Group 2: Aditee Bhattarai, Nam Dang,
Zihan Huang, Jingxin Yao

Credit Risk Analytics

Analysis of risk for Residential Mortgage and Equity Loan

Project Background

Far Reaching Impacts of the Real Estate Industry

2008 Housing Bubble Burst (Magdoff & Yates, 2010)

- Unemployment rate at 4.4%
- Wages rising by 4.2%
- Dow Jones index hit an all-time high
- The real estate market became the core of the economy

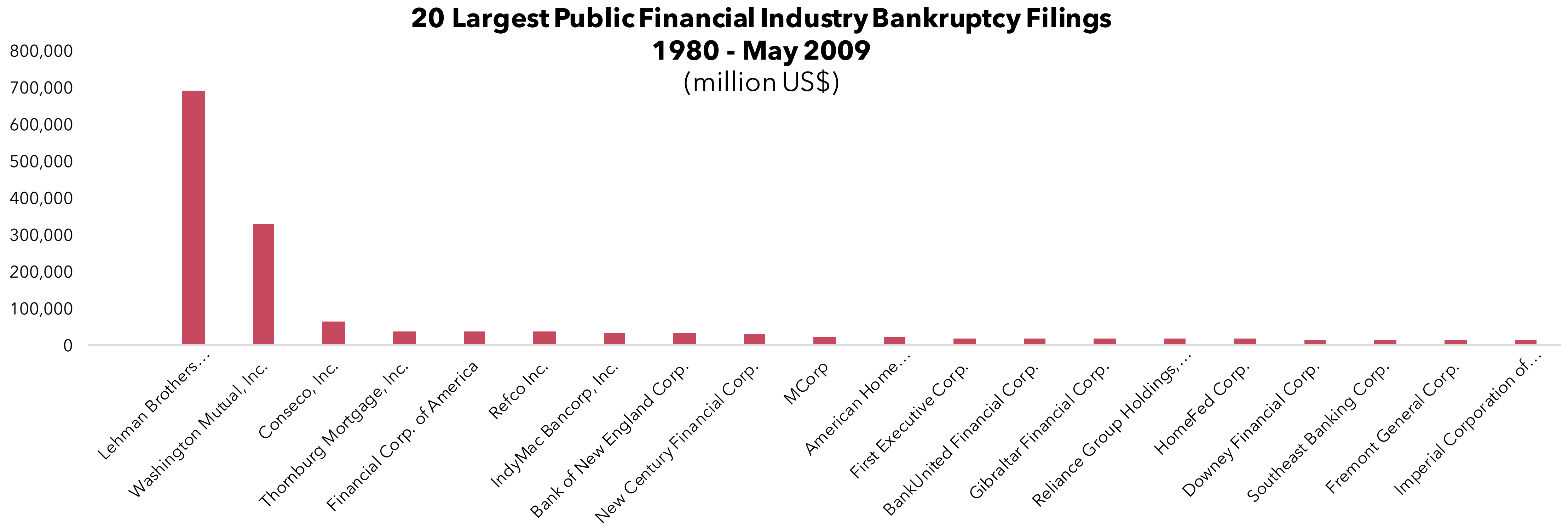
2008-2009

2006-2007

- Unemployment rate at 9.5% (the actual rate could be at 16.5%)
- Housing prices fell by 9.5%
- Dow Jones index witnessed the largest drop in intraday trading

Damage to Companies

- 2008 crisis led to the largest bankruptcies of Lehman Brothers and Washington Mutual.

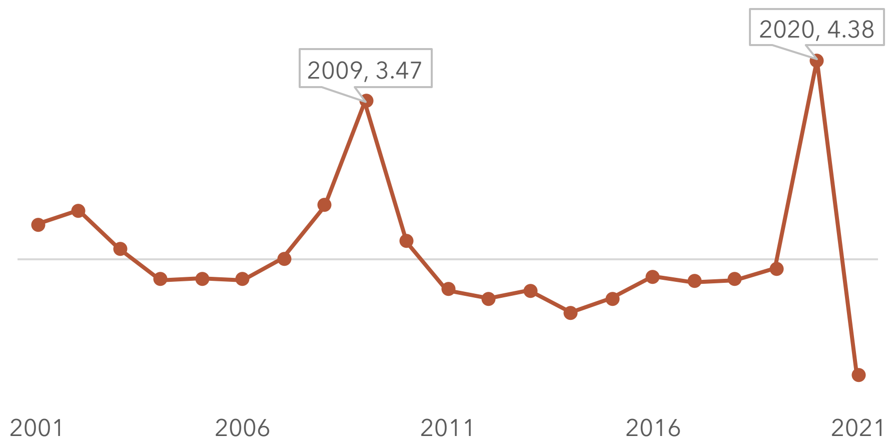


Source: BankruptcyData.Com - Experts in Bankruptcy Research [http://www.bankruptcydata.com/researchcenter2.htm]

Another Crisis?

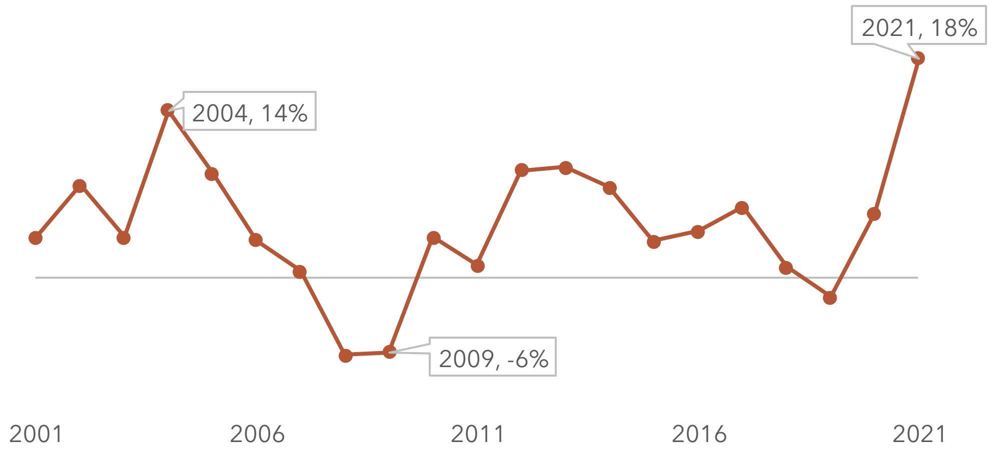
- In 2020, Covid-19 pandemic causing unemployment rate and housing price to spike.
- Pattern becomes similar to 2008 situation.
- Companies should be prudent to mortgage loan risk.

Annual Change in Unemployment Rate
(%)



Source: World Bank

Percentage Change in Median Housing Price
(%)



Source: U.S. Census Bureau and U.S. Department of Housing and Urban Development

Research Goal

Dataset	Description	Research goal
Mortgage Dataset	• Contain many macroeconomic factors such as Interest Rate, GDP, Unemployment Rate, etc.	1. Understand how macroeconomic factors affect loan default. 2. Develop customer profiles with different risk level. 3. Build models to predict default rate.
Home Equity Dataset	• Contain customer information such as Debt-to-Income ratio, Reason for Loan, Occupation.	

Mortgage Loan Dataset

Exploring effects of Macroeconomic and other factors on Mortgage Loans

Dataset Review

8,146 Rows
27 Variables

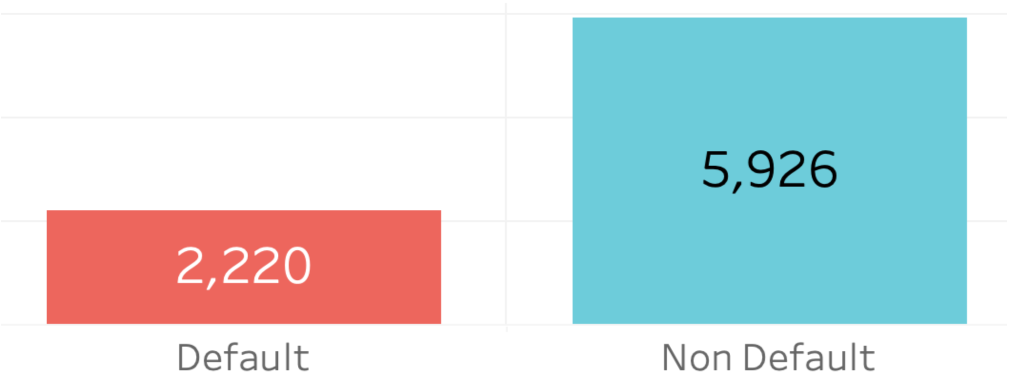
1 Unique ID
5 nominal variables
21 numerical variables

763 Missing Values
12 Numerical Variables with Outliers

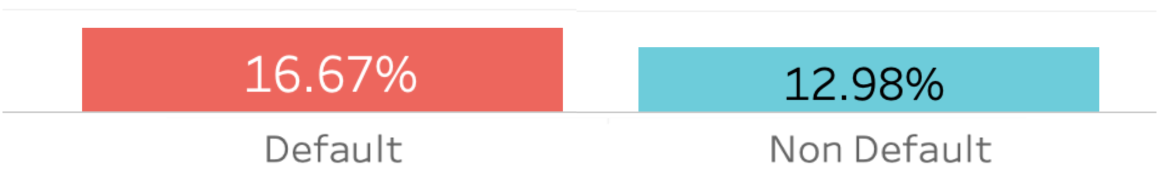
HOME MORTGAGE
LOAN SAMPLE DATASET

Customer Population

The Customer Distribution Chart



Investor Distribution by Groups

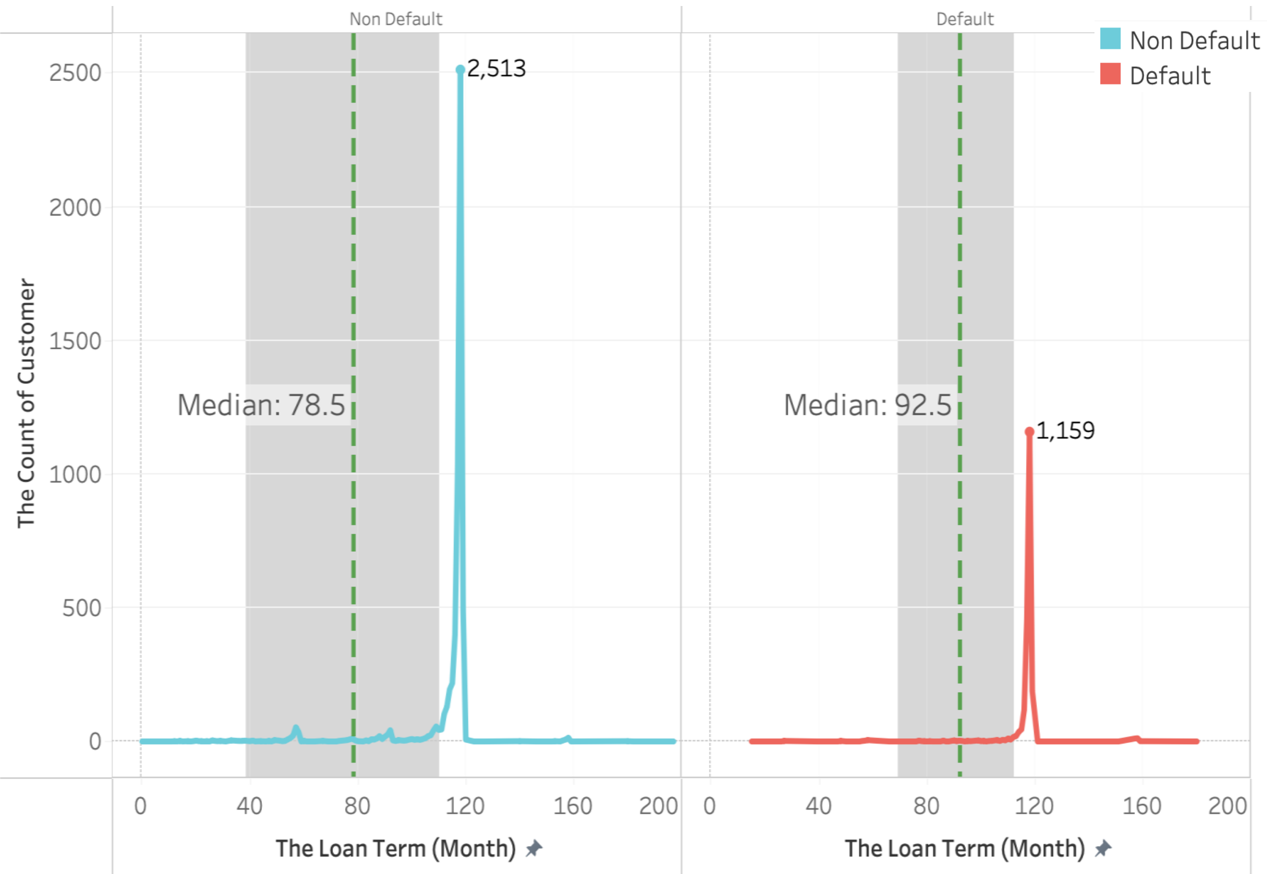


- Default customers account for the minority in the mortgage loan.
- Customers with the Borrower Identity in the Default group populate higher than them in the Non-Default one.

Mortgage Loan Term

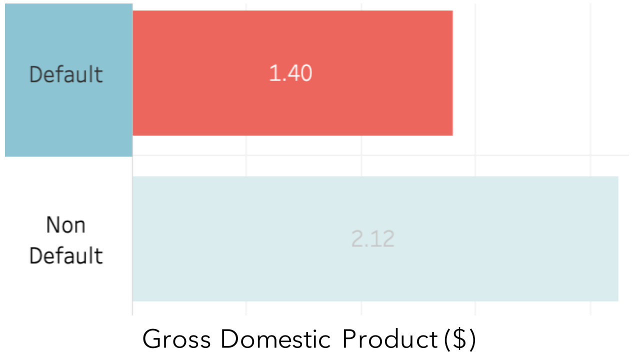
- The median mortgage loan term in the Default group is longer than another one.
- Most customers repaid the loan before 120 (10 years) months for both groups.

The Distribution of Loan Terms by Groups

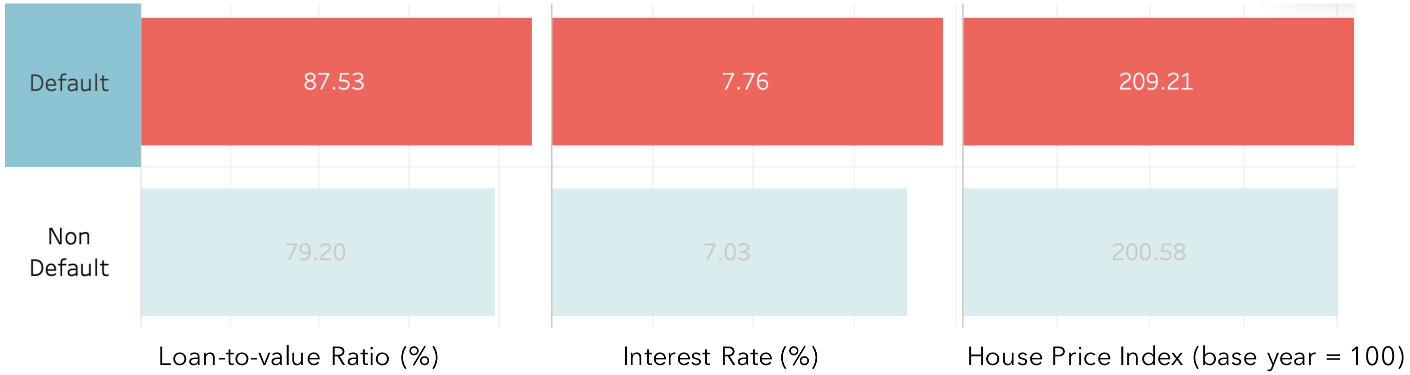


Market Conditions

The Average Variation of GDP Growth



Average Variations among Other Variables



- Slower increments of GDP were happening for customers who defaulted.
- Higher values about the loan structure were shown for customers who defaulted.

Home Equity Loan Dataset

Exploring effects of Customer Behavior on Home Equity Loans

Dataset Review

14 Variables
5962 Observations

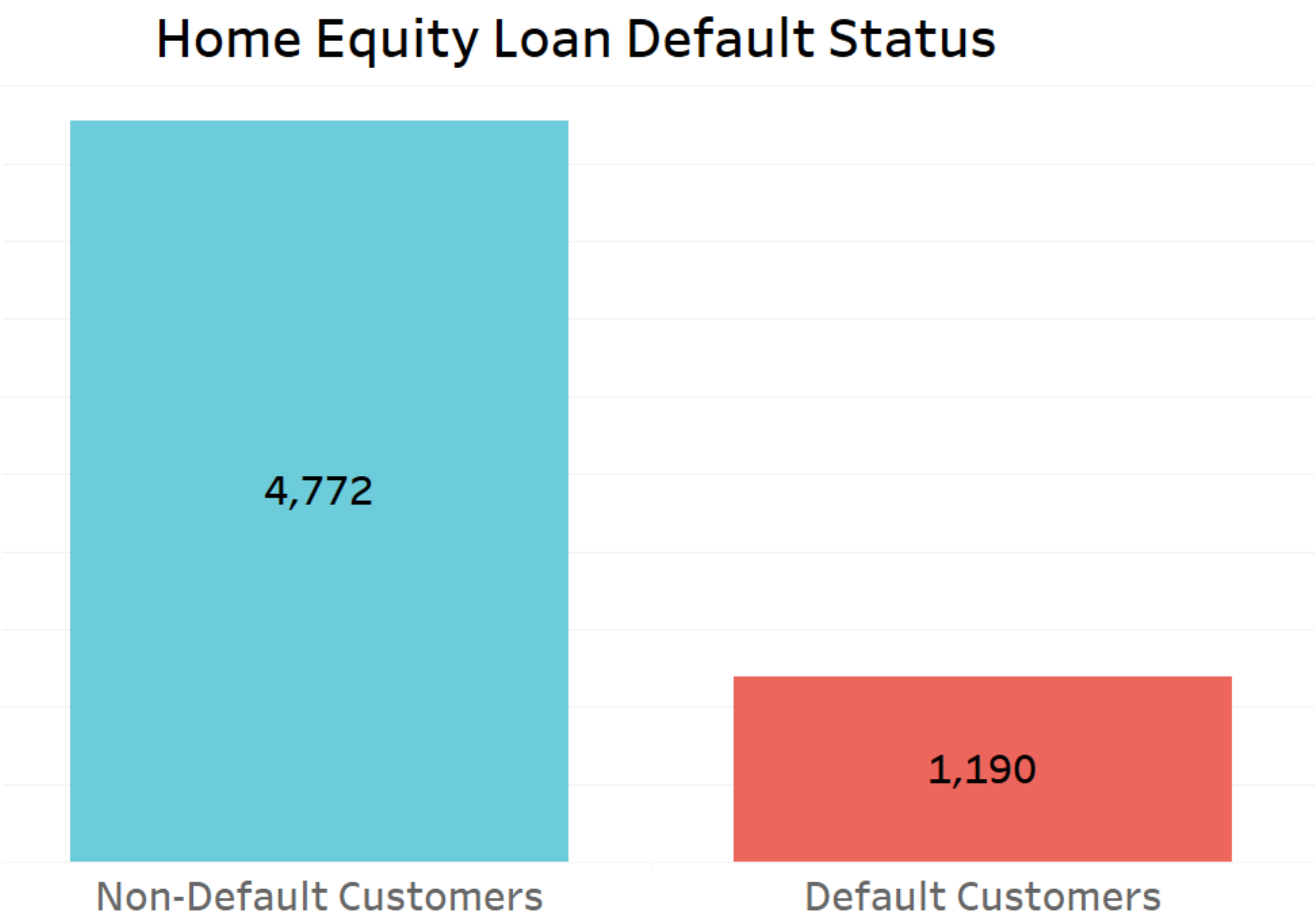
1 Unique ID
2 Categorical Variables
10 Numerical Variables

5277 Missing Values
(2620 rows)
Outliers in all Numerical
Variable

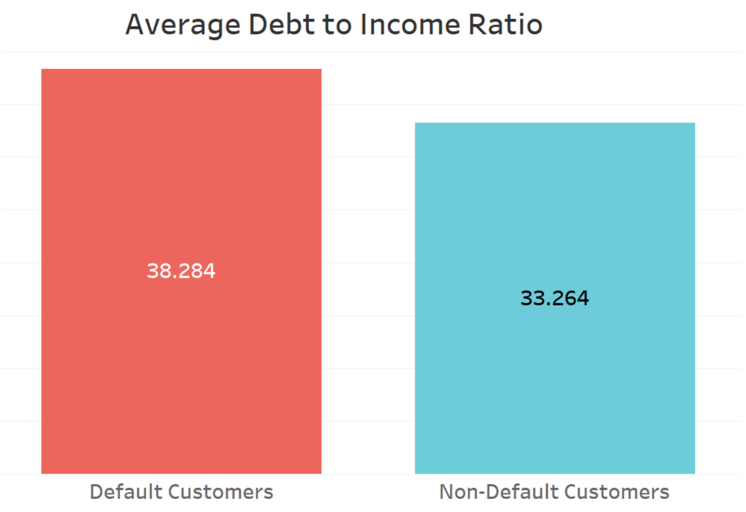
HOME EQUITY LOAN
DATASET

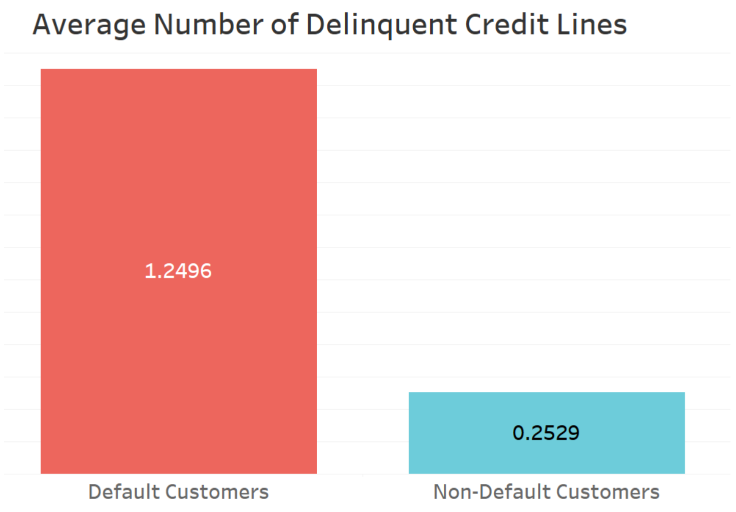
Variable of Interest: Default

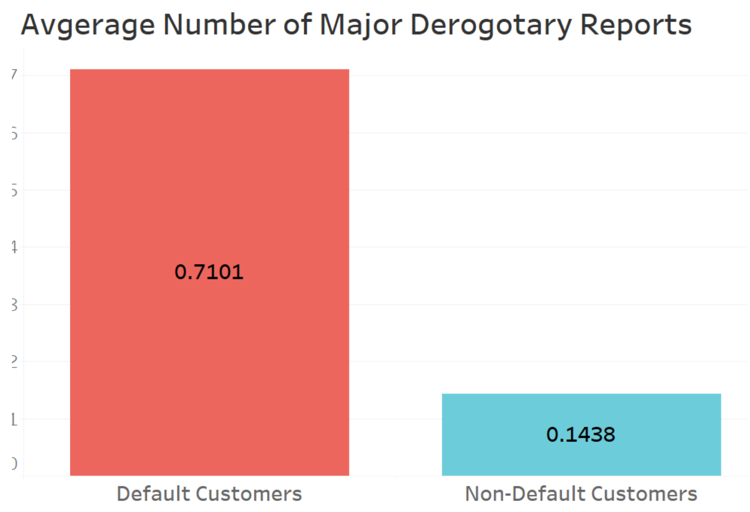
- Target variable: Default
  0: No default
  1: Default
- Class Imbalance exists in Target Variable
- Defaults in only 20% observations



Customer Profile- Home Equity Loan

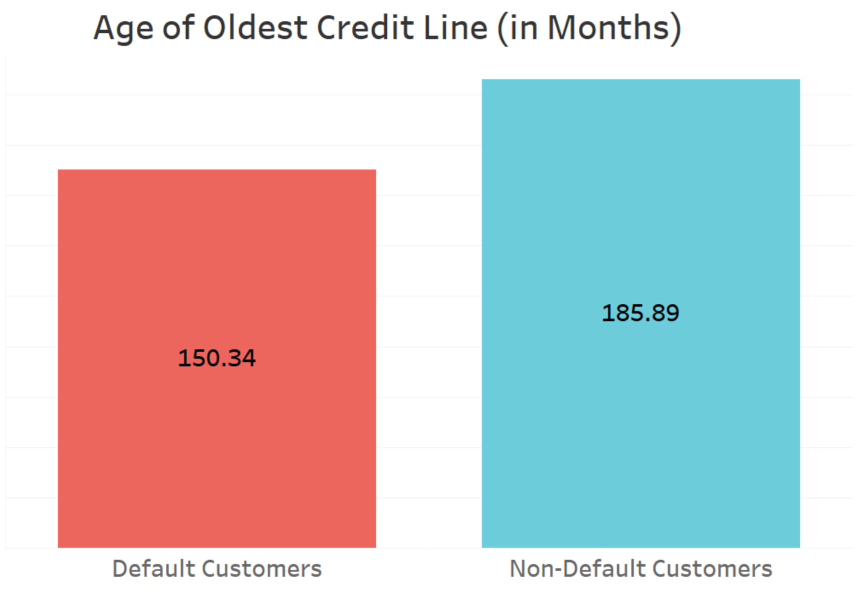


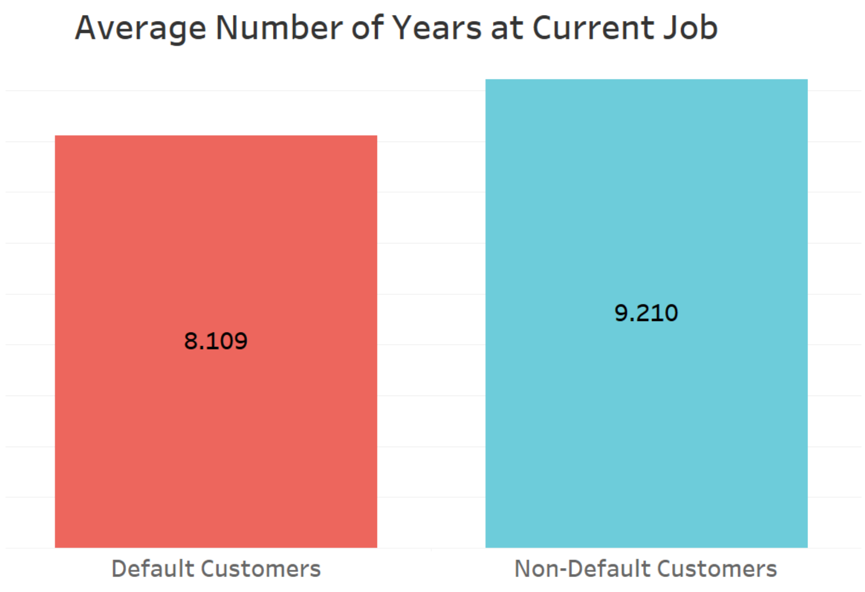




- Debt to income ratio is higher in defaulters.
- People with a higher number of delinquent credit lines have a higher rate of default.
- People with higher-than-average number of derogatory reports also have higher rate of default.

Customer Profile- Home Equity Loan





- Credit line age is lower among people who default.
- Years on current job is lower among people who default.

Clustering Analysis

Clusters 2, 3, 6, 7: (Lower Default Rate)

- Proportion of Default Customers ranging from 12% to 15%
- Higher Loan Amount, Mortgage Balance
- Lower Debt to Income Ratio
- Older Credit Lines
- Higher concentration of Occupation "Professional/ Executives"

Clusters 1, 4, 5, 8: (Higher Default Rate)

- Proportion of Default Customers ranging from 18% to 56%
- Higher Debt to Income Ratio
- Lower Credit Line Age and Years on Job
- More number or derogatory reports and delinquent credit lines
- Higher concentration of Occupation "Others"

Default Prediction

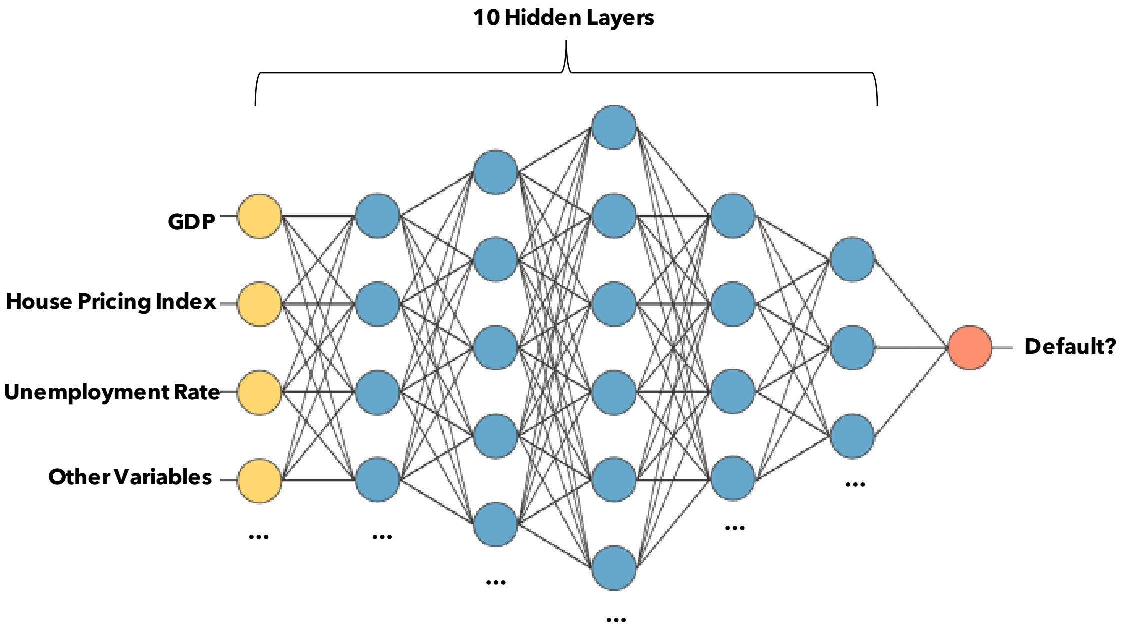
Building machine learning models to support default prediction

Prediction Model

- Research Support
  - Popular Models: SVM and Deep Neural Networks
  - Better Performance
- Deep Learning Model Advantages
  - Feature Generation Automation
  - Better Self-learning Capabilities
  - Cost Effectiveness
  - Better Performance Over Large Scale Of Data

Source: Advantage of Deep Learning, Plus Use Cases and Examples – [https://www.width.ai/post/advantages-of-deep-learning]

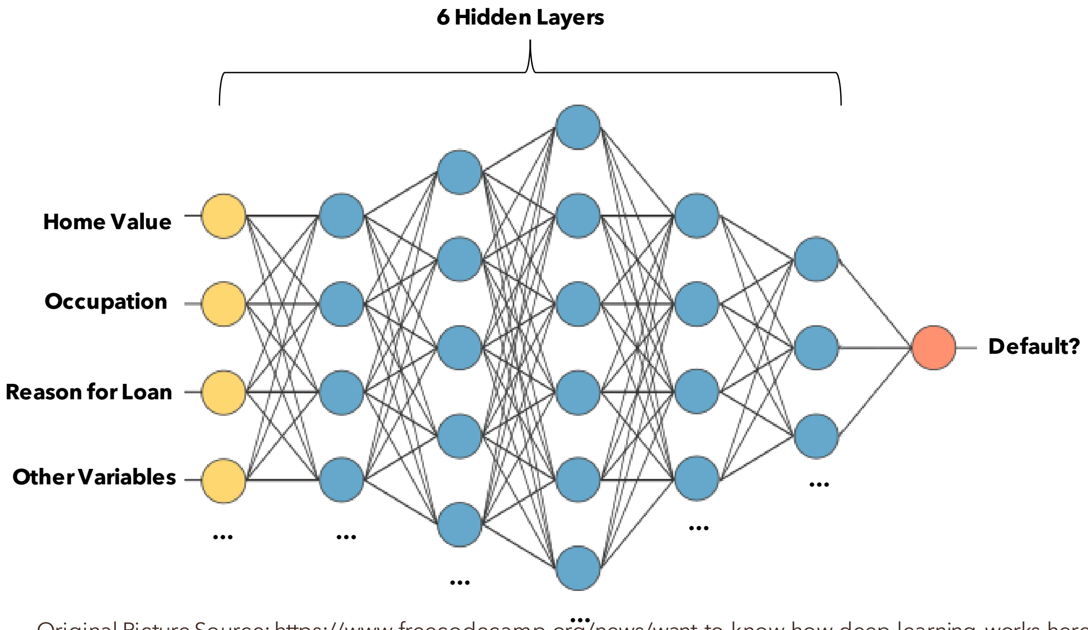
Model Performance – Mortgage Loan



Original Picture Source: https://www.freecodecamp.org/news/want-to-know-how-deep-learning-works-heres-a-quick-guide-for-everyone-1aedeca88076/

Metric	Value	Explanation
Accuracy	**0.81**	Proportion of correct predictions
Kappa	0.61	Agreement between model and random chance
Specificity	**0.80**	Actual non-default correctly classified.
Precision	0.79	Predicted default that are truly default.
Recall	**0.83**	Actual default correctly classified.
F1	0.81	Balance precision and recall.

Model Performance - Home Equity Loan



Original Picture Source: https://www.freecodecamp.org/news/want-to-know-how-deep-learning-works-heres-a-quick-guide-for-everyone-1aedeca88076/

Metric	Value	Explanation
Accuracy	**0.89**	Proportion of correct predictions
Kappa	0.78	Agreement between model and random chance
Specificity	**0.84**	Actual non-default correctly classified.
Precision	0.93	Predicted default that are truly default.
Recall	**0.84**	Actual default correctly classified.
F1	0.89	Balance precision and recall.

Insights & Recommendations

Holistic approach to preventing loan defaults

MORTGAGE LOAN

- Important factors
  - Macroeconomic factors: GDP Growth, HPI, and Interest Rate
  - Investor Borrower Identity Review
- Collect more customer information

HOME EQUITY LOAN

- Strengthen Approval Examination
  - Lower credit line years
  - Higher than average number of derogatory reports, delinquent credit lines, and number of inquiries
- Collect More Information
  - Expand occupation categories
  - FICO Score

- Credit Worthiness Prediction Model: Deep Learning

References

- Pérez-Martín*, A. Pérez-Torregrosa, M. Vaca. (2018), Big Data techniques to measure credit banking risk in home equity loans. Journal of business research, 2018, Vol.89, p.448-454 Retrieved from Big Data techniques to measure credit banking risk in home equity loans - ScienceDirect (drexel.edu)

- Amadeo, K. (2022). When and why did the stock market crash in 2008? The Balance. Retrieved July 8, 2022, from https://www.thebalance.com/stock-market-crash-of-2008-3305535#:~:text=The%20stock%20market%20crash%20of%202008%20occurred%20on%20September%2029,largest%20point%20drop%20in%20history.

- Chad Cowden, Frank J. Fabozzi, and Abdolreza Nazemi. (2019), Default Prediction of Commercial Real Estate Properties Using Machine Learning Techniques. Journal of portfolio management, 2019, Vol.45 (7), p.55-67. Retrieved from Default Prediction of Commercial Real Estate Properties Using Machine Learning Techniques - ProQuest

- Christie, L. (2009, February 12). Home prices in Record Plunge. CNNMoney. Retrieved July 8, 2022, from https://money.cnn.com/2009/02/12/real_estate/Latest_median_prices/#:~:text=The%20median%20price%20for%20a,1.6%25%20between%202006%20and%202007

- Edward N.C. Tong∗, Christophe Mues & Lyn Thomas (2013). A zero-adjusted gamma model for mortgage loan loss given default, *International Journal of Forecasting*, 29 (4), 548-562. Retrieved from https://doi.org/10.1016/j.ijforecast.2013.03.003

- Emad Azhar Ali, Syed; Sajjad Hussain Rizvi, Syed; Lai, Fong-Woon; Faizan Ali, Rao; Ali Jan, Ahmad (2021), Predicting Delinquency on Mortgage Loans: An Exhaustive Parametric Comparison of Machine Learning Techniques. International Journal of Industrial Engineering and Management, 2021, Vol.12 (Issue 1), p.1-13. Retrieved from Predicting Delinquency on Mortgage Loans: An Exhaustive Parametric Comparison of Machine Learning Techniques - ProQuest (drexel.edu)

- Feibelman, A. (2022). BANKRUPTCY AND THE STATE. *Emory Bankruptcy Developments Journal, 38*(1), 1-50. Retrieved from http://ezproxy2.library.drexel.edu/login?url=https://www.proquest.com/scholarly-journals/bankruptcy-state/docview/2637689114/se-2?accountid=10559

- Goolsbee, Austan D., and Alan B. Krueger. 2015. "A Retrospective Look at Rescuing and Restructuring General Motors and Chrysler." *Journal of Economic Perspectives*, 29 (2): 3-24.DOI: 10.1257/jep.29.2.3

- Jackson, J. R., & Kaserman, D. L. (1980). Default risk on home mortgage loans: A test of competing hypotheses: ABSTRACT. *Journal of Risk and Insurance (Pre-1986), 47*(4), 678. Retrieved from http://ezproxy2.library.drexel.edu/login?url=https://www.proquest.com/scholarly-journals/default-risk-on-home-mortgage-loans-test/docview/235110404/se-2?accountid=10559

- Kim, Dong-sup; Shin, Seungwoo, THE ECONOMIC EXPLAINABILITY OF MACHINE LEARNING AND STANDARD ECONOMETRIC MODELS-AN APPLICATION TO THE U.S. MORTGAGE DEFAULT RISK. International journal of strategic property management, 2021, Vol.25 (5), p.396-412. Retrieved from THE ECONOMIC EXPLAINABILITY OF MACHINE LEARNING AND STANDARD ECONOMETRIC MODELS-AN APPLICATION TO THE U.S. MORTGAGE DEFAULT RISK. - Document - Gale Academic OneFile (drexel.edu)

- Magdoff, F., & Yates, M. D. (2010). ABCs of the economic crisis: What Working People Need To Know. Aakar Books. Retrieved from https://books.google.com/books?hl=en&lr=&id=LzlVCgAAQBAJ&oi=fnd&pg=PA9&dq=what+happens+after+economic+crisis&ots=Pd0VyZNGJW&sig=hnlfe1t32HVBmMj1hlpiN7UXPVY#v=onepage&q=what%20happens%20after%20economic%20crisis&f=false .

- Mitrašević, M., & Bardarova, S. (2020). CAUSES OF NON-PAYMENT OF MORTGAGE LOANS: THEORETICAL AND PRACTICAL ASPECTS. *UTMS Journal of Economics, 11*(2), 138-150. Retrieved from http://ezproxy2.library.drexel.edu/login?url=https://www.proquest.com/scholarly-journals/causes-non-payment-mortgage-loans-theoretical/docview/2571154652/se-2

- Song, G. (2022). Large US bank takeovers in 2008: Performance and implications. *Journal of Capital Markets Studies, 6*(1), 33-47. https://doi.org/10.1108/JCMS-06-2021-0021

- Teng, Huei-Wen; Lee, Michael. (2019), Estimation Procedures of Using Five Alternative Machine Learning Methods for Predicting Credit Card Default. Review of Pacific basin financial markets and policies, 2019, Vol.22 (3), p.1950021. Retrieved from Estimation Procedures of Using Five Alternative Machine Learning Methods fo...: EBSCOhost (drexel.edu)

- THE FINANCIAL CRISIS INQUIRY COMMISSION (2011), Final Report of the Causes of the Financial and Economic Crisis in the United States. Retrieved from https://cybercemetery.unt.edu/archive/fcic/20110310173545/http://c0182732.cdn1.cloudfiles.rackspacecloud.com/fcic_final_report_full.pdf

THANK YOU!