Rotman

Master of Management Analytics

CREDIT RISK ASSESSMENT

Analysis on the mortgage market in Nebraska

By Team Calcoholics

Sumit Dargan | Treshan Waas | Keshav Gupta | Michael Adekunle Adelere



Rotman School of Management
UNIVERSITY OF TORONTO

October 13, 2021 Prepared by First Name and Last Name

Credit Risk Assessment

Rotman

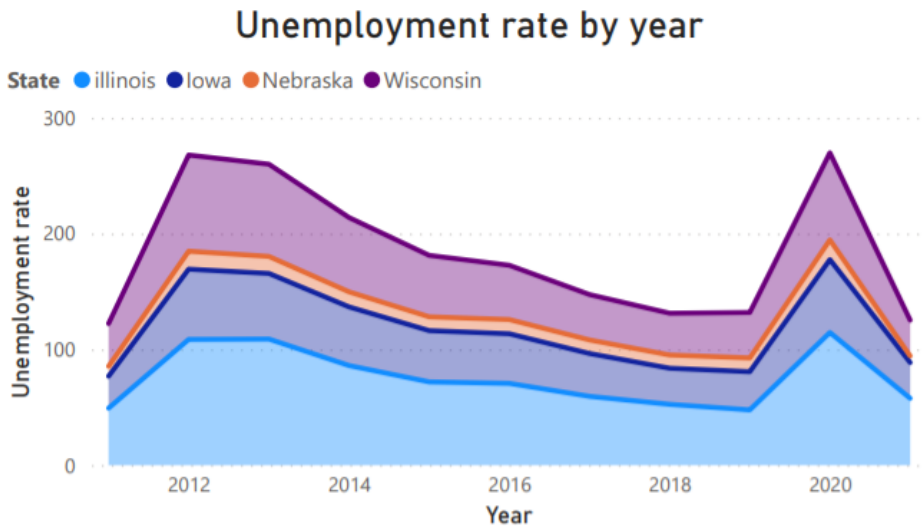
Table of Content

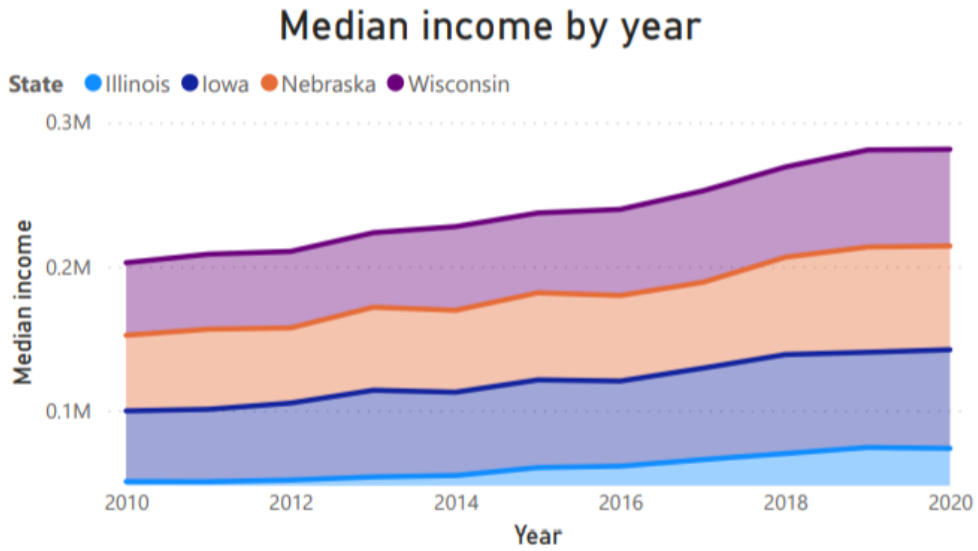
- Current Positioning : Landscape of the mortgage market in Nebraska
- Key target customer segments
- Future Considerations
- Final Recommendations

Credit Risk Assessment

Rotman

MACRO ECONOMIC INDICATORS OF NEBRASKA VS. ILLINOIS, IOWA AND WISCONSIN (1/2)





The following are the key macro economic indicators, that helps the visualize the state of the economy.

A comparative analysis would help us to understand the positioning of **Nebraska** mortgage market relative to where Great Lakes Midwest Bank has it's operations (Illinois, Iowa and Wisconsin)

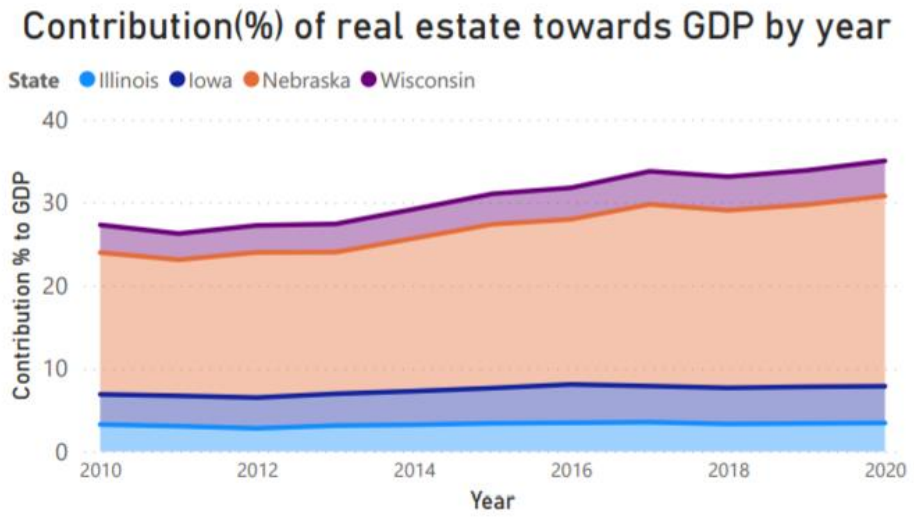
The unemployment rates in Nebraska closely mimics the trends across Iowa – but since we're already catering to Wisconsin with fairly high unemployment rate, Nebraska market **doesn't pose a challenge here**

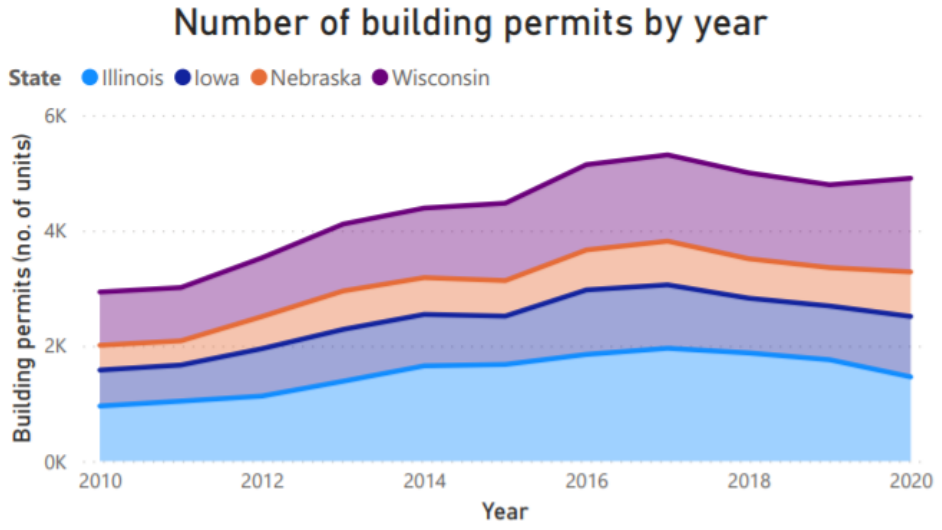
The median income of Nebraska is on an uptrend, with the state having fairly high median incomes as compare to Iowa and Illinois – creating a market place for "large loans"

Credit Risk Assessment

Rotman

MACRO ECONOMIC INDICATORS OF NEBRASKA VS. ILLINOIS, IOWA AND WISCONSIN (1/2)





The following are the key macro economic indicators, that helps the visualize the state of the economy.

A comparative analysis would help us to understand the positioning of **Nebraska** mortgage market relative to where Great Lakes Midwest Bank has it's operations (Illinois, Iowa and Wisconsin)

The contribution of real estate as a sector has been always high and is gradually increasing in Nebraska, which adds further value to be in the mortgage loan market.

Since government support plays a key role in the growth of the real estate economy – a constant increase in "number of building permits" gives an assurance that the real estate sector has enough support from the local authorities

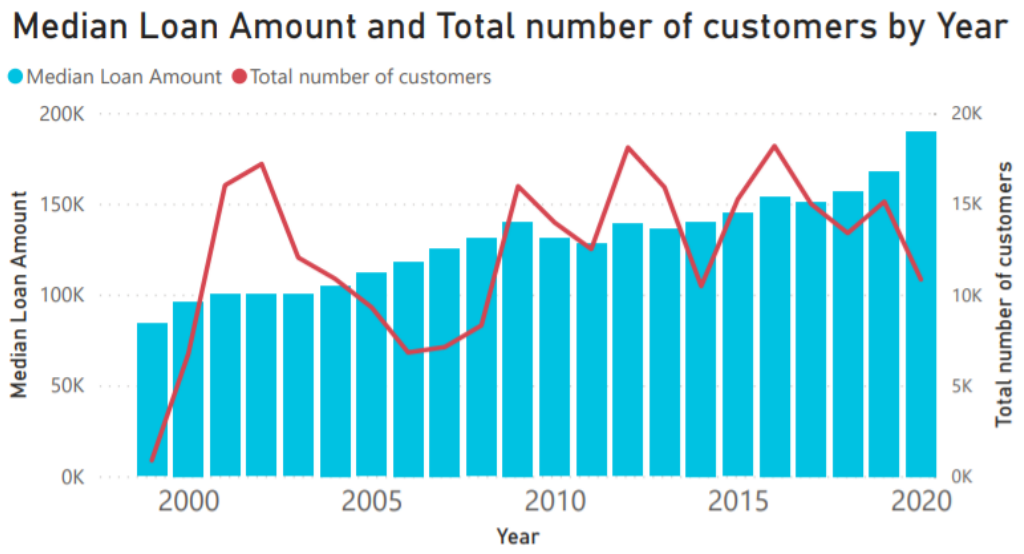
Nebraska is a state of an art economy; with increasing median salaries and contribution of real estate in GDP – the mortgage loan market appears to be in a sweet spot.

Also, this market opportunity is even more lucrative for Great Lakes Midwest Bank since they are already operating in the neighboring states and with close closeness proximity in the macro economic factors – the senior leadership would understand the market well.

Credit Risk Assessment

Rotman

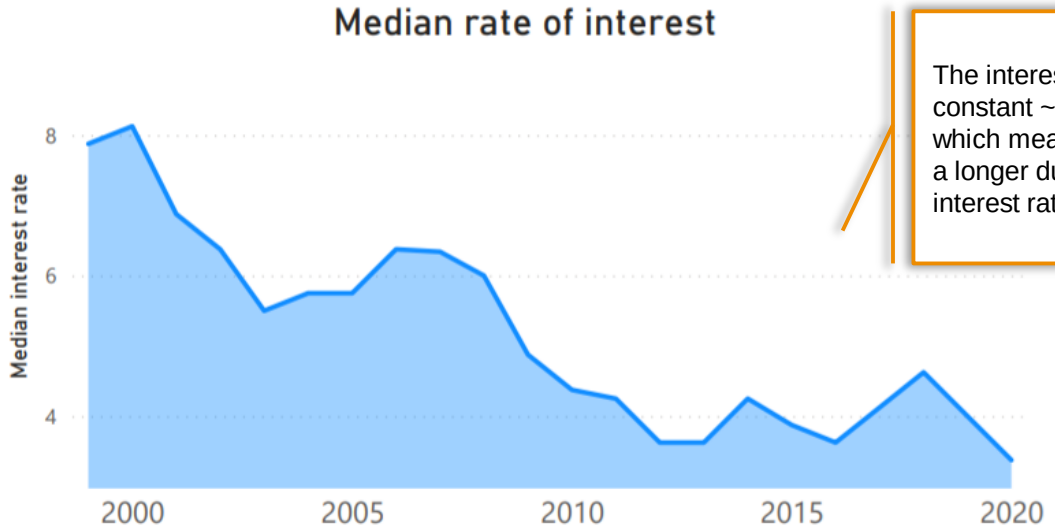
Current Positioning: Mortgage Market Outlook



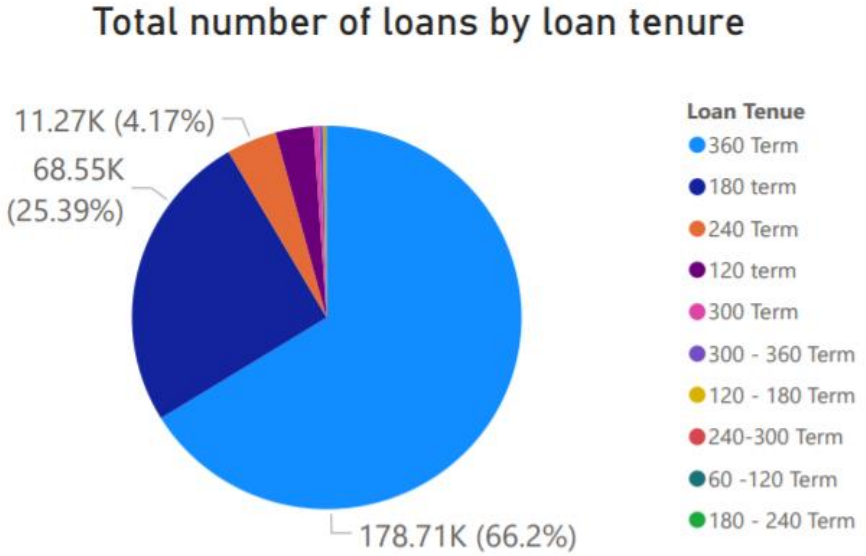
As the total number of customers remain **fairly constant,** Nebraska market appears to be a **saturated market**. But with gradual increase in loan amount, we can see some forward looking **green shoots**

In terms of the tenure of loan, 360 months and 180 months are the preferred segments across the customers.

With most of the people preferring longer duration loans, this gives us business visibility for long term



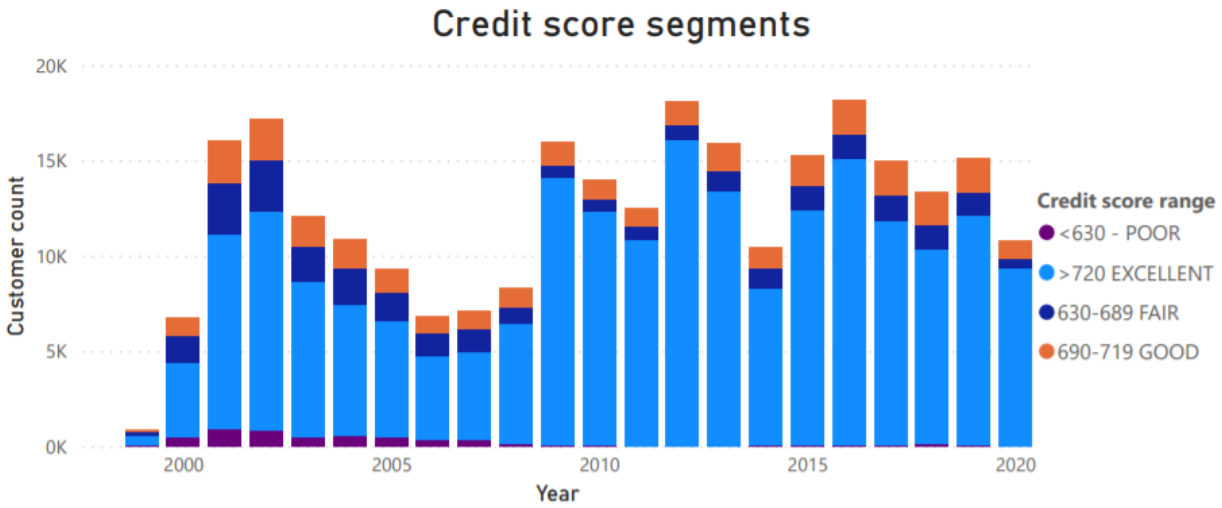
The interest rates have been fairly constant ~4% in the past 10 years, which means if we see the market with a longer duration perspective the interest rates **would vary +/- 10%**



Credit Risk Assessment

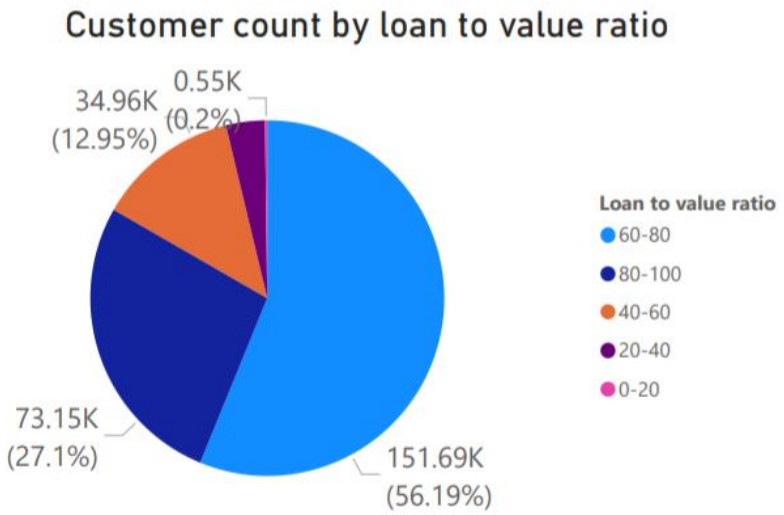
Rotman

Current Positioning: Customer Segment Outlook (1/2)

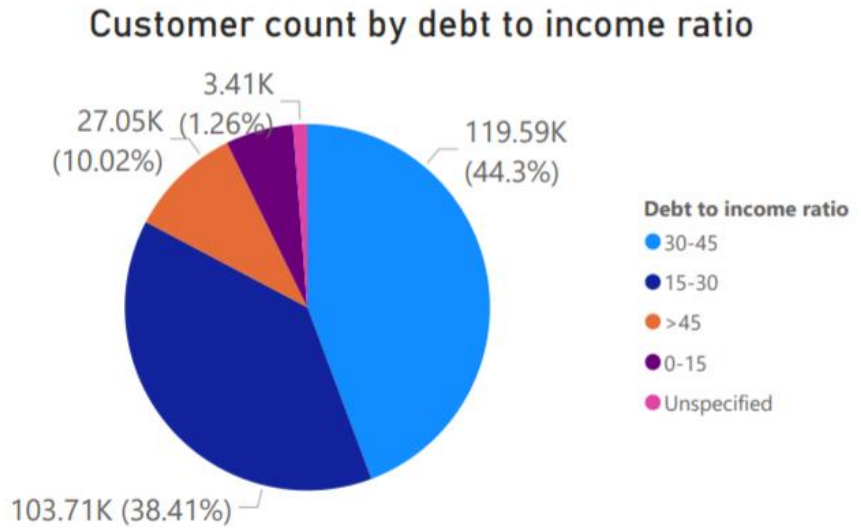


As per credit risk assessment in concerned, the **Credit Score** is an important indicator to give out mortgage loans – since most of the loans over time have been given out to people with **Excellent** and **Good** credit scores

The customers in Nebraska market are in need of credit – as **more than 40%** of all the loans are given to people with Debt to income ratio greater than 30



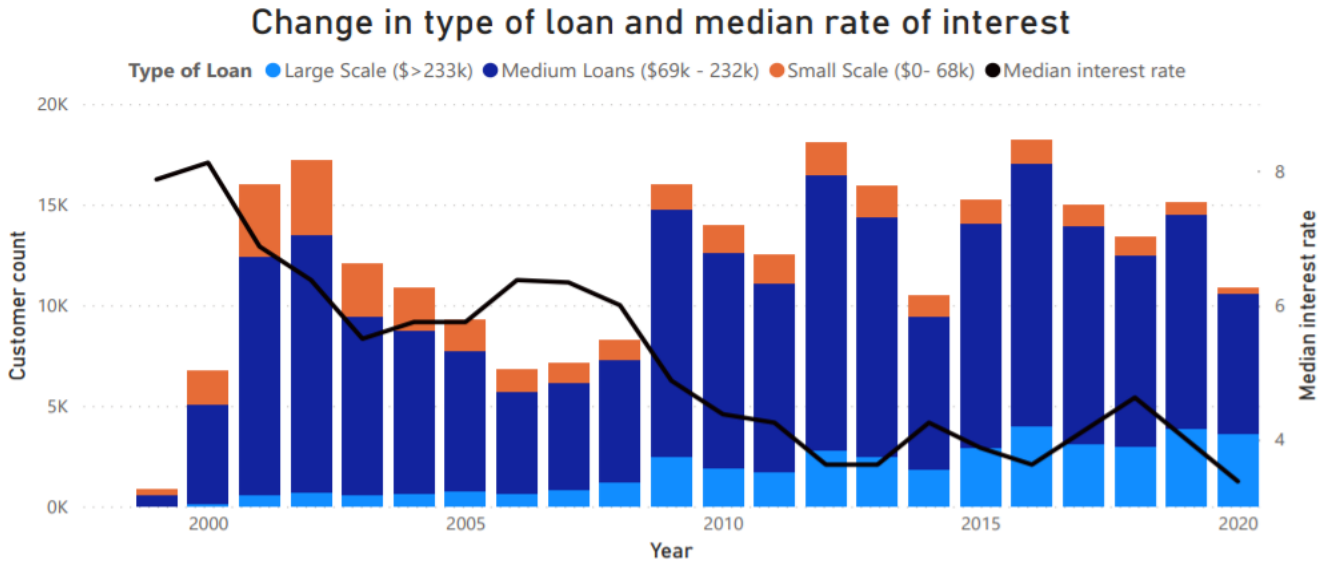
Over time the banks in the market have readily given out loans to customers with fairly high loan to value ratio – since more than 80% of loans are given to customer who have **Loan to value ratio above 60**



Credit Risk Assessment

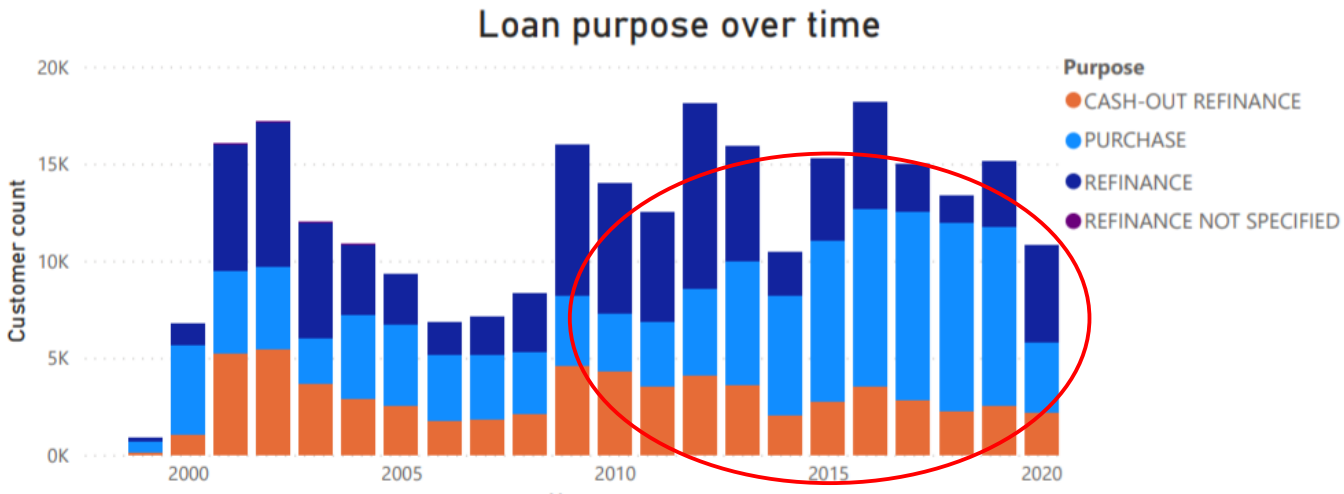
Rotman

Current Positioning: Customer Segment Outlook (2/2)



Over time, as the interest rates have decreased – there has been a **significant uptake** in the number of Large scale loans given in the state.

Coupled with the growth the real estate economy (slide 3), growth in large scale loans make Nebraska a lucrative market opportunity



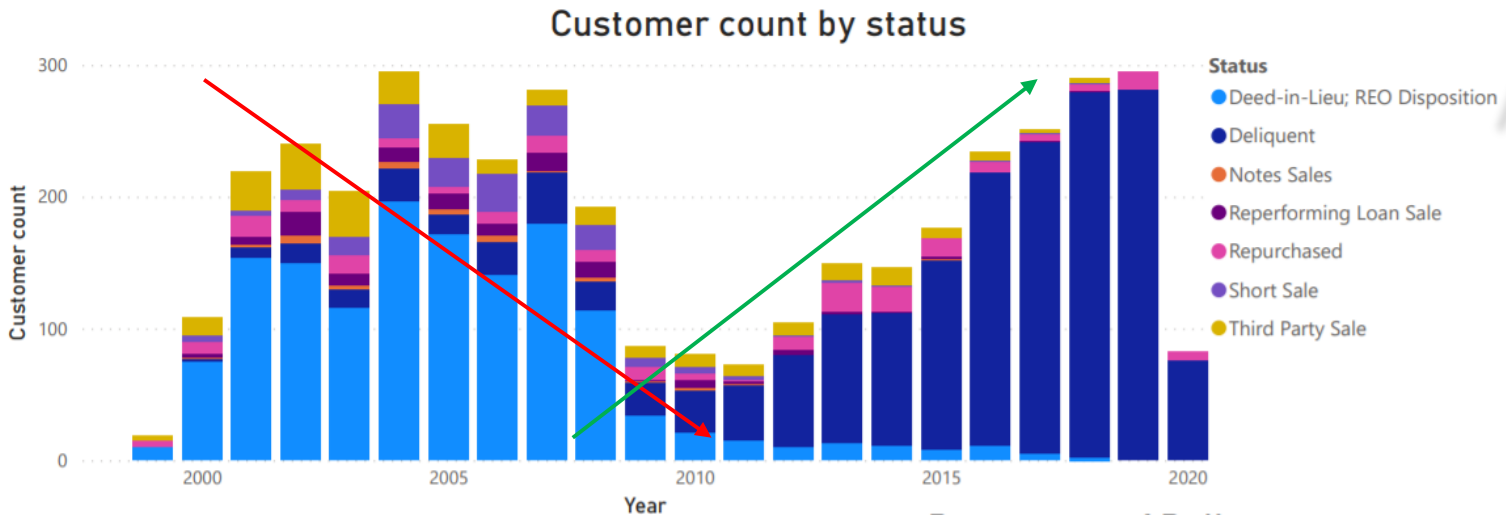
The growth in the real estate economy, is also supported by a gradual increase in the number of “Purchase” mortgage loans taken over time.

The growth in the new purchase loans, makes Nebraska an interesting market.

Credit Risk Assessment

Rotman

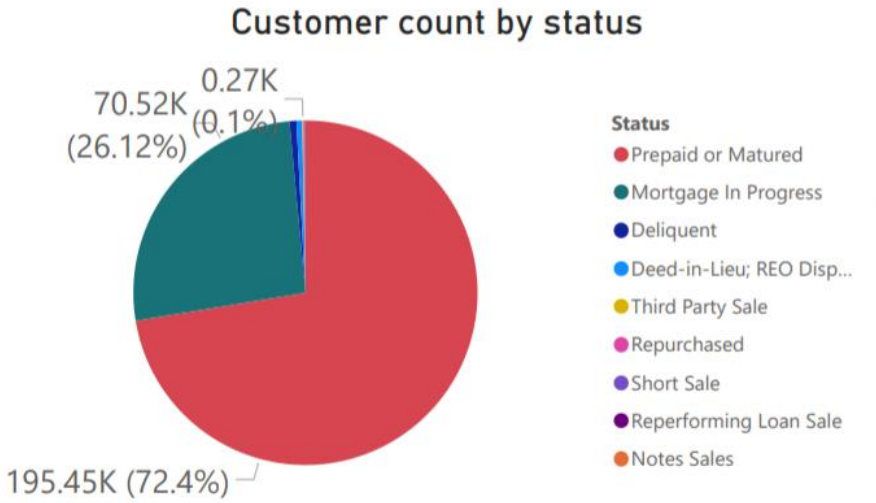
Current Positioning: Delinquent Customers Outlook



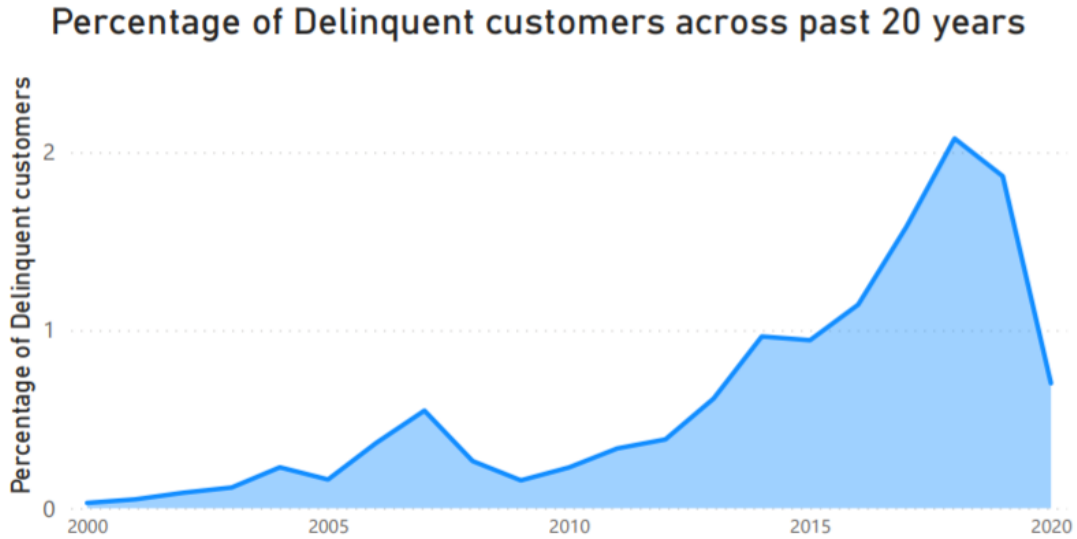
In the past 20 years, there seems to be a clear trend – wherein the number of people under Deed-in-lieu has declines while the number of Delinquent people has consistently gone up.

This shows that there's an internal policy change after 2010, which doesn't give enough advantage to customers to go in for Deed-in-Lieu

The percentage of Delinquent customers in a year is **constantly going up** – which gives an indication that the credit risk assessment of the customers has now become **more important** than ever before



In the past 20 years, the number of prepaid loans have accounted for more than 70% of all loans, giving an assurance about the **credibility** of the customers



Credit Risk Assessment

Rotman

Key Target Customers: Our Approach

Key Business Objectives

Understand how do we define “Risk” in the Nebraska mortgage loan market and decide that who should be given the loan

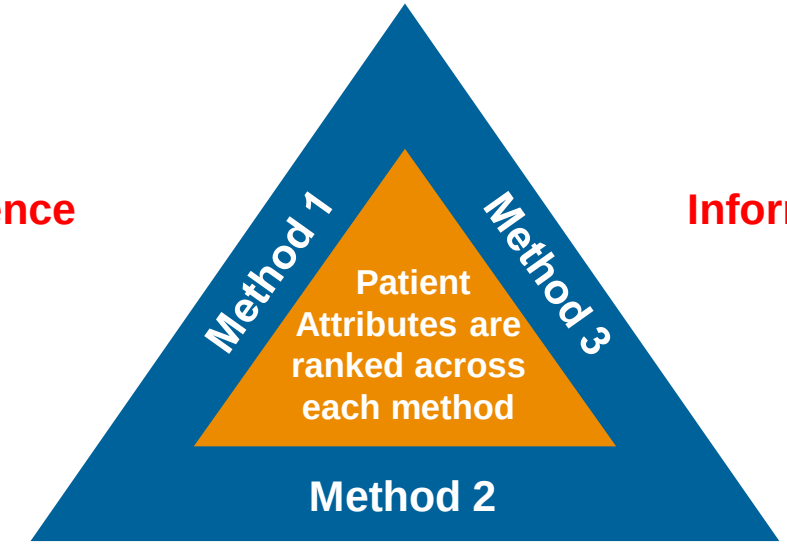
Recommended action plan

- Assess customer behaviour in the Nebraska mortgage loan market and develop a customer segmentation / classification model to decide who should be given a loan
- Evaluate the objective function to maximize the dollar amount given out to the loans, while minimizing the risk (chances to be delinquent)
- Gauge the most critical / important customer attributes to define “Risk” at giving out a loan

The key customer attributes across years, are evaluated through a triangulation of three sophisticated statistical techniques:

Relative Importance by XGBoost

Weight of Evidence



Information Value

An aggregated ranking score from each method is taken into account to come with the **Top 5** Customer attributes that influences / drives the risk associated with a particular customer

Credit Risk Assessment

Rotman

Key Target Customers: Key Attributes

Customer Attribute	Level Description
Credit Score	630-689 FAIR
	<630 - POOR
	690-719 GOOD
	>720 EXCELLENT
Original Combined Loan to Value Ratio (OCLTV)	95-146
	90-94
	80-83
	84-89
	65-71
	76-79
	72-75
	53-64
	4-52

Customer Attribute	Level Description
Debt to Income Ratio (DTI)	42-45
	39-41
	46-64
	35-38
	32-34
	29-31
	26-28
	22-25
	18-21
	1-17
Mortgage Insurance Percentage (mi_pct)	17-22
	23-40
	0-16
Type of Loan	Small Scale ($0- $68k)
	Medium Loans ($69k - $232k)
	Large Scale (> $233k)

The different levels of OCLTV, DTI and mi_pct are governed by "**Weight of Evidence" Methodology**

All the customer attributes are weighted as per the significance in prediction of delinquent customer

Credit Risk Assessment

Rotman

Key Target Customers: Design Process and Sample case

1. Using a triangulation approach of different statistical techniques, **5 key customer attributes** are selected
2. The 5 attributes are weighted as per the "Information Value" Gain provided relative to other attributes
3. All the levels within an attribute are ranked as per the "Weight of Evidence" and then weighted as per the relative rank.
4. All the customers, are then given a "**Loan Qualification**" score – depending on the customer characteristics
5. Looking at the previous "**Loan Qualification**" scores of Delinquent customers the cutoff to give loan is then decided

Sample Customer

	Customer 1
Example Customer 1	
Credit Score	680
OCLTV	75
DTI	25
Mi_pct	0
Type of loan	Small Scale

	Customer 1
Example Customer 1	
Credit Score	720
OCLTV	85
DTI	15
Mi_pct	20
Type of loan	Large Scale

Great Lakes Midwest Bank

The loan qualification score will be calculated for each patient, and on the basis of **pre-defined cutoffs** the loan will be given

Sample Customer

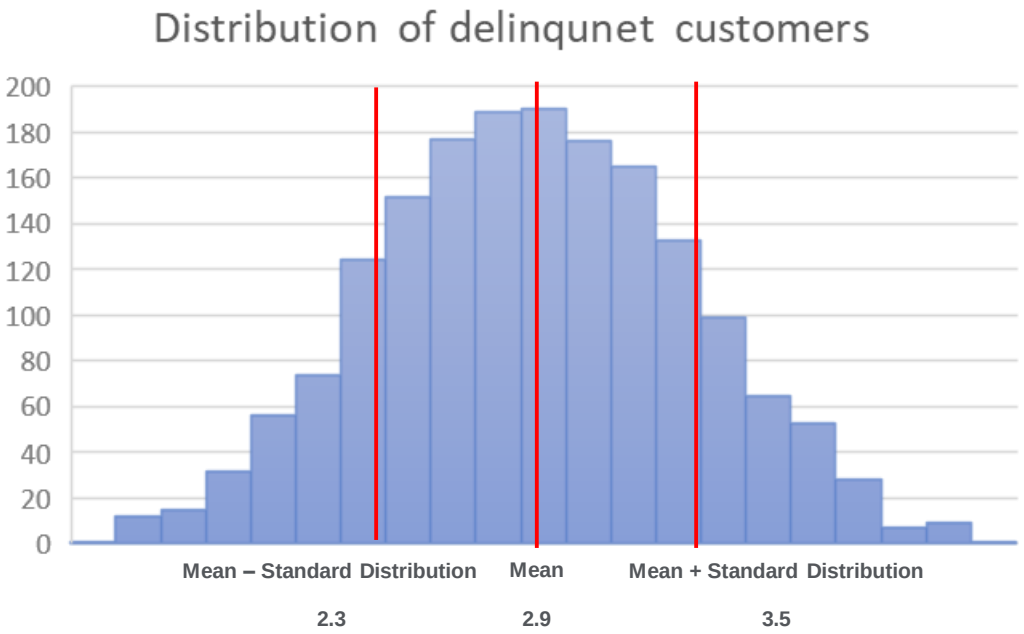
	Customer 1
Example Customer 1	
Credit Score	680
OCLTV	75
DTI	25
Mi_pct	0
Type of loan	Small Scale

	Customer 1
Example Customer 1	
Credit Score	720
OCLTV	85
DTI	15
Mi_pct	20
Type of loan	Large Scale

Credit Risk Assessment

Rotman

Key Target Customers: How to use the approach / Final Recommendations



The "**Loan Qualification**" scores for the previous delinquent customers are utilized to create 3 set cutoff to define risk profiles for the customers

1. Low Risk Profile : Customer who have "Loan Qualification" scores greater than 3.5 (Mean + Standard Deviation)

2. Moderate Risk Profile : Customer who have "Loan Qualification" scores greater than 2.9 (Mean)

3. High Risk Profile : Customer who have "Loan Qualification" scores lesser than 2.3 (Mean - Standard Deviation)

Final Recommendations:

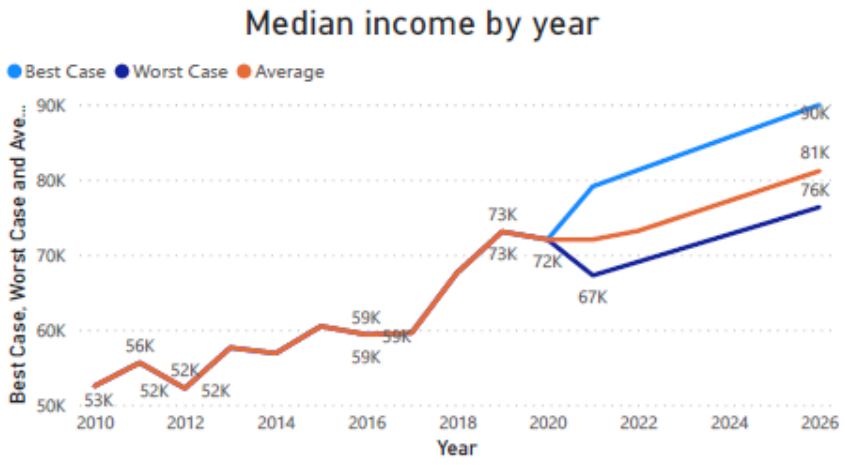
Nebraska which is a state of an art economy, gives both green shoots and some potential concerns to enter the mortgage loan market. However, the macro-economic factors favor and hence the Great Lakes Midwest bank **should enter into the market**.

As the bank enters into the market, the bank should consider giving out loans to customers who have "**Loan Qualification"** scores greater than 2.3 (Mean – Standard Deviation), since like that they would be able to take moderate risk, but at the same time ensuing that the delinquent customers are as less as possible.

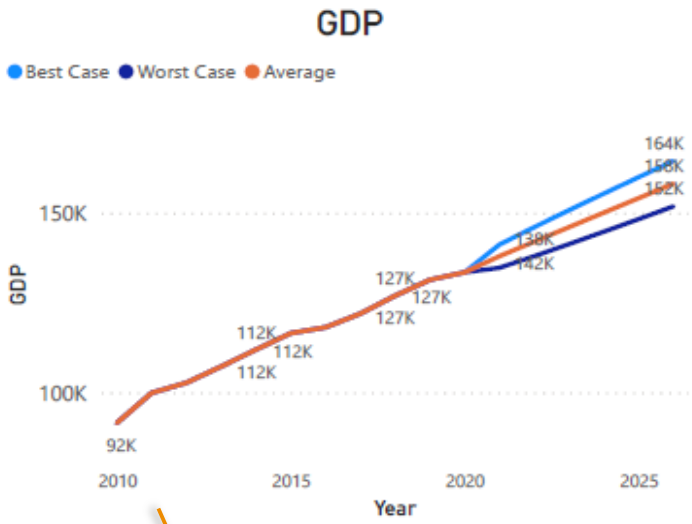
Credit Risk Assessment

Rotman

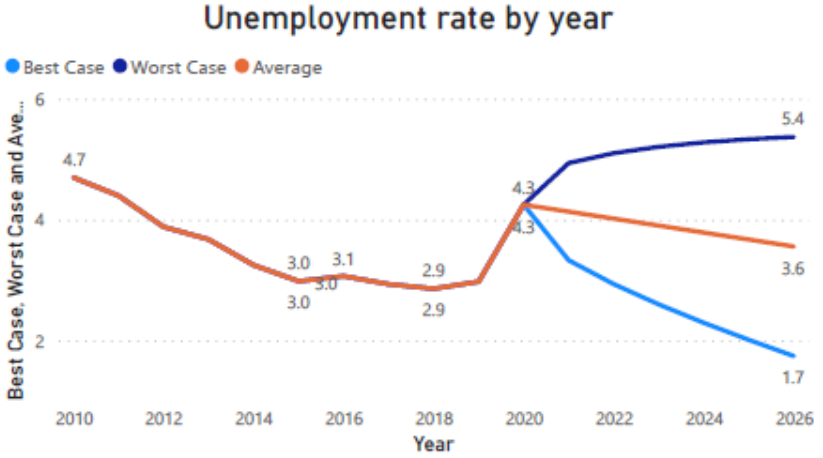
Future Considerations (1/2)



Median income shows a growth in the next 5 years.



GDP growth remains positive in the next 5 years



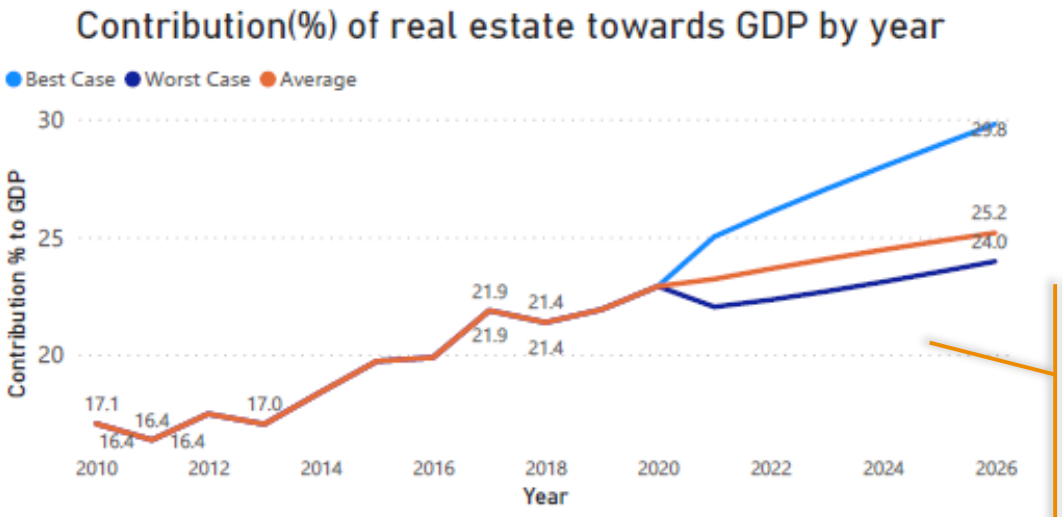
In the best case scenario there will be a rapid drop in unemployment which will help to boost the income levels

- The worst case economic indicators predicted shows that Nebraska is a stable state to grow businesses
- We could expect economic recovery considering the economic trends in the pandemic situation
- Westbank will benefit by a average economic recovery as median income, gdp will grow at pace.
- In the average predicated scenario, pace of recovery cant be expected to all time low as Nebraska state has taken considerable time to over the years to bring down unemployment rates

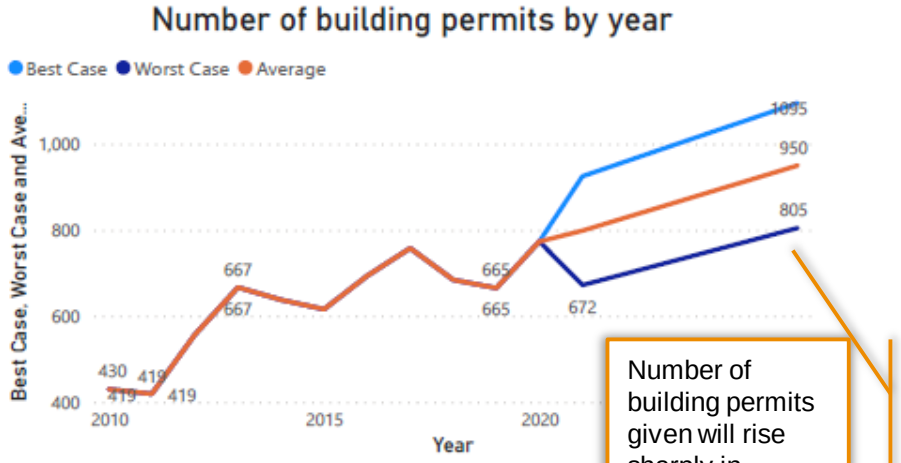
Credit Risk Assessment

Rotman

Future Considerations (2/2)



The contribution % of real estate towards GDP will indicate a strong rise which gives great opportunity for the mortgage market



Number of building permits given will rise sharply in Nebraska

- A key economic indicator in the Nebraska market is the contribution in the real estate market in GDP. Compared to the other states where west bank runs its operations, Nebraska state stands out from the rest due to its significant contribution in the real estate market towards its GDP
- We can predict higher customer counts and mortgages in the coming years. Even at a worst case scenario, the percentage indicate a downturn. This signifies the stability of the Nebraska market.
- We can also predict that with the booming real estate market the median income level will increase. This will further contribute to a multipler effect which will boost the real estate and mortgage market
- The number of building permits issues have increased since 2010 at a significant pace. This is further backed by its real estate contribution towards its GDP

APPENDIX

Credit Risk Assessment

Rotman

Statistical Technique: Weight of evidence (1/3)

What is Weight of Evidence (WOE)?

The weight of evidence tells the predictive power of an independent variable in relation to the dependent variable. Since it evolved from credit scoring world, it is generally described as a measure of the separation of good and bad customers. **"Bad Customers"** refers to the customers who defaulted on a loan. and **"Good Customers"** refers to the customers who paid back loan.

$$WOE = \ln\left(\frac{\text{Distribution of Goods}}{\text{Distribution of Bads}}\right)$$

Steps of Calculating WOE

1.For a continuous variable, split data into 10 parts (or lesser depending on the distribution).
2.Calculate the number of events and non-events in each group (bin)
3.Calculate the % of events and % of non-events in each group.
4.Calculate WOE by taking natural log of division of % of non-events and % of events

Range	Bins	Non events	Events	% of Non-Events	% of Events	WOE	IV
0-50	1	197	20	5.4%	5.9%	-0.0952	0.0005
51-100	2	450	34	12.3%	10.1%	0.2002	0.0045
101-150	3	492	39	13.4%	11.5%	0.1522	0.0029
151-200	4	597	51	16.3%	15.1%	0.0774	0.0009
201-250	5	609	54	16.6%	16.0%	0.0401	0.0003
251-300	6	582	55	15.9%	16.3%	-0.0236	0.0001
301-350	7	386	41	10.5%	12.1%	-0.1405	0.0022
351-400	8	165	23	4.5%	6.8%	-0.4123	0.0095
>401	9	184	21	5.0%	6.2%	-0.2123	0.0025
	Total	3662	338				0.0234

Credit Risk Assessment

Rotman

Statistical Technique: Information Value (2/3)

What is Information Value (IV)?

Information value is one of the most useful technique to select important variables in a predictive model. It helps to rank variables on the basis of their importance. The IV is calculated using the following formula :

$$IV = \sum (\%\ of\ non\text{-}events - \%\ of\ events) * WOE$$

Rules related to Information Value

Information Value	Variable Predictiveness
Less than 0.02	Not useful for prediction
0.02 to 0.1	Weak predictive Power
0.1 to 0.3	Medium predictive Power
0.3 to 0.5	Strong predictive Power
>0.5	Suspicious Predictive Power

- Less than 0.02, then the predictor is not useful for modeling (separating the Goods from the Bads)
- 0.02 to 0.1, then the predictor has only a weak relationship to the Goods/Bads odds ratio
- 0.1 to 0.3, then the predictor has a medium strength relationship to the Goods/Bads odds ratio
- 0.3 to 0.5, then the predictor has a strong relationship to the Goods/Bads odds ratio.
- > 0.5, suspicious relationship (Check once)

Credit Risk Assessment

Rotman

Statistical Technique: Relative importance (3/3)

Feature Importance in Gradient Boosting

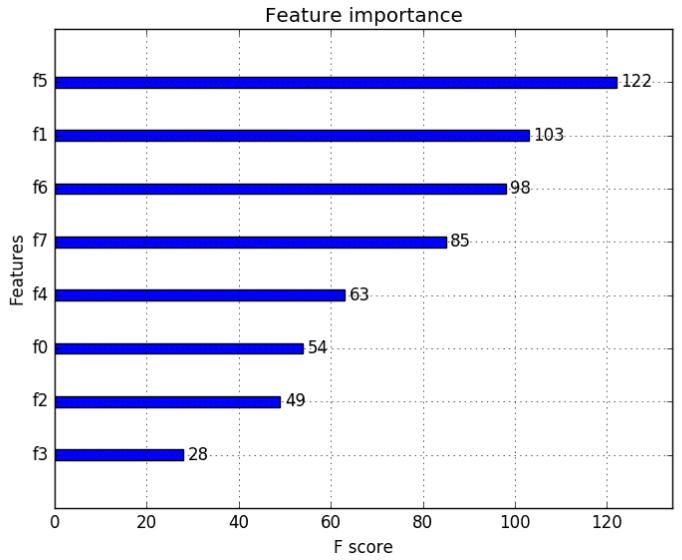
A benefit of using gradient boosting is that after the boosted trees are constructed, it is relatively straightforward to retrieve importance scores for each attribute.

Generally, importance provides a score that indicates how useful or valuable each feature was in the construction of the boosted decision trees within the model. The more an attribute is used to make key decisions with decision trees, the higher its relative importance.

This importance is calculated explicitly for each attribute in the dataset, allowing attributes to be ranked and compared to each other.

Importance is calculated for a single decision tree by the amount that each attribute split point improves the performance measure, weighted by the number of observations the node is responsible for. The performance measure may be the purity (Gini index) used to select the split points or another more specific error function.

The feature importance's are then averaged across all of the the decision trees within the model.



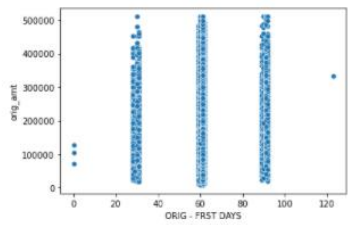
Credit Risk Assessment

Rotman

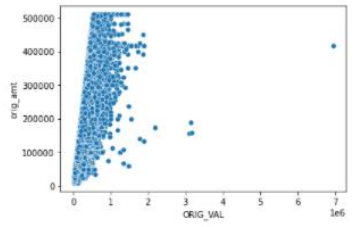
Data Exploration - Python

```
In [2]: import pandas as pd
        import numpy as np
        import seaborn as sns
        import matplotlib.pyplot as plt

In [3]: df = pd.read_csv(r'C:\Users\treshanw\OneDrive - Star Garments\Datathon\dataset_feature_small.csv')

In [4]: df.isnull().sum()

Out[4]: LAST_DTE_YEAR            0
        ORIG_DTE_YEAR            0
        ORIG - FRST DAYS         0
        ORIG-LAST DAYS           0
        EXP_LAST_DATE            0
        LAST_STAT                0
        ORIG_CHN                 0
        SELLER                   0
        SELLER_GROUP             0
        SELLER_SEG               0
        orig_rt                  0
        RATE_SEGMENT             0
        LAST_RT                  0
        orig_amt                 0
        AMT SEGMENT              0
        ORIG_VAL                 0
        DIFF_VAL_AMT             0
        LAST_UPB              6867
        orig_trm                 0
        oltv                     0
        oltv_segment             0
        ocltv                    0
        num_bo                  24
        dti                   3410
        CSCORE_B               736
        CSCORE_SEG             736
        ETHR_FLG                65
```

```
Out[8]: <AxesSubplot:xlabel='ORIG - FRST DAYS', ylabel='orig_amt'>
```



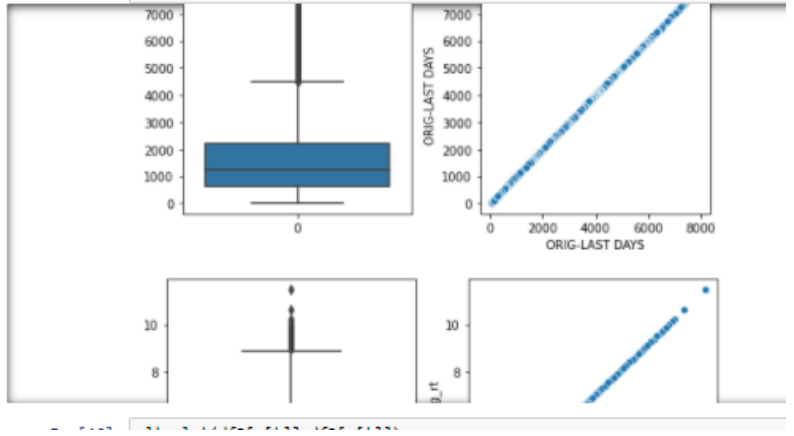

```
In [9]: sns.scatterplot( data=df , x = 'ORIG_VAL' , y ='orig_amt' )
Out[9]: <AxesSubplot:xlabel='ORIG_VAL', ylabel='orig_amt'>
```




```
dtype='object')

In [21]: Q1 = df['orig_amt'].quantile(0.25)
         Q3 = df['orig_amt'].quantile(0.75)
         IQR = Q3 - Q1
         print(IQR)

         a = Q3+(IQR*1.5)
         a

         98000.0

Out[21]: 337000.0

In [23]: df = df[df['orig_amt']<a]

In [26]: Q1 = df['orig_rt'].quantile(0.25)
         Q3 = df['orig_rt'].quantile(0.75)
         IQR = Q3 - Q1
         print(IQR)

         a = Q3+(IQR*1.5)
         a

         2.0

Out[26]: 8.875
```

- Checking missing values
- IQR
- Scatter plots and box plots on each variable to identify correlation and trends

```
In [50]: t=0
         y =0
         t = 0


         for i in x:
             plt.rcParams["figure.figsize"] = [7.00, 3.50]
             plt.rcParams["figure.autolayout"] = True
             f, axes = plt.subplots(1, 2)
             sns.boxplot(data= df3[x[t]], ax=axes[0])
             sns.scatterplot(x = df3[x[t]], y = df3[x[t]],ax=axes[1])
             t = t+1
```



Credit Risk Assessment

Rotman

References

- **R and Python codes: https://github.com/keshavgupta0499/Credit-Risk-Management**
- https://fred.stlouisfed.org
- https://www.ibisworld.com/united-states/economic-profiles/nebraska/
- https://machinelearningmastery.com/feature-importance-and-feature-selection-with-xgboost-in-python/
- https://www.listendata.com/2015/03/weight-of-evidence-woe-and-information.html