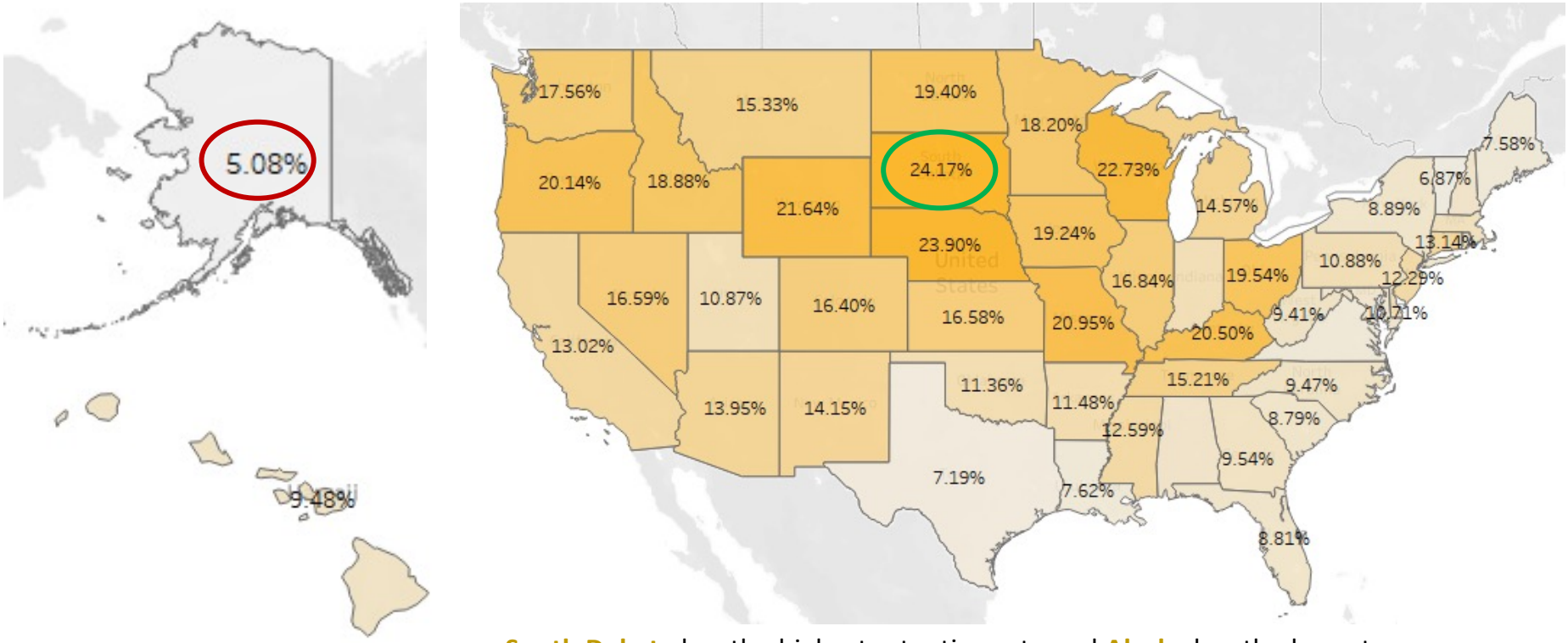
44 independent features

- Original and derived variables US Bank provided
  - Delinquency, Channels, Interest Spread
- External data
  - Demographic & Geographic data

1 dependent feature

- Retention Rate (sum of UPB from refinance loans / sum of UPB)

Retention Rates Based on States



- **South Dakota** has the highest retention rate and **Alaska** has the lowest

- Retention rates (amount of loans calculation) based on states range from 5.08% ~ 24.17%

Groups of Significance



Retail



Male Percent



Loan per capita



Delinquency status
Current



Education Rate



Delinquency status
30 days

Summary & Recommendations



- Predictors' Forecast Examples
- Which Customers to Target?
- Business Recommendations

Model 1 Forecast 1st Pair Example



SIMILAR FEATURES		
SD	State	SD
Retail	Origination Channel	Retail
4.35	Unpaid Balance	4.53
99.68%	Remaining Balance %	99.68%
3.38%	Annual Interest Rate	3.38%
DIFFERENT FEATURES		
0.68	Loan Age	0.36
D120P	**Dlq Status**	**CURR**



REFINANCE?
2.39% - NO

Threshold = 13.90%

REFINANCE?
33.68% - YES

Model 1 Forecast 2nd Pair Example



REFINANCE?
40.67% - YES

SIMILAR FEATURES		
SD	State	SD
CURR	Dlq Status	CURR
DIFFERENT FEATURES		
Retail	**Origination Channel**	**Wholesale**
1.27	Unpaid Balance	4.05
3.63%	Annual Interest Rate	4.38%
14.01	Loan Age in Years	2.78
113.61%	Remaining Balance %	95.59%

Threshold = 13.90%



REFINANCE?
7.69% - NO

Model 2 Forecast 1st Pair Example



REFINANCE?
10.34% - NO

SIMILAR FEATURES		
IA	State	IA
Same	State Demographics	Same
CURR	Dlq Status	CURR
Retail	Origination Channel	Retail
2.69	Unpaid Balance	2.38
99.06%	Remaining Balance	99.11%
DIFFERENT FEATURES		
$-1,818.80	**Potential Savings**	**$725.48**
1551	Housing Starts	938
463.92	Home Price Index	451.30

Threshold = 13.90%



REFINANCE?
20.31% - YES

Model 2 Forecast 2nd Pair Example



REFINANCE?
64.11% - YES

SIMILAR FEATURES		
CA	State	IA
Same	Economic Variables	Same
CURR	Dlq Status	CURR
Retail	Origination Channel	Retail
DIFFERENT FEATURES		
8.09	Unpaid Balance	3.81
77.71%	Remaining Balance	99.34%
$16,062.59	**Potential Savings**	**$419.61**
466	**Number of Branches**	**70**
33.90%	Education Rate	28.60%
37.00	Median Age in State	38.50

Threshold = 13.90%



REFINANCE?
12.94% - NO

0.081192

Model 3 Example: South Dakota VS. Alaska



Retention Rate 22.45%

Significant Positive Features		
47.96%	Retail	2.6%
0.24%	Loans Per Capita	0.13%
50.81%	Male Percent	52%
98.42%	Delinquency: Current	97.49%
Significant Negative Features		
0.67%	Delinquency: 30 days	1.67%
17.7%	Education Rate	29.6%



Retention Rate 8.12%

Recommendations

Target Loans



Acquired through Retail



High Potential Savings



Current Loan



State Matters!

Recommendations

Target States



High % of Retail customers



High Loan per Capita



High % of Current Loans

Business Recommendations



Improve Presence through Increasing # of Branches

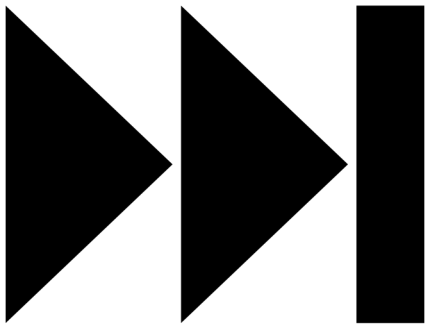


Be Proactive by Recapturing using Predictive Models



Foster Better Relationships by Employee Training

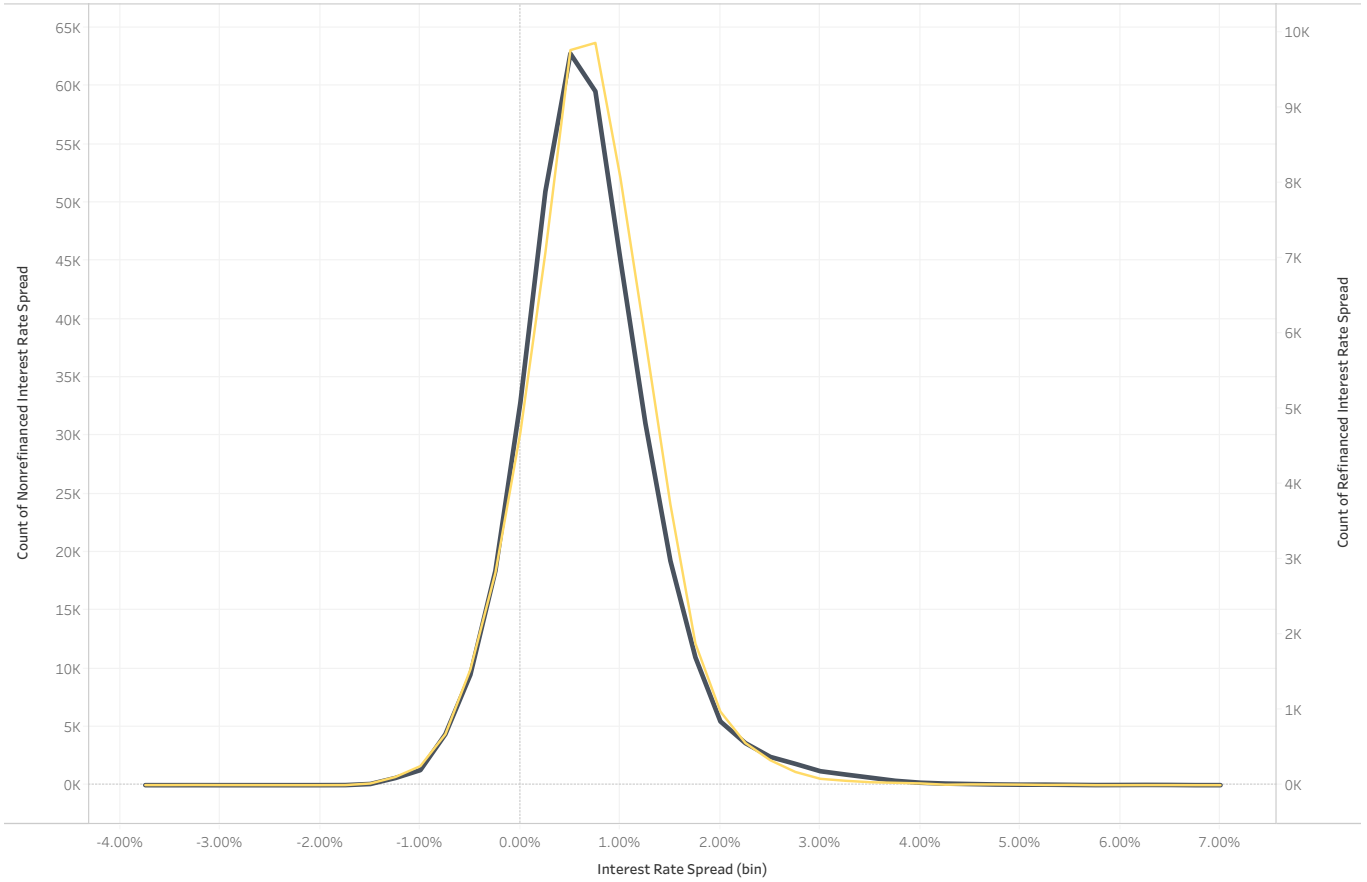
Further Analysis



Interest Rate Spread

Strengths &
Weaknesses

Interest Rate Spread Analysis

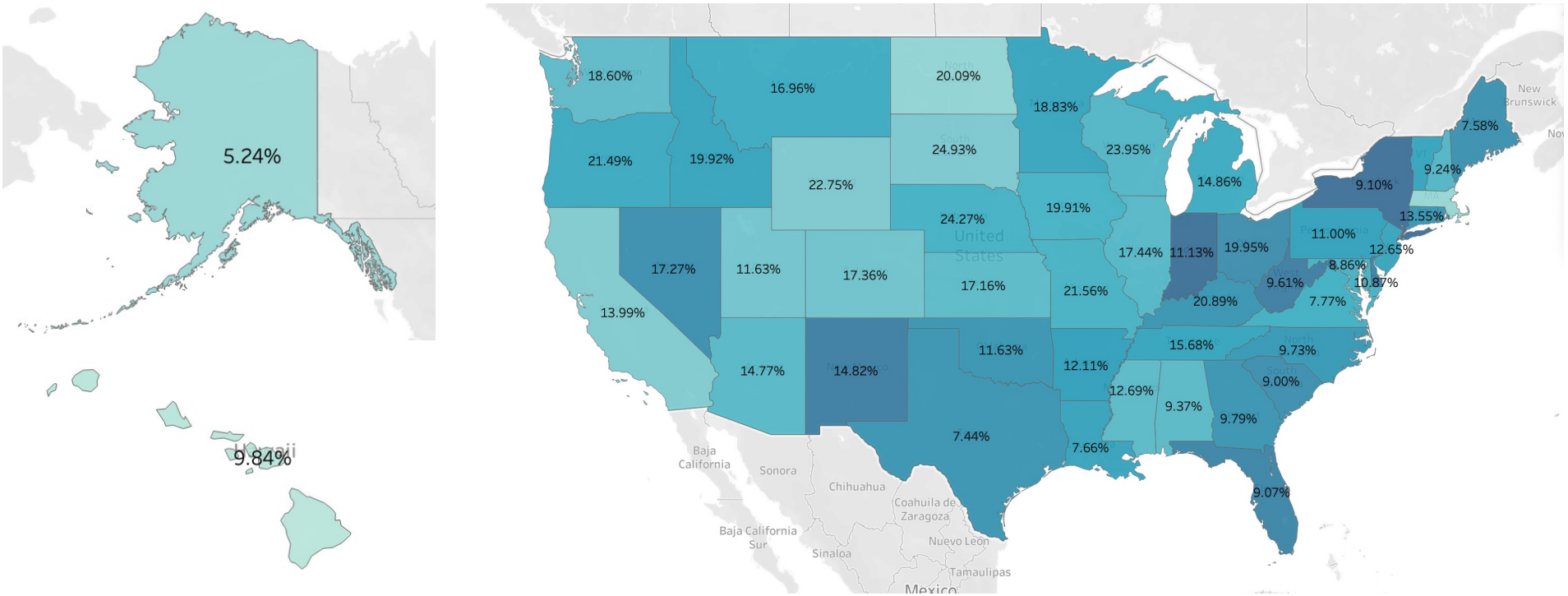


Avg Int Rate Spread

Refinanced => 0.9247%

Not Refinanced => 0.8120%

Interest Rate Spread on all Refinanced Loans



Color => Retention Rate
Label => Interest Rate Spread

Strengths and Weaknesses

STRENGHTS

- Able to identify SIGNIFICANT FEATURES, predicting retention
- Tested wide variety of predictors

WEAKNESSES

- Not accounting for profit and cost while estimating the THRESHOLD
- Data is SKEWED towards the COVID-19 Pandemic/Recession

Questions?

Appendix: Model 1 Coefficients

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
UPB_in_HundredThousands	-0.0313	0.0052	-6.0437	0.0000	-0.0415	-0.0212
ANNUAL_INTEREST_RATE	-15.4740	0.7099	-21.7968	0.0000	-16.8654	-14.0826
loan_age_in_years	-0.0704	0.0020	-35.9848	0.0000	-0.0742	-0.0666
RemainingBalancePercentage	0.2302	0.0313	7.3672	0.0000	0.1690	0.2915
Dlqstatus_D120P	-3.0845	0.2146	-14.3700	0.0000	-3.5052	-2.6638
Dlqstatus_D30	-1.0974	0.0942	-11.6542	0.0000	-1.2819	-0.9128
Property_Alpha_State_Code_AL	-0.3251	0.0659	-4.9335	0.0000	-0.4542	-0.1959
Property_Alpha_State_Code_AR	-0.3963	0.0563	-7.0355	0.0000	-0.5067	-0.2859
Property_Alpha_State_Code_AZ	-0.3256	0.0339	-9.5934	0.0000	-0.3922	-0.2591
Property_Alpha_State_Code_FL	-0.5369	0.0423	-12.6823	0.0000	-0.6199	-0.4540
Property_Alpha_State_Code_GA	-0.2130	0.0439	-4.8540	0.0000	-0.2991	-0.1270
Property_Alpha_State_Code_IN	-0.2874	0.0553	-5.1970	0.0000	-0.3958	-0.1790
Property_Alpha_State_Code_MA	-0.2101	0.0350	-6.0065	0.0000	-0.2787	-0.1416
Property_Alpha_State_Code_MD	-0.2403	0.0451	-5.3332	0.0000	-0.3287	-0.1520
Property_Alpha_State_Code_MI	-0.1462	0.0462	-3.1675	0.0015	-0.2367	-0.0557
Property_Alpha_State_Code_MN	0.1150	0.0178	6.4502	0.0000	0.0800	0.1499
Property_Alpha_State_Code_MO	0.0561	0.0222	2.5306	0.0114	0.0127	0.0996
Property_Alpha_State_Code_NC	-0.1942	0.0365	-5.3280	0.0000	-0.2657	-0.1228
Property_Alpha_State_Code_OK	-0.1656	0.0629	-2.6303	0.0085	-0.2889	-0.0422
Property_Alpha_State_Code_PA	-0.2313	0.0503	-4.5970	0.0000	-0.3299	-0.1327
Property_Alpha_State_Code_SC	-0.3303	0.0668	-4.9423	0.0000	-0.4613	-0.1993
Property_Alpha_State_Code_TX	-0.5188	0.0383	-13.5495	0.0000	-0.5938	-0.4437
Property_Alpha_State_Code_UT	-0.1689	0.0390	-4.3275	0.0000	-0.2454	-0.0924
Property_Alpha_State_Code_VA	-0.3516	0.0357	-9.8399	0.0000	-0.4217	-0.2816
Origination_Channel_Bulk Transferred	-1.4546	0.0270	-53.8612	0.0000	-1.5075	-1.4016
Origination_Channel_Consumer Direct	-0.5567	0.0183	-30.4077	0.0000	-0.5926	-0.5208
Origination_Channel_Correspondent	-1.3419	0.0123	-108.6899	0.0000	-1.3661	-1.3177
Origination_Channel_Wholesale	-1.7637	0.0464	-37.9937	0.0000	-1.8546	-1.6727

Appendix: Model 2 Coefficients

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
UPB_in_HundredThousands	-0.0344	0.0063	-5.4781	0.0000	-0.0467	-0.0221
RemainingBalancePercentage	1.6267	0.0367	44.2718	0.0000	1.5547	1.6988
Housing_Starts	-0.0003	0.0001	-5.1523	0.0000	-0.0004	-0.0002
Home_Price_Index	0.0031	0.0007	4.6076	0.0000	0.0018	0.0044
Home_Vacancy_Rate	0.4328	0.0734	5.8980	0.0000	0.2890	0.5766
New_Home_Sales	-0.0003	0.0001	-4.1167	0.0000	-0.0005	-0.0002
Potential_Savings	0.0001	0.0000	29.5430	0.0000	0.0001	0.0001
Median_Age	-0.0972	0.0073	-13.3090	0.0000	-0.1115	-0.0829
Number_of_Branches	0.0010	0.0001	9.0206	0.0000	0.0008	0.0012
Poverty rate	-8.5497	0.8025	-10.6538	0.0000	-10.1225	-6.9768
Education rate	-2.8819	0.2687	-10.7238	0.0000	-3.4086	-2.3552
Male Percent	-5.1594	0.7629	-6.7629	0.0000	-6.6547	-3.6642
Female Percent	10.5434	1.3042	8.0841	0.0000	7.9872	13.0996
% Renter occupied housing units	-2.5656	0.2535	-10.1203	0.0000	-3.0625	-2.0687
Dlqstatus_D120P	-3.2780	0.2146	-15.2769	0.0000	-3.6985	-2.8574
Dlqstatus_D30	-1.1497	0.0941	-12.2127	0.0000	-1.3342	-0.9651
Dlqstatus_D60	-2.9677	0.3562	-8.3325	0.0000	-3.6657	-2.2696
Property_Alpha_State_Code_AL	-0.2286	0.0687	-3.3251	0.0009	-0.3633	-0.0938
Property_Alpha_State_Code_CA	-0.4583	0.0573	-7.9921	0.0000	-0.5706	-0.3459
Property_Alpha_State_Code_GA	-0.2158	0.0488	-4.4198	0.0000	-0.3115	-0.1201
Property_Alpha_State_Code_IA	-0.1555	0.0363	-4.2883	0.0000	-0.2266	-0.0845
Property_Alpha_State_Code_ID	-0.3559	0.0479	-7.4273	0.0000	-0.4498	-0.2620
Property_Alpha_State_Code_IN	-0.5257	0.0606	-8.6711	0.0000	-0.6445	-0.4068
Property_Alpha_State_Code_KY	0.1795	0.0397	4.5242	0.0000	0.1017	0.2572
Property_Alpha_State_Code_MD	-0.3375	0.0519	-6.5093	0.0000	-0.4392	-0.2359
Property_Alpha_State_Code_MN	-0.1883	0.0277	-6.8099	0.0000	-0.2425	-0.1341
Property_Alpha_State_Code_OR	0.2522	0.0334	7.5622	0.0000	0.1869	0.3176
Property_Alpha_State_Code_SC	-0.2788	0.0693	-4.0220	0.0001	-0.4147	-0.1429
Property_Alpha_State_Code_TX	-0.5768	0.0463	-12.4657	0.0000	-0.6675	-0.4861
Property_Alpha_State_Code_UT	-1.0553	0.0727	-14.5064	0.0000	-1.1978	-0.9127
Property_Alpha_State_Code_VA	-0.3220	0.0383	-8.4121	0.0000	-0.3971	-0.2470
Origination_Channel_Bulk Transferred	-1.3742	0.0272	-50.4912	0.0000	-1.4275	-1.3209
Origination_Channel_Consumer Direct	-0.4439	0.0186	-23.8112	0.0000	-0.4804	-0.4074
Origination_Channel_Correspondent	-1.2623	0.0128	-98.8522	0.0000	-1.2873	-1.2372
Origination_Channel_Wholesale	-1.6291	0.0465	-35.0263	0.0000	-1.7202	-1.5379