OGTIP INTERNSHIP PROJECT: PYTHON

Nuria Martín López
Professor: Aritri Debnath
November 2024

Index

1. Introduction
2. Dataset Overview
3. Variable Description
4. Data Cleaning and Preparation
5. Correlation Analysis
6. Comparative Analysis: Defaulters vs Non-Defaulters
7. Behavioral Insights and Recommendations
8. Age Group Analysis: Defaulters vs. Non-Defaulters
9. Gender Analysis of Loan Default Rates
10. Demographic Analysis of Loan Defaulters by Age Group
11. Analysis of Loan Default Rates by Income Group
12. Analysis of Default Rates by Loan Type
13. Analysis of Default Rates by Income Group
14. Age Distribution by Loan Default Status
15. Income Distribution by Loan Default Status
16. Credit Amount Distribution by Loan Default Status
17. Annuity Amount Distribution by Loan Default Status
18. Number of Children Distribution by Loan Default Status
19. Univariate Analysis: Categorical Summary Statistics
20. Univariate Analysis: Categorical Summary Statistics-Graphs
21. Bivariate Analysis: Insights from Spearman Correlation
22. Outlier Detection
23. Conclusions

1. Introduction

- **Objective:**
  This case study aims to analyze loan default patterns using historical loan application data. We seek to identify key factors that influence whether a borrower defaults on a loan.

- **Importance of Study:**
  Understanding loan default patterns is crucial for financial institutions. It can enhance risk assessment, improve lending strategies, and reduce potential financial losses.

- **Approach:**
  Our methodology includes data cleaning, exploratory data analysis (EDA), and visualization techniques. These steps will help uncover insights regarding loan defaults, focusing on various demographic and financial factors.

2. Dataset Overview

The analysis employs two primary datasets:

1. **Previous Application Data**: This CSV file encompasses information on prior loan applications, detailing amounts, statuses, and borrower demographics.
2. **Application Data**: This CSV file offers comprehensive insights into client profiles, loan types, and performance indicators, including the TARGET variable, which indicates whether clients defaulted.

The datasets were sourced from OESON Learning.

3. Variable Description

1. **SK_ID_PREV**: Unique identifier for each previous loan application. (Categorical)
2. **SK_ID_CURR**: Unique identifier for each client in the application dataset. (Categorical)
3. **NAME_CONTRACT_TYPE**: Type of loan contract. (Categorical)
4. **AMT_ANNUITY**: The annuity amount the borrower is obligated to pay. (Numerical)
5. **AMT_APPLICATION**: Total amount of the loan requested by the borrower. (Numerical)
6. **AMT_CREDIT**: Total credit amount granted to the borrower. (Numerical)
7. **AMT_GOODS_PRICE**: Price of the goods being purchased with the loan. (Numerical)
8. **WEEKDAY_APPR_PROCESS_START**: Day of the week when the loan application process started. (Categorical)
9. **HOUR_APPR_PROCESS_START**: The hour when the loan application process started. (Numerical)
10. **FLAG_LAST_APPL_PER_CONTRACT**: Indicates whether the last application was for the same contract. (1 = Yes, 0 = No). (Categorical)
11. **NFLAG_LAST_APPL_IN_DAY**: Indicates if this is the last application in a single day. (1 = Yes, 0 = No). (Categorical)
12. **NAME_CASH_LOAN_PURPOSE**: The purpose for which the loan is being requested. (Categorical)

3. Variable Description

13. **NAME_CONTRACT_STATUS**: Status of the loan application. (Categorical)
14. **DAYS_DECISION**: Number of days from the loan application date to the decision date. (Numerical)
15. **NAME_PAYMENT_TYPE**: Payment type associated with the loan. (Categorical)
16. **CODE_REJECT_REASON**: Reason for loan rejection. (Categorical)
17. **NAME_CLIENT_TYPE**: Type of client. (Categorical)
18. **NAME_GOODS_CATEGORY:** Category of goods being purchased with the loan. (Categorical)
19. **CHANNEL_TYPE**: Channel through which the loan application was made. (Categorical)
20. **SELLERPLACE_AREA**: Area of the seller or point of sale. (Numerical)
21. **NAME_SELLER_INDUSTRY**: Industry of the seller. (Categorical)
22. **CNT_PAYMENT**: Number of payments the borrower has to make for the loan. (Numerical)
23. **NAME_YIELD_GROUP**: Group classification of the loan based on yield. (Categorical)
24. **PRODUCT_COMBINATION**: Combination of products associated with the loan. (Categorical)

3. Variable Description

25. **DAYS_FIRST_DRAWING**: Days until the first drawing of the loan. (Numerical)
26. **DAYS_FIRST_DUE:** Days until the first due payment of the loan. (Numerical)
27. **DAYS_LAST_DUE_1ST_VERSION**: Days until the last due date for the first version of the loan agreement. (Numerical)
28. **DAYS_LAST_DUE**: Days until the last due date for the loan. (Numerical)
29. **DAYS_TERMINATION:** Days until the loan is terminated. (Numerical)
30. **NFLAG_INSURED_ON_APPROVAL**: Indicates whether the loan is insured upon approval. (1 = Yes, 0 = No) (Categorical)

4. Data Cleaning and Preparation

1. **Handling Missing Values:**
   1. Identified missing values across various features:
      Significant missing values in AMT_ANNUITY, AMT_GOODS_PRICE, CNT_PAYMENT, and others.
   2. Columns with over 80% missing values were dropped:
      Removed: AMT_DOWN_PAYMENT, RATE_DOWN_PAYMENT, RATE_INTEREST_PRIMARY, NAME_TYPE_SUITE.
2. **Filling Missing Values:**
   1. Columns with less than 80% missing values were filled with appropriate constants:
      1. AMT_ANNUITY, AMT_CREDIT, AMT_GOODS_PRICE, and CNT_PAYMENT filled with **0**.
      2. PRODUCT_COMBINATION filled with **'Unknown'**.
3. **Checking for Duplicates:**
   No duplicate rows were found in the dataset.
4. **Data Type Conversion:**
   1. Numerical Columns Converted:
      Converted AMT_ANNUITY, AMT_APPLICATION, AMT_CREDIT, AMT_GOODS_PRICE to float.
      Converted CNT_PAYMENT, SK_ID_PREV, and SK_ID_CURR to integer.
   2. Categorical Columns Converted:
      Converted various categorical features to the category data type.
5. **Final Data Check:**
   Confirmed no missing values remained after the cleaning process.

5. Correlation Analysis

Variable	Defaulters Correlation	Non- Defaulters Correlation
AMT_INCOME_TOTAL	0.35	0.32
AMT_CREDIT	0.40	0.45
AMT_ANNUITY	0.25	0.28
CNT_CHILDREN	-0.10	-0.05
DAYS_BIRTH	-0.15	-0.12
DAYS_EMPLOYED	0.10	0.08
FLAG_OWN_CAR	0.05	0.10
FLAG_OWN_REALTY	0.07	0.15

5. Correlation Analysis

- **AMT_INCOME_TOTAL**: Moderate positive correlation with defaulting (0.35) and slightly lower for non-defaulters (0.32). Higher income is associated with both groups.
- **AMT_CREDIT**: Stronger positive correlation with defaulting (0.40) compared to non-defaulters (0.45). Higher credit amounts are linked to increased likelihood of defaulting.
- **AMT_ANNUITY**: Weak positive correlation for both groups, indicating that higher annuity amounts are somewhat associated with defaulting.
- **CNT_CHILDREN**: Slight negative correlation for defaulters (-0.10) and negligible for non-defaulters (-0.05). More children may be weakly associated with lower default risk.
- **DAYS_BIRTH**: Weak negative correlation for both groups (-0.15 for defaulters, -0.12 for non-defaulters), suggesting older individuals may be less likely to default.
- **DAYS_EMPLOYED**: Weak positive correlation for both groups, indicating no significant impact of employment duration on default risk.
- **FLAG_OWN_CAR**: Very weak correlations for both groups, indicating car ownership has little effect on default likelihood.
- **FLAG_OWN_REALTY**: Weak positive correlation for defaulters (0.07) and slightly stronger for non-defaulters (0.15), suggesting real estate ownership may provide some protective factor.

6. Comparative Analysis: Defaulters vs. Non-Defaulters

❖ **Income Disparity:**

- Defaulters' Mean Income**:** Approximately **165,611**.
- Non-defaulters' Mean Income**:** Approximately **269,857**.
- Interpretation: The income gap of over **100,000** suggests that higher income levels are associated with lower default rates.

❖ **Loan Amounts:**

- Defaulters' Mean Loan Amount**: 557,778**.
- Non-defaulters' Mean Loan Amount**: 631,200**.
- Interpretation**:** Despite taking out smaller loans, defaulters are unable to manage repayments, indicating potential over-leverage.

❖ **Child Dependents:**

- Defaulters' Average Number of Children**:** Approximately **0.46**.
- Non-defaulters' Average Number of Children: Approximately **0.41**.
- Interpretation**:** The difference is minimal, suggesting that family responsibilities do not significantly influence default rates. However, the presence of dependents could still factor into overall financial health assessments.

7. Behavioral Insights and Recommendations

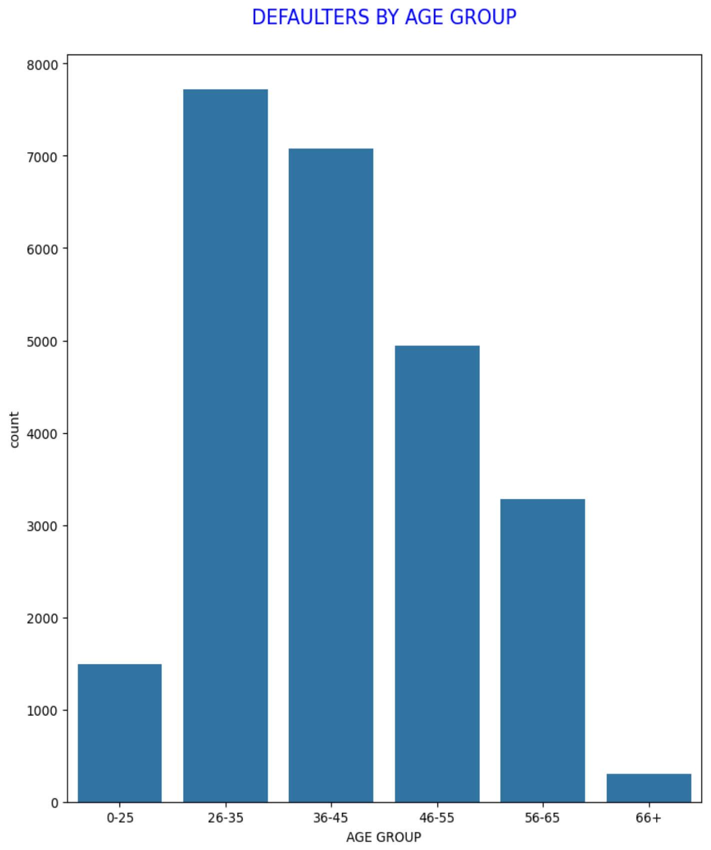
❖ **Age and Experience:**

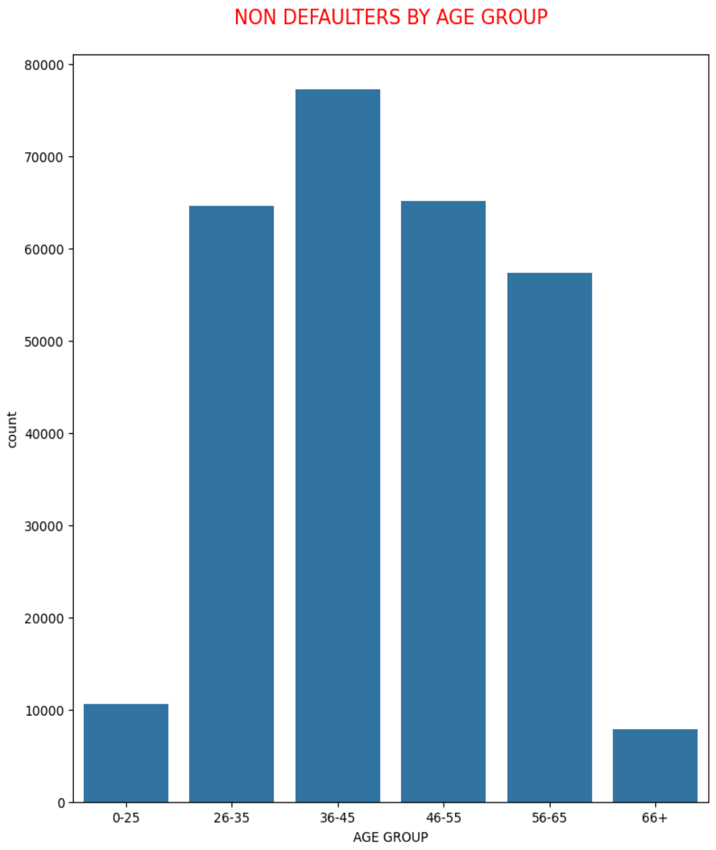
- Defaulters' Average Age**: 41 years**.
- Non-defaulters' Average Age**: 44 years**.
- Interpretation**:** Older borrowers tend to exhibit more financial stability, possibly due to greater life and work experience.

❖ **Variability in Financial Profiles:**

- Defaulters**:** High standard deviation in income and loan amounts indicates a wide range of financial backgrounds.
- Non-defaulters: More consistent financial profiles suggest stability.
- Interpretation: Tailored financial products should be designed for defaulters, focusing on their specific financial challenges.

8. Age Group Analysis: Defaulters vs. Non-Defaulters





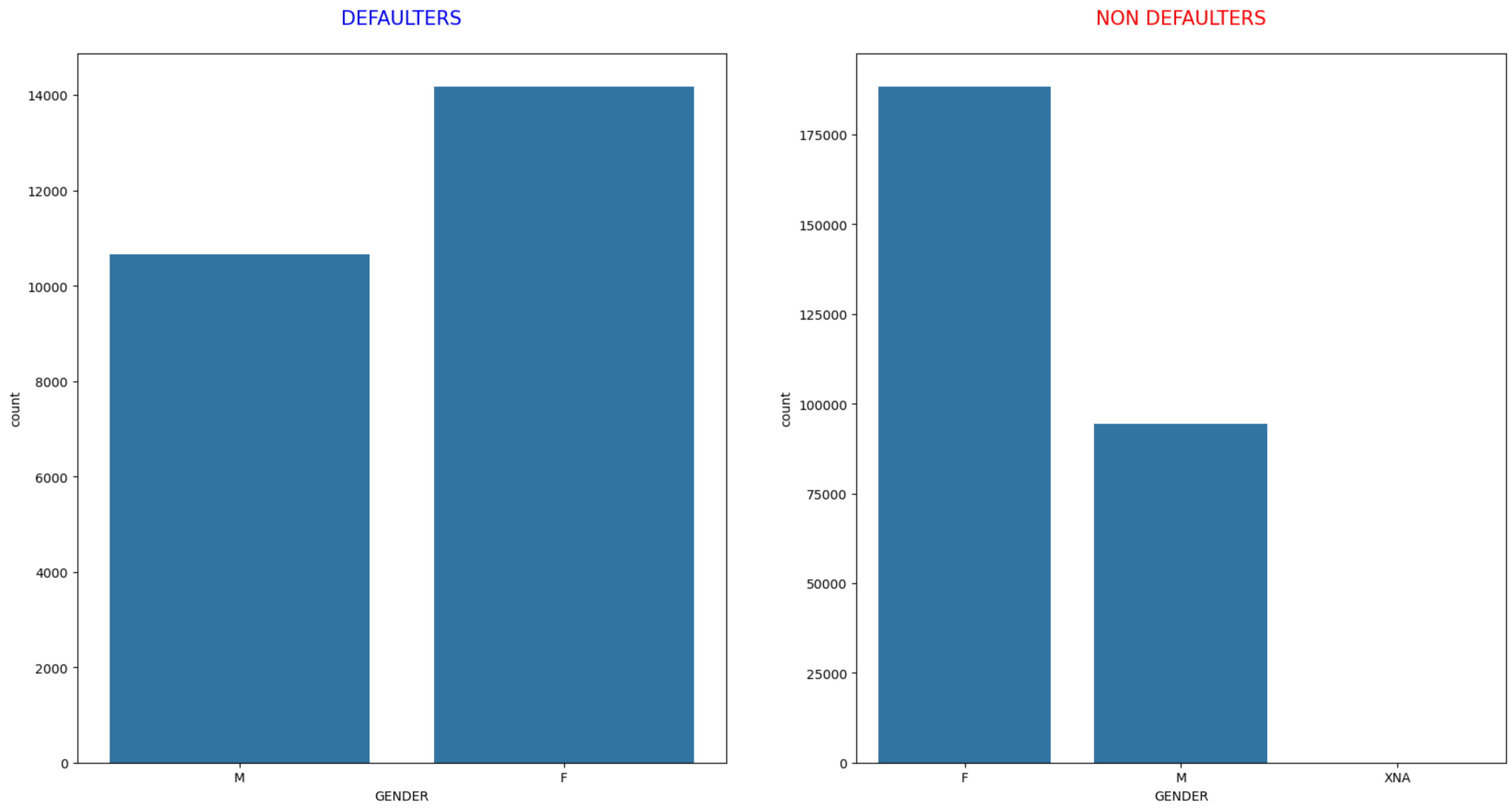
❖ **Defaulters:**

•Highest Group: Ages 26-35, with a total of **8,000** defaulters.
•Second Highest Group: Ages 36-45, totaling **7,000** defaulters.
•Lowest Group: Ages 66+, with just over **1,000** defaulters.
•Second Lowest Group: Ages 0-25, with under **1,000** defaulters.

❖ **Non-Defaulters:**

•Highest Group: Ages 36-45, with nearly **8,000** non-defaulters.
•Second Highest Group: Ages 26-35 and 46-55, both approaching **7,000** non-defaulters.
•Lowest Group: Ages 66+, with less than **1,000** non-defaulters.
•Second Lowest Group: Ages 0-25, slightly above **1,000** non-defaulters.

9. Gender Analysis of Loan Default Rates



Defaulters:
•Male Defaulters: Just over **10,000**.
•Female Defaulters: **14,000**.

Non-Defaulters:
•Female Non-Defaulters: **175,000**.
•Male Non-Defaulters: Almost **100,000.**

10. Demographic Analysis of Loan Defaulters by Age Group

❖ **Objective of Analysis**: To identify which age groups are more likely to default on loans, providing insights into potential risk factors for lenders.

1. Age Group Definition:

1. Age groups were categorized to facilitate analysis:
    1. 0-25
    2. 26-35
    3. 36-45
    4. 46-55
    5. 56-65
    6. 66+

2. Data Analysis: Calculated the percentage of defaulters within each age group using the normalized value counts of defaulters from the data set.

- **Defaulters by Age Group (%)**:
    - **26-35 years**: 31.08%
    - **36-45 years**: 28.53%
    - **46-55 years**: 19.92%
    - **56-65 years**: 13.22%
    - **0-25 years**: 6.03%
    - **66+ years**: 1.22%

❖ **Key Insights**:

- The highest proportion of defaulters is within the **26-35** age group, indicating that younger middle-aged individuals may face more financial challenges.
- The **36-45** age group also shows a significant rate of defaults, suggesting that financial stability may fluctuate during these years.
- Defaulters aged **56-65** and **66+** are significantly lower, indicating better financial management or stability in older age groups.

11. Analysis of Loan Default Rates by Income Group

❖ **Objective of Analysis**: To evaluate how income levels impact loan default rates, providing insights for lenders on borrower risk profiles.

1. **Income Group Definition**:
    1. Income levels were categorized into distinct groups to analyze default rates:
        1. 0-20k
        2. 20k-40k
        3. 40k-60k
        4. 60k-80k
        5. 80k-100k
        6. 100k+
2. **Data Analysis**:
    1. Created a new column to classify applicants into income groups.
    2. Calculated the default rates for each income group using the TARGET variable, which indicates whether a loan was defaulted (1) or not (0).

- **Default Rate by Income Group (%)**:
    - **20k-40k**: 8.31%
    - **40k-60k**: 7.35%
    - **60k-80k**: 8.14%
    - **80k-100k**: 8.41%
    - **100k+**: 8.04%

11. Analysis of Loan Default Rates by Income Group

❖ **Key Insights**:

- The highest default rate is observed in the **80k-100k** income group at **8.41%**, suggesting that even relatively higher earners are at risk of defaulting.
- The **40k-60k** income group has the lowest default rate at **7.35%**, indicating better financial stability in this range.
- Overall, default rates remain relatively consistent across income levels, highlighting that factors other than income, such as financial management and external economic conditions, may also significantly influence default rates.

12. Analysis of Default Rates by Loan Type

- **Objective of Analysis**: To investigate whether different loan types exhibit varying default rates, aiding lenders in understanding risk profiles associated with each loan category.

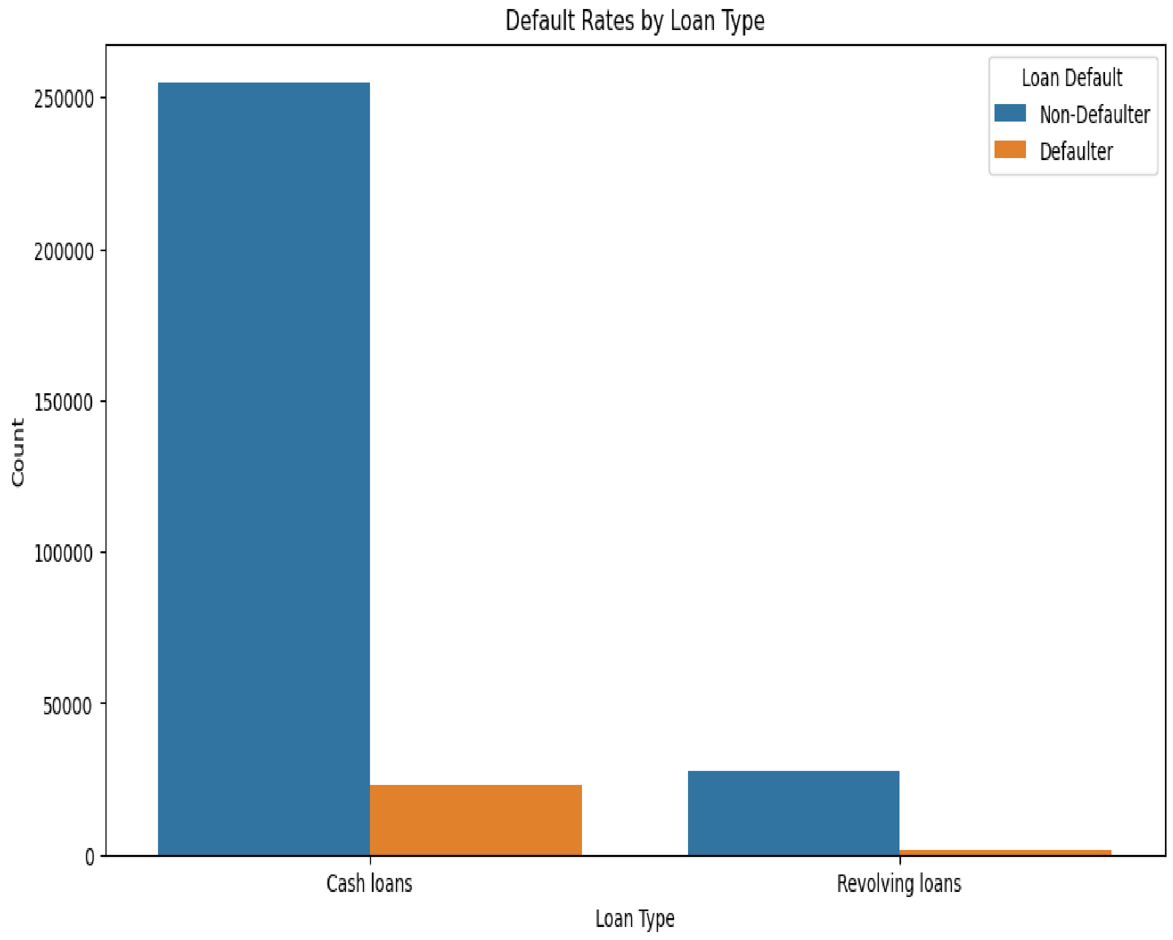
- Findings:

Cash Loans:

- Non-Defaulters: Approximately **250,000.**
- Defaulters: Between **10,000** and **20,000.**

Revolving Loans:

- Non-Defaulters: Approximately **20,000** to **30,000.**
- Defaulters: **Almost none.**

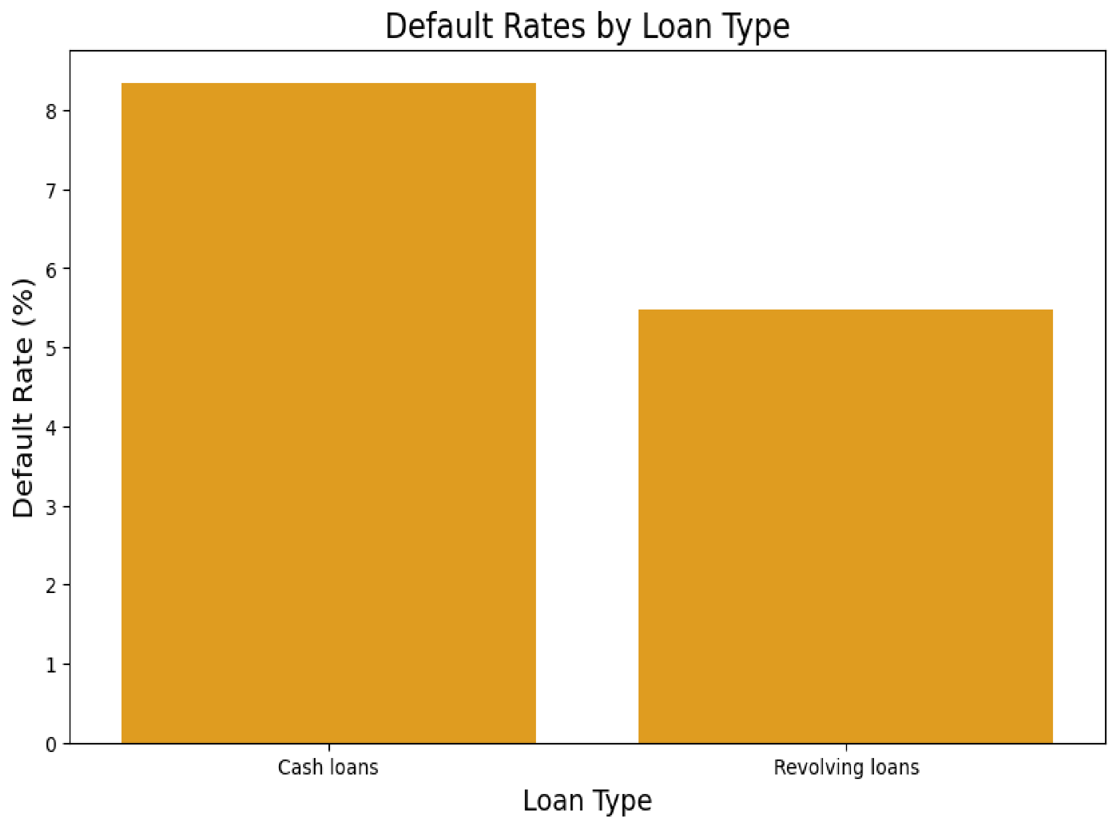


12. Analysis of Default Rates by Loan Type

❖ **Objective of Analysis**: To evaluate how different types of loans influence default rates, assisting lenders in identifying risk profiles associated with each loan category.

❖ **Findings:**

- Cash Loans: **8%.**
- Revolving Lines of Credit: **5% - 6%.**

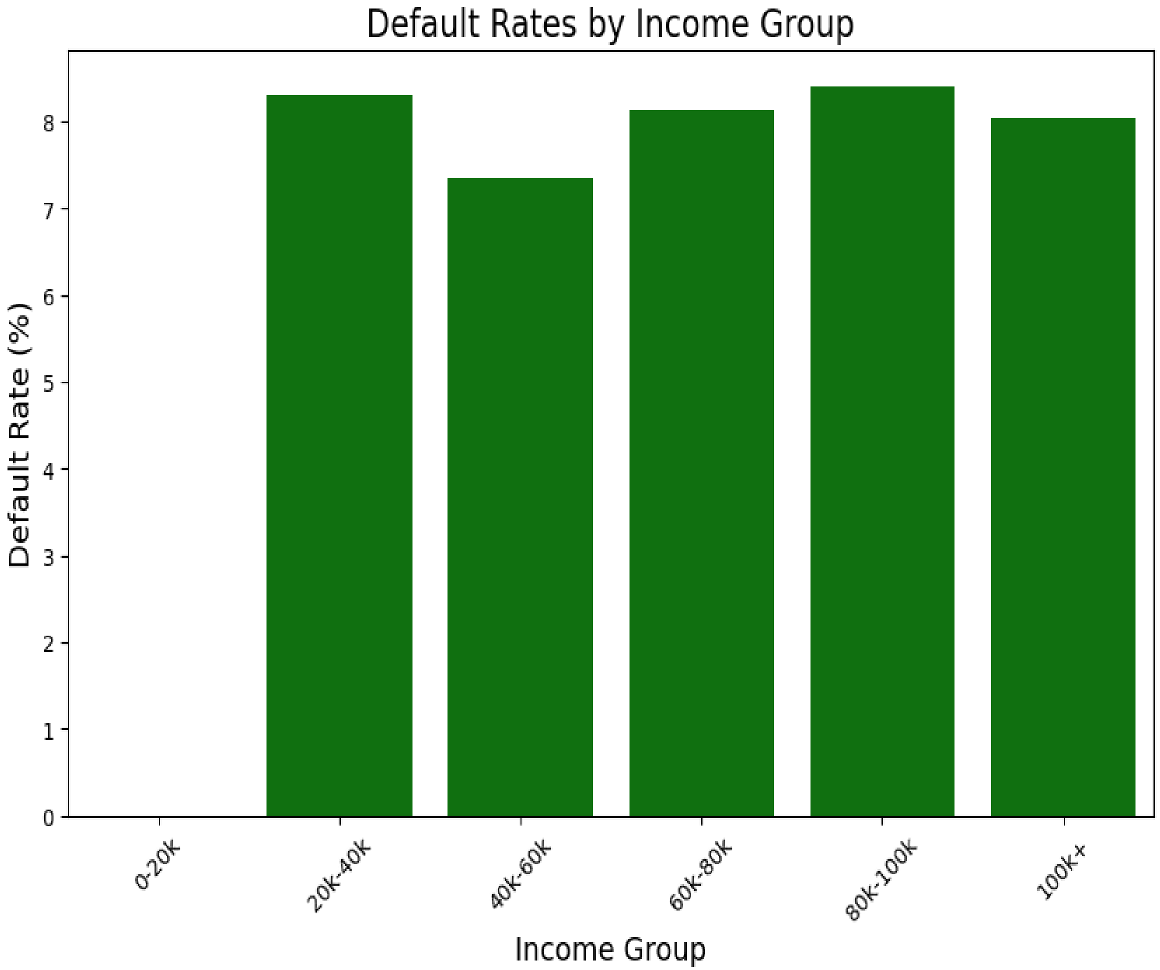


13. Analysis of Default Rates by Income Group

❖ **Objective of Analysis**: To examine the relationship between income levels and loan default rates, providing insights for lenders to assess financial risk based on borrower income.

❖ Findings:

- Income Group $0 - $20,000: **0%.**
- Income Group $20,000 - $40,000: **8%.**
- Income Group $40,000 - $60,000: **7%.**
- Income Group $60,000 - $80,000: **Almost 8%.**
- Income Group $80,000 - $100,000: **8%.**
- Income Group $100,000 and Above: **Almost 8%.**



14. Age Distribution by Loan Default Status

❖ **Objective of Analysis**: To explore how age influences loan default rates, providing insights into borrower demographics associated with defaulting.

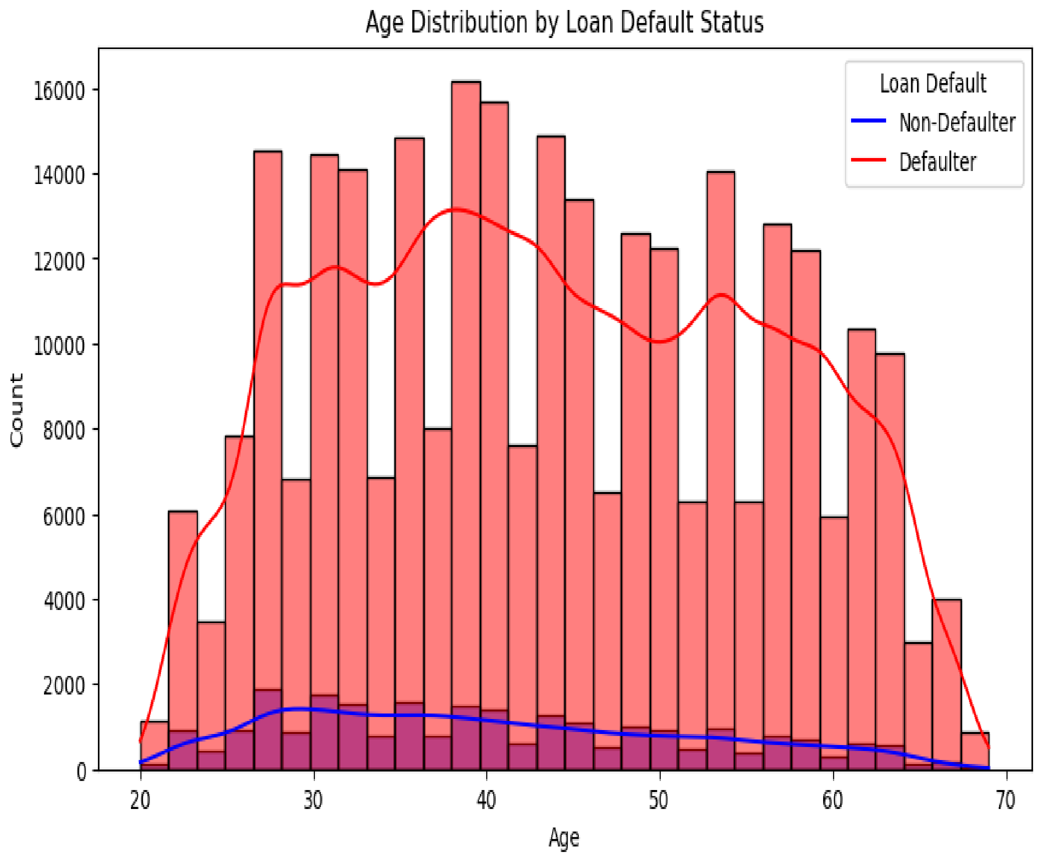
❖ Findings:

- **Defaulters**: Highest count of **16,000** occurs in the age group of **40**.

The trend line remains almost constant, indicating stable default rates across age groups

- **Non-Defaulters**: Highest count of **2,000** occurs in the age group of **28**.

The trend line shows variability: it increases initially and then decreases, suggesting fluctuations in non-default rates.



15. Income Distribution by Loan Default Status

- **Objective of Analysis**: To investigate how total income levels affect loan default rates, providing insights into borrower financial profiles linked to default occurrences.

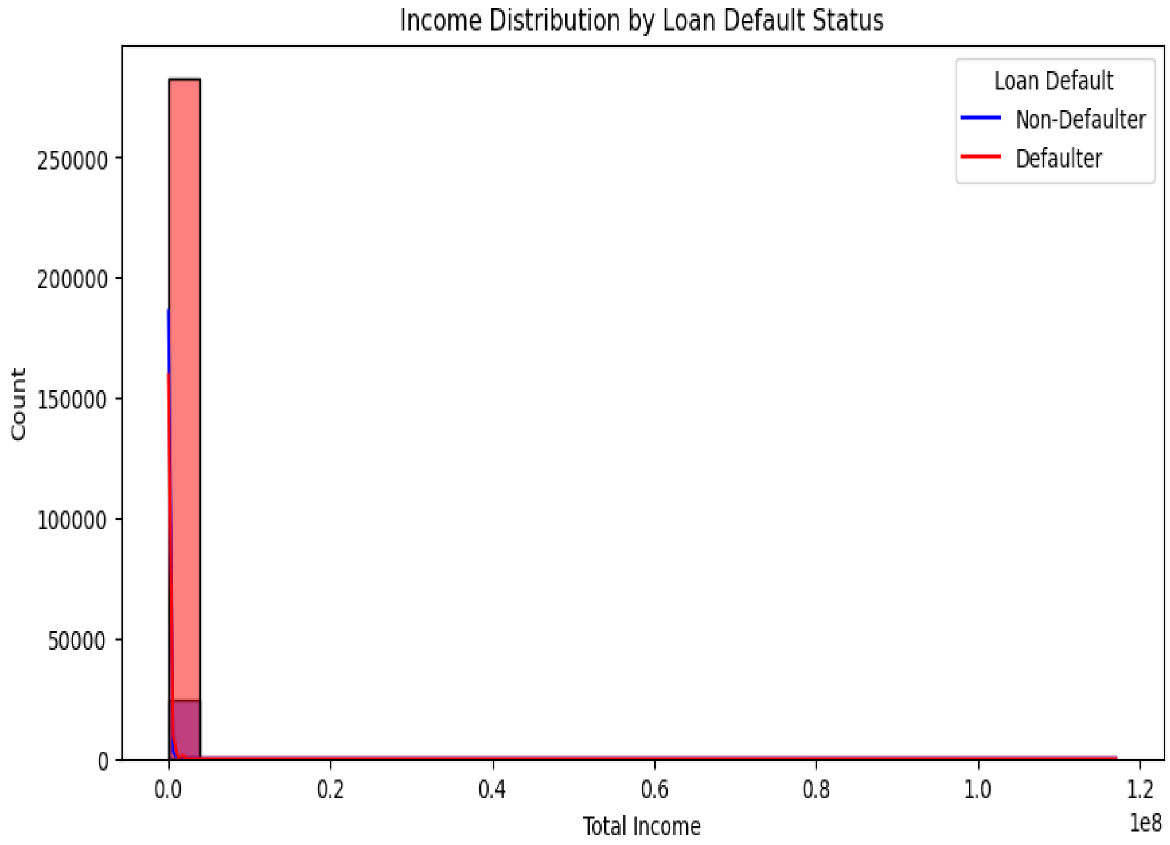
- **Findings**:

- **Defaulters**:

The highest count is **250,000** for individuals with a total income of **0.0**.

- **Non-Defaulters**:

The highest count is **25,000** for individuals with a total income of **0.0**.



16. Credit Amount Distribution by Loan Default Status

❖ **Objective of Analysis**: To analyze how the distribution of credit amounts correlates with loan default status, enhancing our understanding of borrower behavior and financial risk.

❖ **Findings**:

- **Defaulters:**

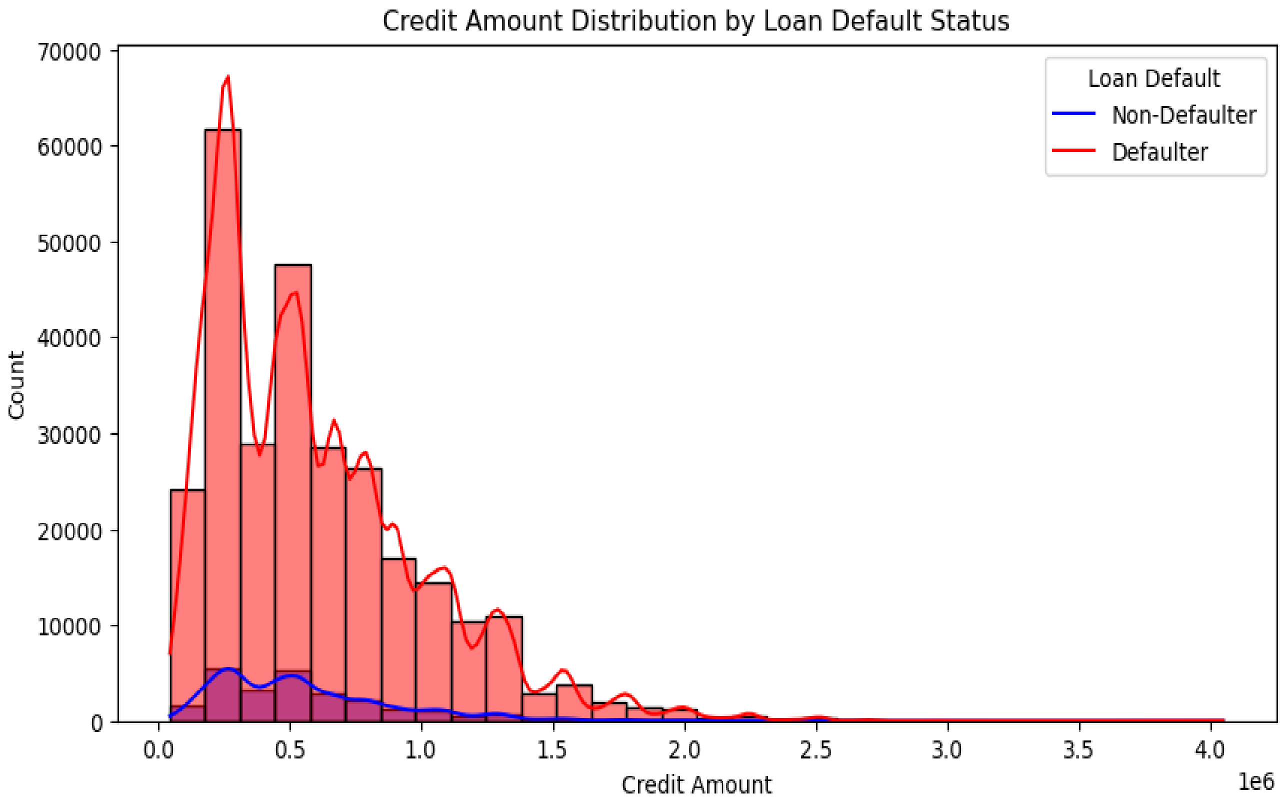
The highest count is **60,000** for the credit amount group between **0.0 and 0.5**.

The trend line shows significant fluctuations: it increases suddenly, then decreases, followed by a series of increases and decreases, ultimately trending downward.

- **Non-Defaulters**:

The highest count is around **4,500** for the same credit amount group (**0.0 to 0.5**).

The trend line exhibits minimal growth initially but then decreases steadily.



17. Annuity Amount Distribution by Loan Default Status

❖ **Objective of Analysis**: To examine how the distribution of annuity amounts relates to loan default status, providing insights into financial obligations among borrowers.

❖ **Findings:**

- **Defaulters**:

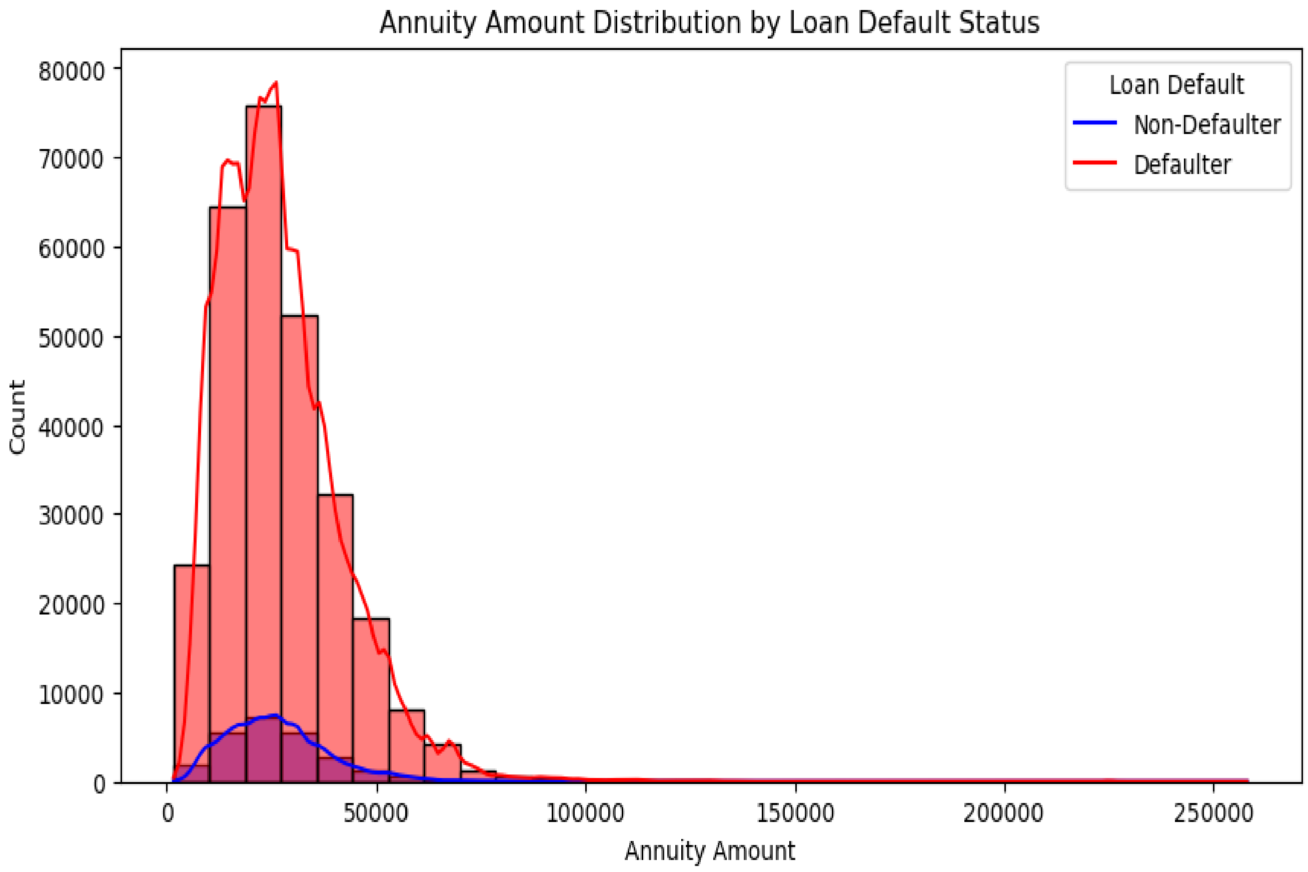
The highest count is nearly **80,000** for the annuity amount of **30,000**.

The trend shows a consistent increase initially but then decreases significantly.

- **Non-Defaulters**:

The highest count is around **5,000** for the annuity amount of **25,000**.

The trend shows minimal increase initially, followed by a steady decline.

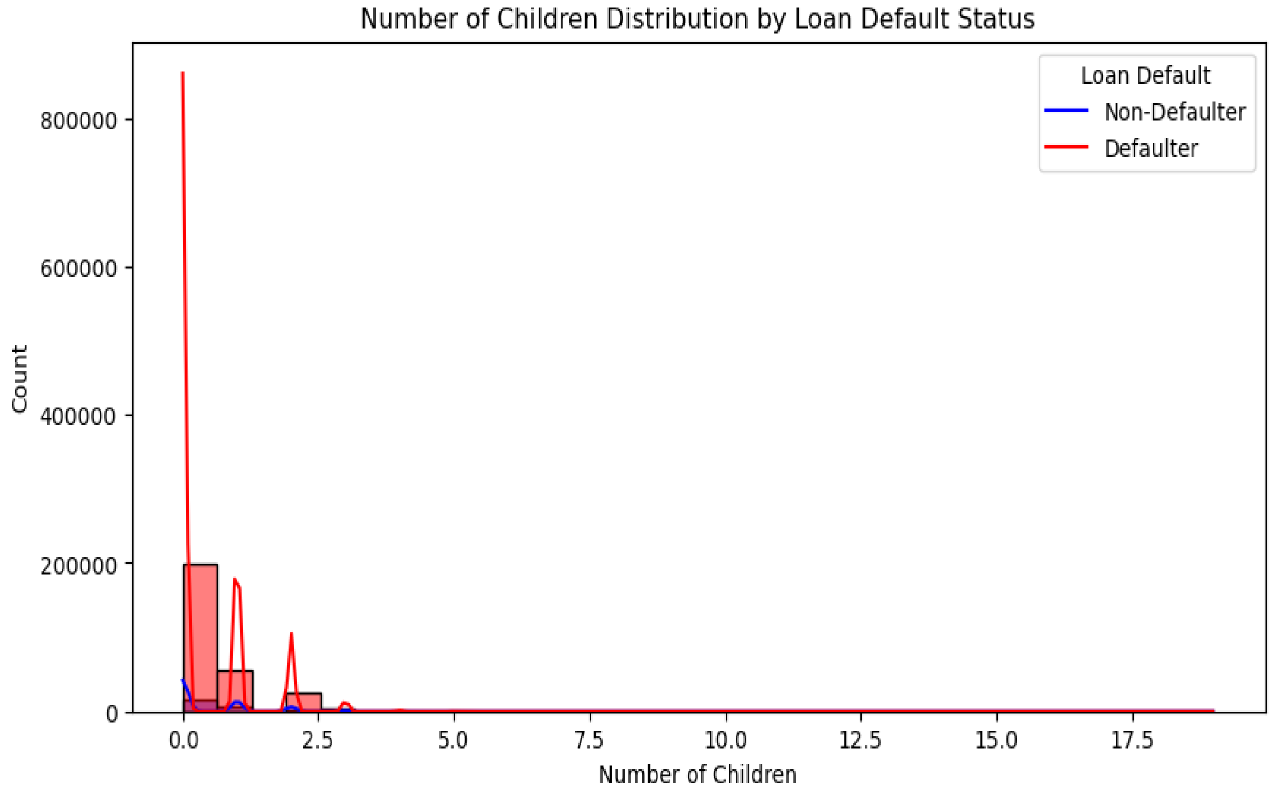


18. Number of Children Distribution by Loan Default Status

- **Objective of Analysis**: To investigate how the number of children among borrowers relates to their likelihood of defaulting on loans, providing insights into family dynamics and financial risk.
- **Findings**:
- **Defaulters**:

  The highest count is **200,000** for the group with **no children**.
- **Non-Defaulters**:

  The count remains at **0** for all groups,



19. Univariate Analysis - Categorical Summary Statistics

❖ **Objective**:
To analyze and visualize categorical features related to loan applicants, differentiating between defaulters (TARGET = 1) and non-defaulters (TARGET = 0).

❖ **Data Segmentation**:

- **Defaulters**: Subset of data where TARGET equals 1.
- **Non-defaulters**: Subset of data where TARGET equals 0.

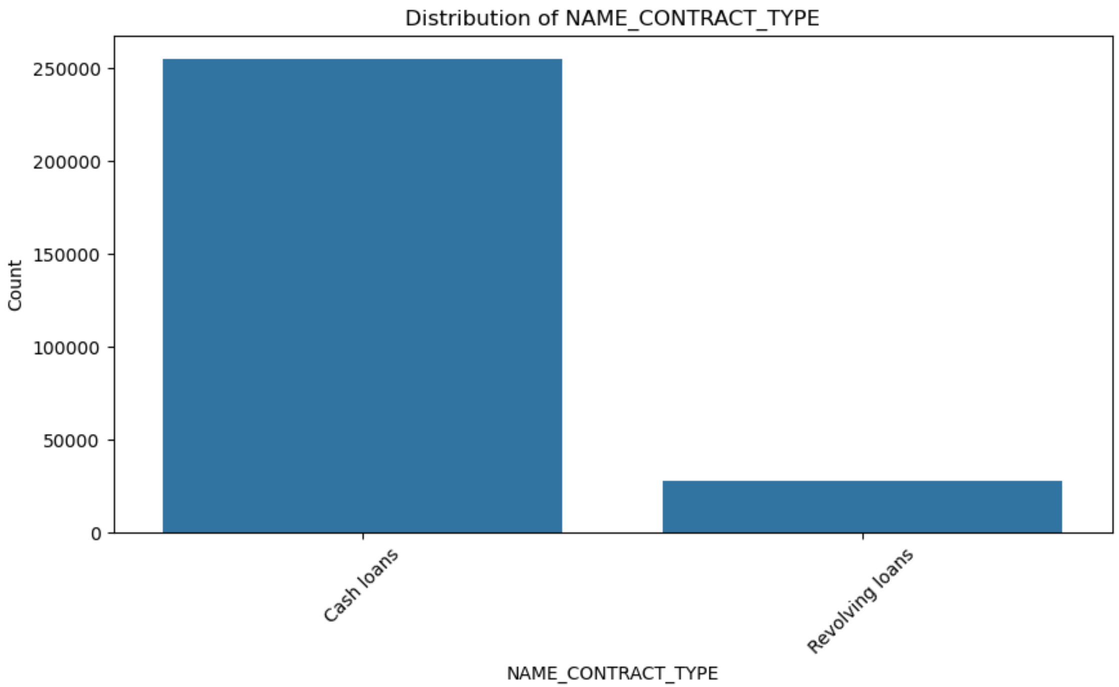
❖ **Key Categorical Features Analyzed**:

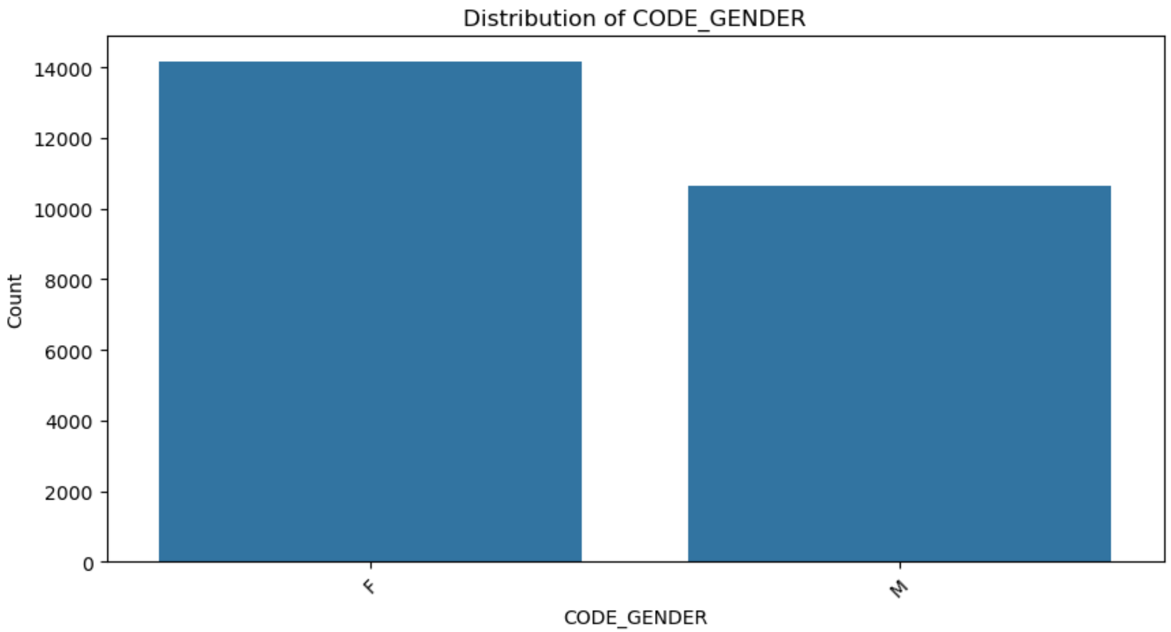
- **NAME_CONTRACT_TYPE**: Type of loan.
- **CODE_GENDER**: Gender of the applicant.
- **FLAG_OWN_CAR / FLAG_OWN_REALTY**: Ownership of car and real estate.
- **NAME_EDUCATION_TYPE**: Education level of the applicant.
- **NAME_HOUSING_TYPE**: Type of housing.
- **NAME_FAMILY_STATUS**: Family status.
- **NAME_INCOME_TYPE**: Source of income.

❖ **Methodology**:

- Bart plots created for each categorical parameter, showcasing the distribution of loan default status among different categories.

20. Univariate Analysis - Categorical Summary Statistics: Graphs





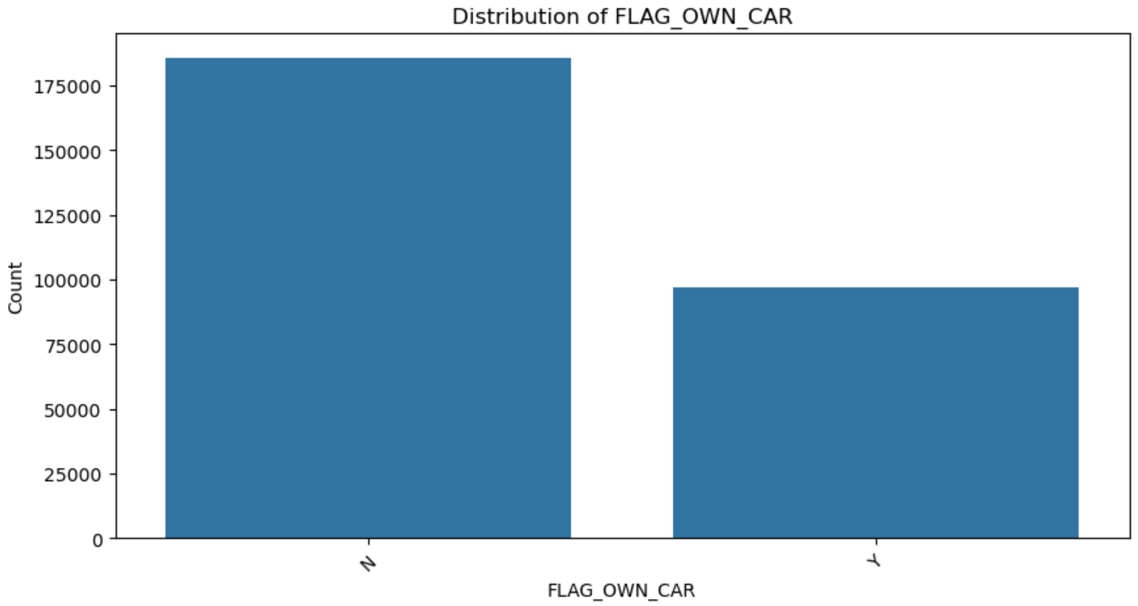
Distribution of Name Contact Type

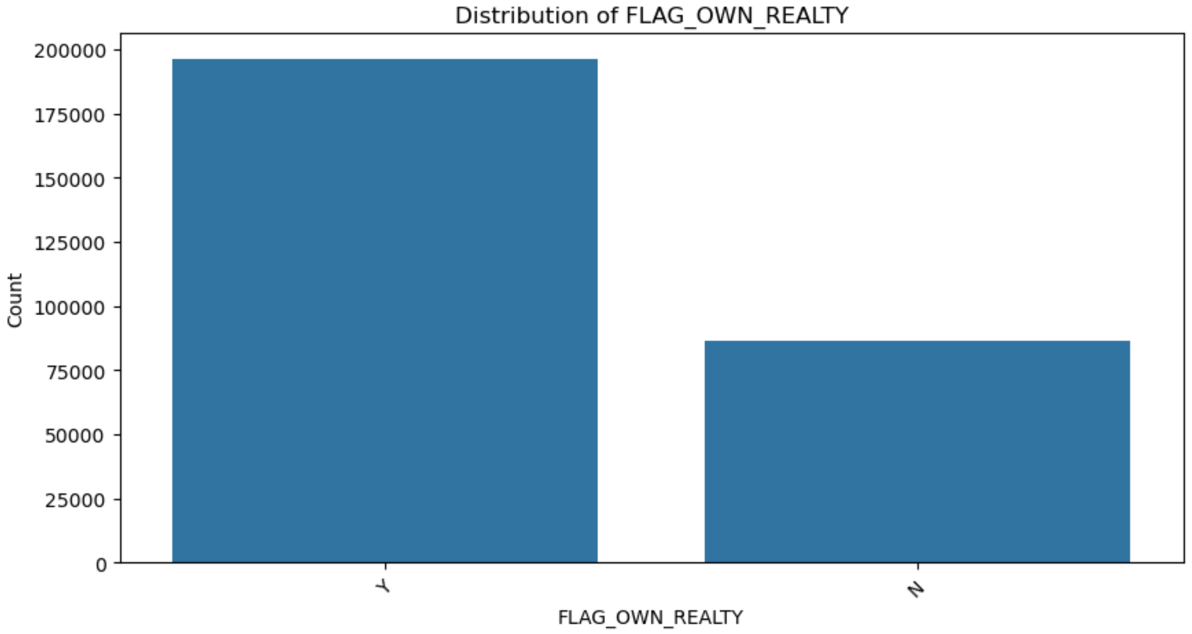
- ❖ **Cash Loans:**
- • Total Contracts: 250,000.
- ❖ **Revolving Loans:**
- • Total Contracts: Approximately 20,000.

Distribution of Contacts by Gender

- ❖ **Female (F)**: 14,000 individuals.
- ❖ **Male (M)**: 11,000 individuals.

20. Univariate Analysis - Categorical Summary Statistics: Graphs





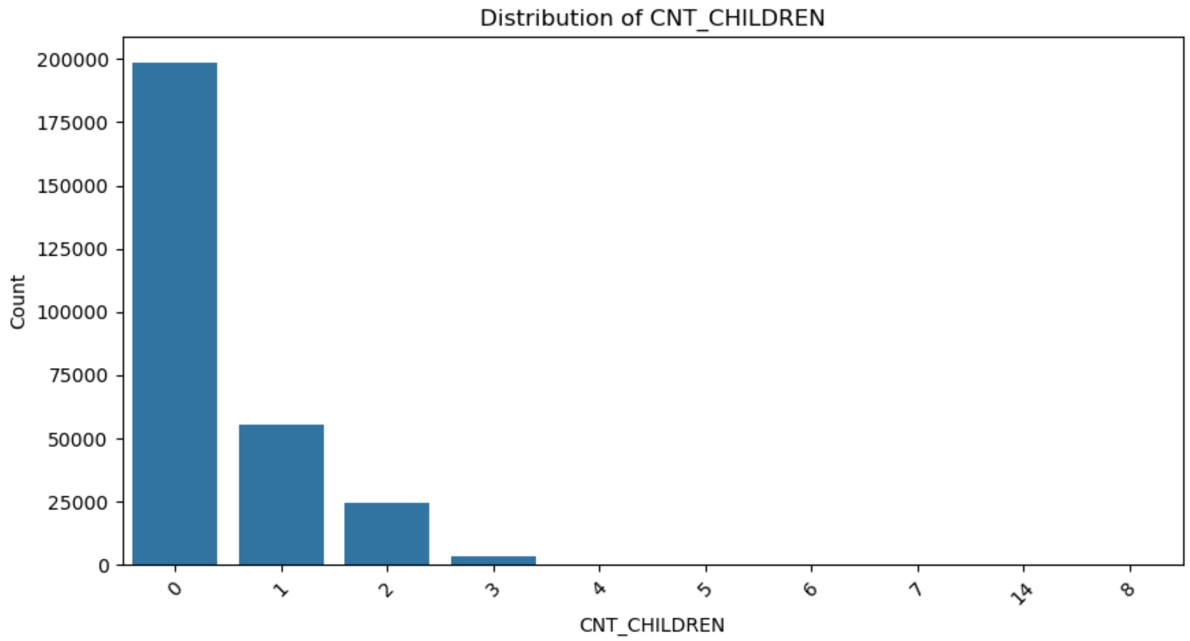
<u>Flag: Own Car</u>

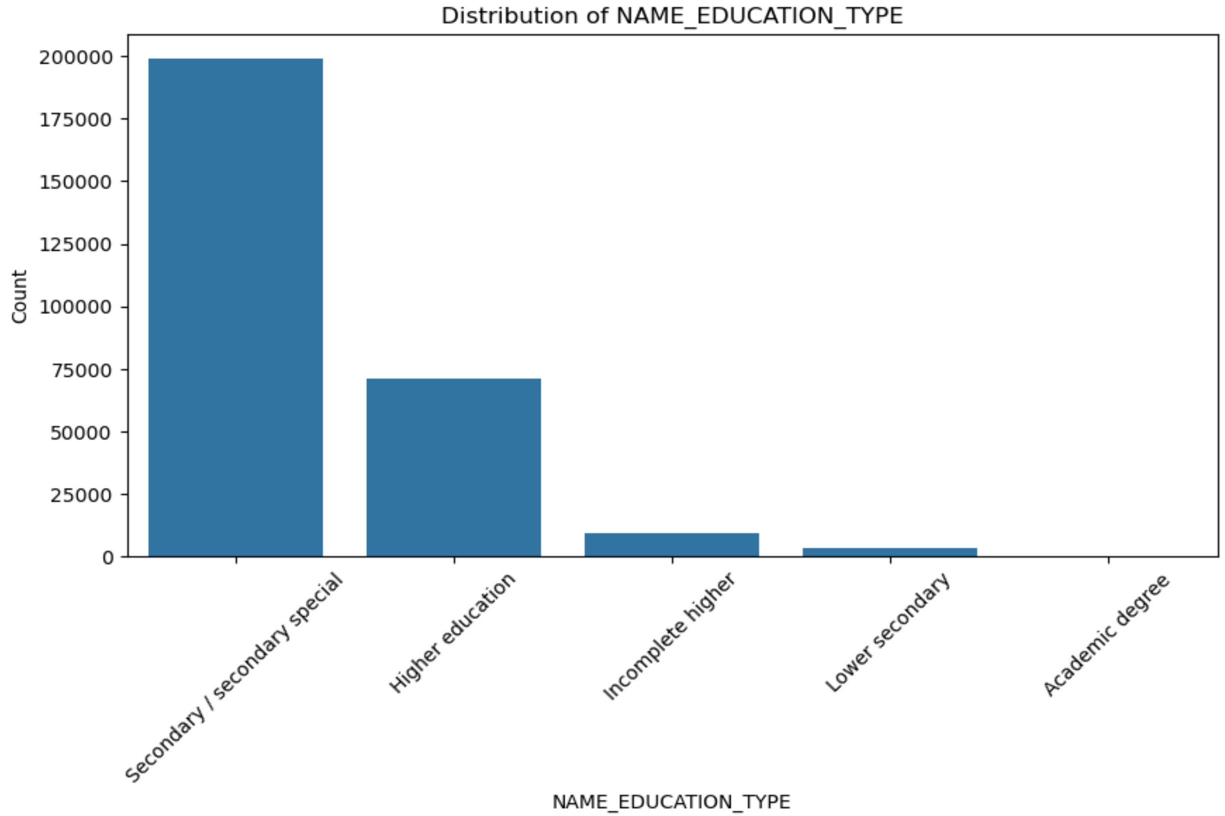
- **No (N)**: Just over 175,000 individuals.
- **Yes (Y)**: Approximately 80,000 individuals.

<u>Flag: Own Realty</u>

- **Yes (Y)**: Almost 200,000 individuals.
- **No (N)**: Just over 75,000 individuals.

20. Univariate Analysis - Categorical Summary Statistics: Graphs





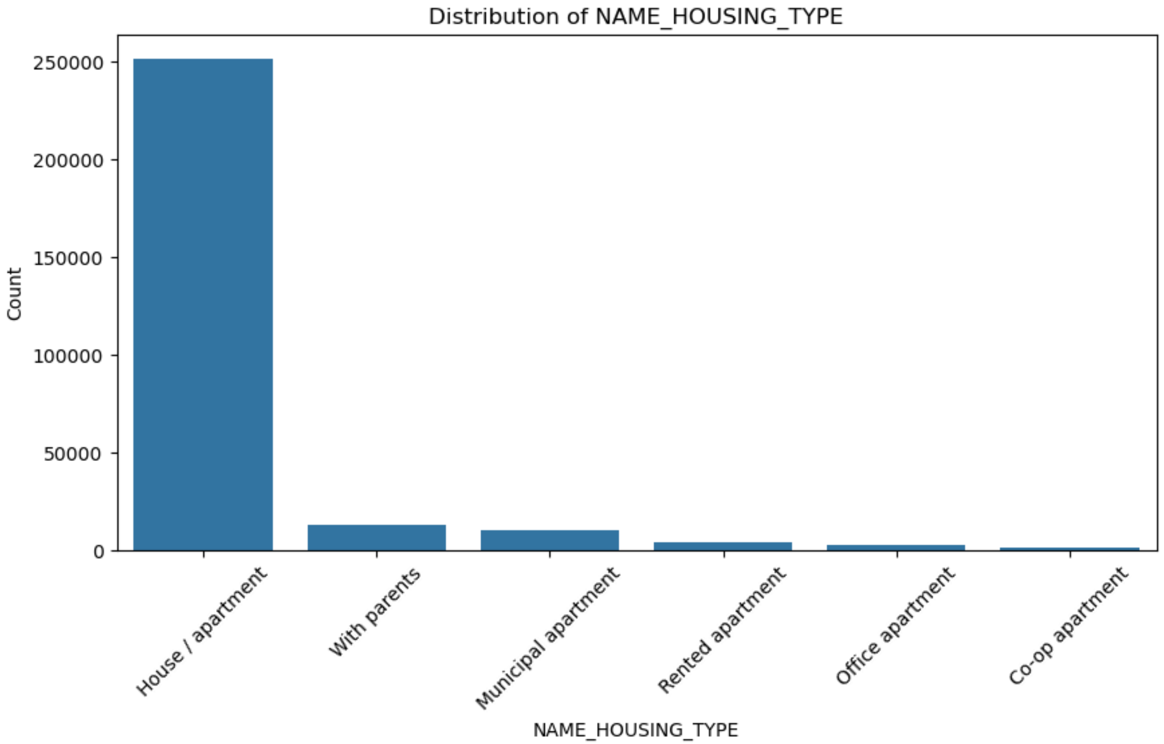
Distribution of CNT Children

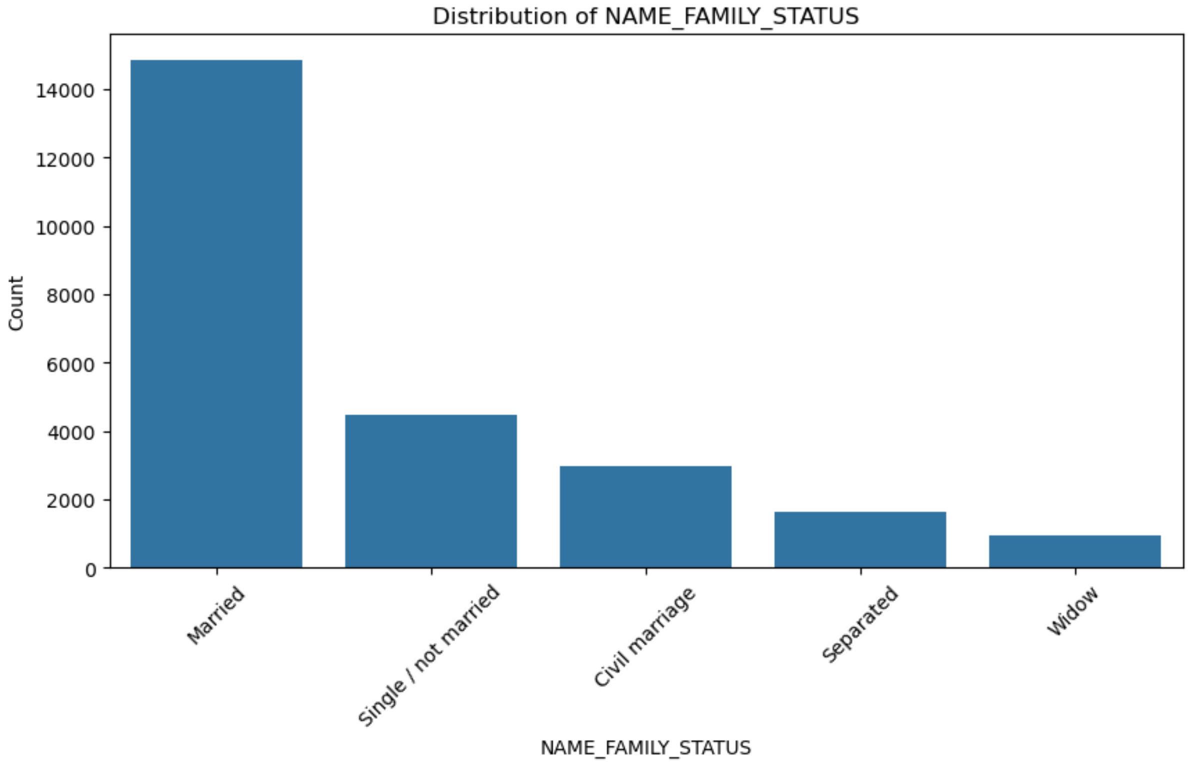
- **No Children (0)**: Almost 200,000 individuals.
- **One Child (1)**: Just over 50,000 individuals.
- **Two Children (2)**: Approximately 25,000 individuals.
- **Three Children (3)**: Almost **0** individuals.
- **Four to Eight Children (4-8)**: **0** individuals.
- **Nine to Fourteen Children (9-14)**: 0 individuals.

Distribution of Name Education Type

- **Secondary/Secondary Special**: 200,000 individuals.
- **Higher Education**: Almost 75,000 individuals.
- **Incomplete Higher Education**: Around 5,000 individuals.
- **Lower Secondary Education**: Around 1,000 individuals.
- **Academic Degree**: 0 individuals.

20. Univariate Analysis - Categorical Summary Statistics: Graphs





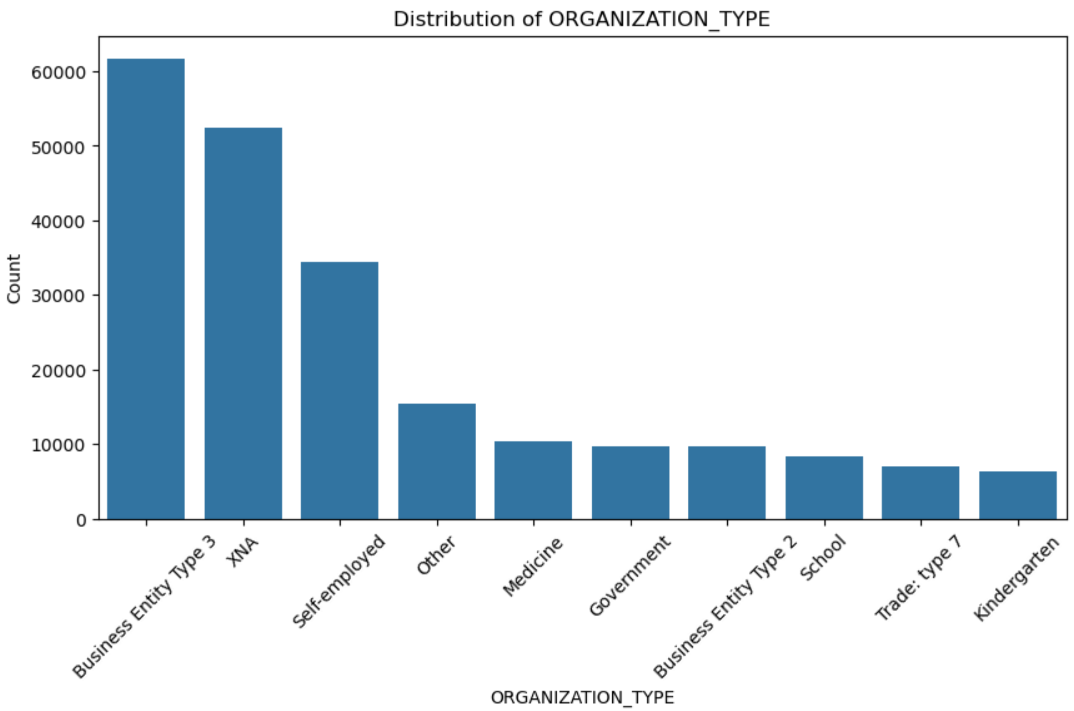
Distribution of Housing Type

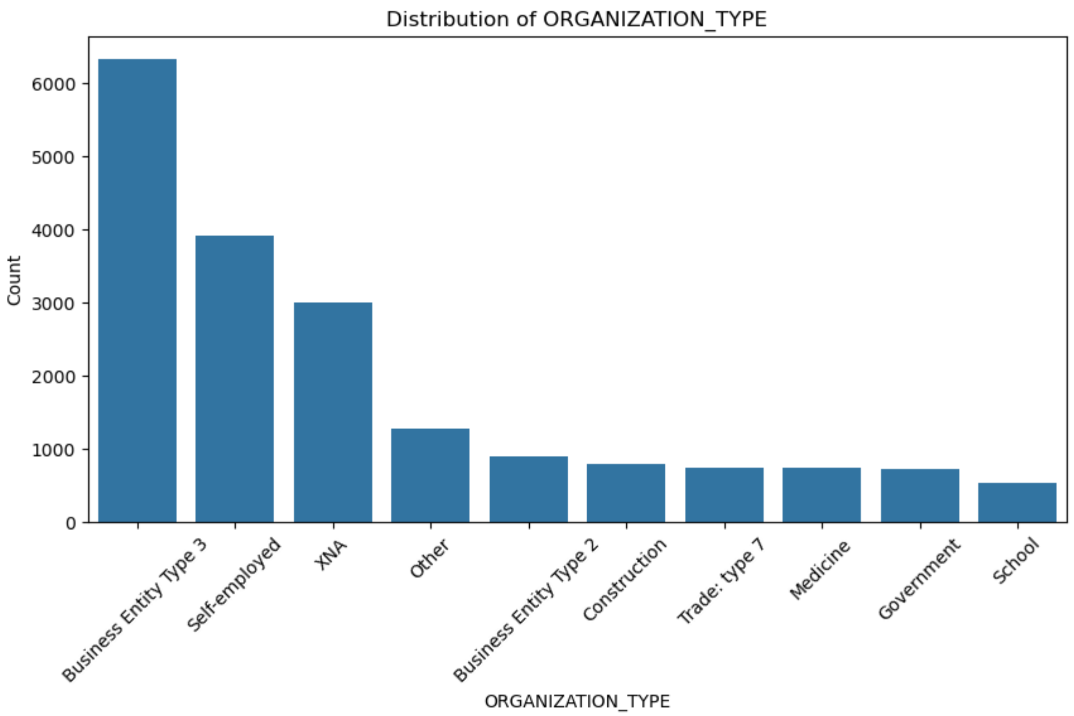
- **House/Apartment**: 250,000 individuals.
- **Living with Parents**: Approximately 20,000 individuals.
- **Municipal Apartment**: Approximately 20,000 individuals.
- **Rented Apartment**: Around 10,000 individuals.
- **Office Apartment**: Almost 0 individuals.
- **Co-op Apartment**: Almost 0 individuals.

Distribution of Family Status

- **Married**: Over 14,000 individuals.
- **Single/Not Married**: 5,000 individuals.
- **Civil Marriage**: 3,000 individuals.
- **Separated**: Almost 2,000 individuals.
- **Widowed**: 1,000 individuals.

20. Univariate Analysis - Categorical Summary Statistics: Graphs





Distribution of Organization Type (Set 1)

- **Business Entity Type** 3: 60,000 organizations.
- **XNA**: A bit over 50,000 organizations.
- **Self-Employed**: 35,000 individuals.
- **Other**: 15,000 organizations.
- **Medicine**: 10,000 organizations.
- **Government**: 10,000 organizations.
- **Business Entity Type 2**: 10,000 organizations.
- **School**: Around 7,000 organizations.
- **Trade Type 7**: Around 6,000 organizations.
- **Kindergarten**: Around 5,000 organizations.

Distribution of Organization Type (Set 2)

- **Business Entity Type 3**: Over 6,000 organizations.
- **Self-Employed**: 4,000 individuals.
- **XNA**: 3,000 organizations.
- **Other**: 1,500 organizations.
- **Business Entity Type 2**: 1,000 organizations.
- **Construction**: 1,000 organizations.
- **Trade Type 2**: 800 organizations.
- **Medicine**: 800 organizations.
- **Government**: 800 organizations.
- **School**: 500 organizations.

21. Bivariate Analysis: Insights from Spearman Correlation

❖ **Key Findings:**

- **AMT_INCOME_TOTAL & AMT_CREDIT:** Correlation of **0.41**

  Moderate positive relationship; higher income often correlates with higher credit amounts.

- **AMT_CREDIT & AMT_ANNUITY:** Correlation of **0.83**

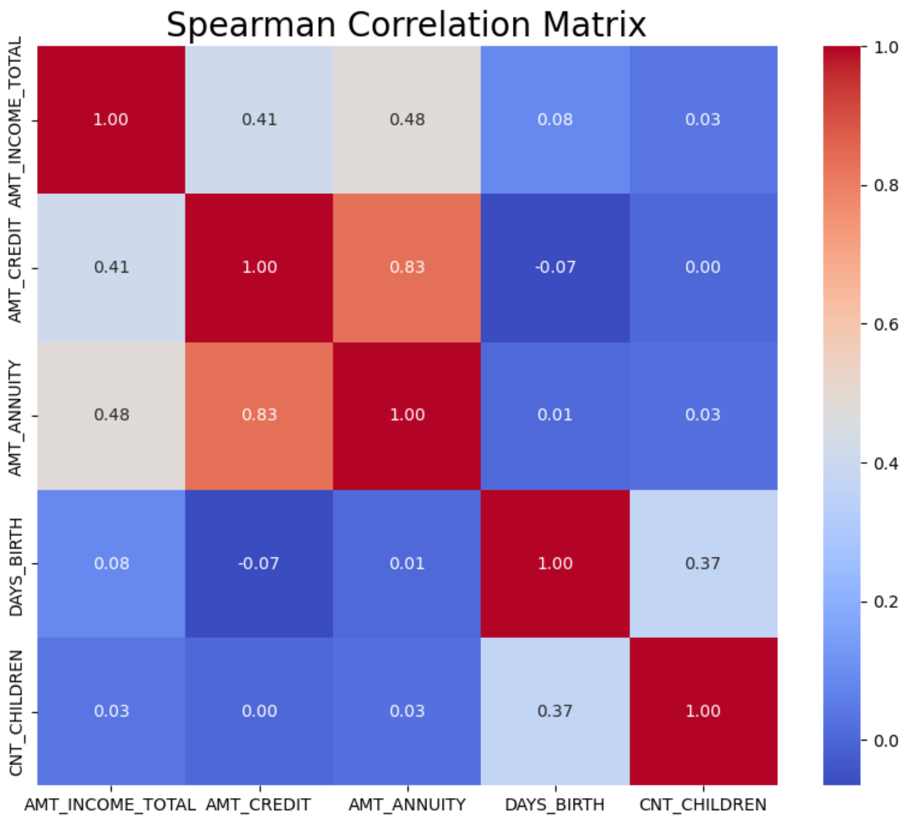
  Strong positive relationship; larger credit amounts typically lead to higher annuities.

- **DAYS_BIRTH & AMT_CREDIT:** Correlation of **-0.07**

  Weak negative correlation; age has minimal effect on credit amount.

- **CNT_CHILDREN & AMT_INCOME_TOTAL:** Correlation of **0.03**

  Very weak positive relationship; number of children has little impact on income.



22. Outlier Detection

❖ **Variables Analyzed:**

- **AMT_CREDIT**
- **AMT_INCOME_TOTAL**
- **AMT_ANNUITY**

❖ **Boxplot Insights:**

- **AMT_CREDIT:**

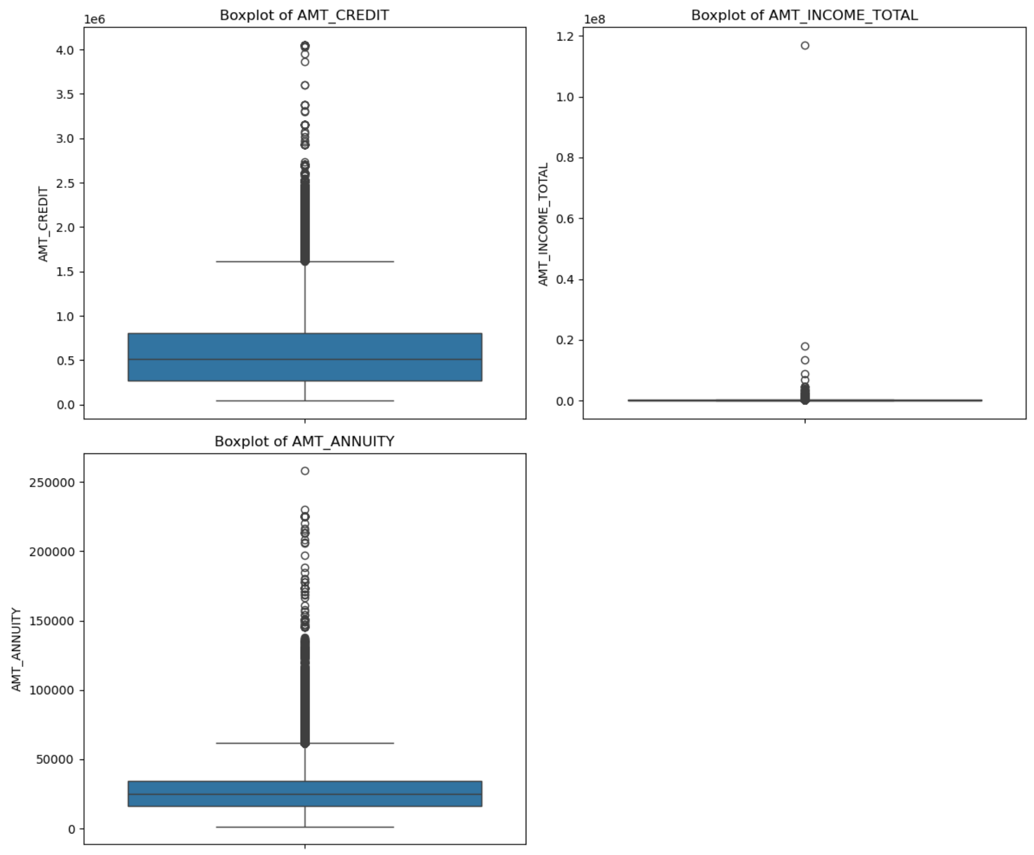
    Presence of high-value outliers; significant range indicating some individuals have very high credit amounts.

- **AMT_INCOME_TOTAL:**

    Outliers present at the higher end, suggesting some individuals report exceptionally high incomes.

- **AMT_ANNUITY:**

    Fewer outliers compared to AMT_CREDIT and AMT_INCOME_TOTAL; most values cluster closely.



23. Conclusions

- ❖ **Defaulter Demographics**:
- **Female Defaulters**: 14,000, surpassing male defaulters (just over 10,000).
- **Non-Defaulter Disparity**: Female non-defaulters (175,000) significantly outnumber male non-defaulters (almost 100,000).
- This gender disparity highlights a potential area for further investigation into the factors influencing loan defaults across genders.
- ❖ **Income Group Performance**:
- **No Defaults**: The $0 - $20,000 income group demonstrates strong financial stability with no default rates.
- **Default Rates**: The $20,000 - $40,000 group shows an 8% default rate, while the $40,000 - $60,000 group has a slightly lower rate at 7%.
- ❖ **Key Takeaway**:
- Default rates remain relatively consistent across higher income brackets ($60,000 - $100,000), suggesting that factors beyond income, such as financial management and economic conditions, play a significant role in default risk.

23. Conclusions

❖ **Default Rates by Loan Type**:

- **Cash Loans**: Default rate stands at 8%, indicating higher risk associated with this loan type.
- **Revolving Lines of Credit**: Show lower default rates (5-6%), suggesting better borrower management of flexible repayment options.

❖ **Borrower Dynamics**:

- Majority of cash loan borrowers (250,000) are non-defaulters, indicating effective management.
- Revolving loans have a notably low number of defaulters, reflecting a favorable risk profile.

❖ **Demographic Patterns**:

- Over 175,000 individuals do not own cars, while nearly 200,000 own real estate, indicating investment stability.
- The majority report no children (200,000), suggesting demographic trends toward smaller family sizes.

❖ **Age and Default Rates**:

- The 26-35 age group has the highest default rate (31.08%), indicating financial challenges in younger middle-aged individuals.
- Older demographics (56+) show lower default rates, implying better financial management.

23. Conclusions

- **Gender Disparities in Defaults**:
  - Female defaulters outnumber males, while female non-defaulters significantly lead. This suggests a need for gender-focused risk assessments and targeted lending strategies.
- **Income Insights**:
  - No defaults in the lowest income group highlight financial stability, but rising default rates in middle-income brackets suggest that factors beyond income influence repayment capabilities.
- **Loan Type Implications**:
  - Cash loans carry higher default rates compared to revolving lines of credit, indicating that borrowers manage flexible repayment options more effectively. This can inform product development and risk management strategies.
- **Demographic and Age Patterns**:
  - Significant proportions of the population do not own cars, yet many own real estate, reflecting investment stability. Younger age groups face higher default rates, underscoring the importance of tailored financial education and support.