### **1. Selection of Credit Model**

For this project, we selected **personal loans** as the credit product for our analysis. Our goal was to develop a traditional credit scoring model that can assign a credit score to any new applicant based on a set of predefined variables.

**Cleaning process and EDA**

As part of the data cleaning process, we conducted the following steps. Some of this analysis was not later used or taken into consideration for our models, but they were made to ensure good practice and to leave the bases for a possible project continuation and/or improvement.

1. **Handling Missing Values** – We checked for null values across all variables.
2. **Variable Classification** – We categorized the variables into nominal, ordinal, interval, and ratio scales.
3. **Cardinality Analysis** – For nominal and ordinal variables, we examined unique values to assess whether they exhibited high or low cardinality.
4. **Proportional Distribution** – We analyzed the proportion of each unique value within the nominal and ordinal variables.
5. **Duplicate Detection** – We checked for duplicate records to ensure data integrity. The result was 0 duplicates.
6. **Histogram Analysis –** To ensure we selected the best variables for our Traditional Credit Model, we plotted each of our ratio ratio variables in which each of the three Scorecard groups (bad, standard, good) was contrasted. This way, an analysis was made to determine the best possible predicting variables.

*See attached document for a more detailed analysis*

**2. Traditional Credit Scoring Model Report**

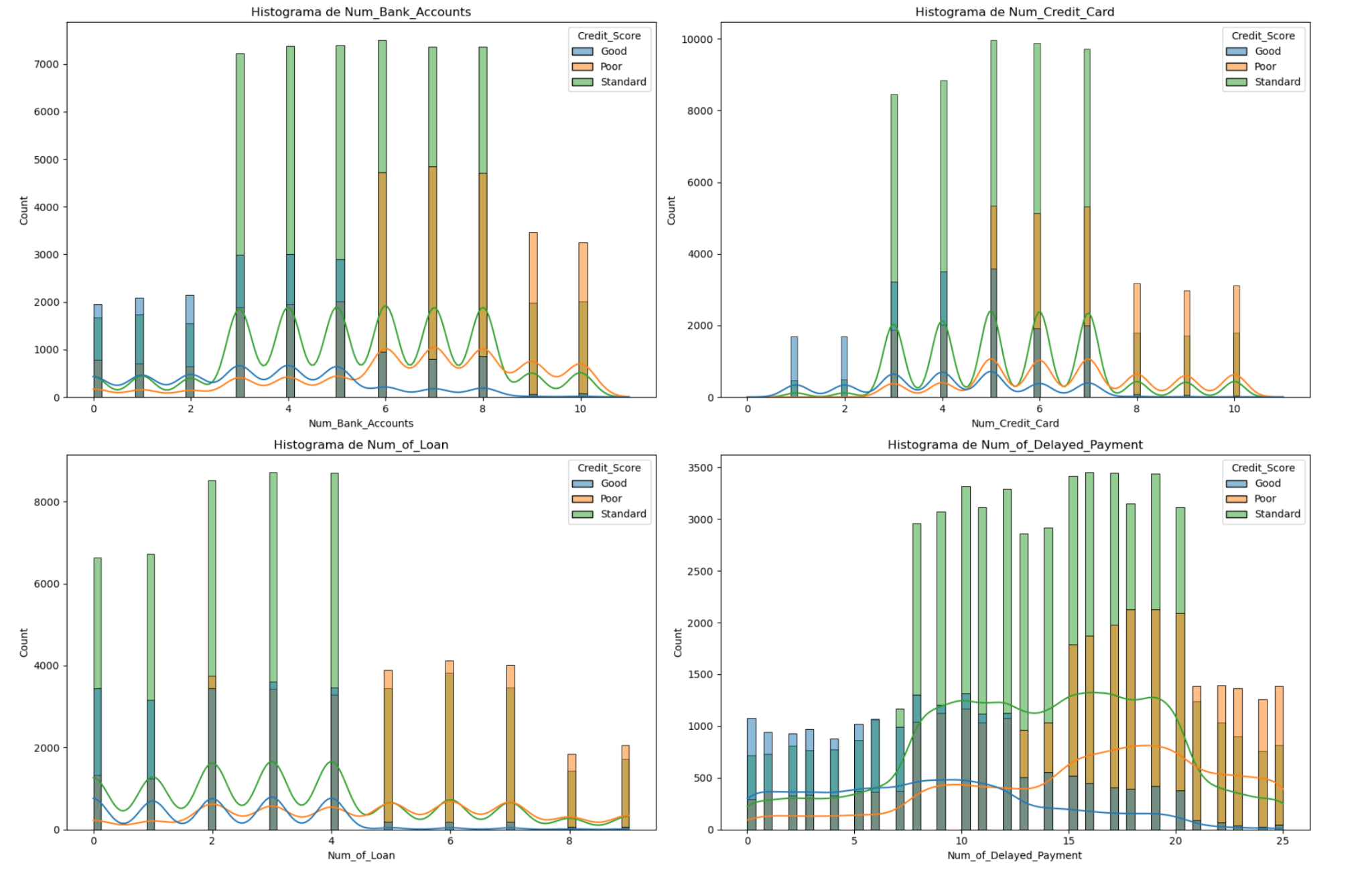
**Introduction and Approach**

This model was built using a methodology inspired by the **Altman Z-score**, where we manually assigned weights to selected features and computed a final credit score. The output of the model categorizes applicants into three possible credit score classifications: **Bad, Standard, and Good**.

### **Theoretical Framework**

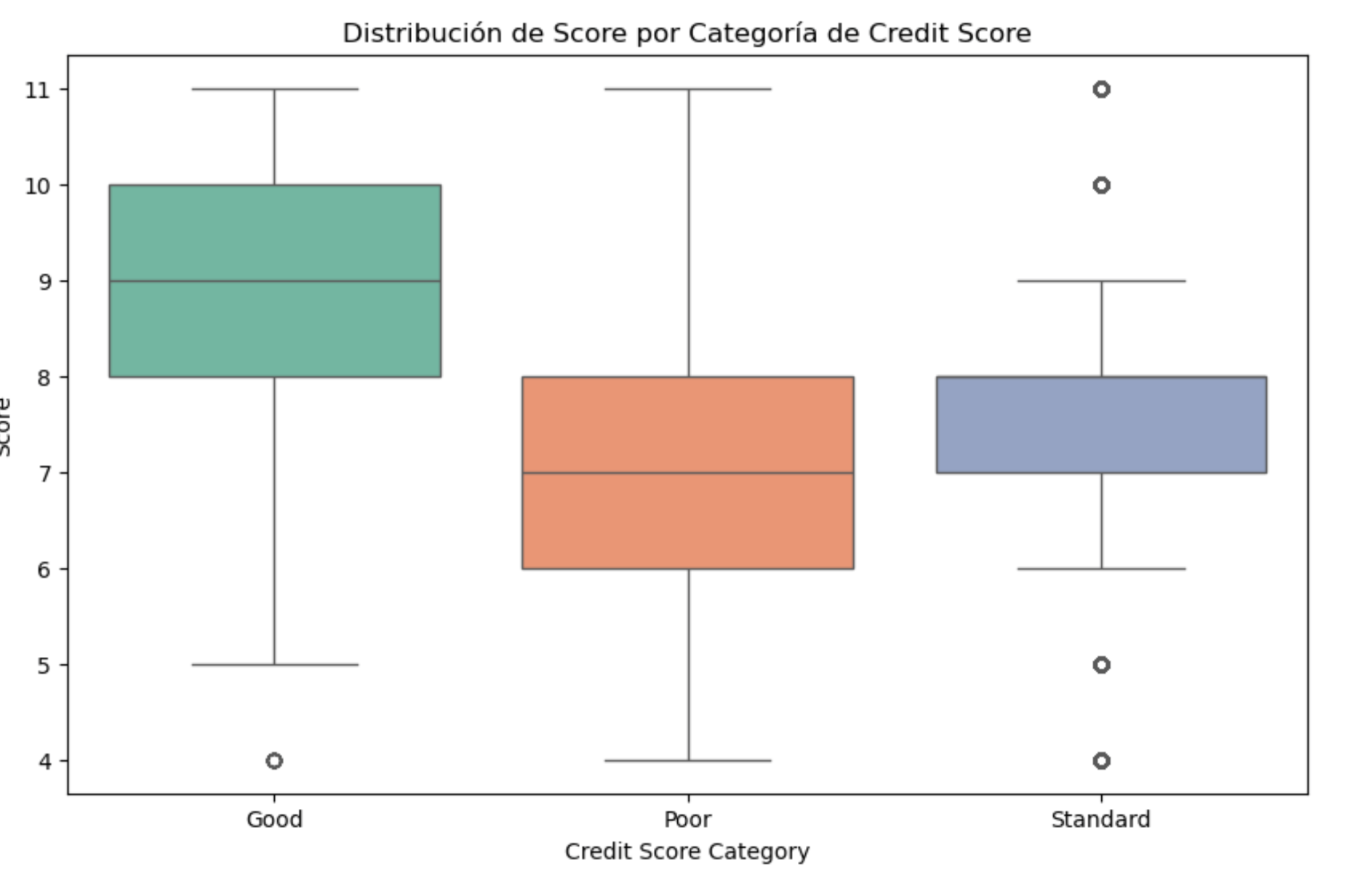
Our scoring model relies on four key variables that influence a borrower’s creditworthiness:

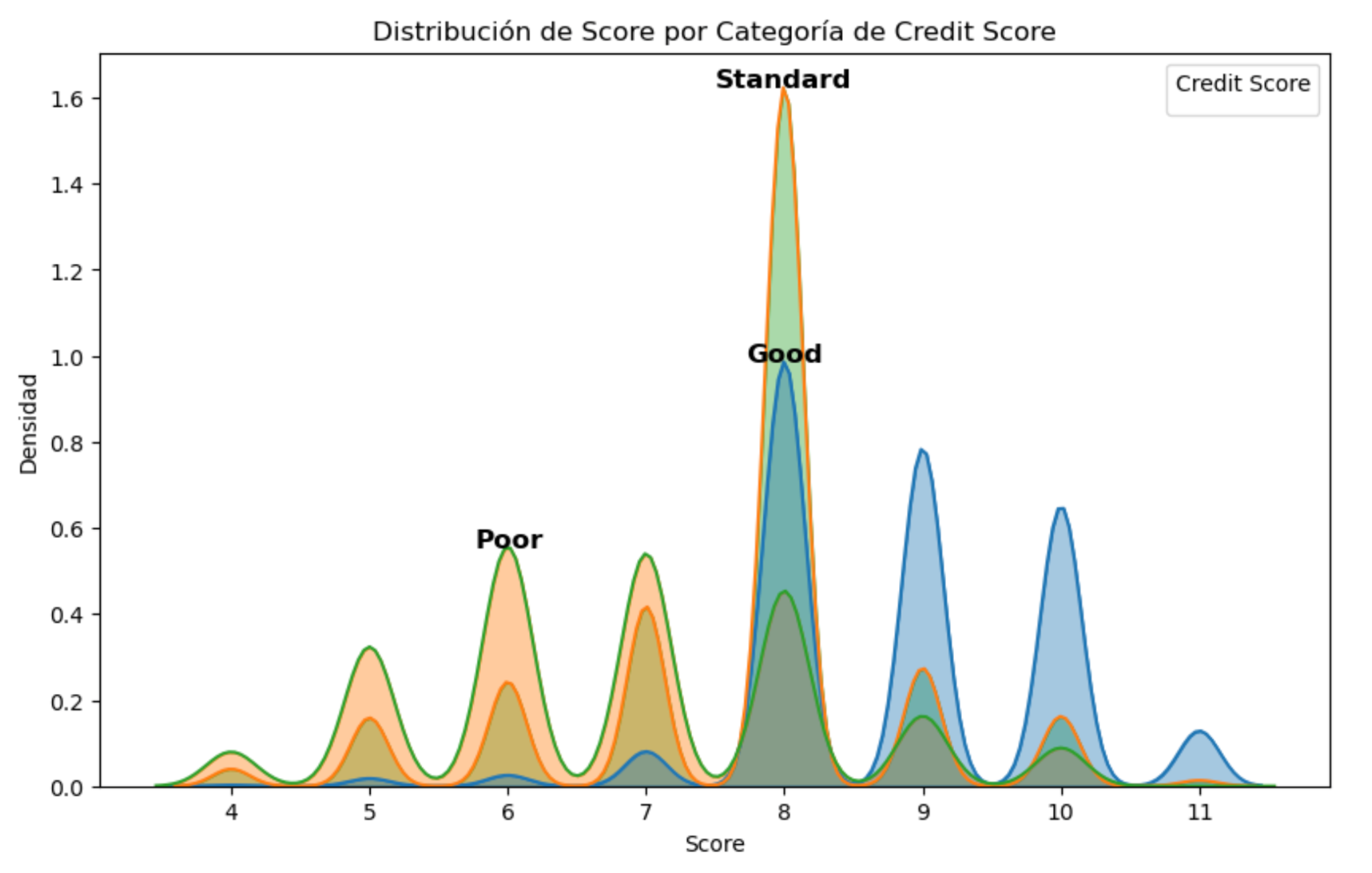
1. **Num\_Bank\_Accounts** – The number of bank accounts held by the applicant.
2. **Num\_Credit\_Card** – The number of credit cards owned by the applicant.
3. **Num\_of\_Loan** – The total number of loans taken by the applicant.
4. **Num\_of\_Delayed\_Payment** – The total number of delayed payments recorded for the applicant.



Each of these variables was assigned a weight based on the visual analysis of the distributions of each variable and the values acquired by them on each of the three credit scores. The final credit score is calculated as a weighted sum of these variables, and applicants are classified into one of the three credit score categories accordingly.

A distribution analysis of the scores assigned was made to be better assured the final groups were being assigned correctly:





This analysis was made with the one hundred thousand records found in the training dataset. So, an assumption that was made is that the distributions of this training dataset used are the same or similar to the ones of the test dataset (fifty thousand records) which we didn't use.

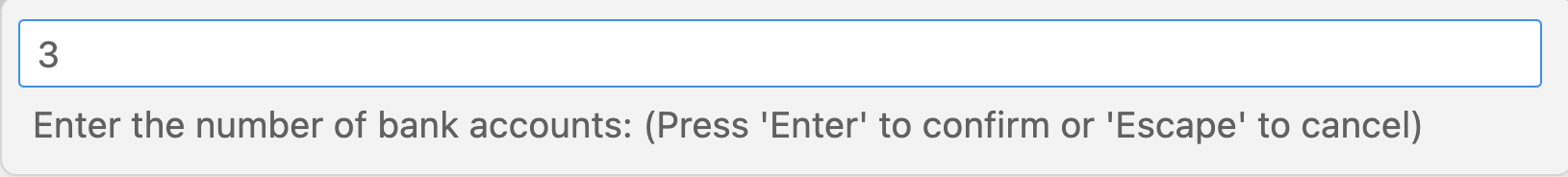
### **Results**

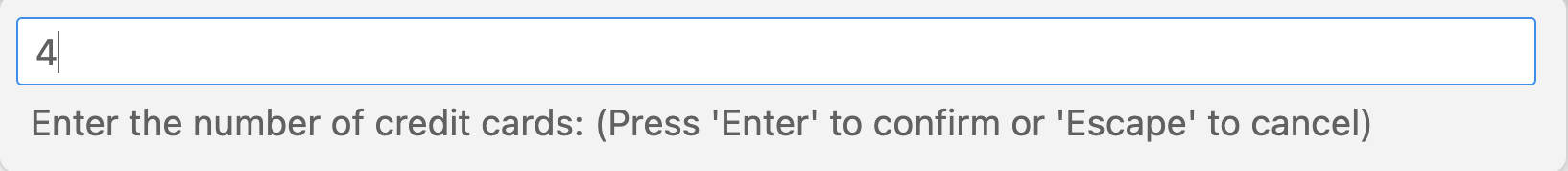
*The final scores and classifications made by our model are as follows:*

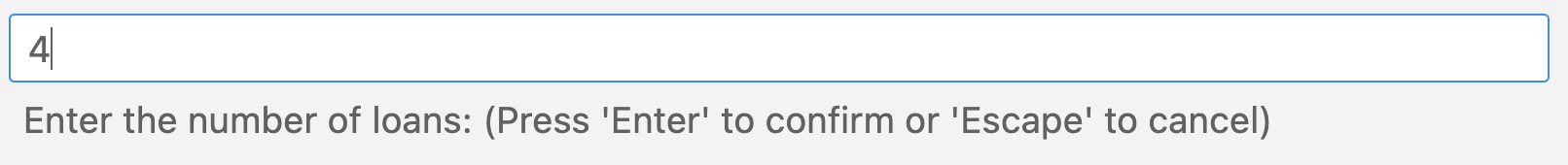
* Poor: 4 to 6 points
* Standard: 7 to 8 points
* Good: 9 to 12 points

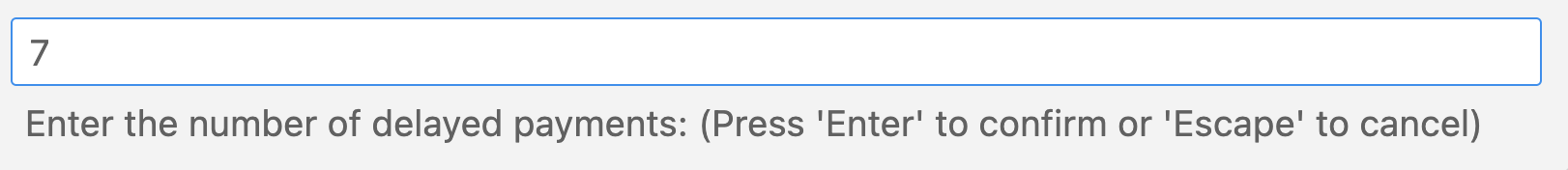
**New credit Score decisions demonstration**

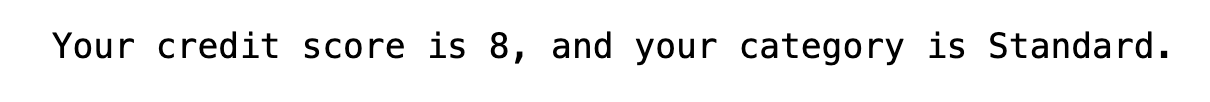
As part of the development of this model, an adaptation was made so that new “applicants” can enter their information and get a Credit Score along with their classification. Here is an example:

****

****

****

****

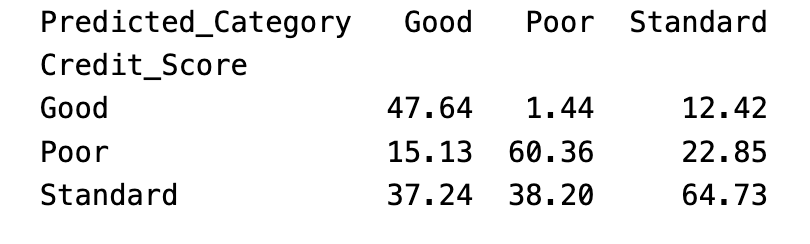
****

#### **Model Performance**

To assess the performance of our credit scoring model, we compared its accuracy against a benchmark classification. The key findings are as follows:

* **Our model’s accuracy:** **60%**
* **Benchmark accuracy:** **64%**

**Confusion Matrix:**

****

While our model performed slightly below the benchmark, it still demonstrates a good level of predictive power given the simplicity of the manually assigned weightings. However, this model underperformed significantly to our final Credit Risk Model shown in section 4 (Analysis on that section).

Looking at our Confusion Matrix, we appreciate that there is a strong concentration of Standard values classified as Good, although this does not represent a big economic problem since, as we saw in our graphs, these two groups share a lot of similarities in their behaviours and incorrectly giving credit to a Standard person would not represent a big problem. On the other hand, there is a really small percentage of less than 16% of clients predicted as Good when they should have been considered Poor, so, this small percentage should not represent a major concern.

#### **Analysis of Possible Improvements**

This analysis is made with the supposition that we continue with this manual method of assigning scores, since it is more than evident that the most obvious way to improve it would be with a Machine or Deep learning approach or simply using another traditional credit model like Merton.

Several potential improvements could enhance the accuracy and reliability of our credit scoring model:

1. **Trying new ratio variables:** A way to possibly improve the scorecard would be to try to incorporate, reduce, or try completely new variables in the attempt to get better results.
2. **Feature Engineering:** Transforming existing variables like the nominal or interval variables that were left out and using them for the scorecard could lead to better classification results.

**Conclusion**

The traditional credit scoring model presented in this section offers a simple yet interpretable approach to classifying credit risk, inspired by the principles of the Altman Z-score. By manually assigning weights to four key financial variables—number of bank accounts, credit cards, loans, and delayed payments—we were able to create a scorecard that segments applicants into three categories: Bad, Standard, and Good.

Despite its relatively modest accuracy of 60%, slightly below the 64% benchmark, the model provides valuable insights and a solid foundation for credit risk assessment, particularly in settings where transparency and interpretability are critical. The confusion matrix analysis revealed that most misclassifications occurred between Standard and Good applicants, a less concerning outcome given the behavioral similarities observed between these groups. Furthermore, the model showed strong reliability in minimizing the misclassification of high-risk (Bad) applicants as Good, with this error occurring in fewer than 16% of cases.

While the performance falls a little short when compared to our first benchmark (compared to more advanced models in section 4), this scoring system demonstrates the feasibility of a manually constructed model and offers a baseline for further development. This initial approach lays the groundwork for a transparent and easily explainable credit assessment framework.

**Reflection Axel:**

First, when we used the Traditional Scorecard model, it was valuable to see how these are formed in the most basic way possible, even more when we realize that even nowadays companies use such simple equations to determine whether to give credits or not. Then, building more complex models with the help of AI was a big step forward to understand how such powerful tools can help us ace difficult tasks, but that it is extremely important to understand what those things actually do in order to make new or fix/improve current models.

2 conclusiones!!!!

# 3. Interest Rate Formation Analysis

## Loan Types and Market Distribution

To effectively evaluate credit offerings and tailor risk and pricing models, it’s essential to understand the different types of loans available in the market. Among the various options, **personal loans stand out** for their flexibility and broad consumer appeal. While we briefly review the main loan types, the analysis here centers on **personal loans**.

### Overview of Loan Types

The credit market includes the following common loan categories:

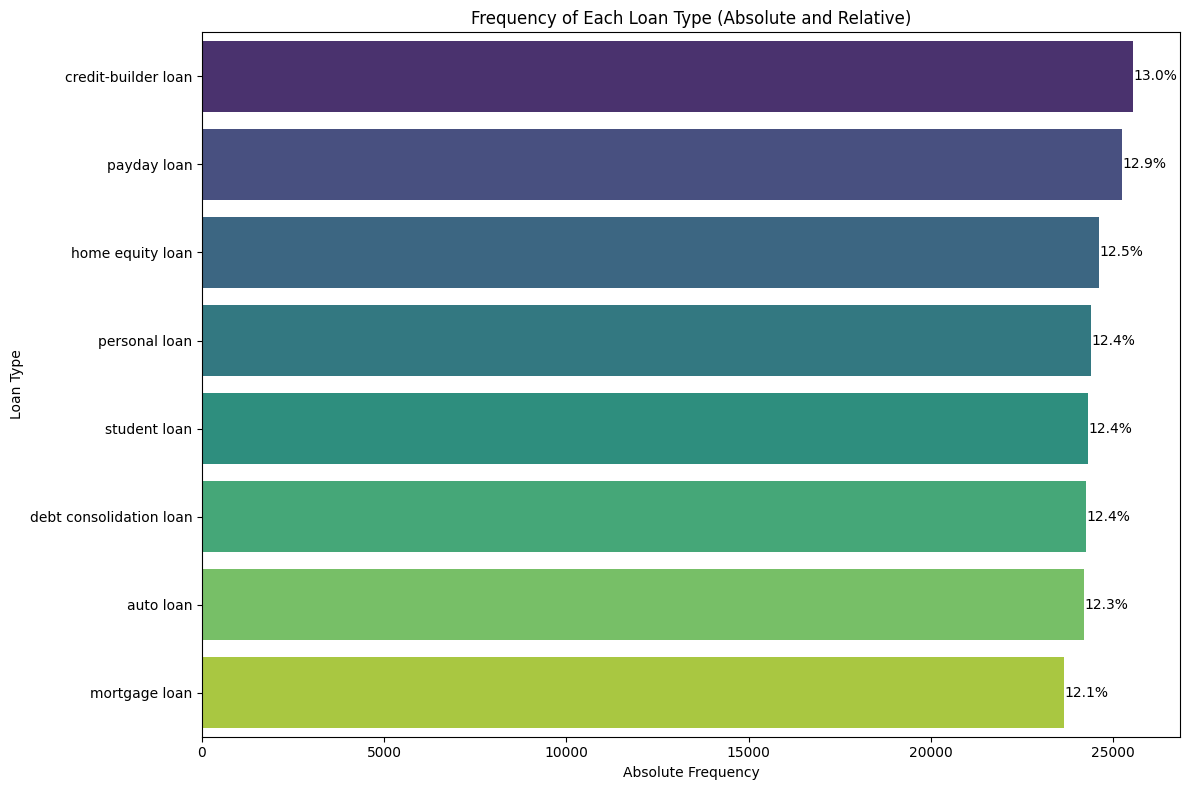
| **Type of Loan** | **Description** |
| --- | --- |
| **Credit-builder loan** | Small loans designed to help individuals build or rebuild credit history. |
| **Payday loan** | Short-term, high-interest loans typically due on the borrower’s next payday. |
| **Home equity loan** | A lump-sum loan secured by the borrower's home equity, typically for large expenses. |
| **Personal loan** | Unsecured loans for various personal expenses, such as medical bills or travel. |
| **Student loan** | Loans designed to cover education-related costs, often with lower interest rates. |
| **Debt consolidation loan** | Combines multiple debts into one, usually with a lower overall interest rate. |
| **Auto loan** | Secured loans used to purchase vehicles. |
| **Mortgage loan** | Long-term loans used to buy real estate, secured by the property itself. |

### Credit Product Distribution and Interpretation

Loan types in the dataset are fairly evenly distributed:

| **Type of Loan** | **Relative Frequency (%)** |
| --- | --- |
| **Credit-builder loan** | 13.02 |
| **Payday loan** | 12.86 |
| **Home equity loan** | 12.54 |
| **Personal loan** | 12.43 |
| **Student loan** | 12.39 |
| **Debt consolidation loan** | 12.36 |
| **Auto loan** | 12.34 |
| **Mortgage loan** | 12.05 |

And can be further illustrated by the following figure:



Although **personal loans** are not the most frequent, they hold a consistent and strong presence in the market. Their versatile nature makes them attractive for a wide range of consumers.

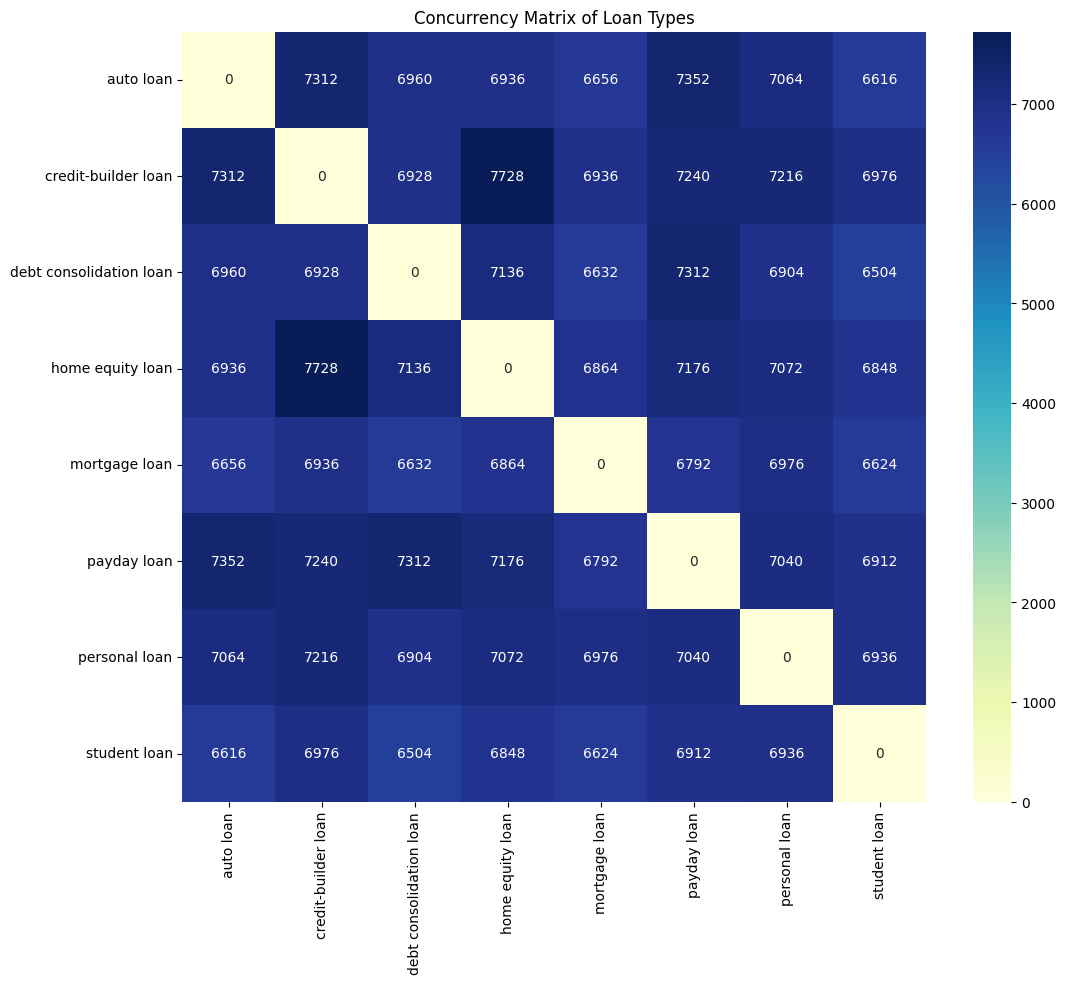
**Why Focus on Personal Loans?**Personal loans are among the most **accessible and flexible** credit products. They are unsecured, require no collateral, and serve various purposes—from medical bills and home repairs to debt consolidation. This makes them ideal for a diverse group of borrowers, especially those with decent credit scores but no assets to pledge.

#### **Common Pairings**

In our data, personal loans are often **paired with student loans**, indicating that borrowers who invest in education may later rely on personal loans for related expenses or financial challenges.

On the contrary, personal loans are **rarely combined with mortgage or debt consolidation loans**, likely because of differing borrower profiles and use cases.

| **Loan Type** | **Reasons to Choose** | **Potential Drawbacks** |
| --- | --- | --- |
| **Personal Loan** | - No collateral required  - Versatile usage  - Competitive interest rates | - Rates may be higher for low-credit borrowers |
| **Payday Loan** | - Fast cash access | - Extremely high interest rates, risk of debt cycle |
| **Student Loan** | - Enables education investment | - Long repayment terms; interest can accumulate quickly |
| **Home Equity Loan** | - Lower rates for secured borrowing | - Risk of losing home if defaulted |
| **Debt Consolidation** | - Simplifies payments, improves credit if managed well | - May encourage overspending if not disciplined |



### 

## Deep diving into Personal Loans

### Rate Components for Personal Loans

Setting interest rates for personal loans involves a multifaceted evaluation of macroeconomic indicators, borrower-specific risk, and market competition. The following components are key to understanding how rates are established for personal loans.

#### 1. Rate Analysis

In this section, it will be discussed how a rate is defined, what are it’s most important factors and how we may determine ours.

| **Component** | **Definition** | **Effect in Personal Loans** | **Unit used to measure** |
| --- | --- | --- | --- |
| Base Rate | The base rate in the U.S. is typically the federal funds rate, set by the Federal Reserve. It's determined by economic conditions like inflation, employment, and growth targets. This rate influences all other interest rates in the economy, including those for loans. Lenders use it as a benchmark, adding premiums to cover risks and costs, so changes in the base rate directly affect the cost of borrowing for consumers and businesses (Investopedia, 2025). | As of April 3rd 2025, the federal funds effective rate is 4.33% (Federal Reserve Board, n.d.).  For rational investors, this rate serves as a benchmark: it is essentially a risk-free return. Therefore, it would be irrational to invest in assets or extend credit yielding less than this rate, given the additional risks and administrative burdens involved in private lending. | Federal Reserve Base Rate |
| Inflation Premium | Compensation added to the base rate to offset the eroding effect of inflation on the purchasing power of money over time. | With inflation currently at 2.8% as of March 2025 (US Inflation Calculator, 2025), lenders incorporate an inflation premium to ensure that the real returns on loans are not diminished by rising prices. This premium protects lenders from the adverse effects of inflation during the loan term. | Inflation (in this case, year to year) |
| Credit Risk Premium | Additional charge to compensate lenders for the risk that a borrower may default on the loan. | Personal loan interest rates vary significantly based on creditworthiness. For instance, borrowers with good credit (690 to 719 credit score) face average APRs of 16.61% (Millerbernd, 2025), while those with fair credit may encounter higher rates. This premium reflects the lender's assessment of the borrower's likelihood to repay the loan. | Credit Score |
| Liquidity Premium | Extra yield demanded by lenders for loans that are not easily tradable or convertible to cash without a significant loss in value. | Personal loans are typically illiquid assets, meaning they cannot be quickly sold or exchanged for cash without potential loss. To compensate for this lack of liquidity, lenders may add a liquidity premium to the interest rate. This premium ensures that lenders are compensated for the potential difficulty in converting the loan into cash if needed. | This is less intuitive, but income statements, comparison of income to amount of loans, etc.  Anything to measure the available amount of money a borrower may have. |
| Administrative Costs | Expenses incurred by lenders in processing, underwriting, and servicing loans. | Lenders charge origination fees to cover these administrative costs. For example, some lenders may charge an administration fee of up to 10% (U.S. News, 2025). These fees are often incorporated into the overall cost of the loan, affecting the annual percentage rate (APR) offered to borrowers. | Administrative Fee/ Origination Fee |
| Profit Margins | The portion of the interest rate that contributes to the lender's profitability after covering all associated costs and risks. | Lenders set interest rates not only to cover costs and risks but also to achieve a desired profit margin. This margin varies among institutions based on their business models, competition, and market conditions. | Profit Margin (as a percentage) applied to the APR after all other expenses are considered. |
| Risk Factors | Market | The risk of reduced profitability or competitiveness due to changes in the external lending environment, especially competitors’ rates. | Rates of other lenders available to consumers. |
| Regulatory | Changes in lending laws, caps on interest rates, or stricter consumer protection rules can affect loan profitability or restrict lending practices. For instance, state-level usury laws may limit the maximum allowable APR, reducing flexibility for risk-based pricing.  This risk is almost negligible due to the usually low rates Personal Loans are issued at. | Caps on interest rates, max APR, etc. |
| Technological | The risk of loss or disruption due to technological failure, cyber threats, or inadequate digital infrastructure. | IT expenditure (server maintenance, cybersecurity auditories, etc.) integrated into administrative costs. |

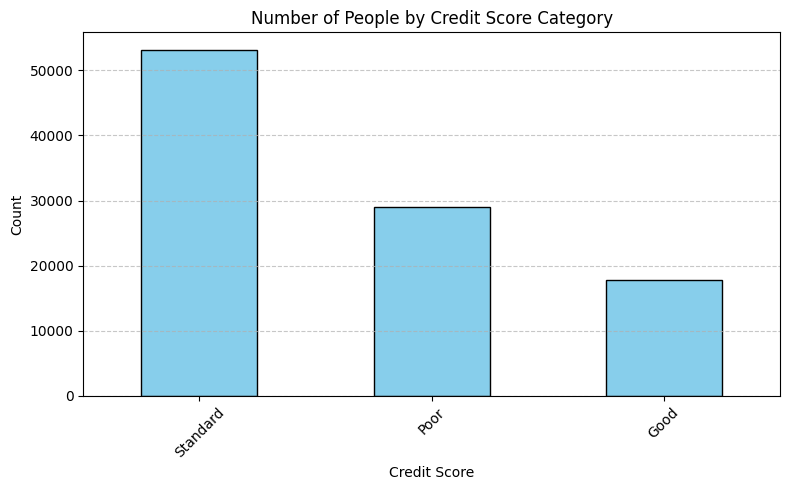
#### 2. Market and demand

As any other product, albeit a financial one, the larger the user-base, the more the economy at scale can help financial institutions offer lower costs while keeping a steady profit margin.

The market can be defined by the individuals that follow the criterion:

1. Being an individual in the US with an active SSN.
2. Having more than 18 years.
3. Achieve the minimum credit requirements in terms of score.

Besides reaching adulthood,being able to legally obtain a credit and having minimal financial health, the market is very wide; according to the Office of Homeland Security Statistics, there are 258.3 million U.S. residents over 18 years (Miller, 2024). And considering that the US population at large follows the trend observed in the dataset which returns that 71% of the population has a “Good” or “Standard” credit score, distributed as seen in the figure:

Where 71% of people have either Standard or Good Scores, this leaves a remaining almost 184 million U.S. adult residents capable of acquiring personal loans.

#### 3. Regulation

Personal loans in the U.S. are subject to both federal and state regulations aimed at protecting borrowers and ensuring transparency in lending practices.

At the federal level, key laws include the Truth in Lending Act (TILA), which mandates clear disclosure of loan terms such as interest rates and repayment schedules; the Equal Credit Opportunity Act (ECOA), which prohibits discrimination in lending; and the Fair Credit Reporting Act (FCRA), which regulates how lenders access and use consumer credit information. Oversight is led by the Consumer Financial Protection Bureau (CFPB), which enforces these protections and monitors lenders for unfair or deceptive practices (CFPB, n.d.).

States have additional authority, especially over interest rate caps, lender licensing, and fees. For example, among the 45 states (plus D.C.) that cap APRs on a $500 six-month loan, rates range from 16% to over 300%, with a median of 39.5%. For a $2,000 two-year loan, the caps are generally lower, ranging from 17% to 59%, with a median of 32.5% (*NCLC*, 2023). This variation highlights how a borrower’s location significantly affects their potential loan costs. Some states also regulate how much can be borrowed, repayment terms, and whether lenders must be state-licensed.

This figure illustrates these limits better.

In cases of default, lenders typically attempt collections and may sell the debt to agencies or pursue legal action. State laws govern actions such as wage garnishment or asset seizure, and bankruptcy protections under federal law may apply in severe cases (FTC, 2020).

#### 4. Market rates

Personal loan interest rates are slightly more favorable for borrowers this week, with the average rate dipping to 20.86%, down from 20.97% the previous week. Personal loans offer flexible financing options for everything from weddings and dental procedures to home improvements and tax bills—but their most common use is debt consolidation. In fact, some lenders like Happy Money and Reach Financial specifically cater to borrowers looking to consolidate debt. While interest rates can vary widely based on credit score, lender, and loan type, those with excellent credit have access to significantly lower APRs, starting as low as 5.99% (Acevedo, 2025).

| **Measurement** | **Rates up to Mar/31** |
| --- | --- |
| Average overall rate | 20.86% |
| Average low rate | 11.52% |
| Average high rate | 30.20% |
| Highest rate | 99.99% |
| Lowest rate | 5.99% |

## Rate determination

We need benchmarks to determine first, a minimum rate, at which there is no loss to lend money, and then maximize our profits considering a competitive rate.

### Minimum Rate

**Benchmarks to consider (minimum rate):**

1. Base Rate
2. Inflation (annualized)
3. Administrative cost (per loan successfully issued), usually called “Origination Fee”

While the Base Rate and Inflation are easily obtainable, this is not the case for an administrative cost, the following is an attempt at determining ours:

Components of the Administrative Cost, according to pwc:

1. Evaluating the prospective borrower’s financial condition (25%)

2. Evaluating and recording guarantees, collateral, and other security arrangements (30%)

3. Negotiating loan terms (15%)

4. Preparing and processing loan documents (10%)

5. Closing the transaction (5%)

A standard Origination Fee value is 1% of the amount lended.

The weights of each component were decided according to the resources usually assigned to the task.

We are able, through the use of models such as the one described in the first section of this document, to minimize and automate evaluation costs and processing of documents, cutting them to a fraction of the time (down to less than an hour, considering the process of extracting the data from the documentation of the costumer) a regular analysis for a personal loan would take, which is 2 to 5 business days on average (Axis Bank, 2024).

Adding now two considerations, we can now calculate our Origination Fee:

1. First, the first two evaluation processes can be cut to about 20% of the original resources spent, since there still are the costs related to paying analysts and servers to run the models, but reducing the costs of bankers and intermediaries.
2. As for the processing of documents, only a very small amount of the original resources is spent since this was also considered during evaluation, amounting to approximately 30% of the original resources needed.
3. Finally, let’s consider the standard 1% of the total amount lended to be the Origination Fee.

Acknowledging the previous points would give us the following process:

| **Component** | **Original weight** | **Optimized weight** |
| --- | --- | --- |
| Evaluation of Financial condition | 25% | 5.00% |
| Evaluation of guarantees, collaterals and security arrangements | 35% | 7.00% |
| Negotiation of terms | 15% | 15.00% |
| Preparation and processing of documents | 15% | 4.50% |
| Closing (Issuance of loan) | 10% | 10.00% |
| **Total** | 100% | 41.50% |
| **As 1%** | **0.01** | **0.00415** |

So now, we can calculate the minimum rate with the following function

where:

= rate of credit

= Base rate or risk-free rate

= Consumer Price Index annualized

= Value of our administrative costs (Origination fee)

For example, calculating this value as of April 6th, 2025:

We know the value of the Base rate (4.33%) and of the annualized consumer price index (2.81%) thanks to FRED (FRED St. Louis, 2025).

This gives us a minimum rate of **7.555%** as of April 6th, 2025.

### Final Rate

**Benchmarks to consider:**

1. Credit Score
2. Liquidity of borrower, relative (as a percentage of overall income)
3. Liquidity of borrower, absolute (to cover payments)
4. Profit Margin
5. Maximum allowed APR
6. Maximum allowed rate

#### Credit Score

Our internal scoring model rates users from 4 to 11, in the following categories:

| **Credit Category** | **Rating** |
| --- | --- |
| Poor | A rating less or equal than 6 |
| Standard | A rating greater than 6 and less or equal than 8 |
| Good | A rating greater than 8 |

#### Liquidity of borrower

**Relative:** The percentage of overall income that the borrower can spend (since it is not compromised in other credits)

**Absolute:** So that the borrower’s absolute free income is at least three times greater than the total monthly loan payments.

#### Profit Margin

Known as NIM (Net Interest Margin).

This is a simple rate, the average in the industry is around 3% (Bloomenthal, 2024).

#### Maximum rates

This value is dependent on the state, since the lowest value in general is for Arizona, at 16%, that will be the selected maximum rate for

### 

### Rate assignment

The final is determined by the previously described factors, in the following manner:

#### Minimum rate

Determined by:

Currently **7.555%**

#### NIM - Profit Margin

This was previously determined to be 3% of the total value of the loan

#### Score rate

Depending the score, there is an addition to the final rate:

| **Credit Category** | **Additional rate** |
| --- | --- |
| Poor | Loan not approved |
| Standard | 1.435% |
| Good | 0.08% |

#### Liquidity rate

Relative liquidity = annual expendable income / total annual income

| **Relative liquidity** | **Additional rate** |
| --- | --- |
| Less than 20% | 3.2% |
| Equal or greater than 20% and less than 50% | 2.1% |
| Equal or greater than 50% | 1% |

Absolute liquidity = annual income - annual expendable income

Yearly value of loan = total loan amount / years

| **Absolute liquidity** | **Additional rate** |
| --- | --- |
| If ~33% of absolute liquidity is less than the yearly value of the loan | Loan not approved |
| If more than ~33% of absolute liquidity is less than the yearly value of the loan and if ~67% of absolute liquidity is less than yearly value of loan | 0.8% |
| If ~67% of absolute liquidity is greater than the yearly value of the loan | No additional rate |

The greater yearly absolute liquidity a user has, the more likely they are to keep up with their payments.

### **Final Takeaway**

Setting personal loan interest rates is a strategic balancing act that blends economic fundamentals, borrower-specific risk, regulatory landscapes, and technological efficiencies. Lenders must account for the base rate, inflation, creditworthiness, liquidity, administrative costs, and profit margins—all while remaining competitive and compliant. Advances in automation and data modeling now allow for significant cost reduction, particularly in underwriting and documentation, making it possible to maintain profitability with lower origination fees. With over 180 million U.S. adults potentially eligible for personal loans and market rates ranging from 5.99% to nearly 100%, there’s both opportunity and complexity in designing lending models that are fair, scalable, and resilient.

**References**

Acevedo, S. (2025, April 1). *Average Personal Loan Interest Rates in 2025*. Business Insider. Retrieved April 6, 2025, from https://www.businessinsider.com/personal-finance/personal-loans/average-personal-loan-interest-rates

Axis Bank. (2024, March 4). *Personal Loan Approval Process & Time: 4 Tips to Improve*. Axis Bank. Retrieved April 6, 2025, from https://www.axisbank.com/progress-with-us-articles/money-matters/borrow/personal-loan-approval-process-and-time

Bloomenthal, A. (2024, Jun). *What Is Net Interest Margin? Overview, Formula, and Example*. Investopedia. Retrieved April 6, 2025, from https://www.investopedia.com/terms/n/netinterestmargin.asp

CFPB. (n.d.). *Compliance resources*. Consumer Financial Protection Bureau. https://www.consumerfinance.gov/compliance/compliance-resources/

Federal Reserve Board. (n.d.). *H.15 - Selected Interest Rates (Daily) - April 04, 2025*. Federal Reserve Board. Retrieved April 6, 2025, from https://www.federalreserve.gov/releases/h15/

FRED St. Louis. (2025, Apr). *CPIAUCSL*. Consumer Price Index for All Urban Consumers: All Items in U.S. City Average. https://fred.stlouisfed.org/series/CPIAUCSL

FTC. (2020). *Credit and Loans*. Federal Trade Comission. https://www.ftc.gov/business-guidance/credit-finance/credit-loans

Miller, S. (2024, September). *Estimates of the Lawful Permanent Resident Population in the United States and the Subpopulation Eligible to Naturalize: 2024 and Revised 2023*. Population Estimates. https://ohss.dhs.gov/sites/default/files/2024-11/2024\_1108\_ohss\_lawful\_permenent\_resident\_population\_estimate\_2024\_and\_revised\_2023.pdf

Millerbernd, A. (2025, March 12). *Average Personal Loan Interest Rates for March 2025*. NerdWallet. Retrieved April 6, 2025, from https://www.nerdwallet.com/article/loans/personal-loans/average-personal-loan-rates?utm\_source=chatgpt.com

PWC. (2025, Mar 12). 4.4 Loan origination fees and costs. *us Loans & investments guide*, (2025).

*State Rate Caps for $500 and $2,000 Loans - NCLC*. (2023, November 15). National Consumer Law Center. https://www.nclc.org/resources/state-rate-caps-for-500-2000-loans/

US Inflation Calculator. (2025, Mar 01). *Current US Inflation Rates: 2000-2025*. Inflation Calculator. Retrieved April 6, 2025, from https://www.usinflationcalculator.com/inflation/current-inflation-rates/

U.S. News. (2025, fEB 28). *Best Personal Loan Rates of April 2025*. Best Personal Loans. https://money.usnews.com/loans/personal-loans/best-personal-loans-versionb