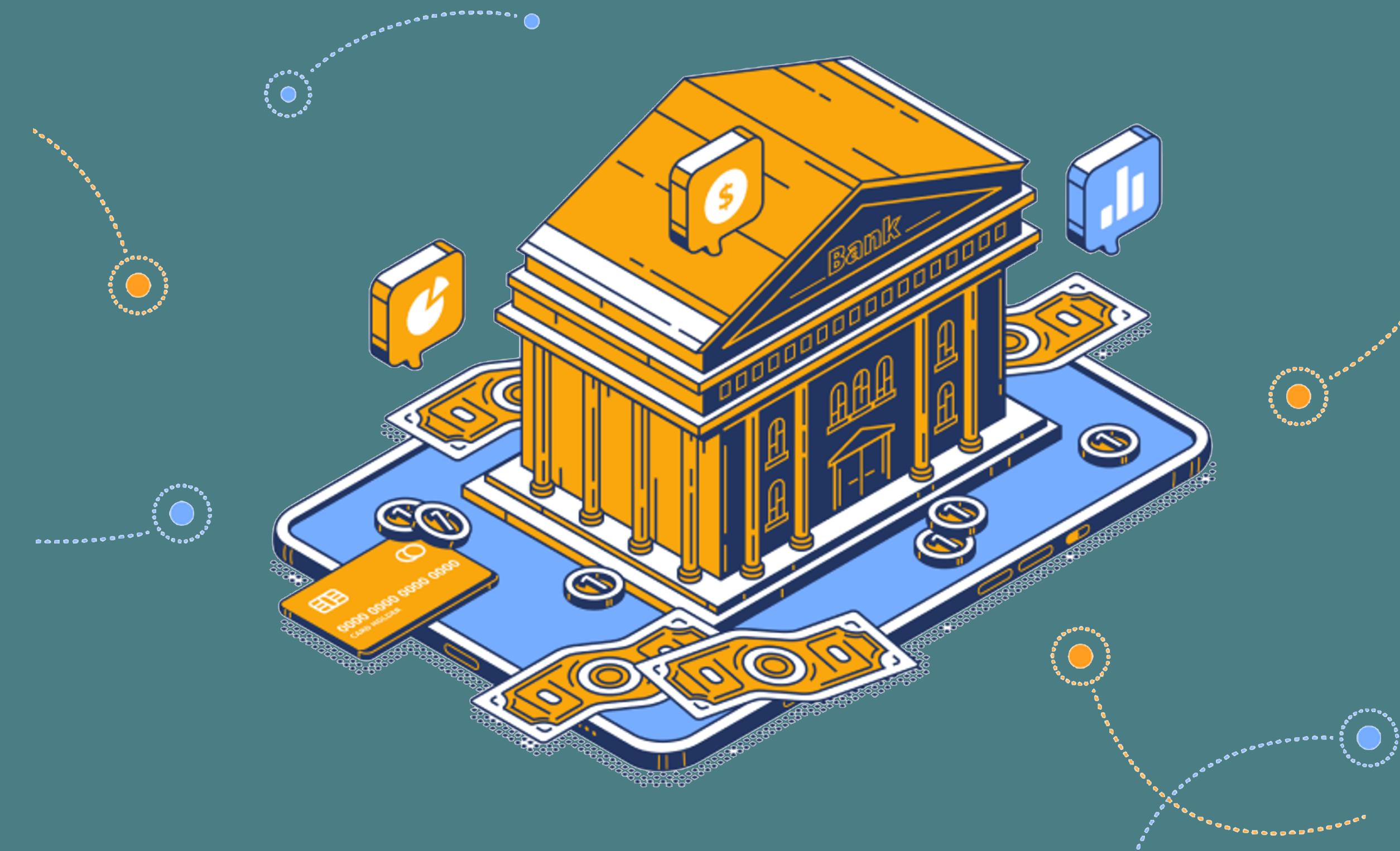


ANALYSIS : RICHARD PANG

Loan Risk Analysis for Banks

Evaluating Key Risk Indicators in Loan Applications



Introduction

- Clean up: Clean the application and previous application datasets using Python.
- Factors Influencing Applications: Analysis of demographic and loan category impacts.
- Occupation Trends: Effects of occupation on loan applications.
- Loan Type Variation: Differences in applications by loan type.
- Gender Influence: Impact of gender on borrowing behavior.
- Family Structure: Effects of family dynamics on applications.
- Education Correlation: Relationship between education level and approval rates.
- Homeownership: Influence of owning a home on borrowing.
- Income Role: Impact of income on loan eligibility.
- Borrower Behavior: Insights into behaviors across demographics.
- Default Rates: Correlation of default rates with borrower profiles.
- Repayment Patterns: Observations on repayment behaviors.
- Objective: Enhance understanding of loan applications and associated risks.

DataSet Summary

Dataset Structure:

- The previous application dataset contains 37 columns and 1,670,214 rows.
- The current application dataset contains 122 columns and 307,511 rows.
- Combined, total dataset comprises 160 columns and 1,977,725 rows.

Key Application Data:

- The “**TARGET**” column categorizes clients:
 - 1 indicates default clients.
 - 2 represents all other cases.

Key Columns:

“Contract Status”:

- Approved loan applications
- Refused loan applications
- Canceled loan applications
- Unused loan applications

ABOUT ME



Richard Pang

Business Analyst - Data Scientist - UX designer

Areas of Expertise

- SQL Query Development
- Python Data Analysis
- Tableau Reporting Skills
- Business Intelligence Analytics
- Data Mining Techniques
- Market Trend Analysis
- Data Warehousing Management
- Team Leadership
- Figma
- Mural
- Marvel
- Adobe Substance painter
- Adobe Photoshop
- Blender
- Autodesk Maya
- Foundry Mari

Context

1

Univariate analysis

- Examining Loan Application Patterns Across Different **Occupation Types**
- Analyzing Loan Application Trends by Various **Contract Types**
- Gender Comparison in Loan Applications: **Male vs. Female Borrowers**
- Gender Comparison in Loan default Applications: **Male vs. Female Borrowers**
- Analyzing Loan Default Patterns in Relation to the **Number of Children**
- Distribution of Loan Applications by **Education Level**
- Distribution of Loan Applications Based on **Housing Type**

2

Bivariate analysis

- Monthly Payment Distribution Across **Education Levels**
- Monthly Loan Payment Analysis by Job Distribution in **Loan Applications**
- Analyzing Contract Status Distribution Compared to **Historical Credit Applications**
- Analyzing Median Application Amounts by **Family Status**
- Assessing the Average Down Payment Amount by **Family Status**

3

Multivariate analysis

- Monthly Loan Payment Trends by **Car Ownership Status**
- Previous Credit Applications Analyzed by **Education Status**
- Contract Status Distribution Across **Occupation Types**
- Correlation Analysis of Key Features in **Default Dataset**
- Correlation Analysis of Key Features in **Non-Default Dataset**

Univariate analysis

Evaluating Individual Risk Factors for Loan Approval and Default Prediction

Examining Loan Application Patterns Across Different **Occupation Types**

Landlords

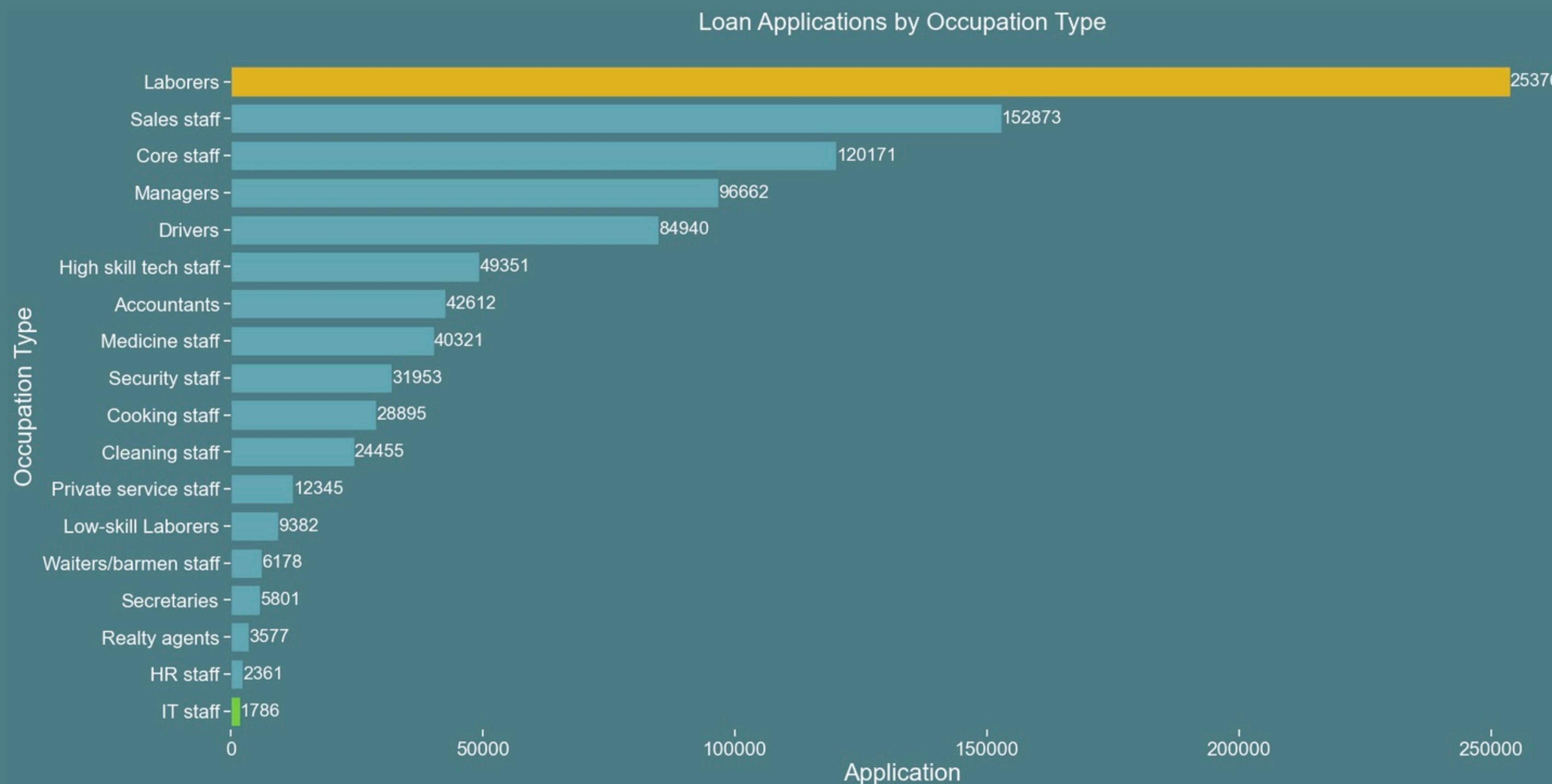
- Landlords have the highest number of loan applications.
- Their lead is significant compared to other occupation types.

Mid-range value

- Sales staff have the second-highest loan applications, followed by core staff and managers.
- Landlords lead sales staff by 100,888 applications.
- Sales staff surpass core staff by 32,702 applications.

IT staff

- IT staff have the lowest number of loan applications at 1,786.
- This is 575 applications fewer than HR staff, who have 2,361 applications.



Analyzing Loan Application Trends by Various **Contract Types**

Cash loans

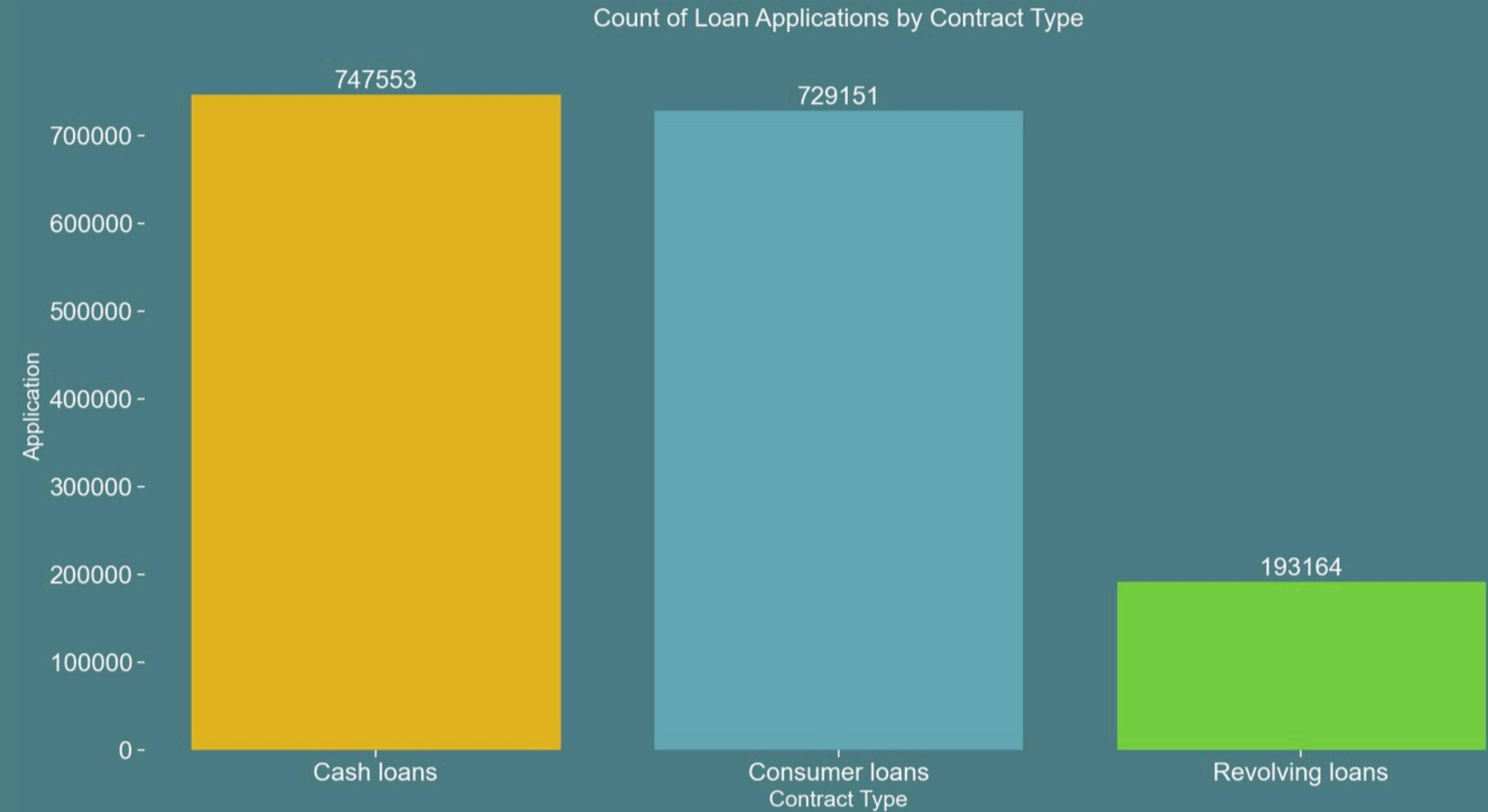
- Cash loans represent the highest number of loan applications.
- The difference between cash loans and consumer loans is only 18,402.

Consumer loans

- Consumer loans have a significantly higher number of loan applications.
- They exceed revolving loans by a margin of 535,987 applications.

Revolving loans

- Revolving loans have the lowest number of applications.
- They total 554,389 applications, a significant difference compared to cash loans.



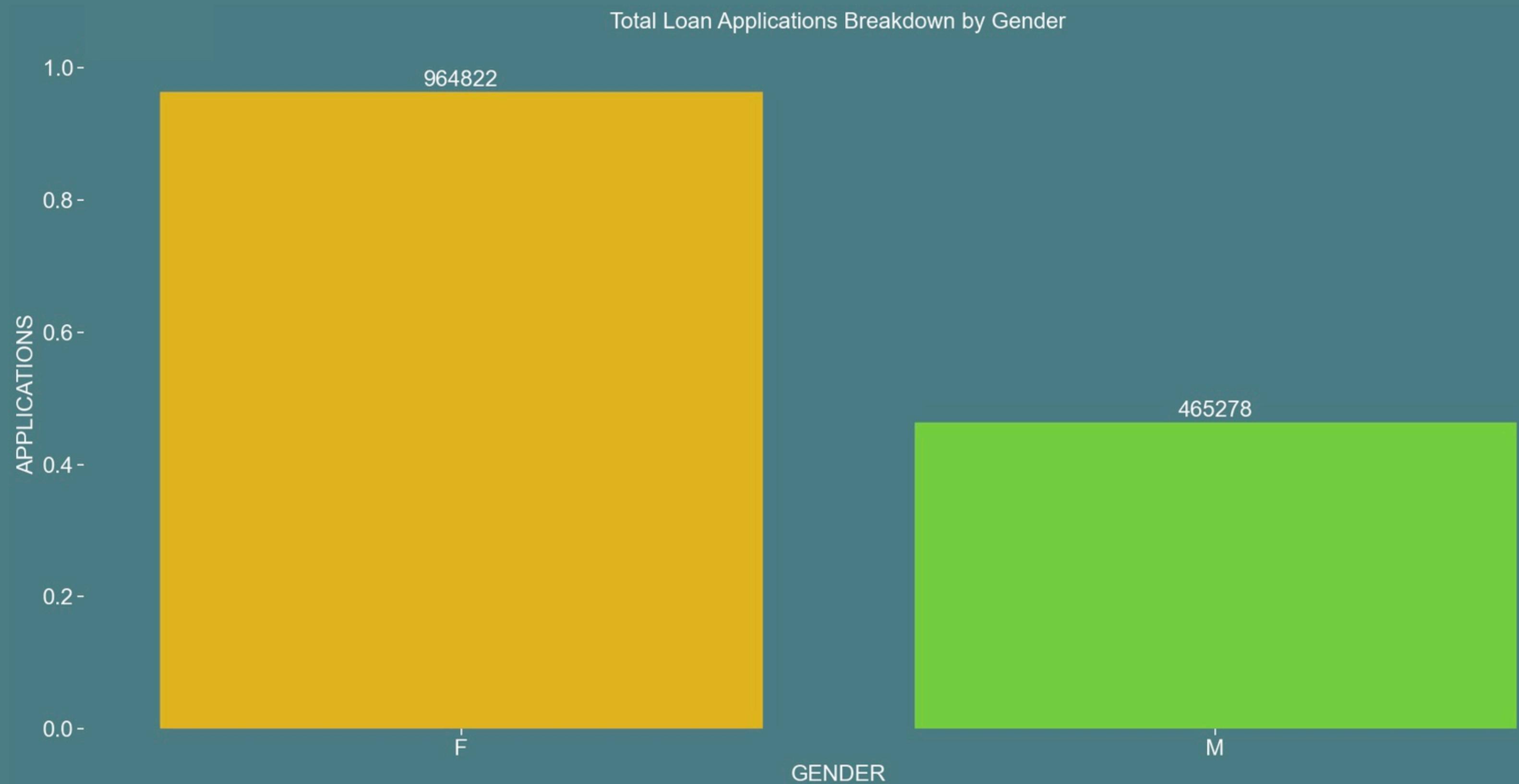
Gender Comparison in Loan Applications: **Male vs. Female Borrowers**

Female

- The number of female loan applications significantly exceeds that of male applications.
- There are 499,544 more applications from females compared to males.

Male

- Male loan applications total only 465,278, compared to 965,822 applications from females.



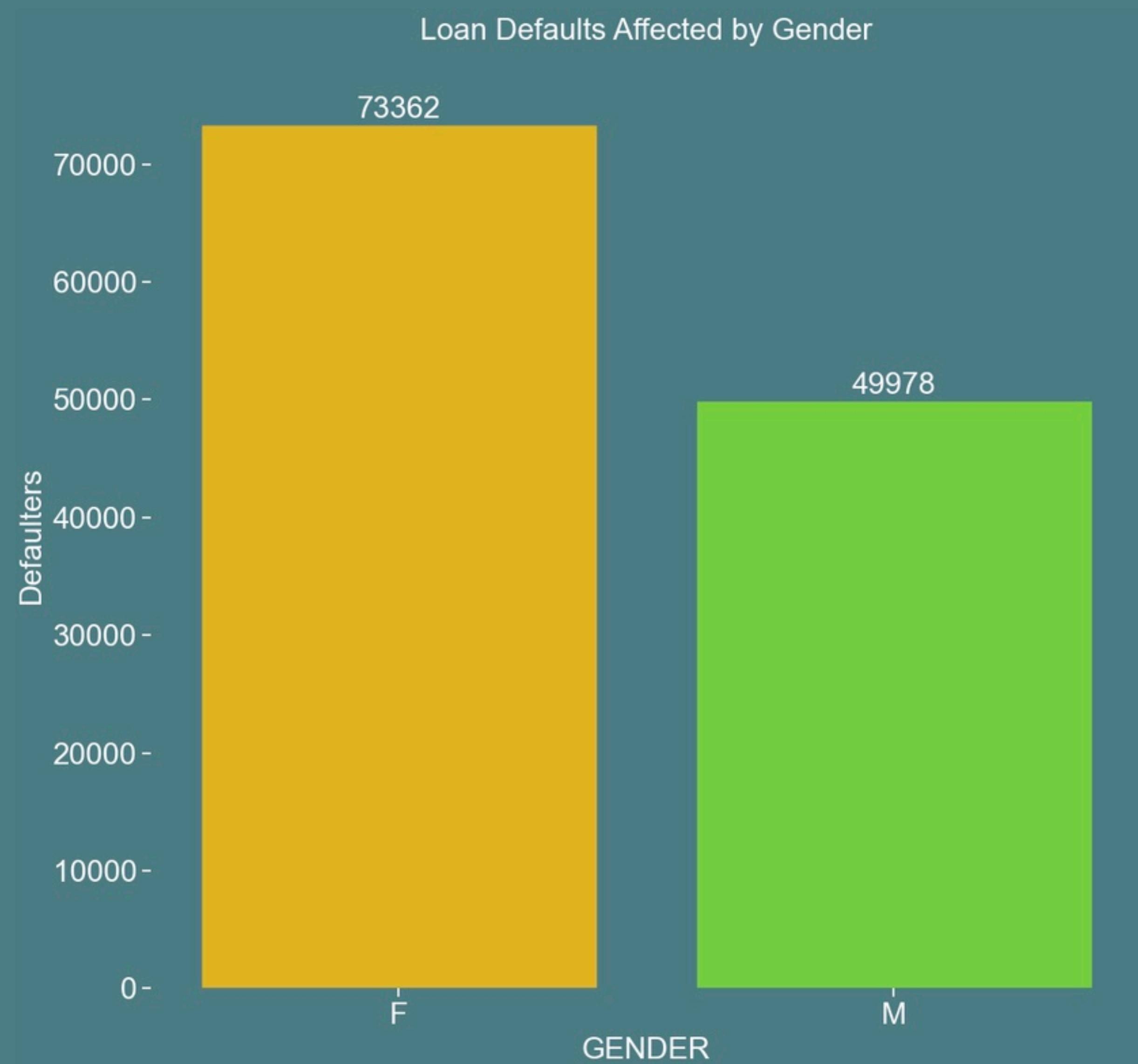
Gender Comparison in Loan default Applications: **Male vs. Female Borrowers**

Female

- The high volume of female loan applications has led to 73,362 defaults out of 965,822 total applications.
- On average, 7.60% of female loan applications result in default.

Male

- Out of 465,278 male loan applications, 49,978 have defaulted.
- On average, 10.74% of male applications result in default.
- This is 3.14% higher than the default rate for female applications.



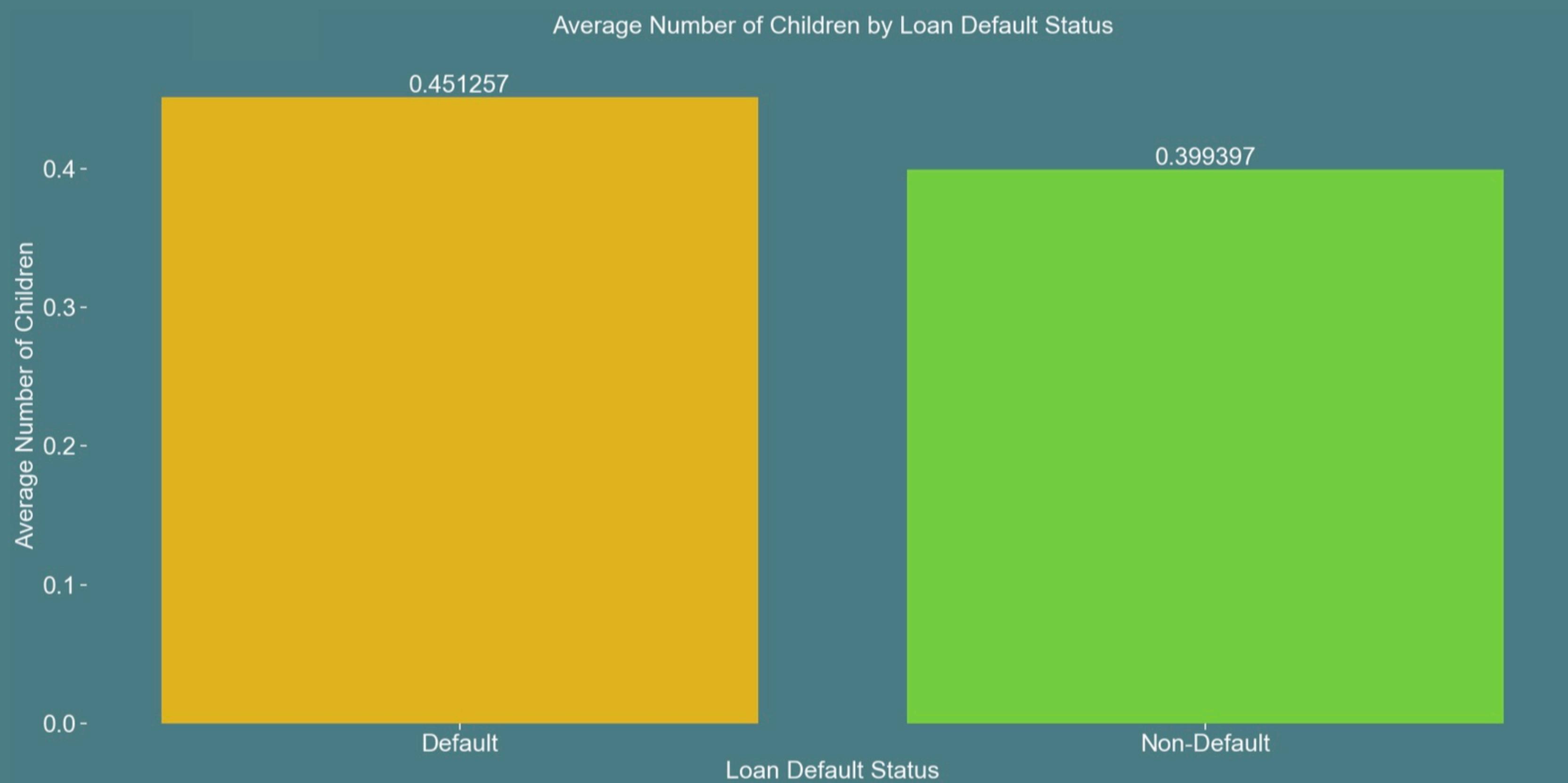
Analyzing Loan Default Patterns in Relation to the Number of Children

Default

- Higher numbers of children in loan applications correlate with increased loan default rates.
- Applicants with more children have a default rate of 0.451257, compared to 0.399397 for non-default loans, indicating higher default rates.

Non - Default

- There is an 11.49% difference between non-default loans and default loans as the number of children increases.



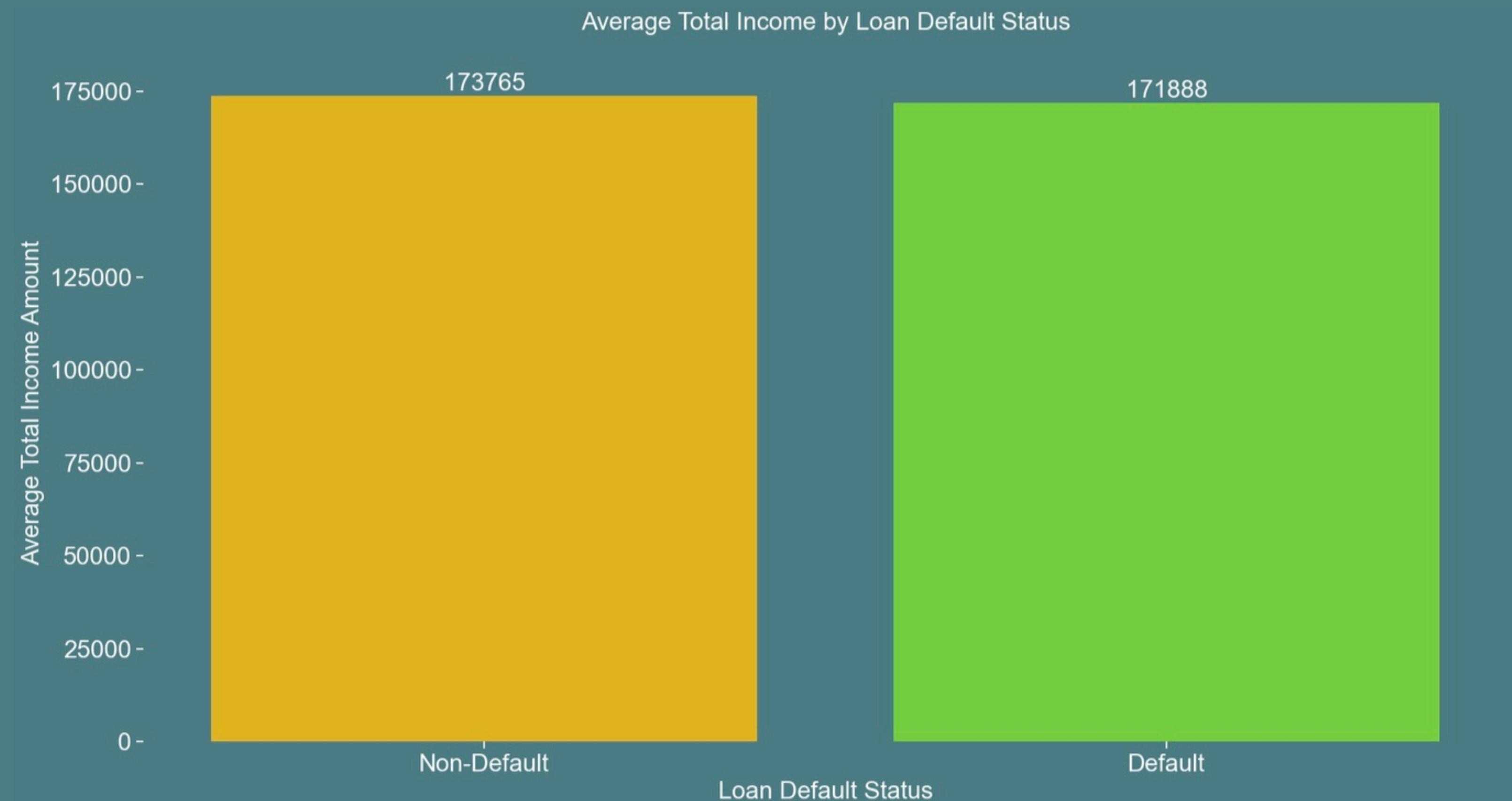
Analyzing Loan Application Patterns: Default and Non-Default Across Income Ranges

Default

- Higher incomes are associated with non-default loans.
- There are 173,765 non-default loans compared to 171,888 default loans, showing a minimal difference.

Non - Default

- There is only a 1.08% difference between non-default and default loans based on total income.



Distribution of Loan Applications by **Education Level**

Secondary / secondary special

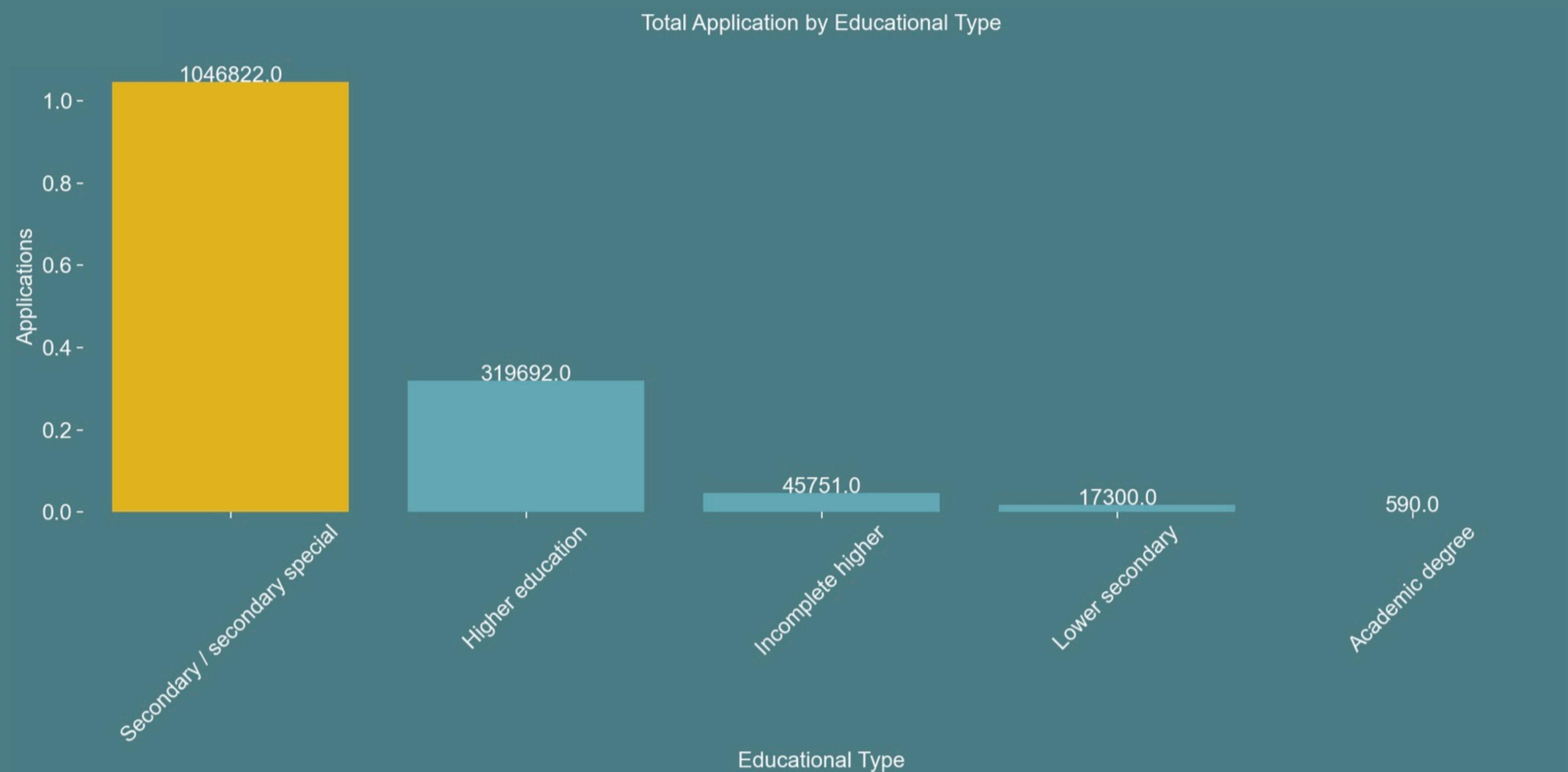
- Secondary and secondary special education levels have the highest loan applications, totalling 1,046,822.
- Higher education ranks second with 319,692 applications.
- The difference in applications between secondary and higher education is 30.54%.

Mid-range value

- Consumer loans have a significantly higher number of loan applications.
- They exceed revolving loans by a margin of 535,987 applications.

Academic degree

- Academic degree holders have the lowest number of loan applications.
- Total applications from this group: 590.
- Overall loan applications: 1,430,155.
- Academic degree holders represent just 0.04% of total applications.



Distribution of Loan Applications Based on **Housing Type**

House / apartment

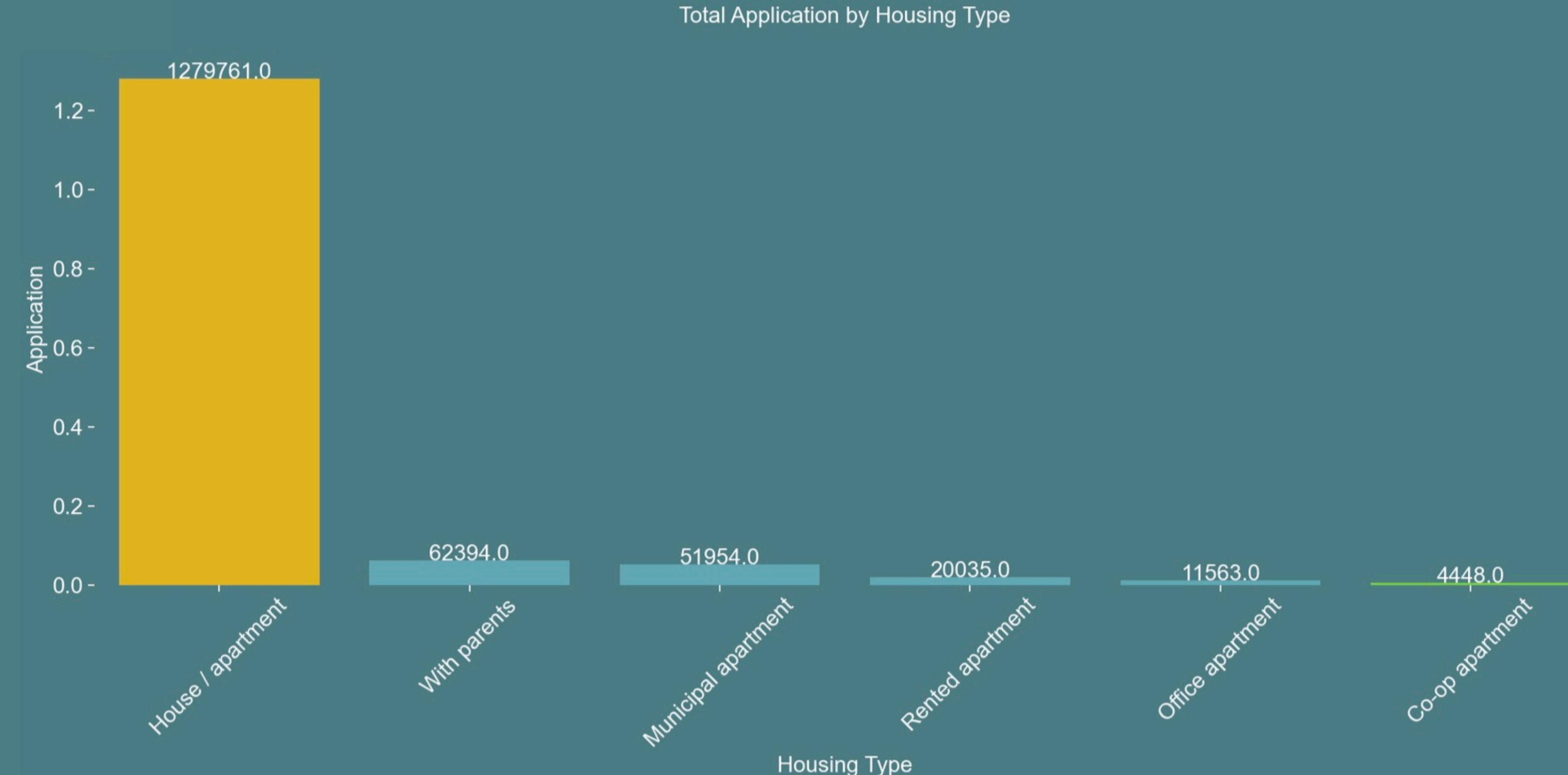
- House/apartment owners submitted the highest number of loan applications.
- Total applications from this group: 1,279,761.
- Overall loan applications: 1,430,155.
- House/apartment owners represent 89.48% of total applications.

Mid-range value

- Mid-range categories (With Parents, Municipal Apartment, Rental Apartment, and Office Apartment) account for 145,946 applications.
- Overall loan applications: 1,430,155.
- These categories represent only 10.2% of total applications.

Co-op apartment

- Co-op apartments have the lowest number of loan applications, with just 4,448 submissions.
- This represents only 0.3% of the total 1,430,155 applications.



Bivariate analysis

Exploring Relationships Between Risk Factors to Predict Loan Outcomes

Monthly Payment Distribution Across **Education Levels**

Academic degree

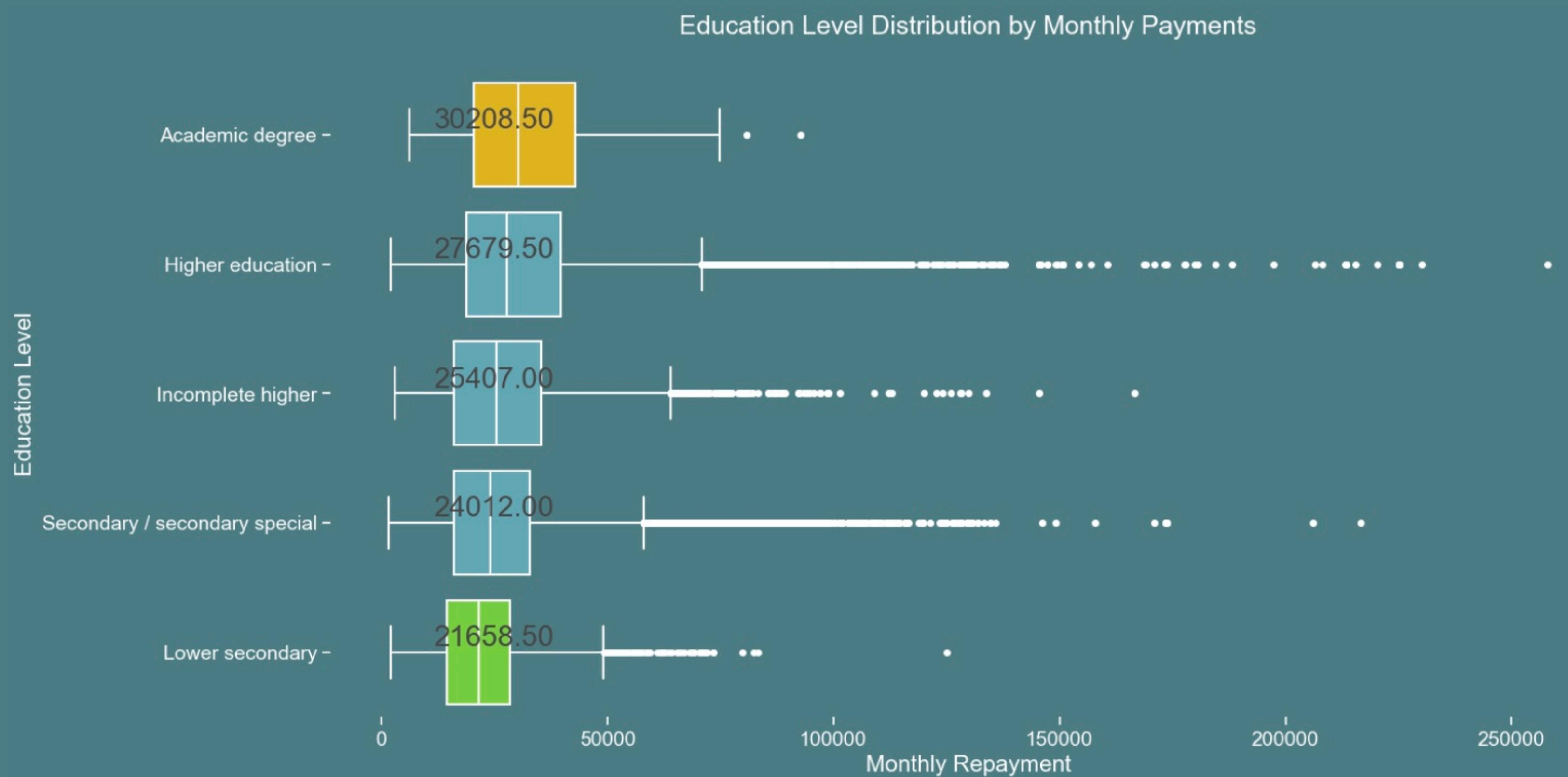
- Academic degree holders have the highest median monthly loan payment per region: 30,208.50.
- They account for 23% of the total median payments across all education levels.

Mid-range value

- Higher Education has a median loan payment of 27,679.50, making up 21% of regional totals.
- Outliers indicate high payments due to unique situations.
- Secondary/Secondary Special outliers show low payments from small loans or favorable terms.

Lower Secondary

- Lower secondary education has the lowest median monthly loan payment per region at 21,658.50.
- This level accounts for only 17% of the total median monthly payments across all education levels per region.



Monthly Loan Payment Analysis by Job Distribution in **Loan Applications**

Businessman

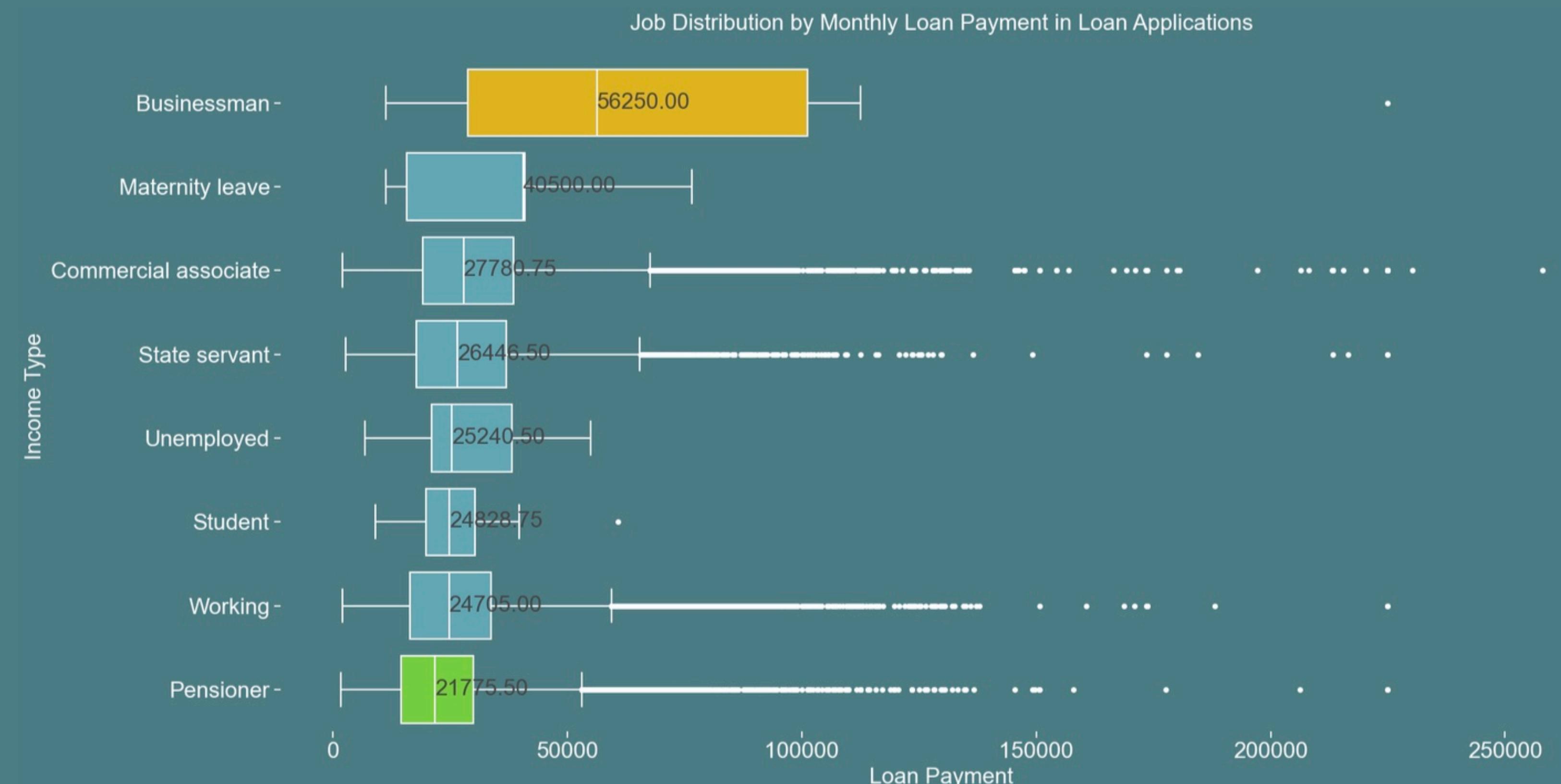
- Business owners hold the highest median monthly loan repayment.
- Median monthly repayment for business owners is 56,250.
- Business loans account for 25% of all loan applications.

Mid-range value

- Maternity leave has the second-highest median monthly loan repayment.
- Maternity leave income accounts for 40,500, representing 18% of total monthly loan payments.
- Commercial associates, state servants, unemployed, students, and workers make up 65% of median monthly loan repayments.

Pensioner

- Pensioners hold the lowest monthly loan application share, making up 9.7% of the median monthly loan repayment.



Analyzing Contract Status Distribution Compared to Historical Credit Applications

Refused

- The previous refusal rate is 54.34% higher than the approval rate.
- Refused applications have a median of 330,345.15, representing 54% of the combined median for all contract applications.

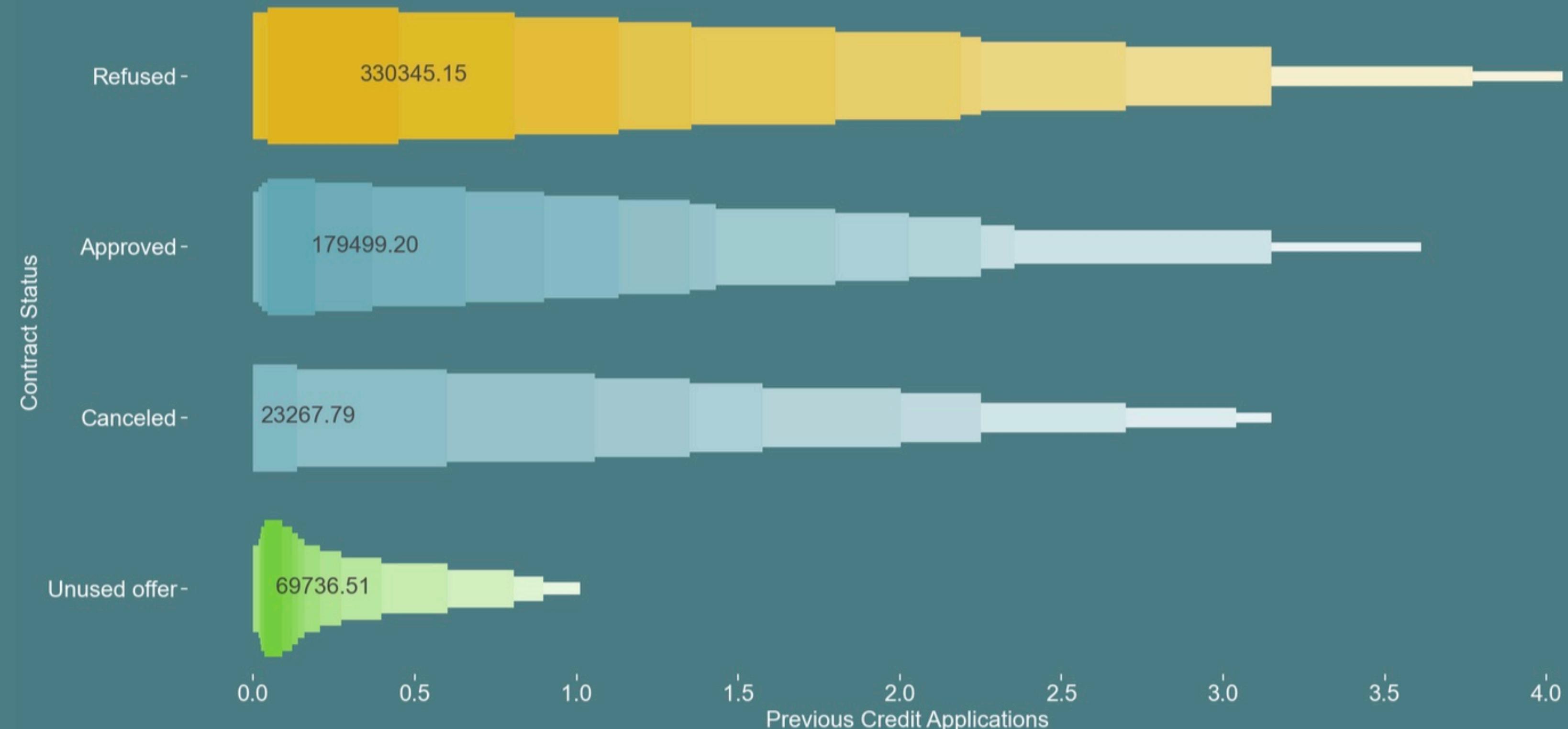
Approved and Canceled

- Approved and canceled credit applications account for only 33% of the total median for previous applications.
- The approval rate is 54.34% lower, indicating it is significantly harder to get a loan approved.

Unused offer

- Unused offers have the lowest median credit application at 69,736.51.
- Used offer holders represent 11% of total previous contract credit applications.

The Means of Evaluating Contract Status Distribution Against Historical Credit Applications



Analyzing Median Application Amounts by **Family Status**

Widow

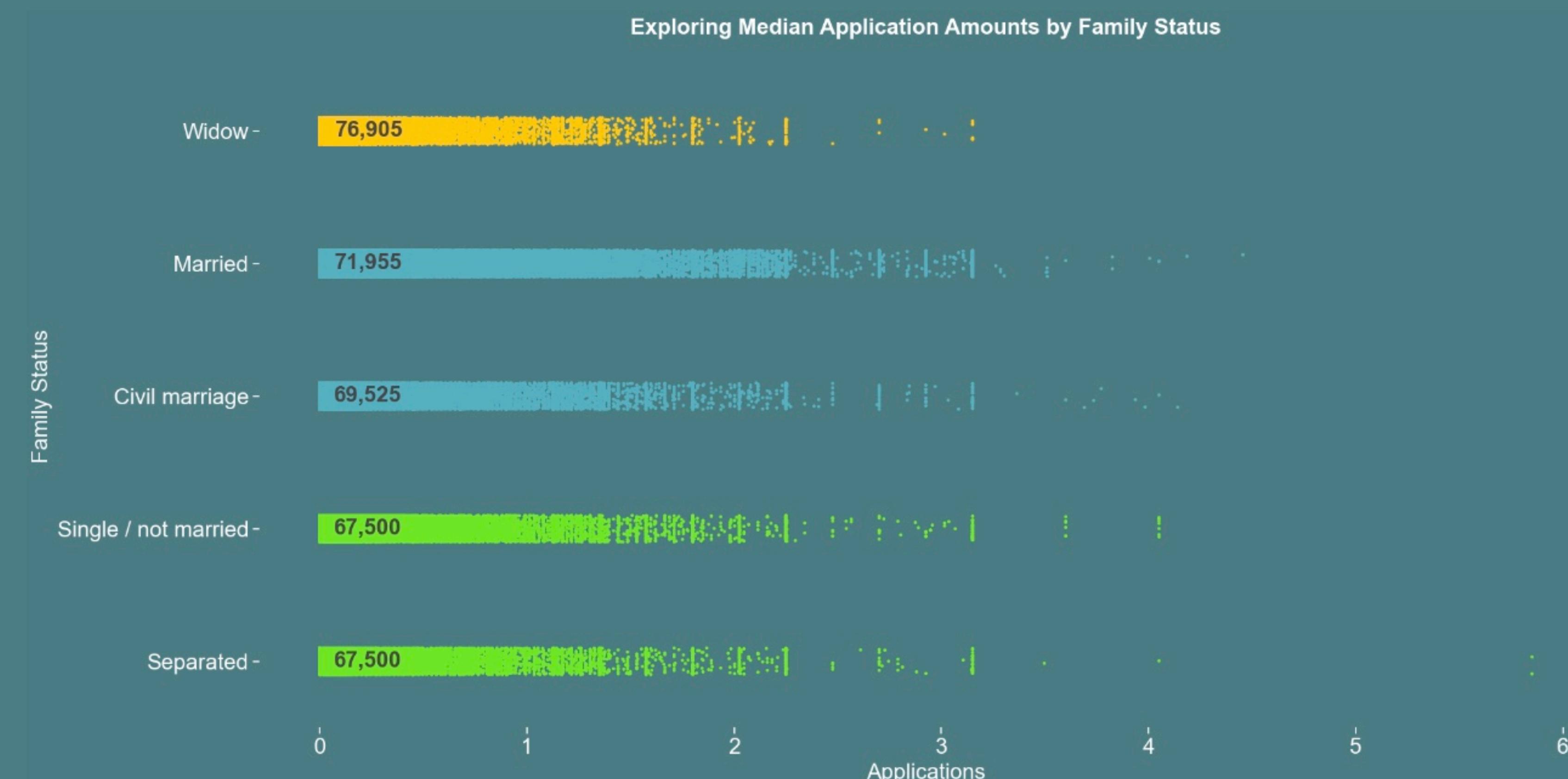
- Widows are 11% above the average median amount for family status applications.
- They hold the highest median applications among all family status categories.

Married and Civil marriage

- The median amount for married applicants is 3.5% above of the average median application amount.
- The median for married application amounts is very close to the highest amount for widows.

Single / not married and Separated

- The median application amounts for singles, not married, and separated are 3% below the average.
- Some outliers in separated applications may have slightly increased the median amount.



Assessing the Average Down Payment Amount by Family Status

Single / not married

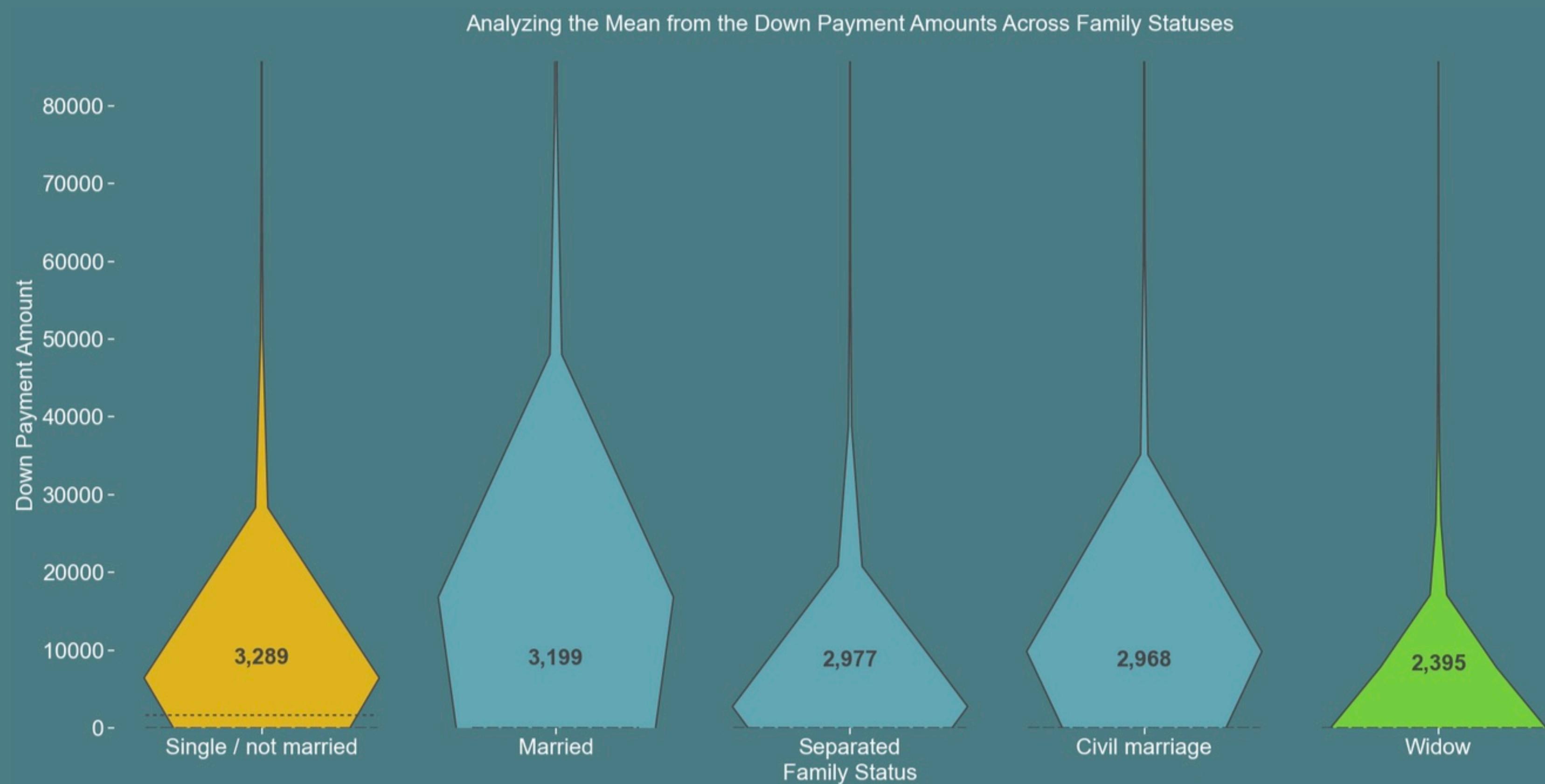
- Single / not married individuals have a mean down payment that is 10.5% above the total average.
- They hold the highest mean amount for down payments.

Mid Range Values

- The mean down payment for married individuals is 3,199, 7.5% above the average of 2,977, and the second highest in the group.
- The mean down payment for separated individuals is 2,977, equal to the average.
- The mean down payment for civil marriages is 2,968, which is 1.3% below the total average.
- There are significantly more applications from married individuals than from single / not married individuals overall.

Widow

- The mean down payment for widows is 2,395, which is 19.5% below the overall down payment application average.
- Widows hold the lowest down payment amount in the group.



Multivariate analysis

Analyzing Multiple Risk Factors Together to Predict Loan Outcomes

Monthly Loan Payment Trends by Car Ownership Status

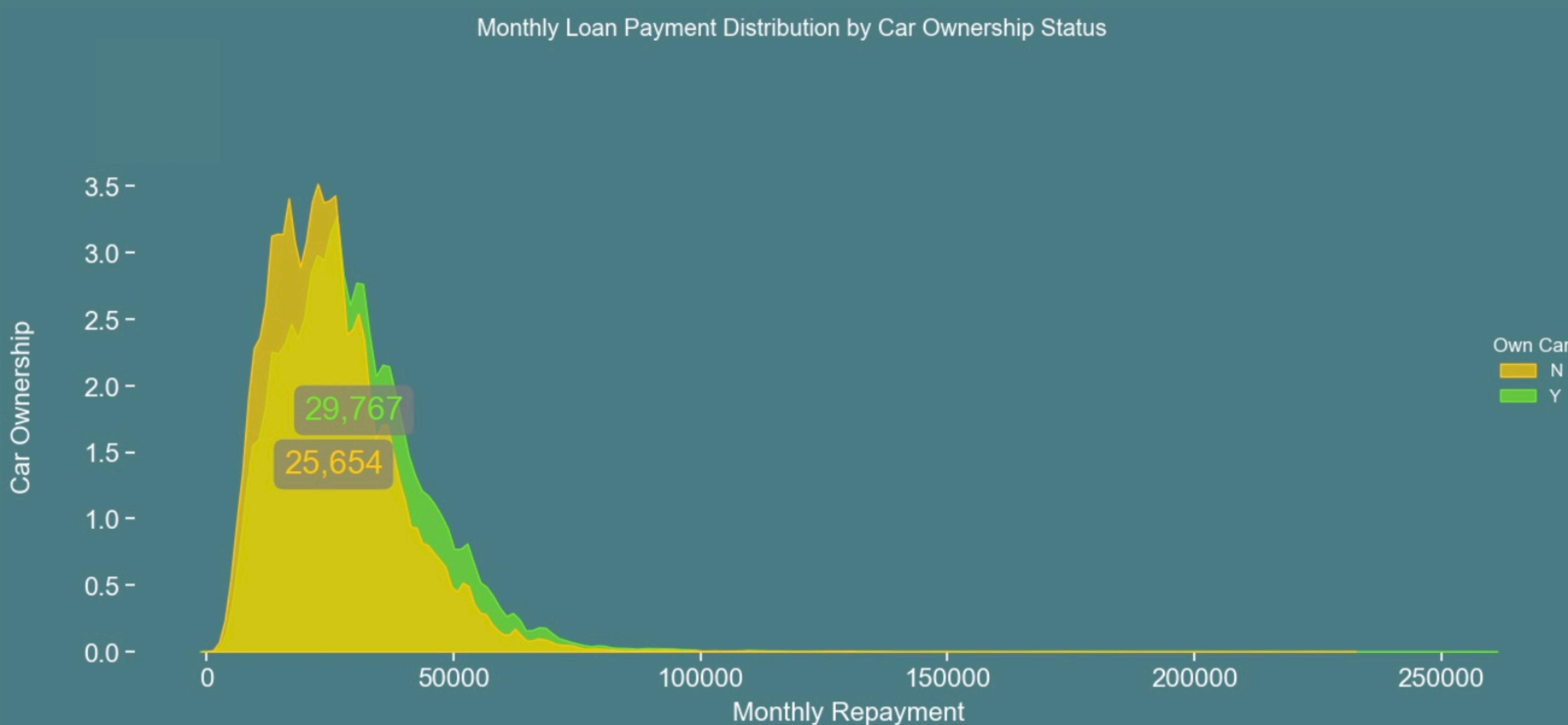
Yes Car

- Car owners pay higher monthly repayments.
- Average monthly repayment for car owners: 29,767.
- This is 7.42% above the overall mean repayment across car ownership statuses.

No Car

- Non-car owners average a monthly payment of 25,654.
- This is 7.54% lower than the overall average.
- There are slightly more non-car owner applicants than car owners.

Monthly Loan Payment Distribution by Car Ownership Status



Previous Credit Applications Analyzed by **Education Status**

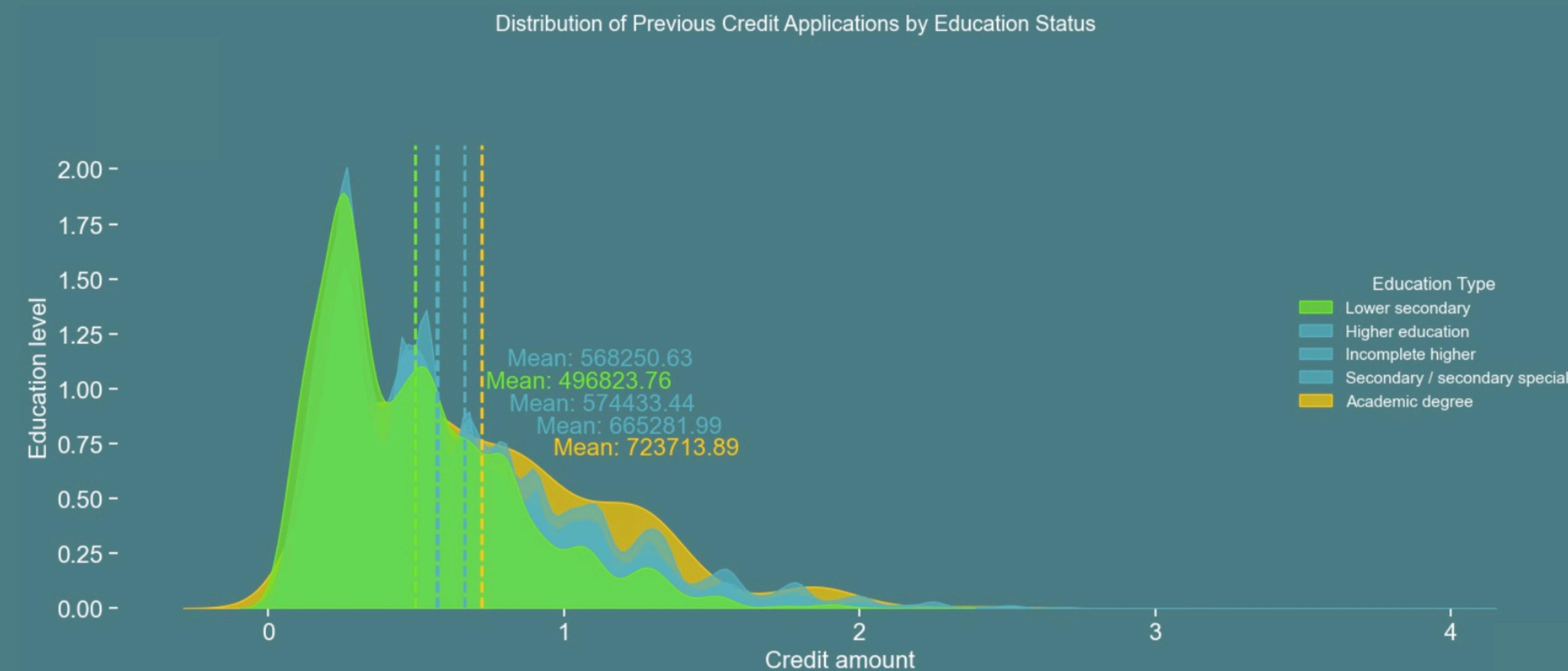
Academic degree

- The academic degree has the highest mean value of 723,713.89,
- Which is 19.8% above the overall average.
- However, it has fewer applications than other education types.

Lower secondary

- Lower secondary has the lowest mean credit amount at 496,823.76.
- This is the lowest mean among all education types.
- Lower secondary is 17.76% lower than the average mean of all education types.
- This indicates a significant disparity in credit amounts compared to other education levels.

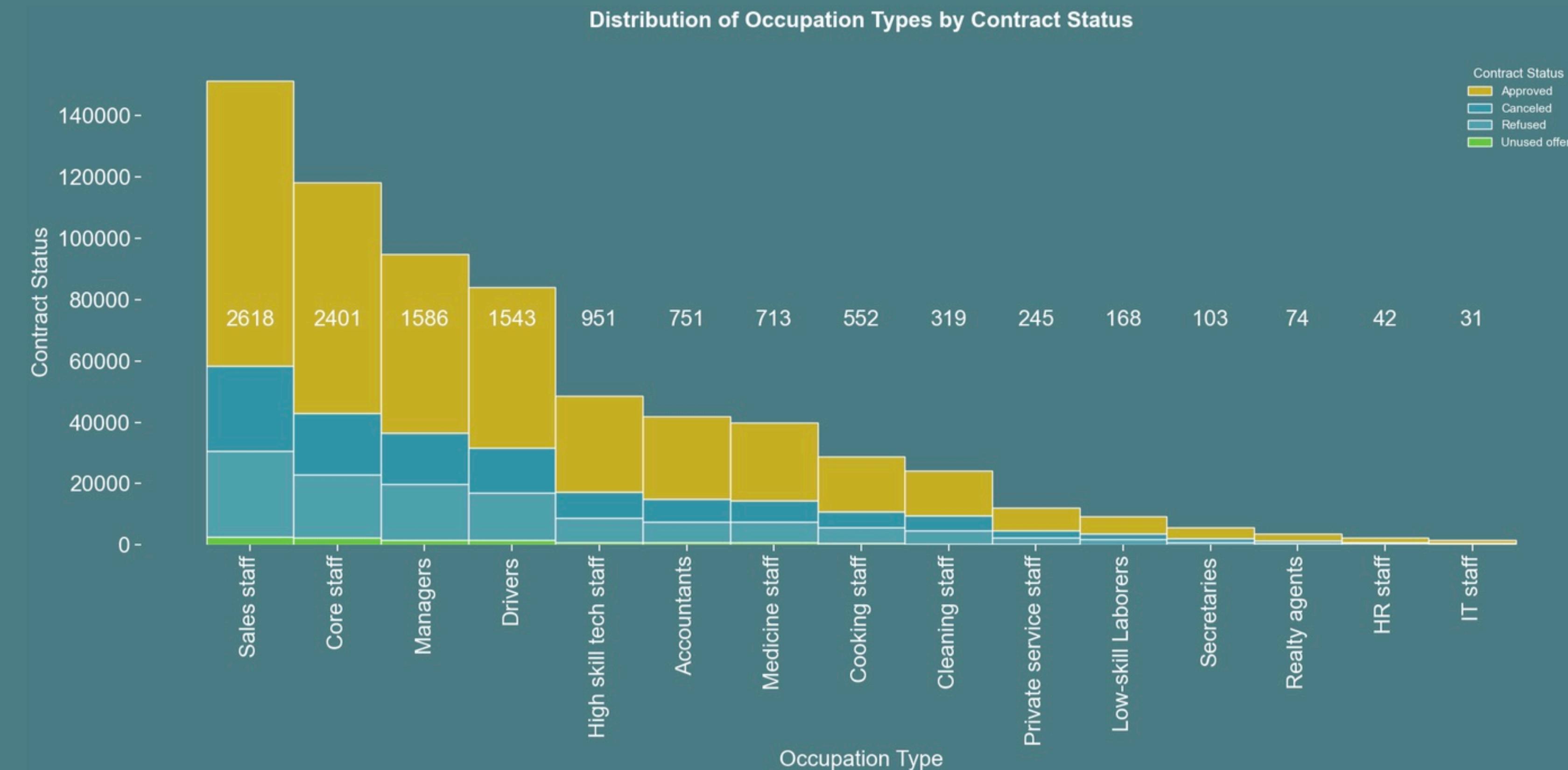
Distribution of Previous Credit Applications by Education Status



Contract Status Distribution Across **Occupation Types**

Contracts

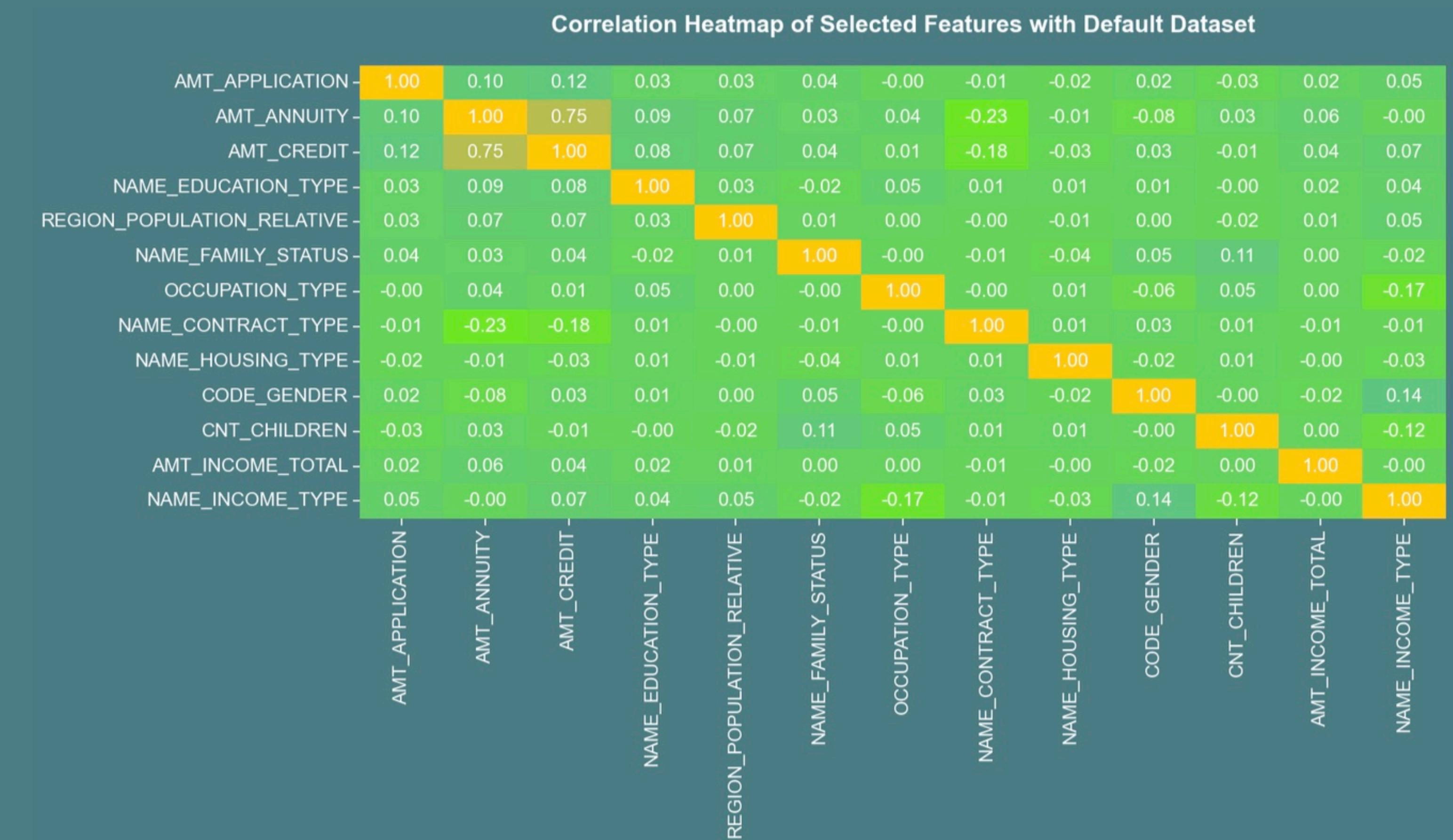
- Sales Staff: Lead with the highest number of contract applications, totaling 2,618, representing 21.6% of all applications.
- Core Staff: Rank second with 19.8% of the total applications, making them significant contributors to contract volume.
- Managers: Hold third place, accounting for 13% of the overall applications.
- Drivers: Close behind in fourth position with 12% of all applications.
- IT Staff: Have the fewest applications, with only 31 out of 12,097, representing 0.26% of the total.



Correlation Analysis of Key Features in **Default Dataset**

Correlation for default application

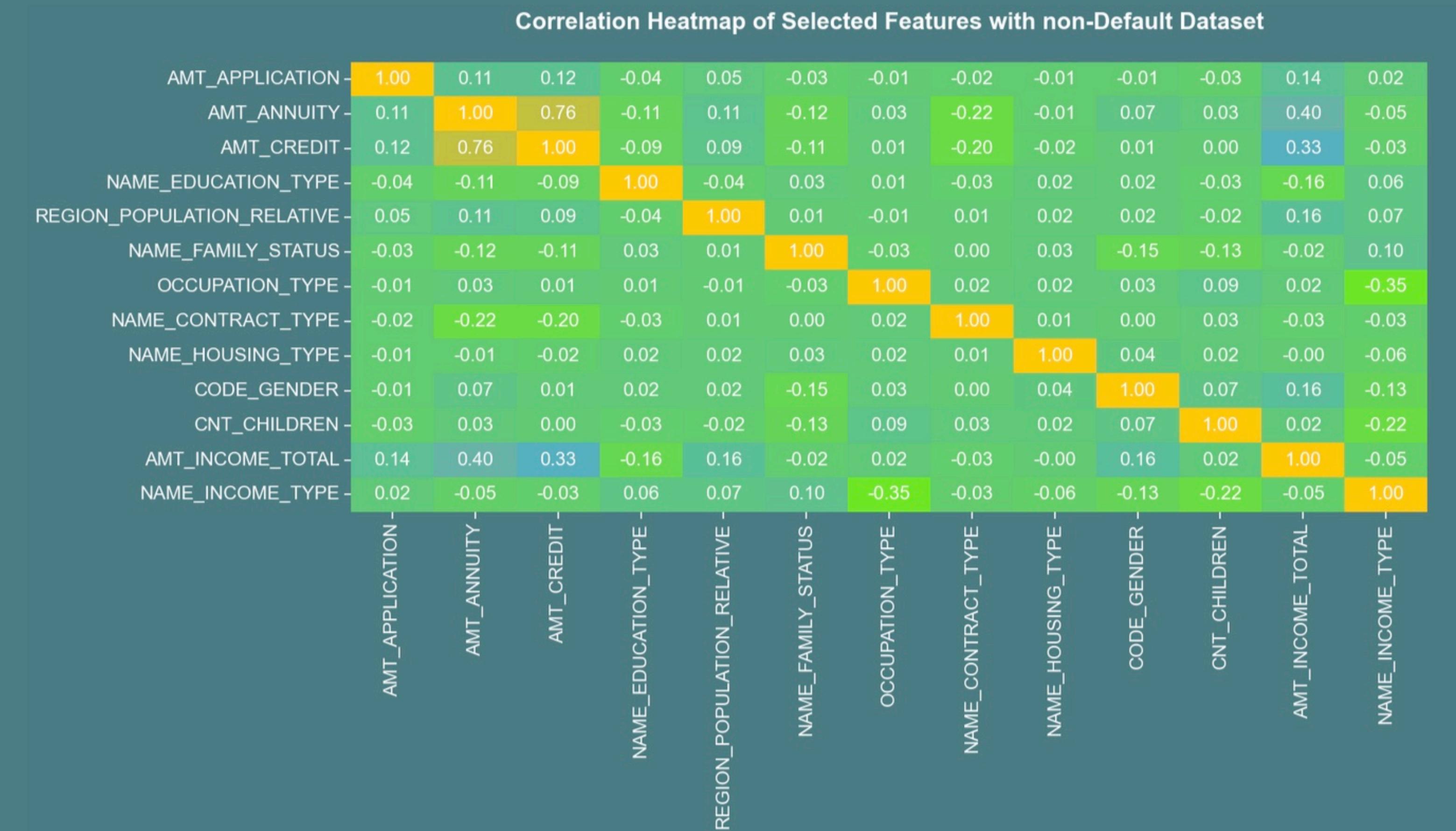
- AMT_ANNUITY & AMT_CREDIT: Exhibit the highest correlation in the dataset.
- NAME_INCOME_TYPE & CODE_GENDER: Hold the second-highest correlation.
- AMT_APPLICATION & AMT_CREDIT: Display the third-highest correlation.
- Loan Annuity & Credit Amount: Strong correlation indicates repayment aligns closely with loan size.
- Income Type & Client Gender: Strong correlation suggests specific income types are closely associated with certain genders.
- Application & Final Credit Amount: Strong correlation suggests clients often receive credit amounts close to what they initially requested.
- These correlations indicate predictable patterns: loan annuity aligns with loan size, income types often associate with specific genders, and credit approvals are close to requested amounts.



Correlation Analysis of Key Features in **Non-Default Dataset**

Correlation for non-default application

- AMT_ANNUITY & AMT_CREDIT: Highest correlation in the dataset.
- AMT_ANNUITY & AMT_INCOME_TOTAL: Second-highest correlation.
- AMT_CREDIT & AMT_INCOME_TOTAL: Third-highest correlation.
- Loan Annuity & Credit Amount: Strong correlation suggests repayment amounts align closely with loan size.
- Loan Annuity & Client Income: Strong correlation suggests that repayment amounts are proportionate to client income.
- Credit Amount & Client Income: Strong correlation suggests loan sizes are adjusted based on client income.
- These correlations suggest loan and repayment amounts are tailored to client income and loan size, indicating income-based loan adjustments and repayment planning.



Conclusion

In conclusion, the most promising candidates for loan approval exhibit the following characteristics:

- Demographic Profile: Female homeowners with secondary education and higher income levels.
- Financial Stability: Single individuals who demonstrate strong down payments.
- Business Ownership: Business owners with a proven track record of timely repayments.
- Loan Preferences: Applicants interested in cash or consumer loans.

By prioritizing these attributes, lenders can increase the chances of approving loans with a reduced risk of default.