

# LOAN DEFAULT

Understanding of risk analytics in banking and financial services

By

Sankaran Meenakshi Sundaram

sankaransms94@gmail.com

# AGENDA

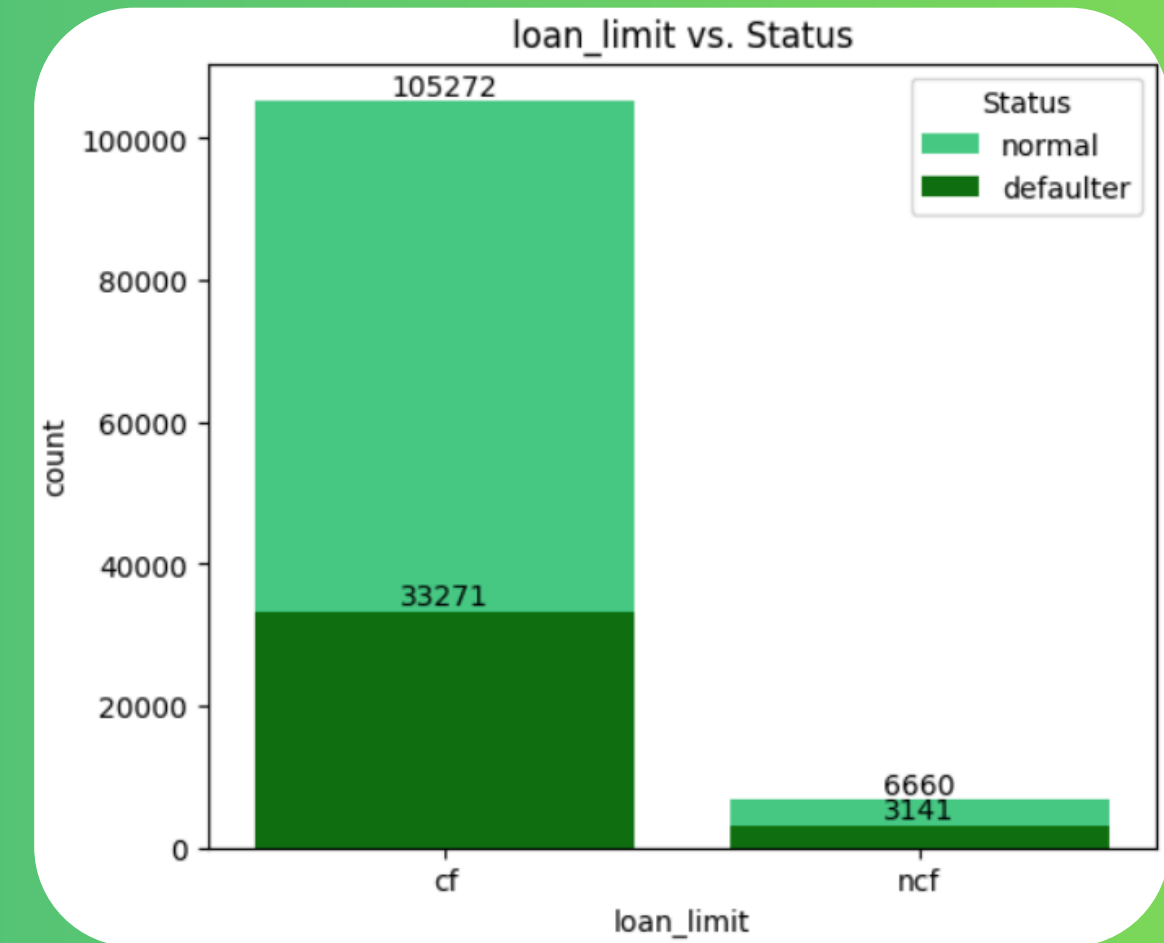
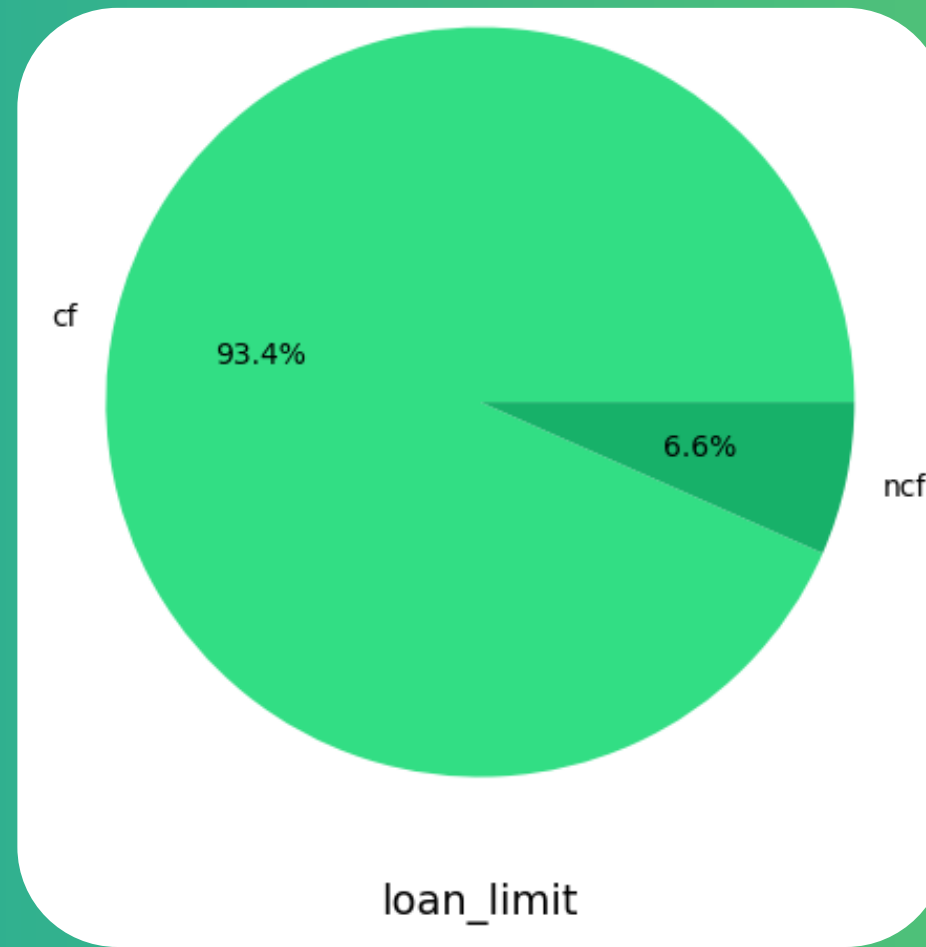
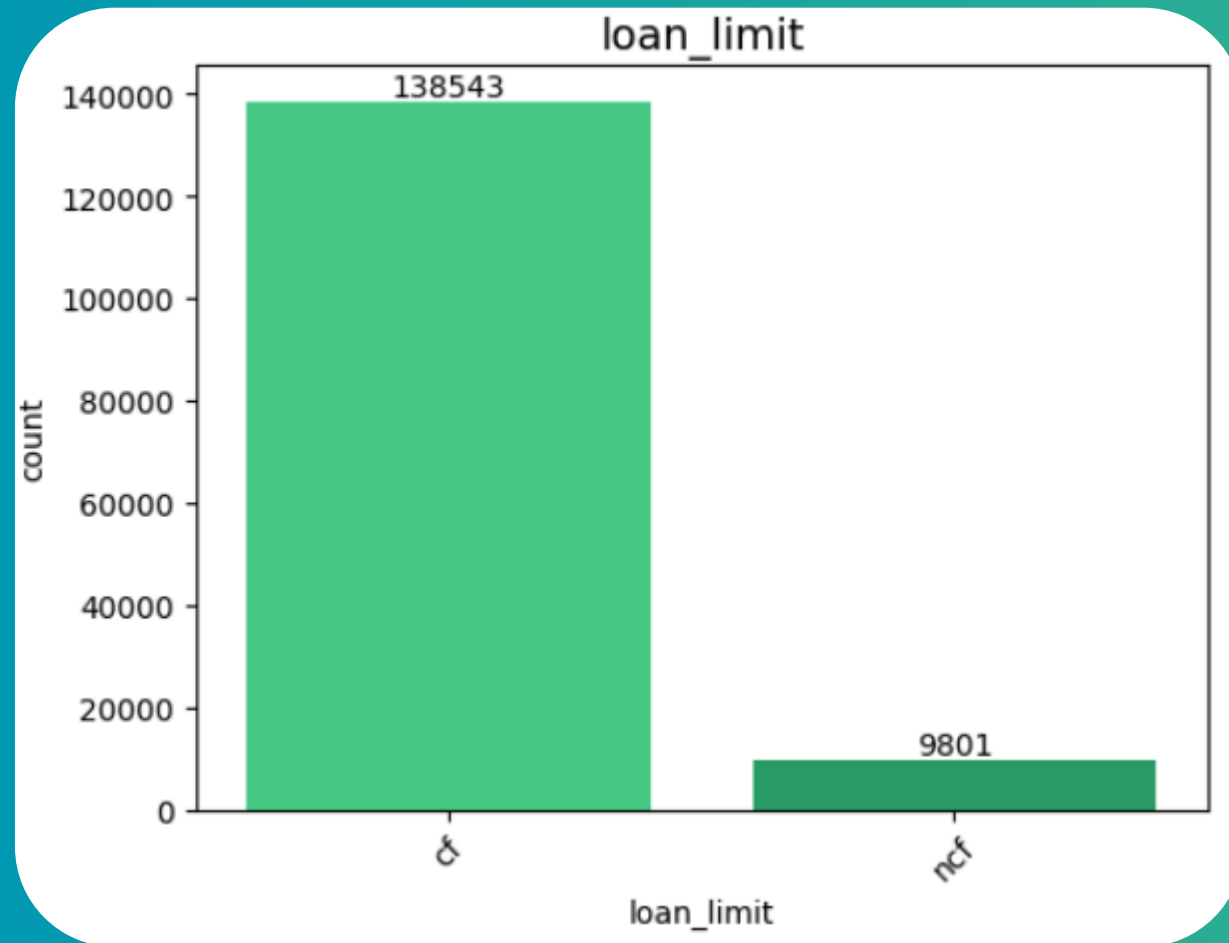
- About the Project
- Introduction to the Project
- Insights and Recommendations based on the Analysis
- Insights and recommendations based on Hypothesis testing

# ABOUT THE PROJECT

This project focuses on **risk analytics within the banking sector**, aiming to identify **key factors that contribute to loan defaults**. By analyzing variables such as loan amounts, LTV (Loan-to-Value) ratios, property values, and borrower demographics, **we uncover insights that can help predict and mitigate default risks**. Our analysis utilizes statistical techniques, including Chi-Square and Mann-Whitney U tests, to highlight relationships between these factors and loan status. **The goal** is to provide **actionable insights for improving risk assessment strategies and making more informed lending decisions to minimize defaults**.

# INTRODUCTION TO THE PROJECT

- **Objective:** Identify **key factors influencing loan defaults** and provide actionable insights to reduce the risk.
- **Dataset Overview:** Analyzed **148,670 records** with key variables such as loan amount, property value, LTV, and credit score.
- **Methods Used:** Exploratory **Data Analysis**, Hypothesis Testing and descriptive statistics.

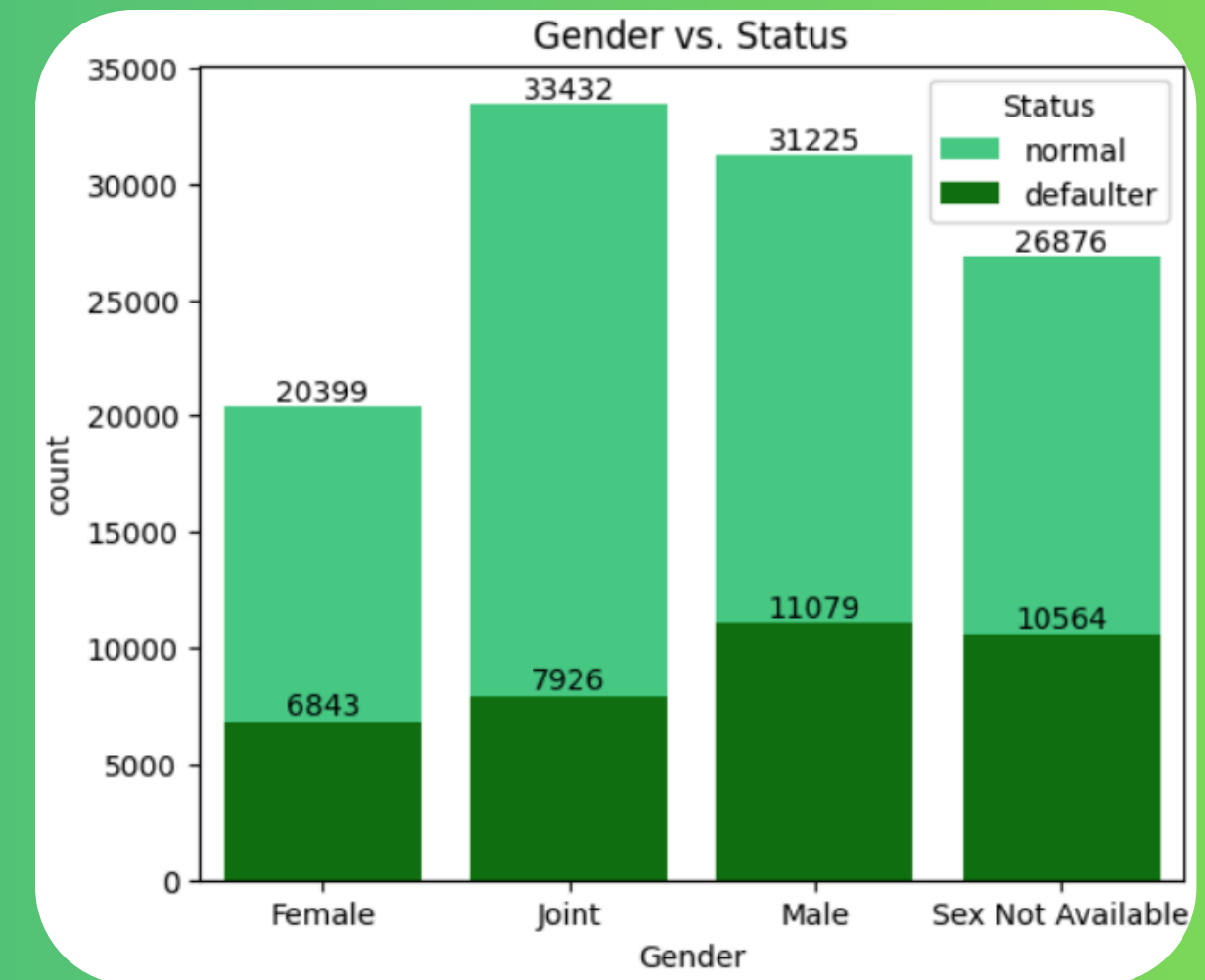
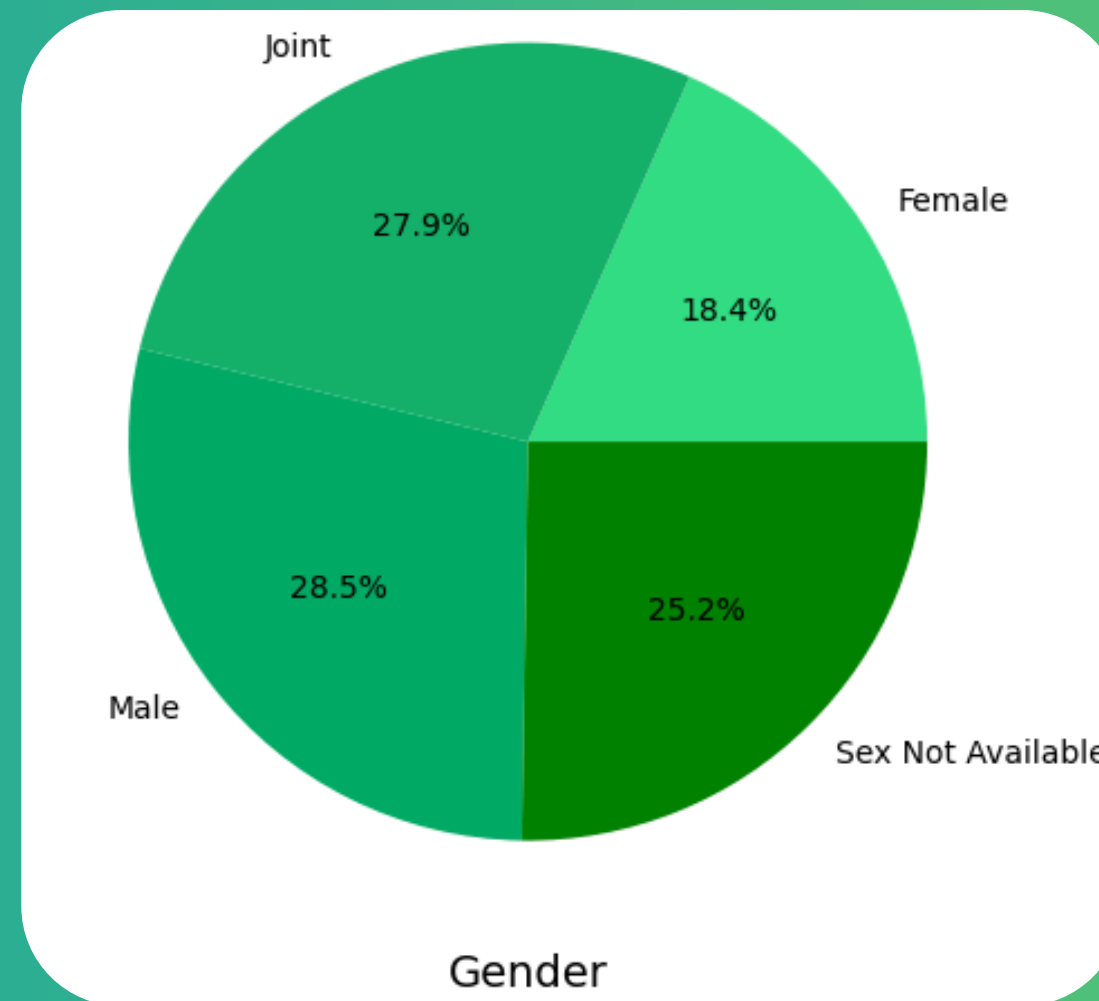
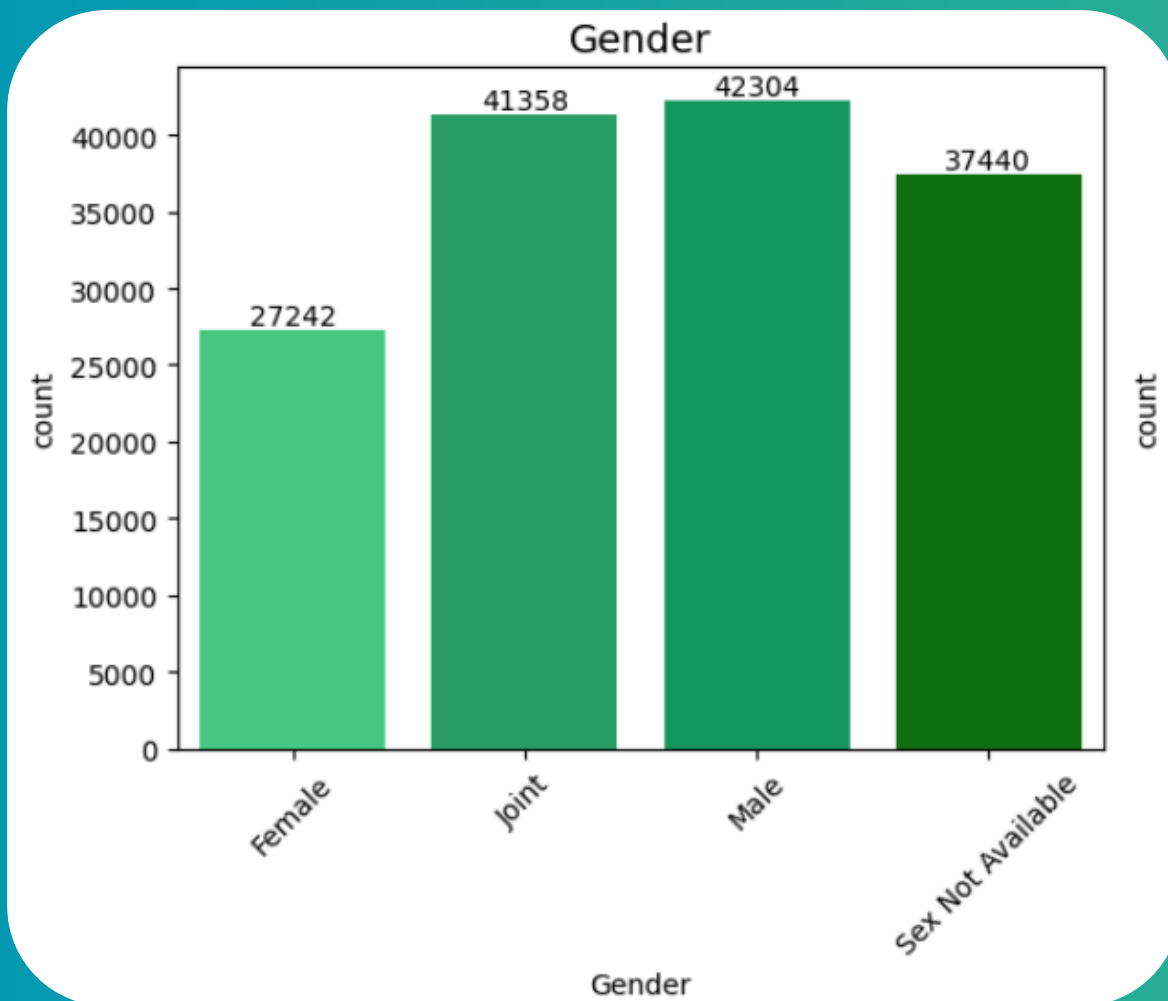


## LOAN LIMIT ANALYSIS

- The **majority** of loans fall under the **conforming category**, indicating a preference or higher eligibility for loans that meet standard criteria.
- **Higher Stability** in **Confirmed Loan Limits**.
- **Increased Risk** in **Non-Confirmed Loan Limits**

### Recommendation:

The significant number of defaults among non-confirmed borrowers highlights an opportunity for lenders to **improve their assessment procedures** and support mechanisms for these clients

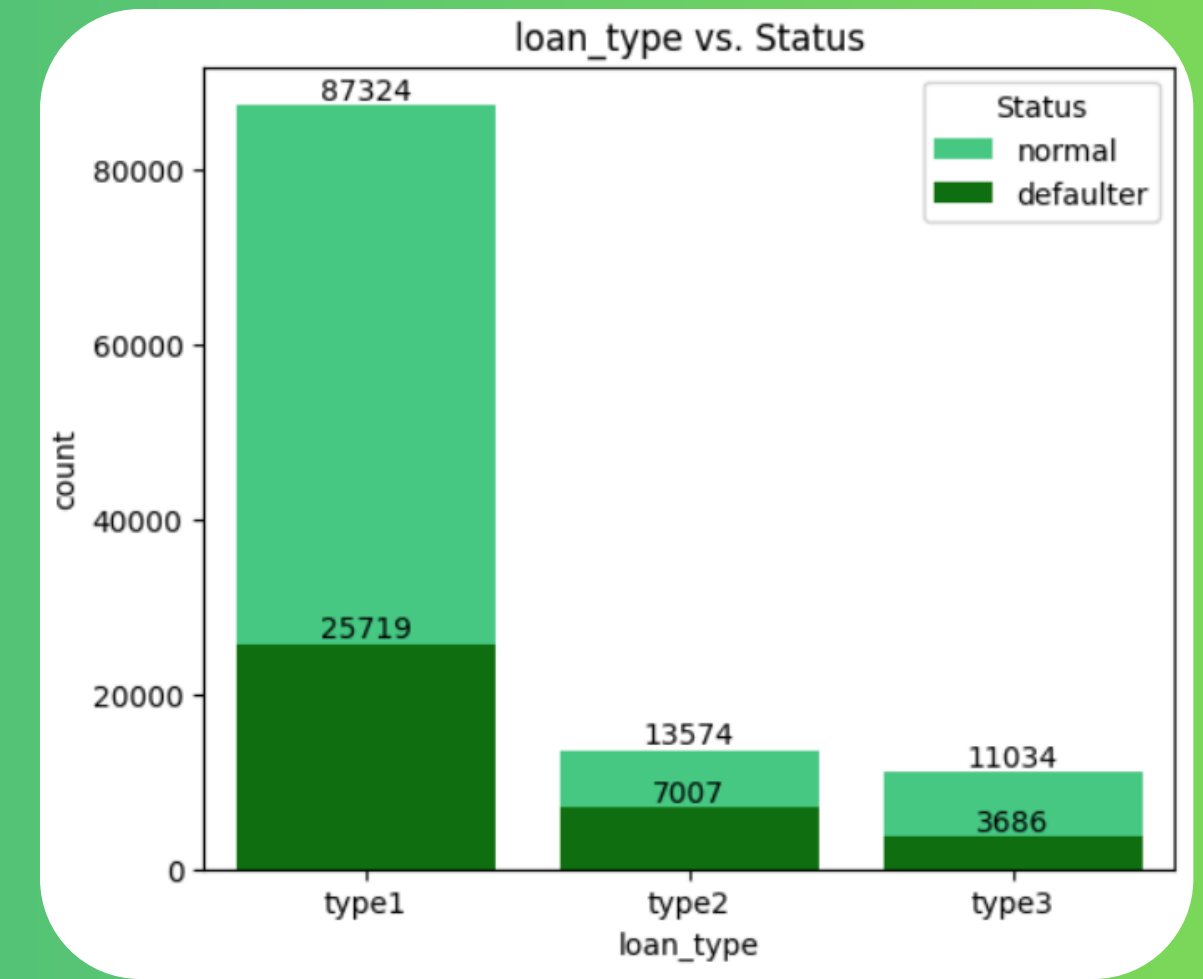
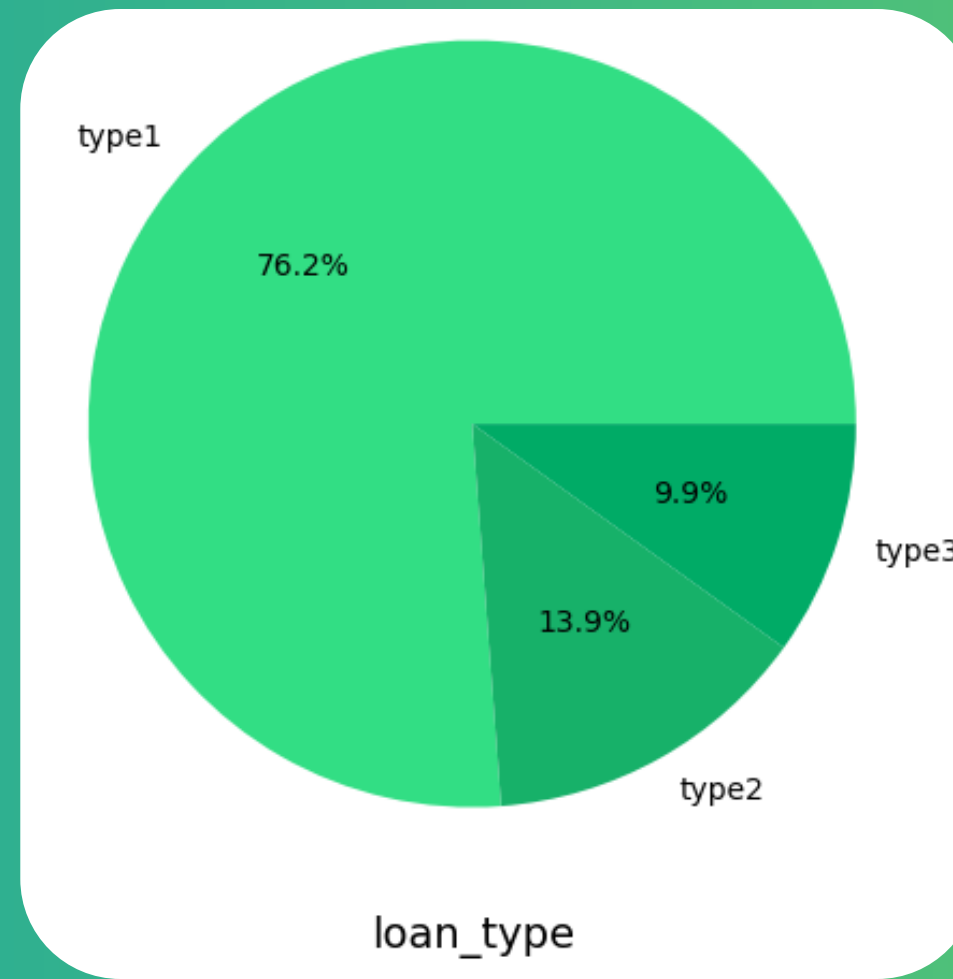
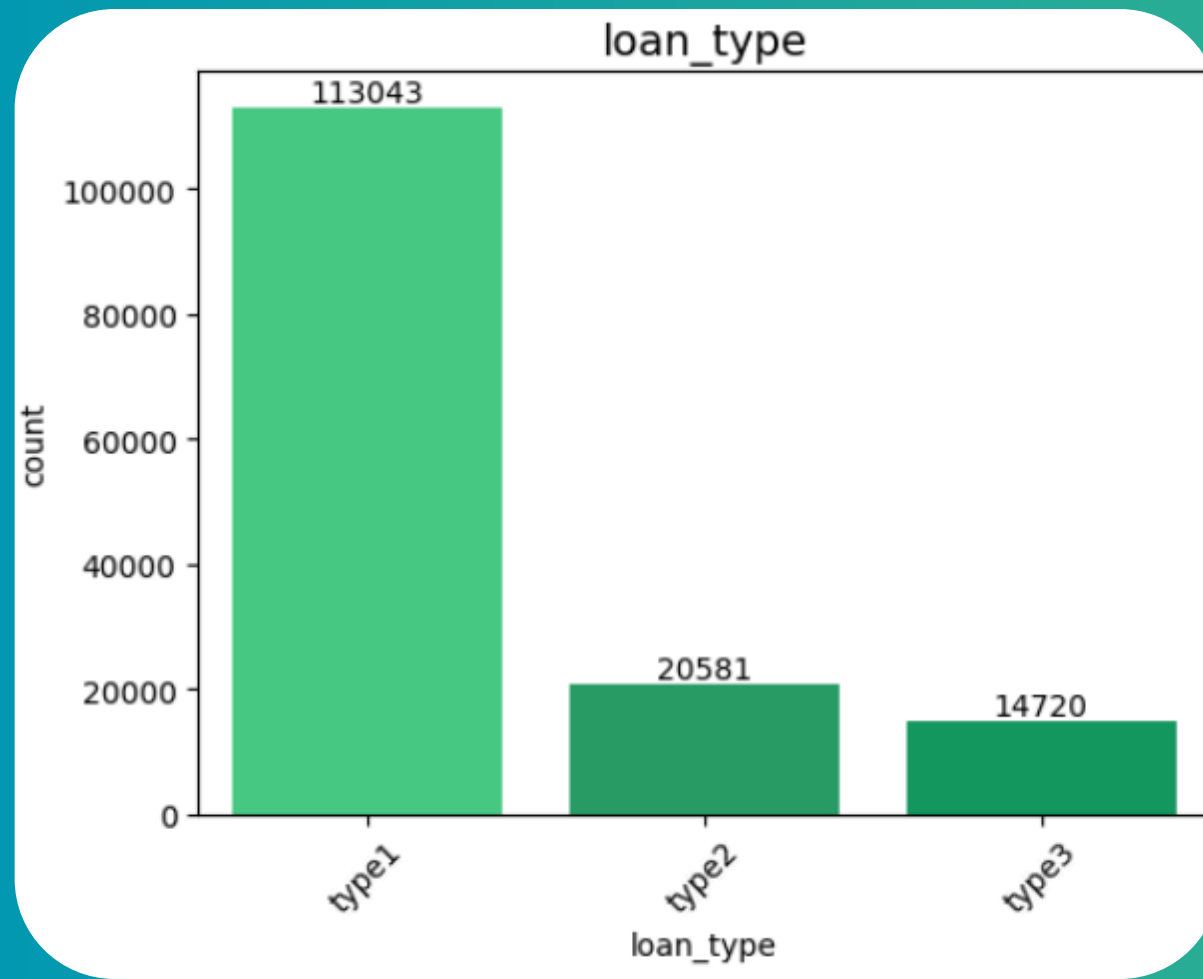


## GENDE ANALYSIS

- There is a **near-even split** between **male and joint applications**, with a **relatively high number of cases** where **gender information is unavailable**. Female applicants represent the smallest group.
- **Joint borrowers** show a relatively **lower default rate**.
- Borrowers with **unspecified gender** also exhibit a **high default rate**.

### Recommendation:

The significant number of defaults among non-confirmed borrowers highlights an opportunity for lenders to **improve their assessment procedures** and support mechanisms for these clients

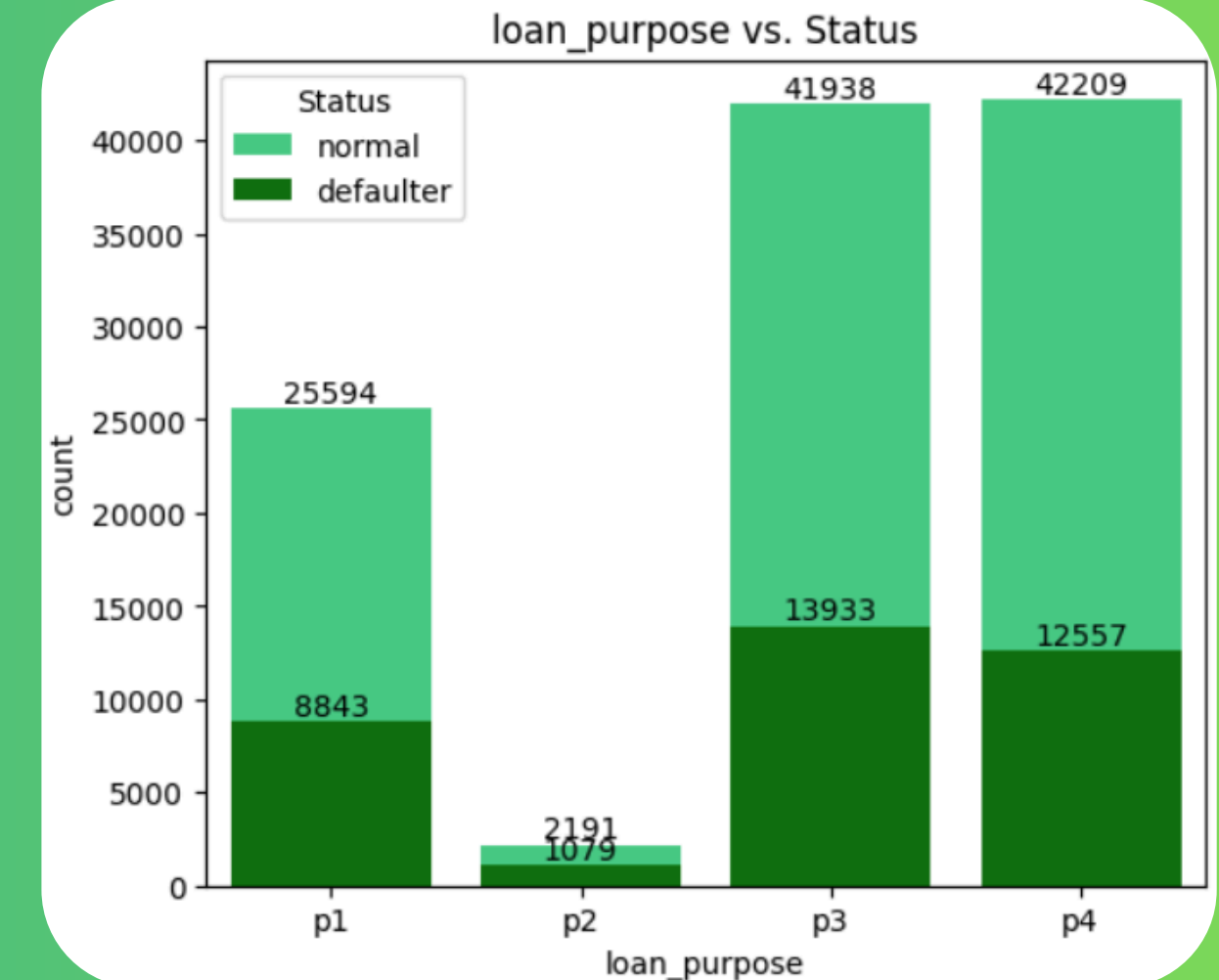
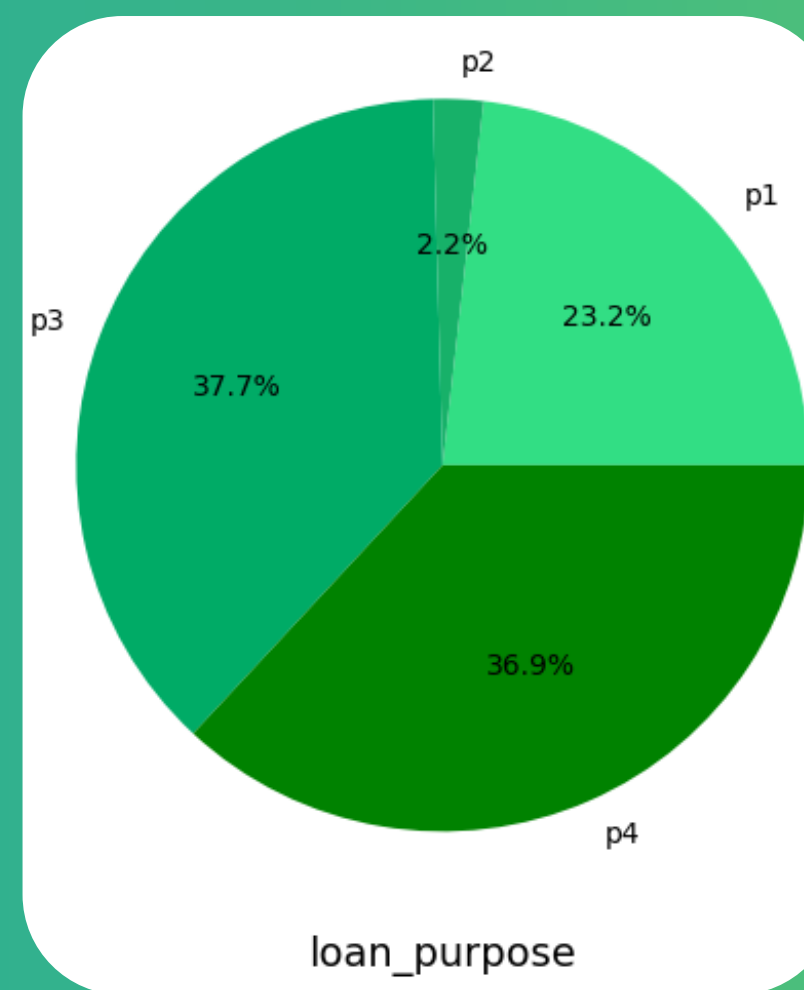
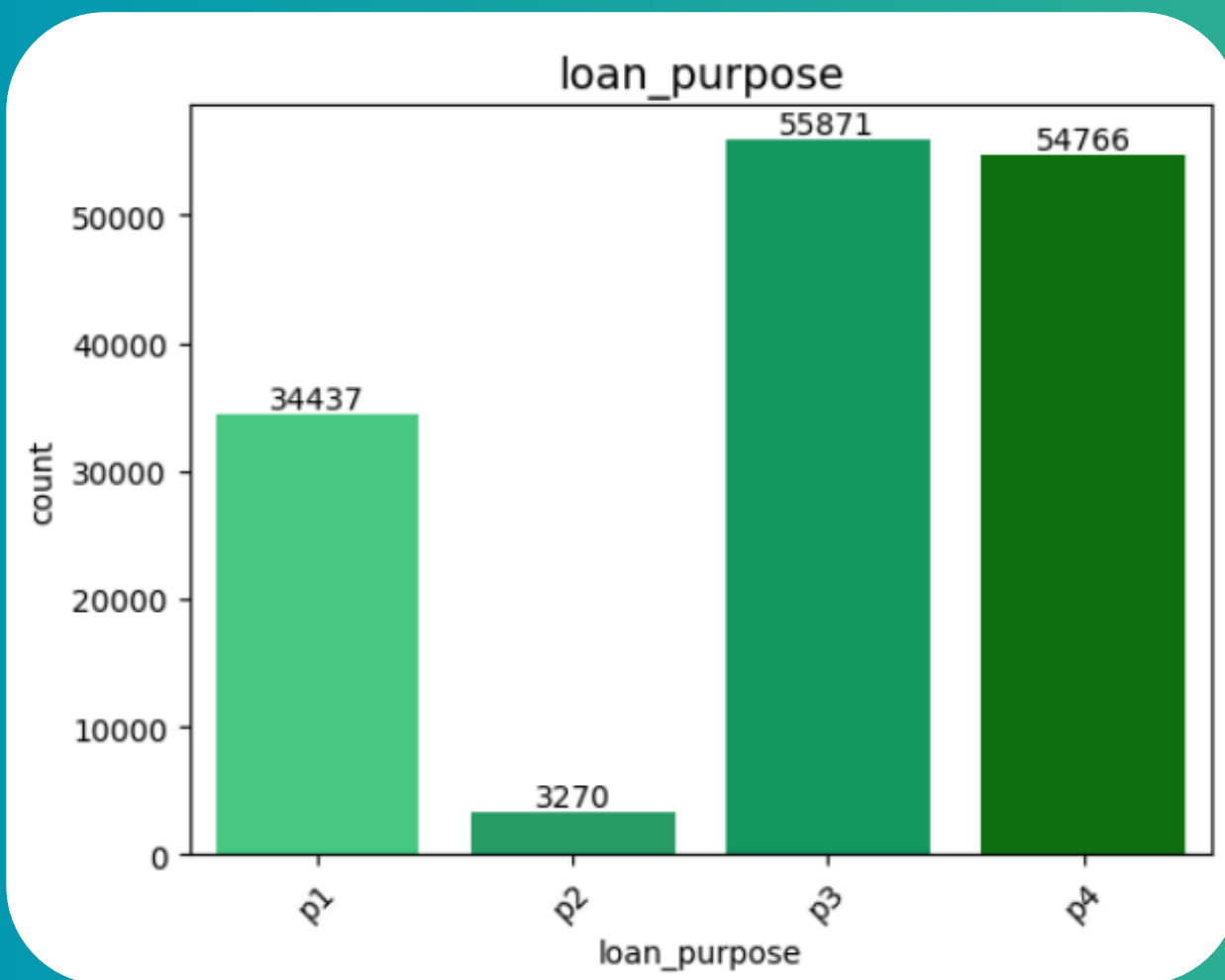


## LOAN TYPE ANALYSIS

- The **vast majority** of loans belong to loan **type 1**, showing a strong concentration in this category.
- Loan **Type 1** stands out as the most favorable in terms of **low defaults**
- Loan **Type 2** presents **significant risks** with a **high default rate**.

### Recommendation:

Revise the lending criteria and **conduct a thorough review** of the terms and borrower profiles for **Loan Type 2**, as it shows a **significantly higher default rate compared to other loan types**, to mitigate risk and improve overall loan performance.



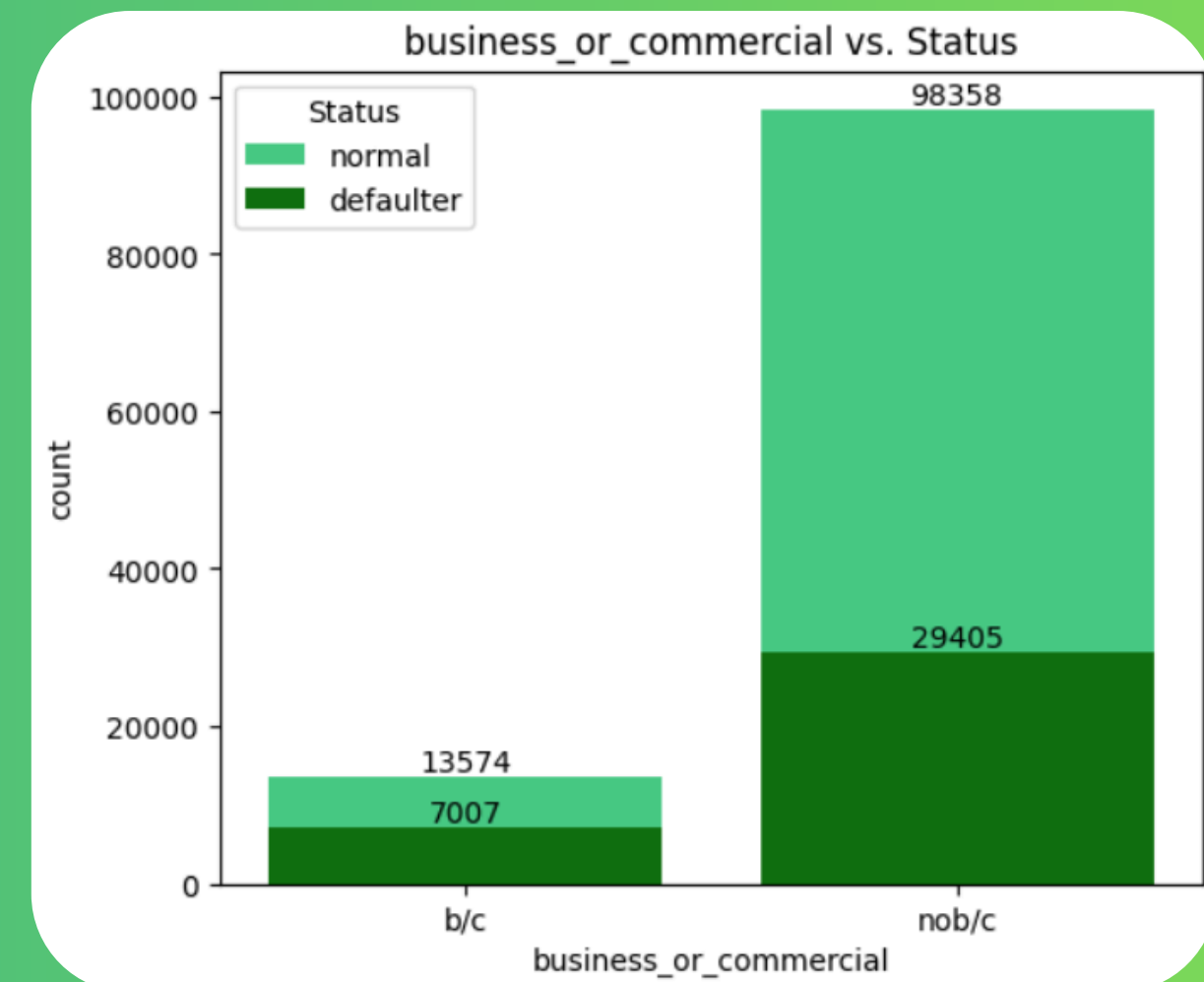
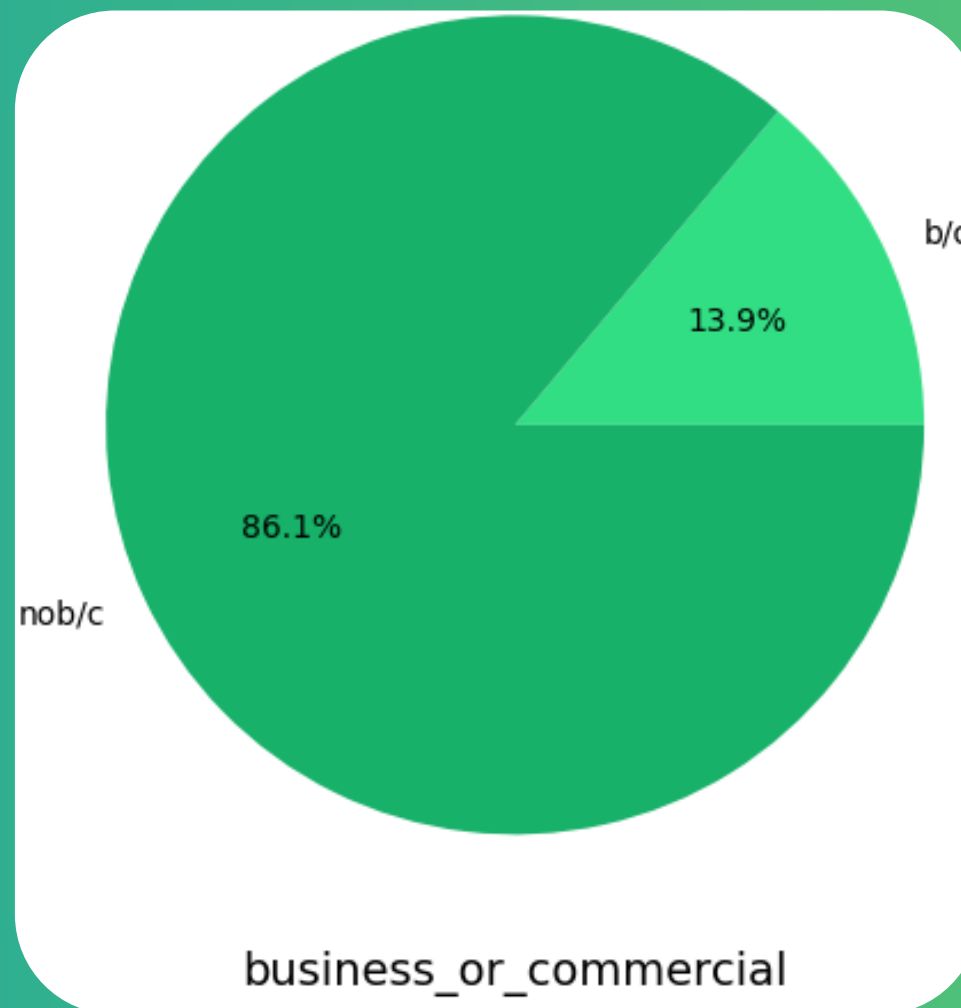
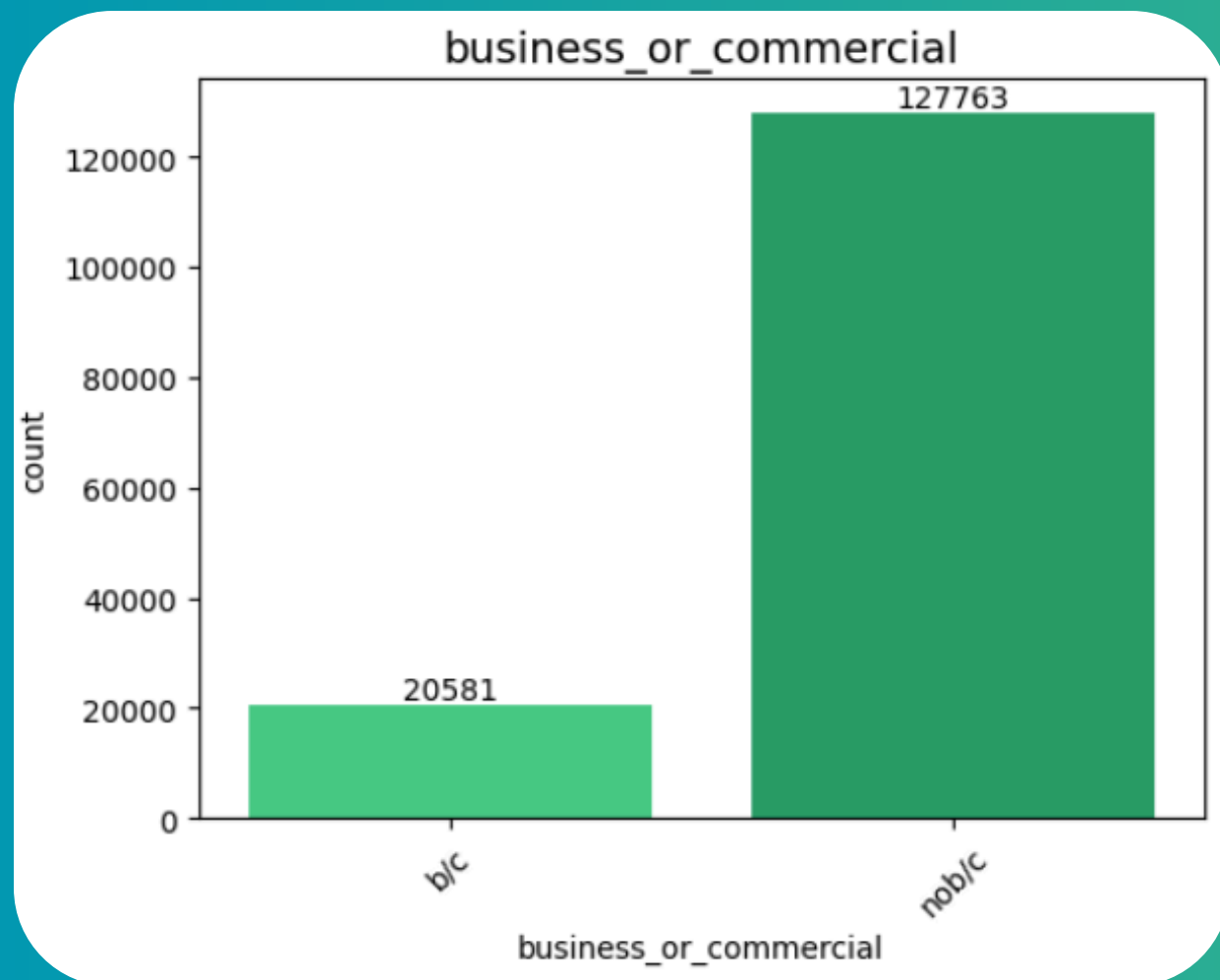
## LOAN PURPOSE ANALYSIS

- The **majority of loans** are associated with **purposes p3 and p4**, indicating a focus on certain use cases for loans.
- Loan **Purposes P1 and P2** show **high default rates**.
- Loan **Purpose P4** has the best performance with the **lowest default rate**.

### Recommendation:

**Conduct a detailed analysis** of Loan **Purpose P1** and Loan **Purpose P2**, as they exhibit high default rates, to **identify underlying risk factors** and develop targeted strategies to mitigate risks associated with these loan purposes.



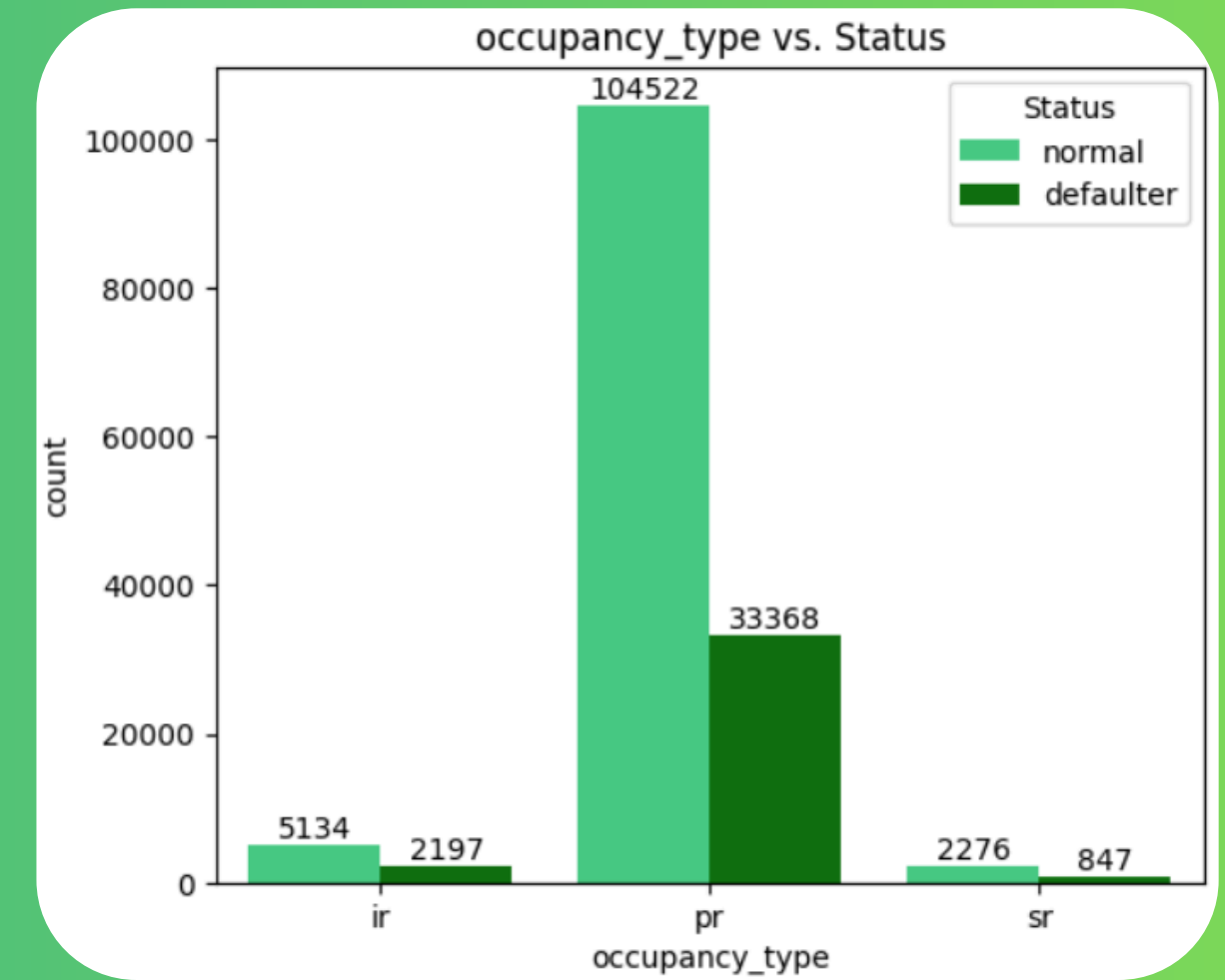
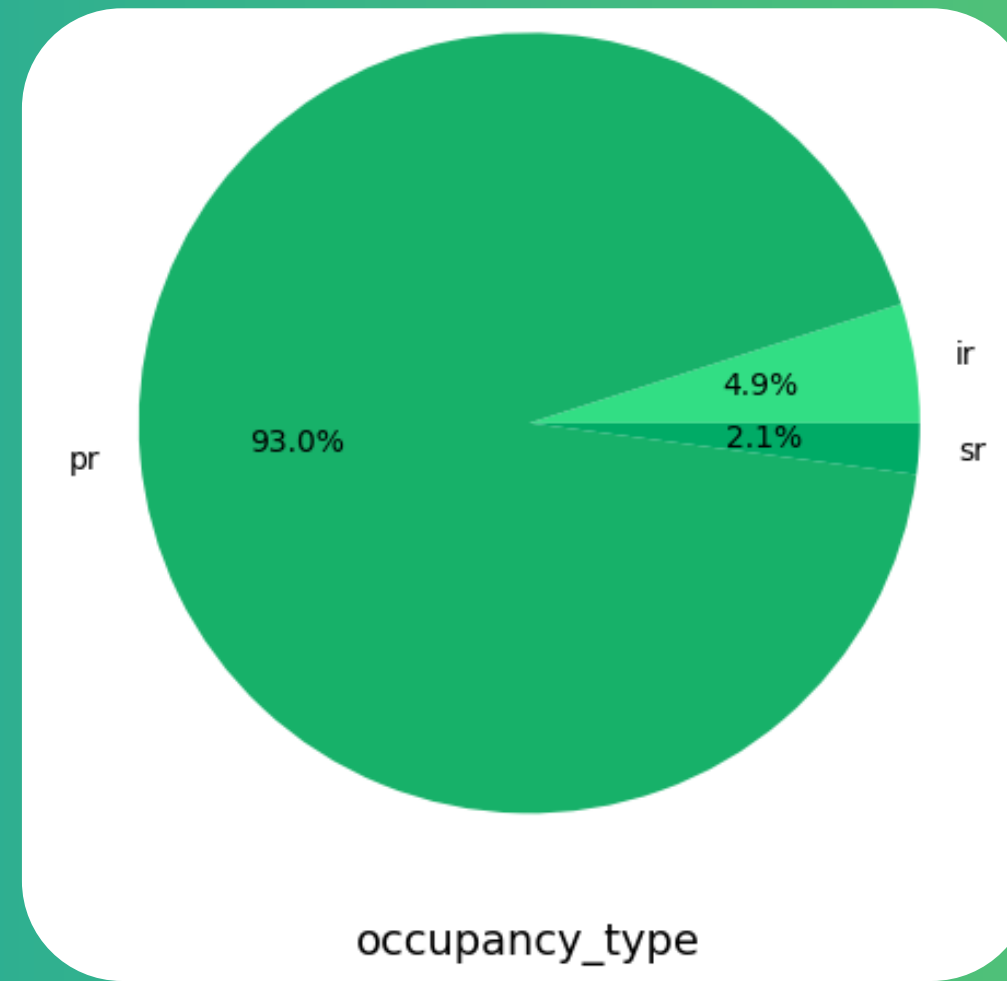
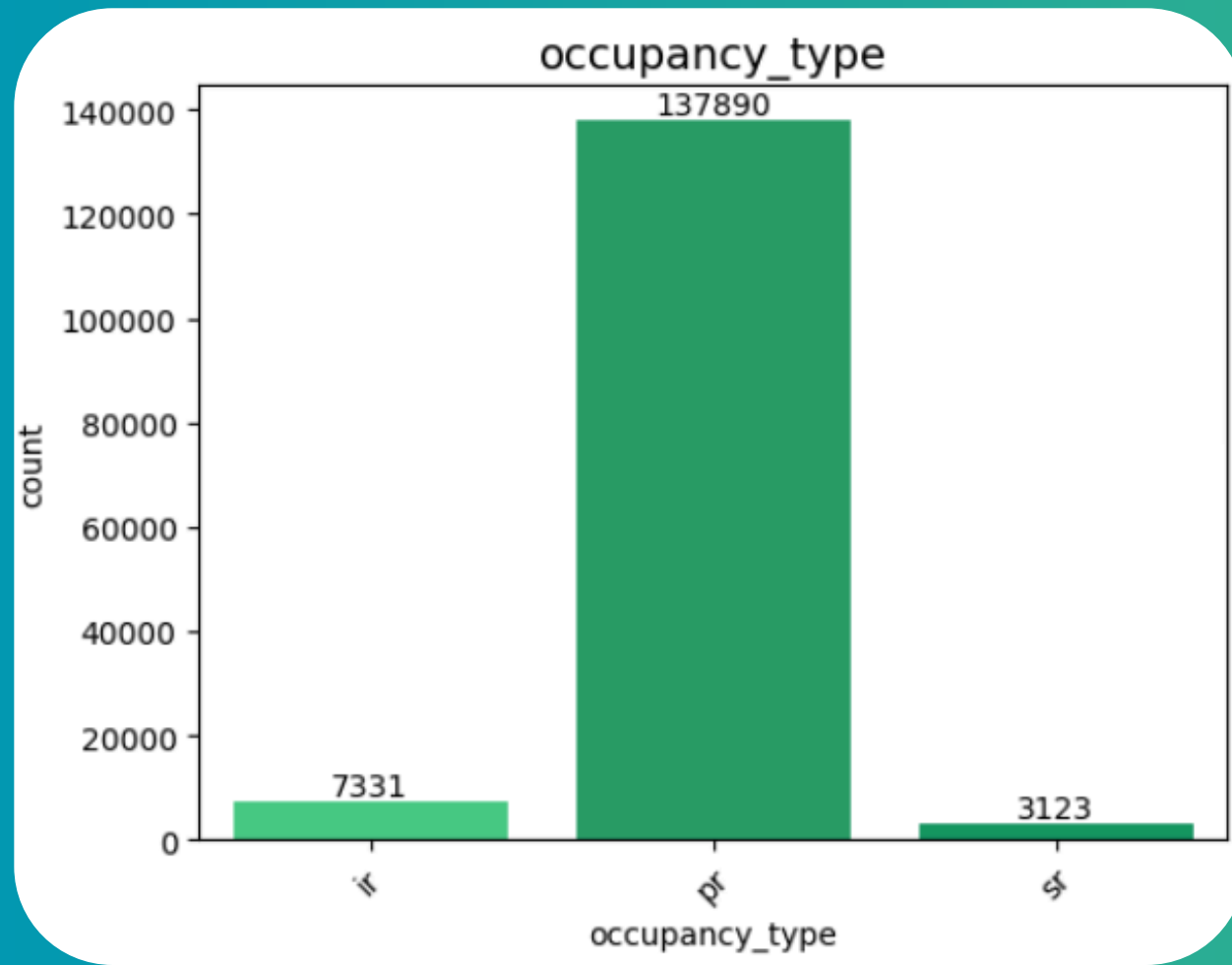


## BUSINESS (OR) COMMERCIAL LOAN ANALYSIS

- The **most loans are for personal or non-business purposes**, with a small proportion allocated for commercial needs.
- Loans given to **business/commercial establishments (b/c)** exhibit a **higher default rate**.
- **Non-commercial or personal loans (nob/c)** demonstrate stronger performance, with a higher proportion of normal loans and a **lower default rate**.

### Recommendation:

**Enhance risk assessment protocols for business/commercial loans** due to their higher default rate, and consider implementing more stringent lending criteria or providing additional support to borrowers in this category to reduce default risks.

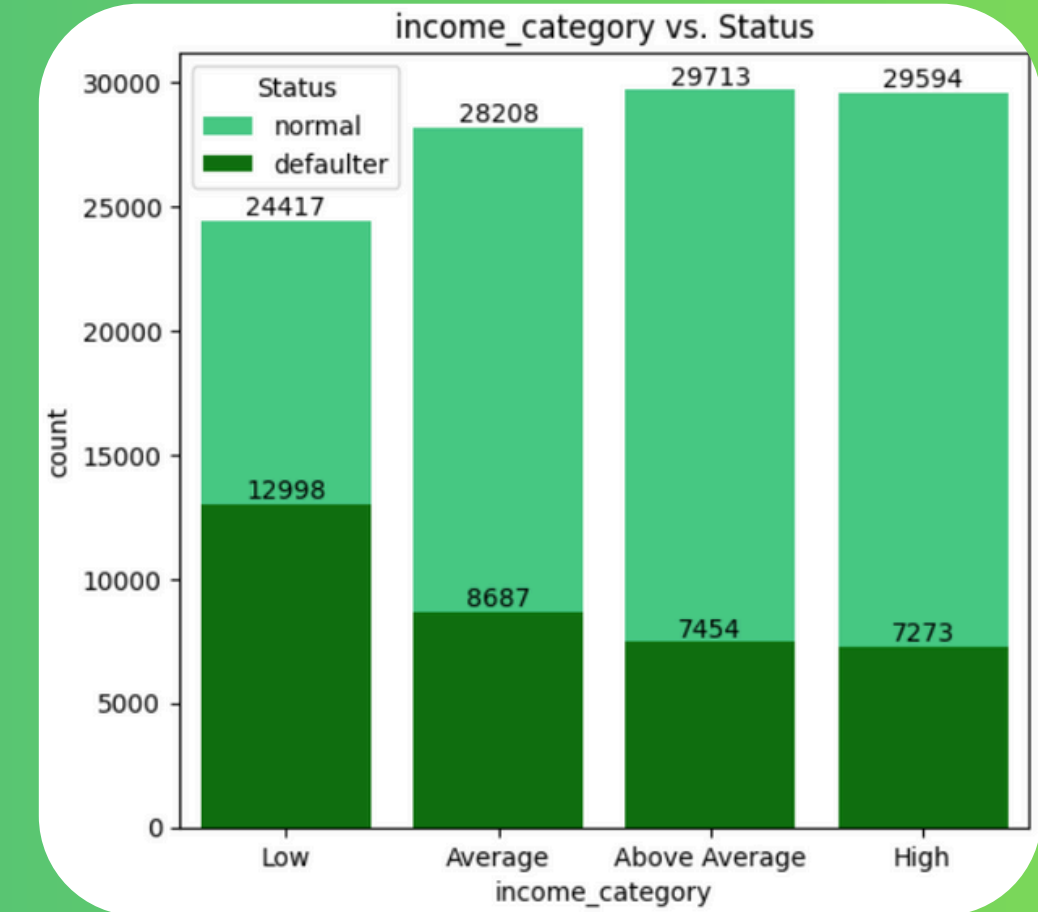
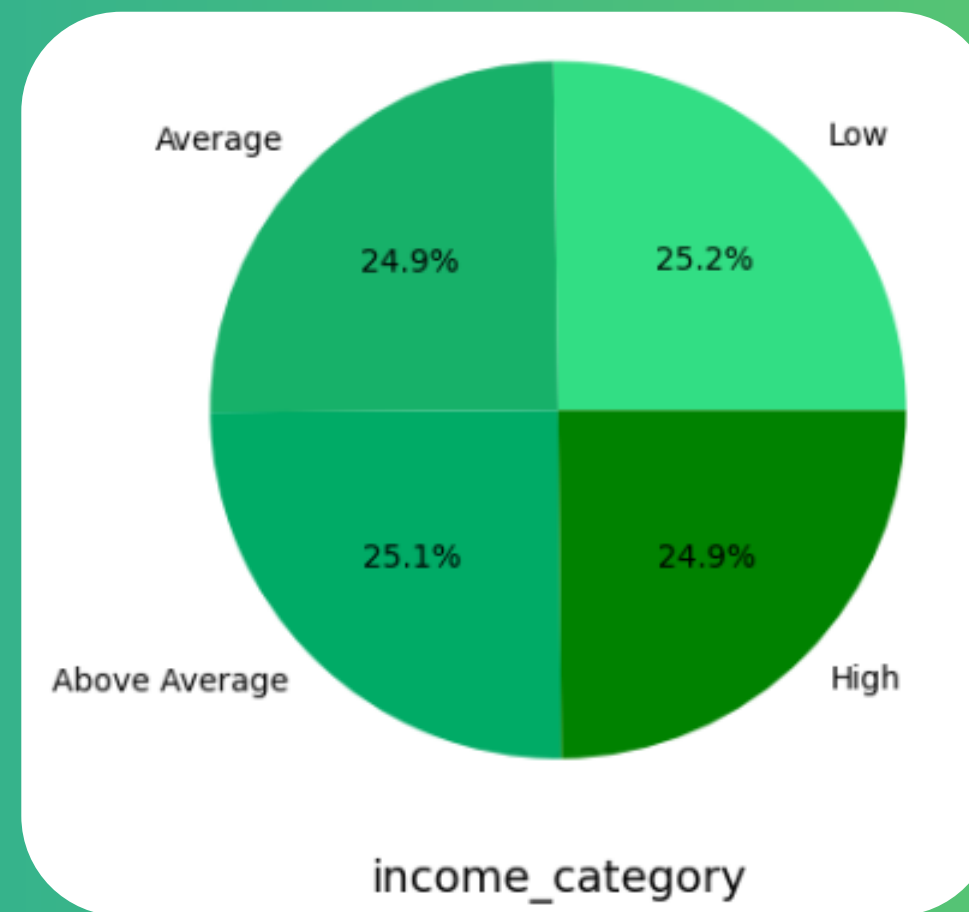
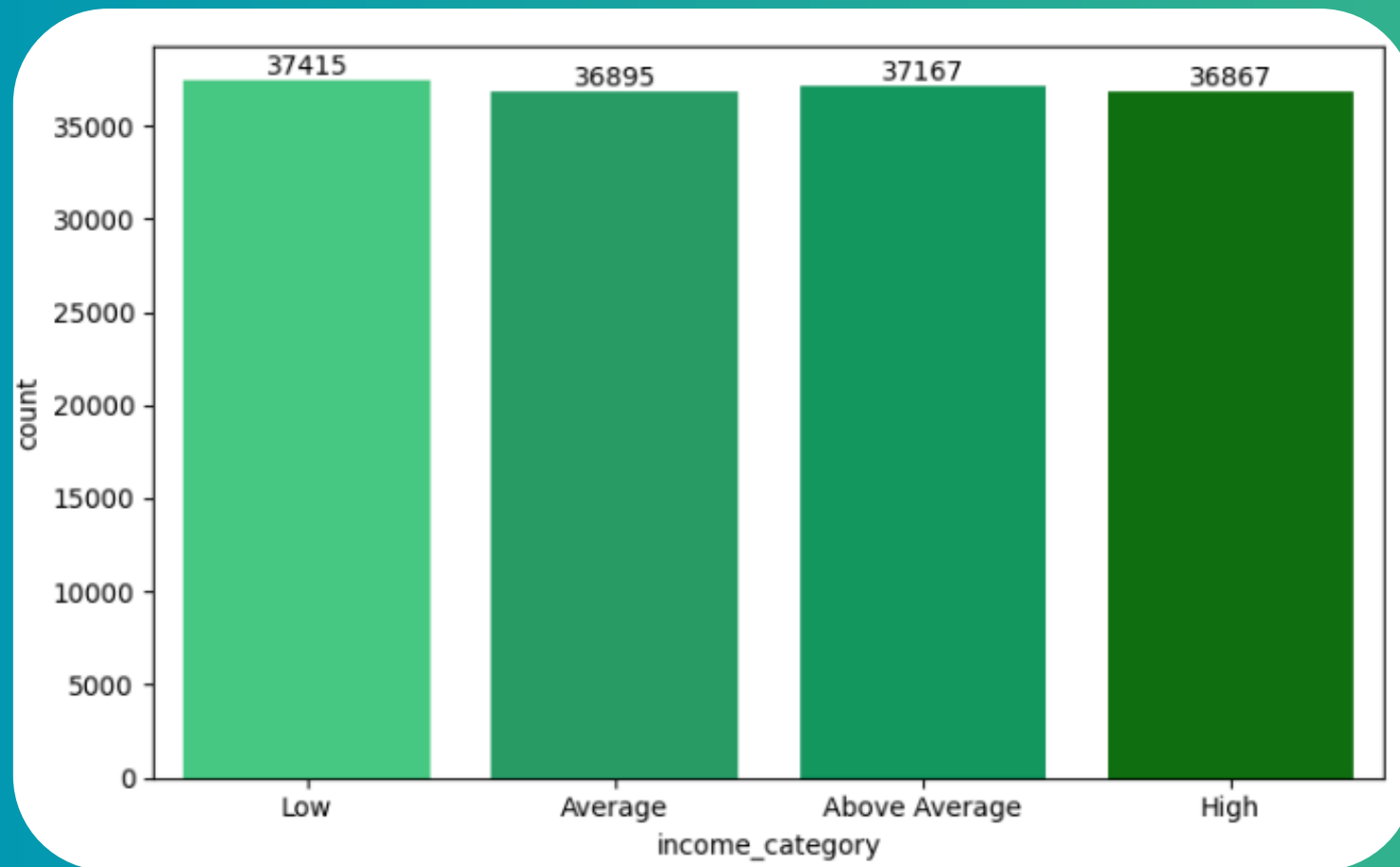


## OCCUPANCY TYPE ANALYSIS

- The **vast majority of loans** are tied to **primary residences**, with investment and secondary residences making up only a small fraction.
- **Higher Default** Rates for **Investment Residences**.
- Stability in Secondary Residences

### Recommendation:

**Implement stricter credit assessments and risk management strategies** for **investment residences** (occupancy type 'ir') to mitigate the high default rates, while maintaining supportive measures for primary and secondary residence borrowers to ensure overall portfolio stability.

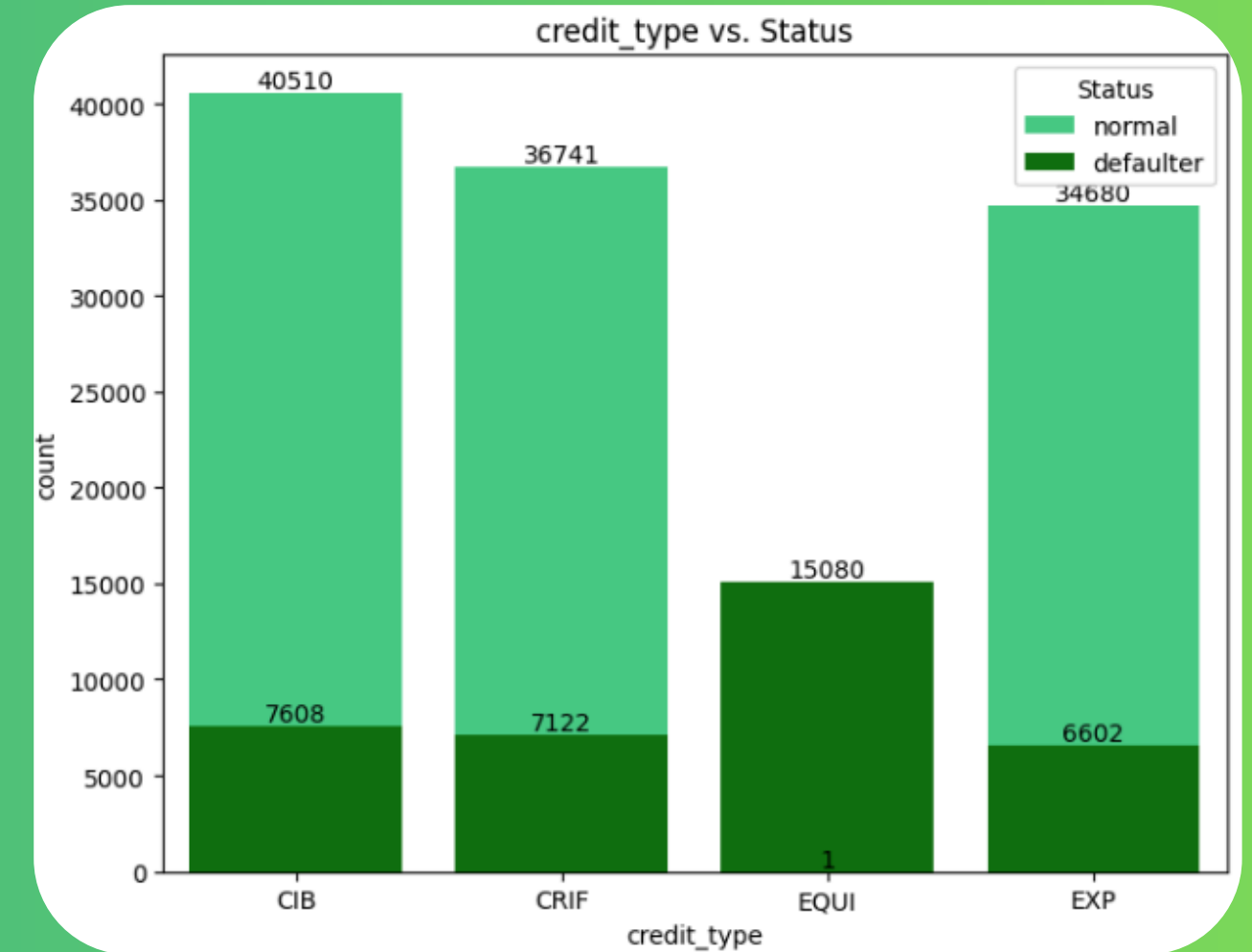
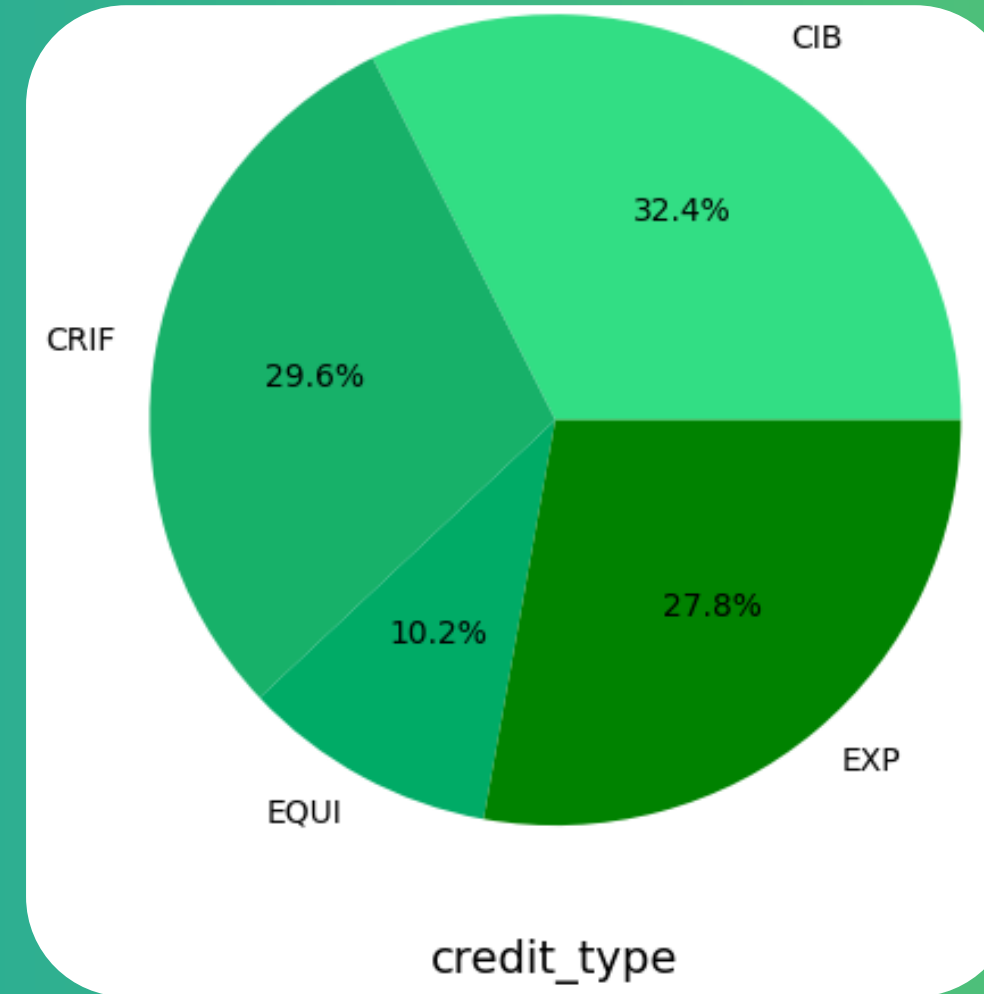
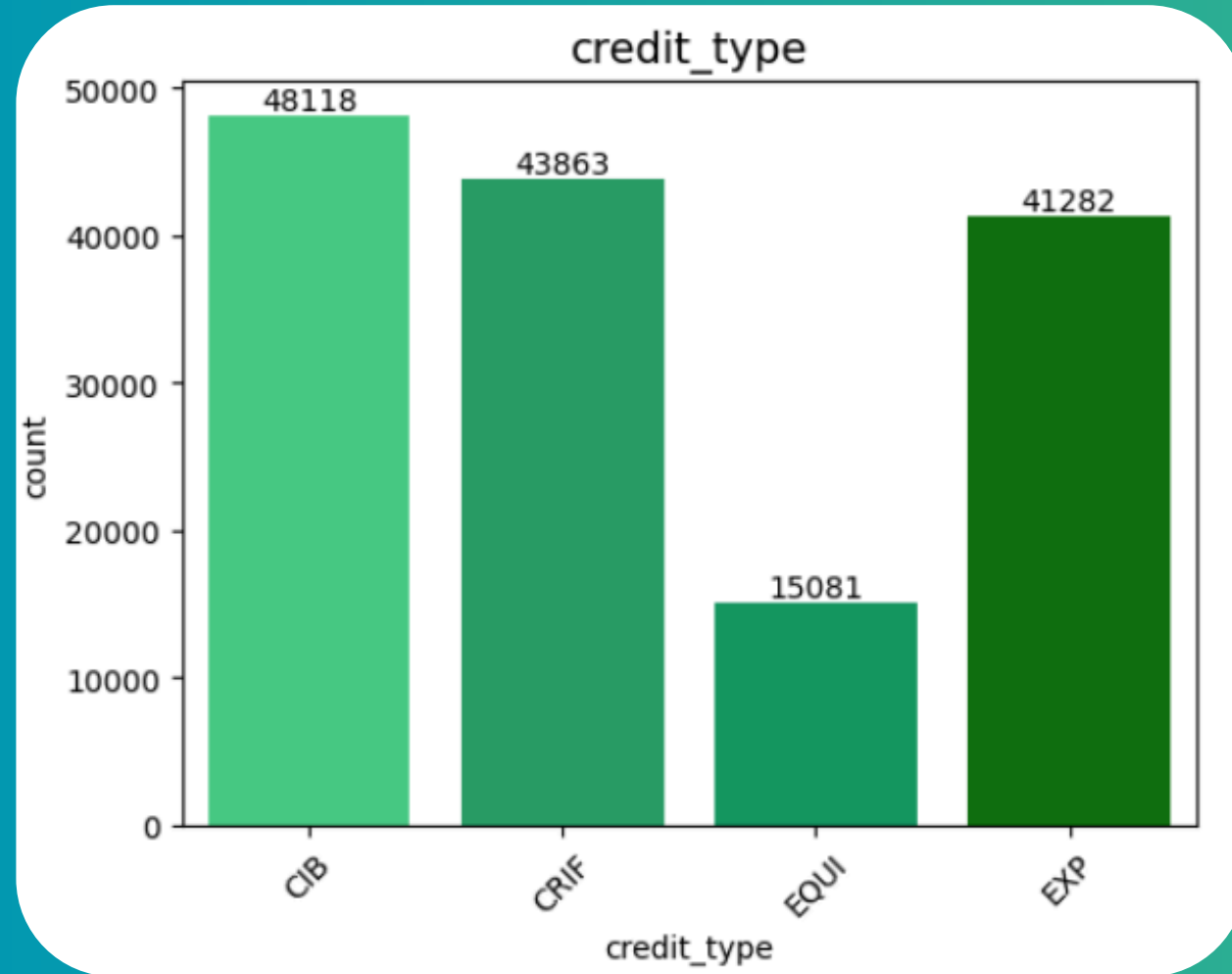


## INCOME CATEGORY ANALYSIS

- The **counts of borrowers** are relatively **balanced across the four income categories**
- The **Low income category** having the **highest number of borrowers (37,415)**, followed closely by Above Average (37,167), Average (36,895), and High (36,867)
- As **income increases from Low to High, the number of defaulters decreases**. The ratio of normal to defaulter improves, showing that **higher income groups tend to default less frequently**.

### Recommendation:

**Implement targeted financial assistance and risk management strategies for low-income borrowers,** as they exhibit a slightly higher default rate compared to higher-income groups

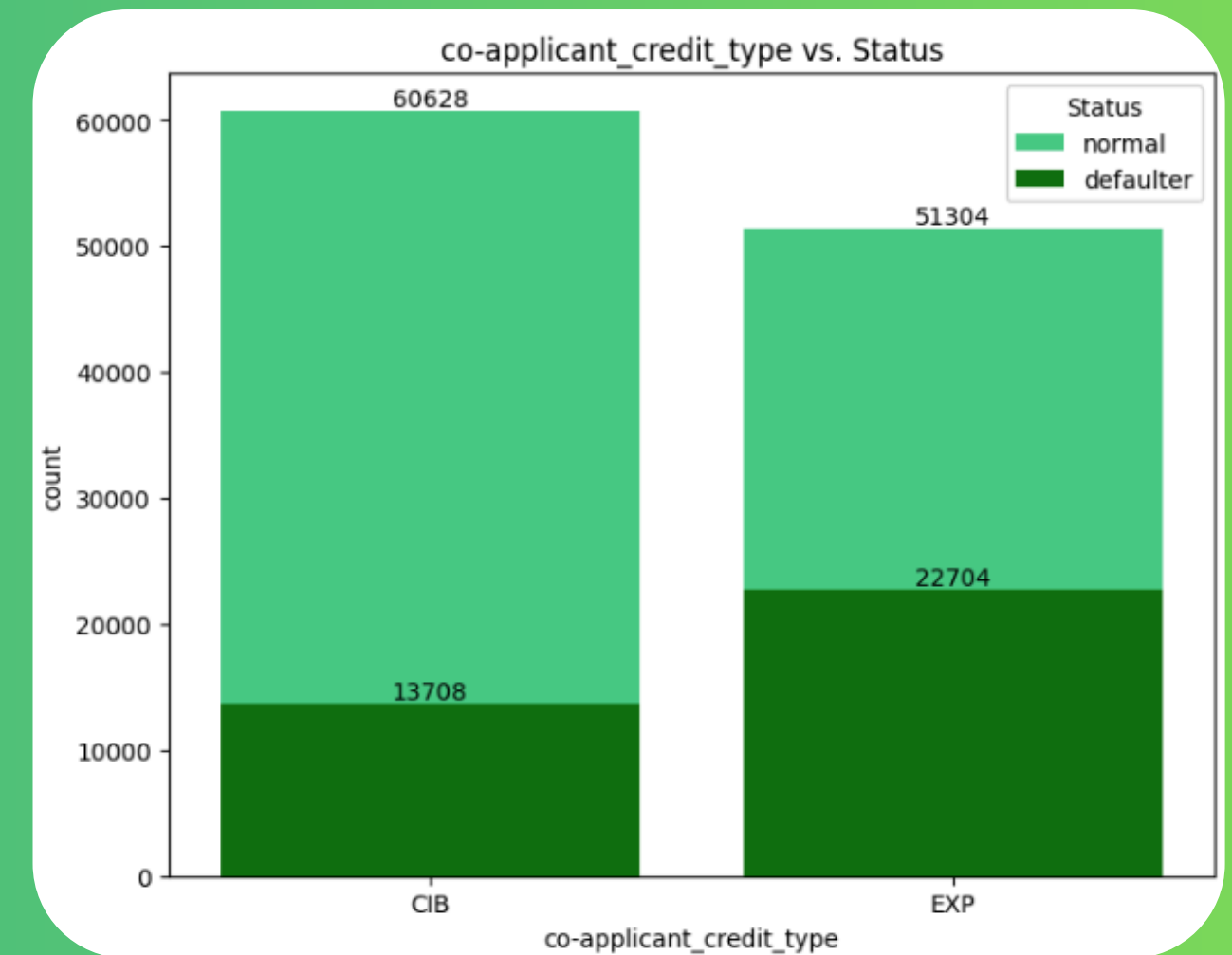
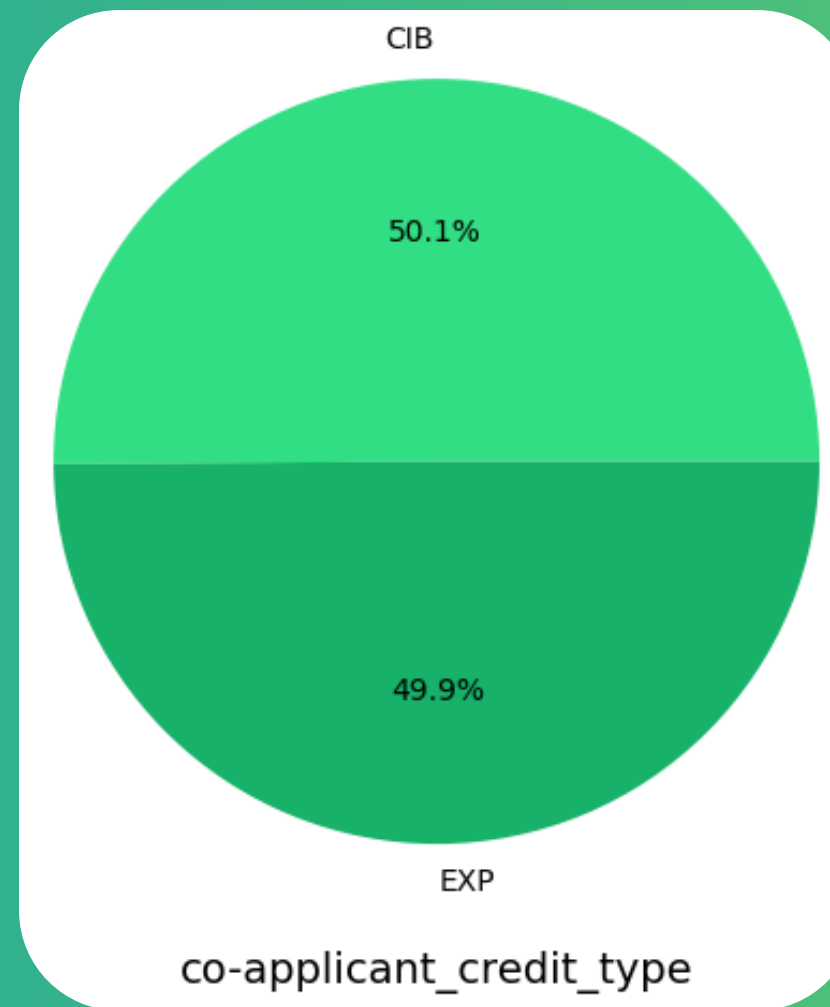
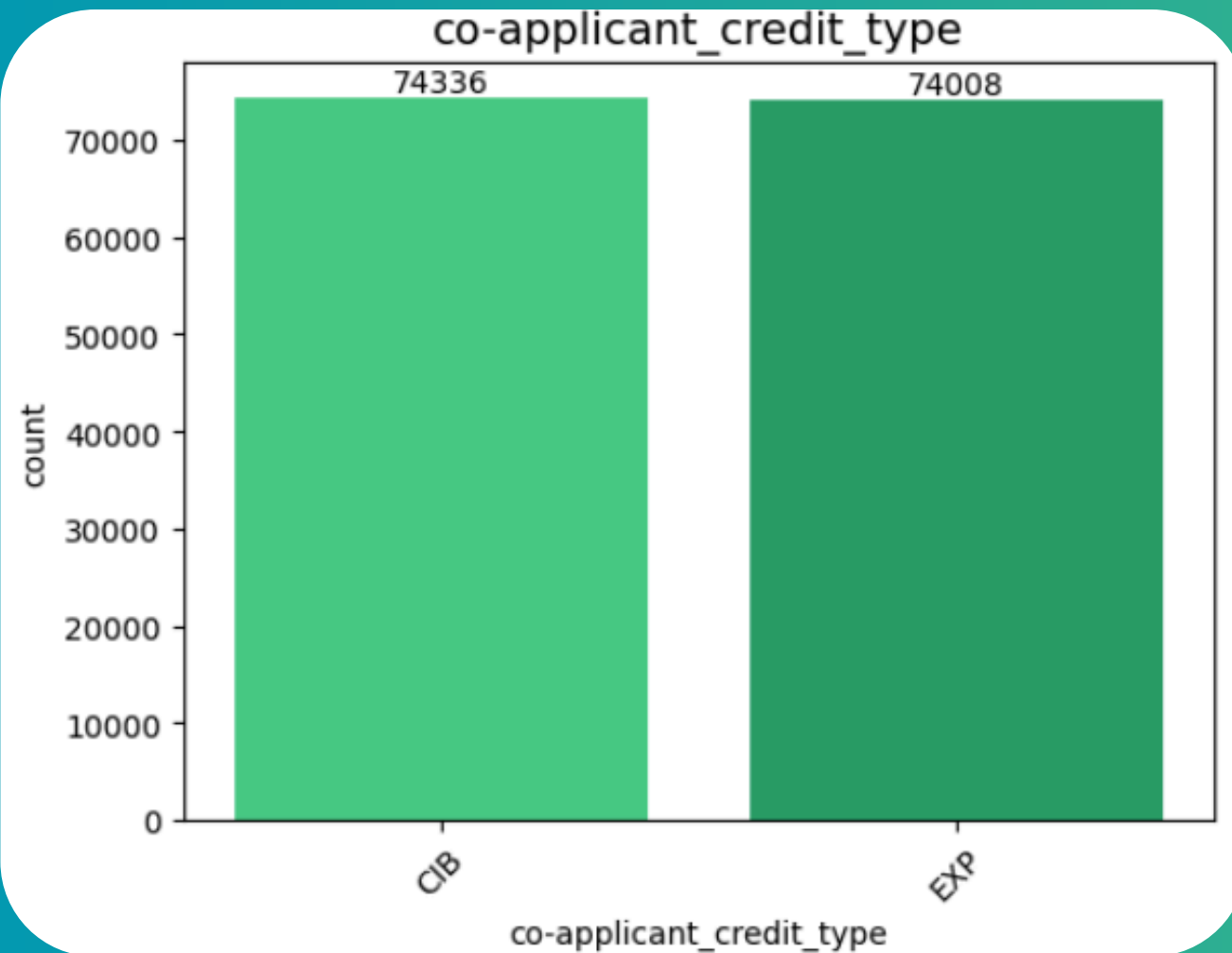


## CREDIT TYPE ANALYSIS

- The distribution of **credit types** is **somewhat balanced**, with **CIB leading**, but a significant portion of applicants also use CRIF and EXP.
- **CIB credit type** has a relatively strong track record with the **lowest default rate among credit types**.
- The **EQUI credit type** has only **1 normal borrower and 15,080 defaulters**. This suggests that EQUI credit is predominantly linked to defaults and poses a **significant risk to lenders**.

### Recommendation:

**Tighten lending criteria for EQUI credit types** to mitigate the overwhelming risk of defaults



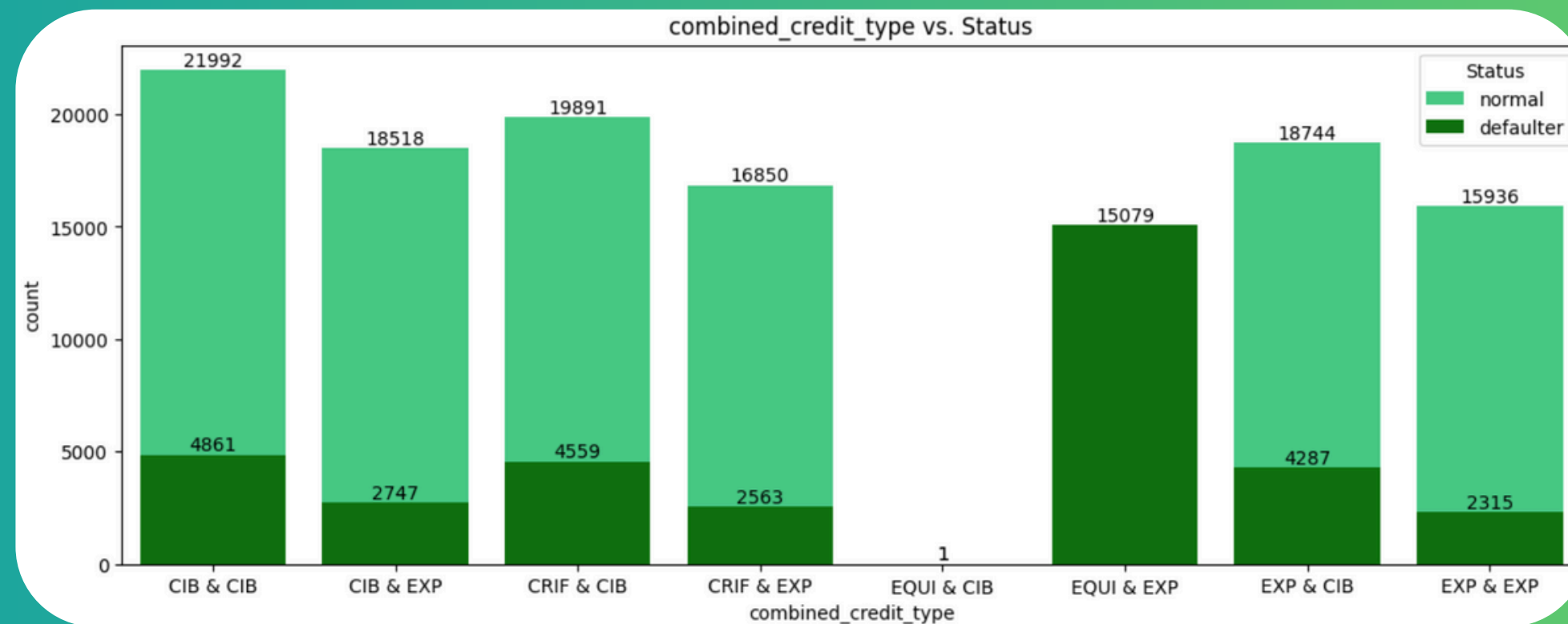
## COAPPLICANT CREDIT TYPE ANALYSIS

- There is **no significant preference for one credit type** over the other when it comes to co-applicants.
- Loans with **CIB as the co-applicant credit type** show a good proportion of **normal status**, although slightly higher default risk compared to primary credit types.
- Loans with **EXP as the co-applicant credit type** have a significantly **higher default rate**, suggesting co-applicants with EXP credit may represent a higher risk.

### Recommendation:

**Strengthen co-applicant assessments**, especially when EXP is the co-applicant credit type, as this significantly increases the likelihood of default.



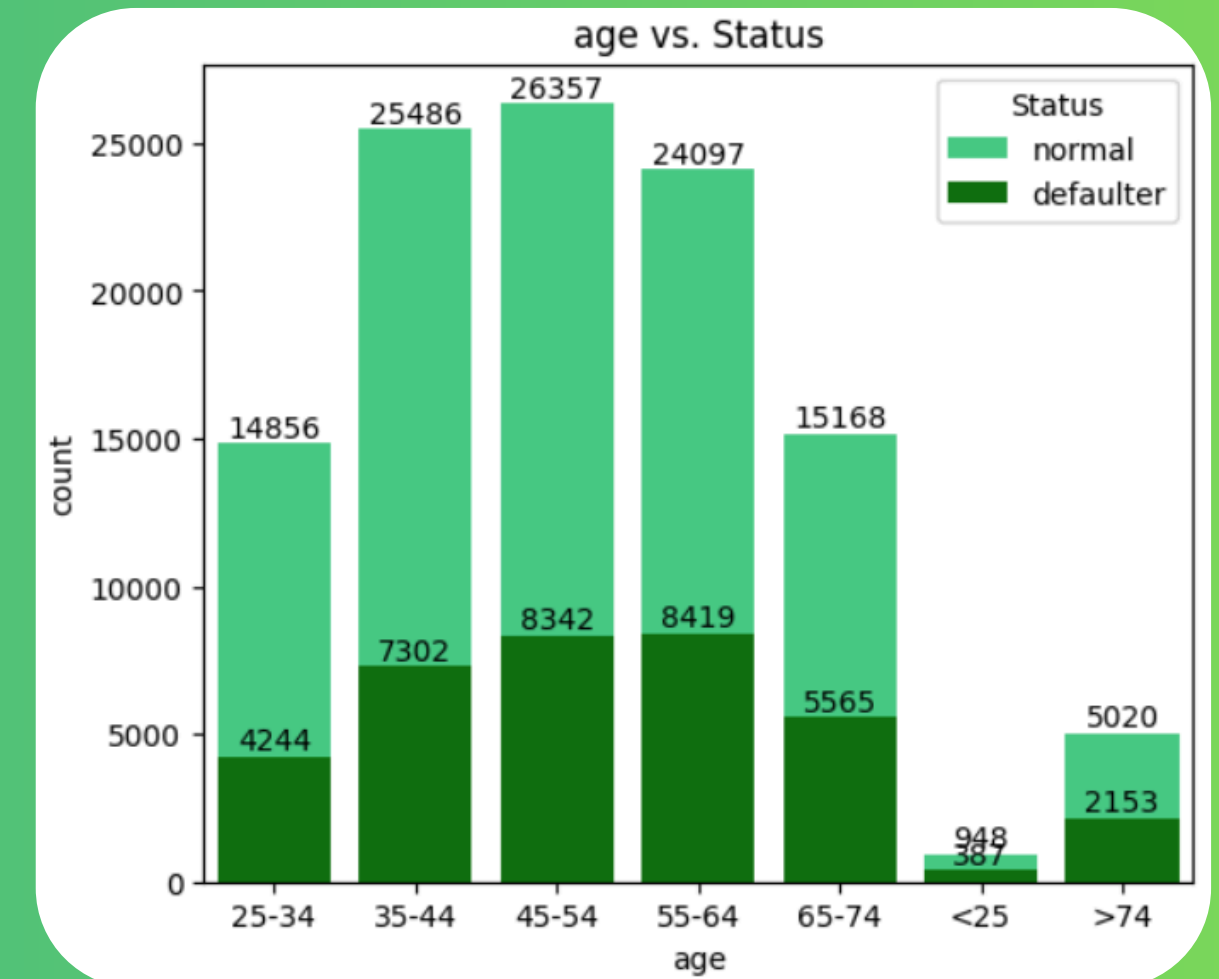
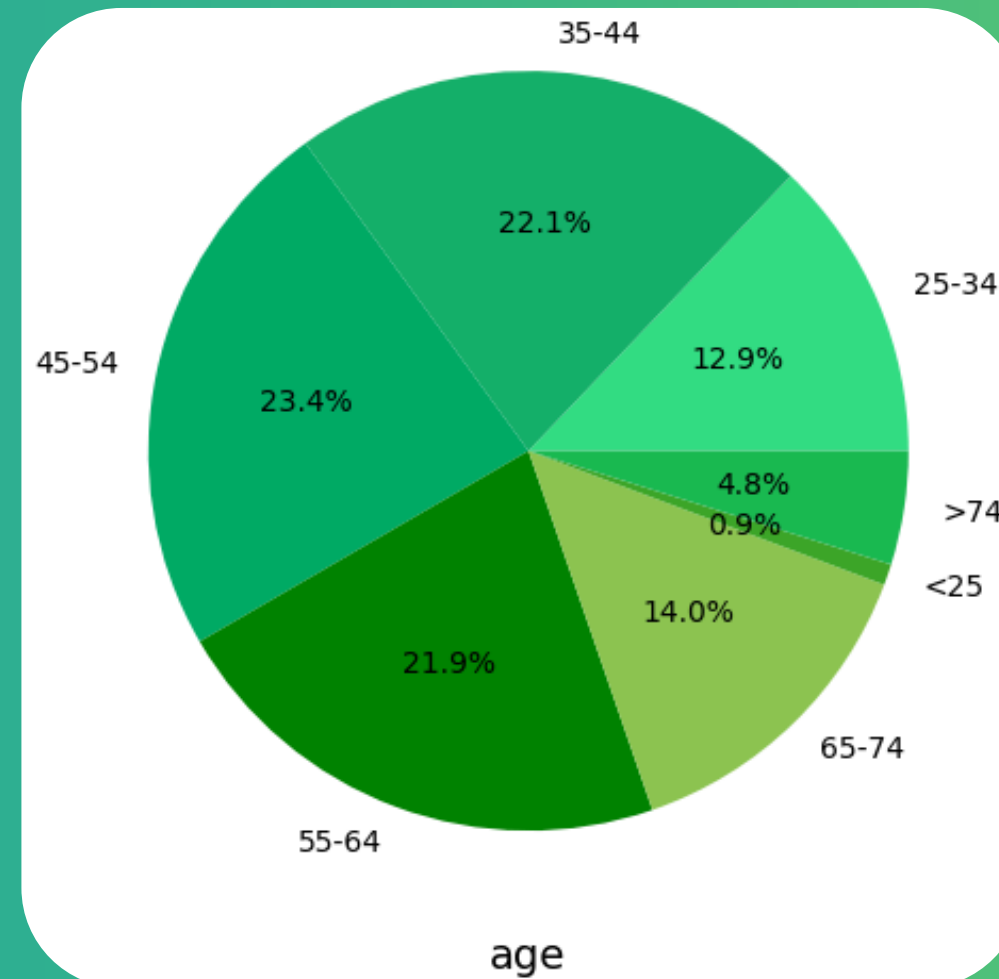
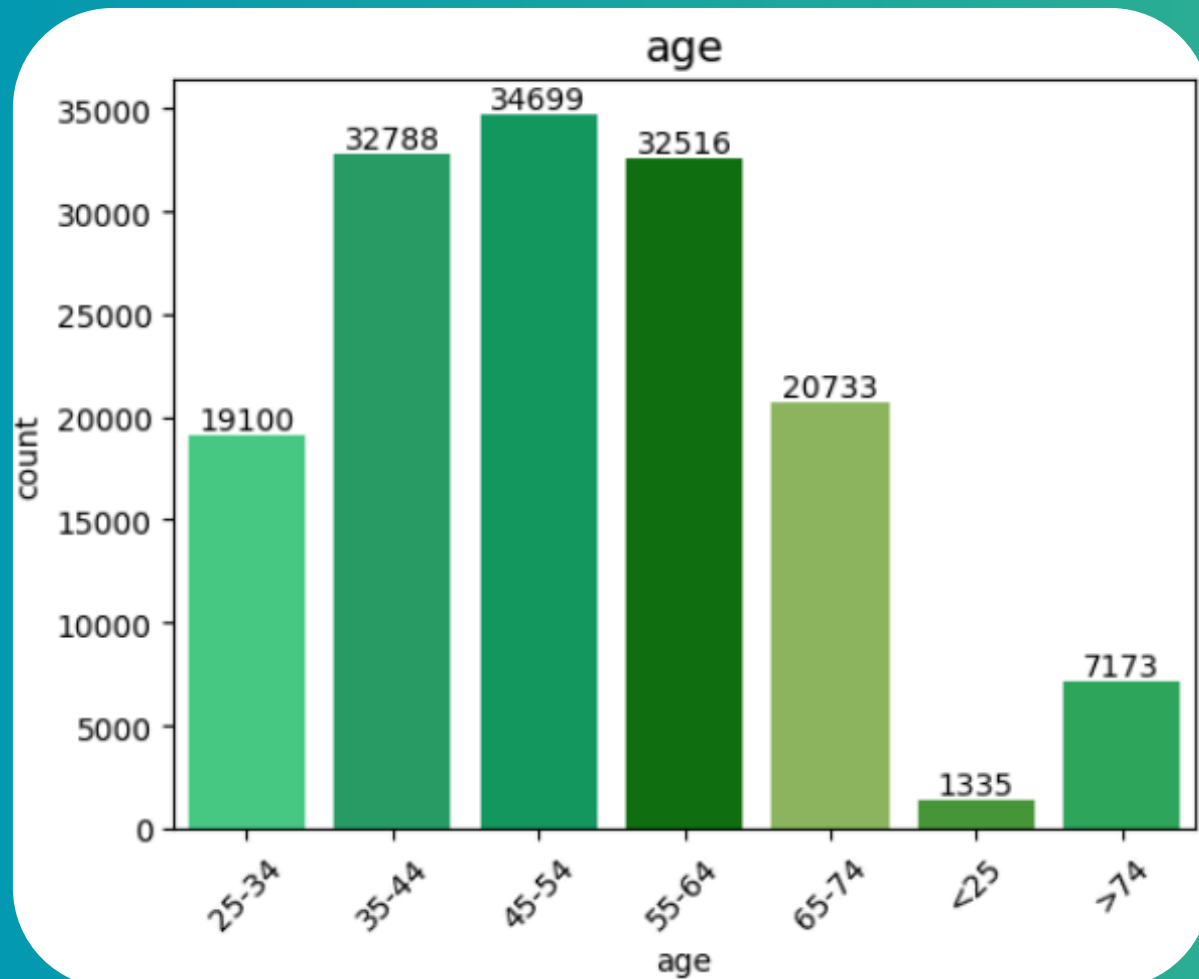


## COMBINED CREDIT TYPE ANALYSIS

- **CIB as a co-applicant credit type** consistently **reduces the default rate**, making it a more reliable choice across different primary credit types (CIB, CRIF, EXP).
- **In most cases, EXP as a co-applicant** results in a higher success rate, showing **lower default rates**.
- **EQUI is highly risky**, especially when **paired with EXP as the co-applicant credit type**, which leads to **100% default rates**. This combination should be avoided.
- The **combination of EXP as both credit type and co-applicant** performs surprisingly well, with a **low default rate (12.86%)**. This shows that EXP credit type can be reliable when both borrower and co-applicant have the same credit type.

### Recommendations:

**Avoid lending to borrowers with EQUI credit type**, especially when **paired with EXP as the co-applicant credit type**, as this combination has an extremely high default risk.

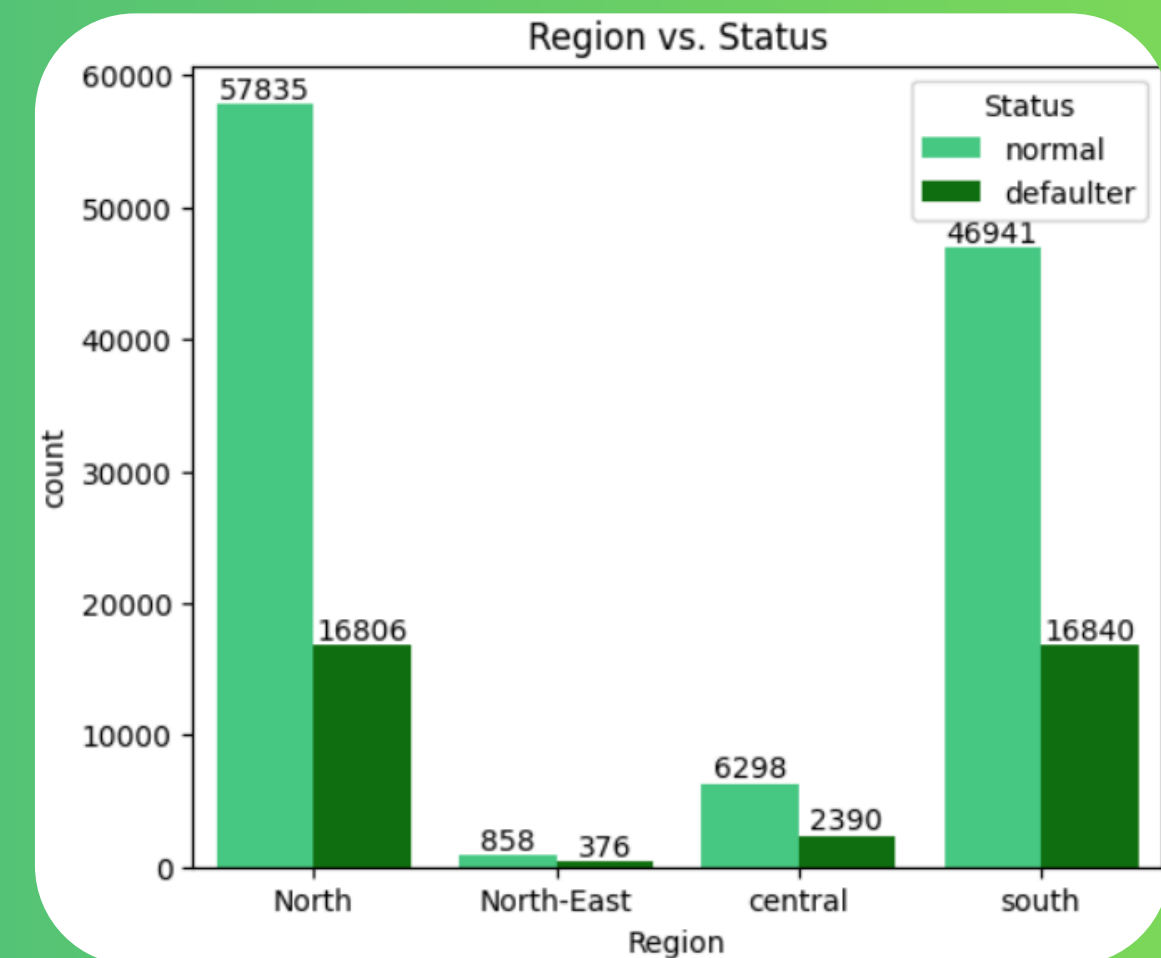
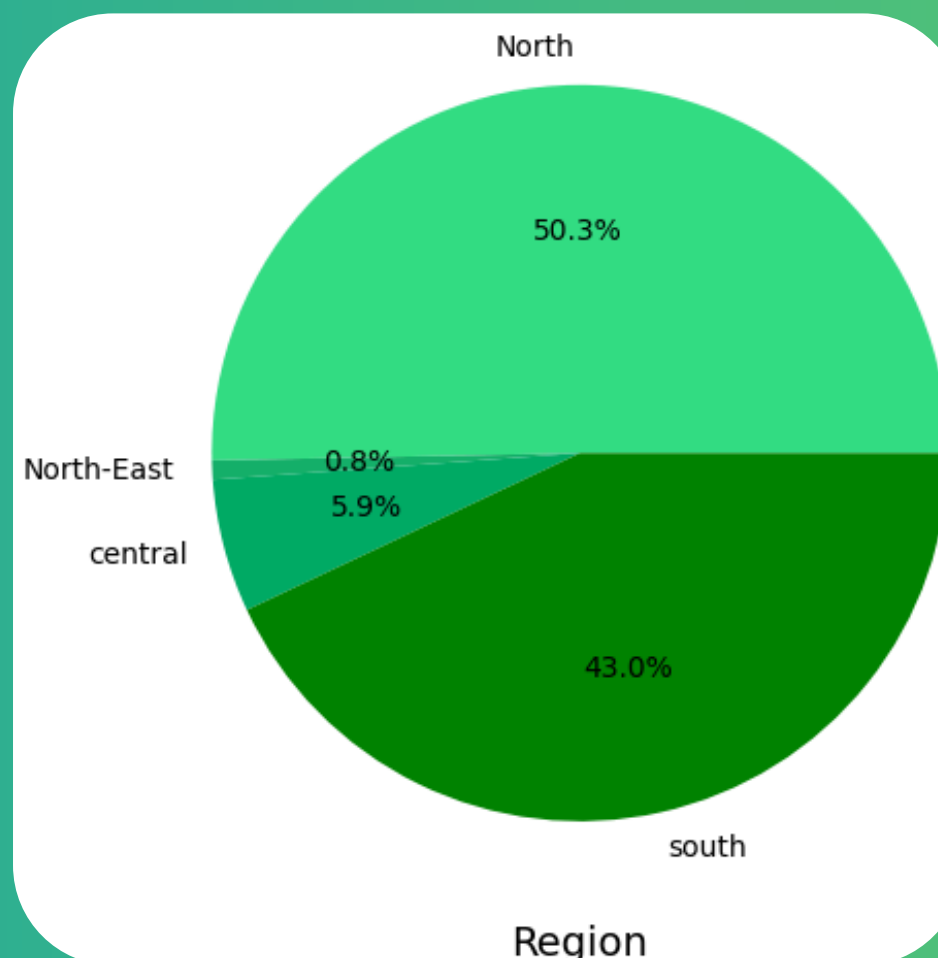
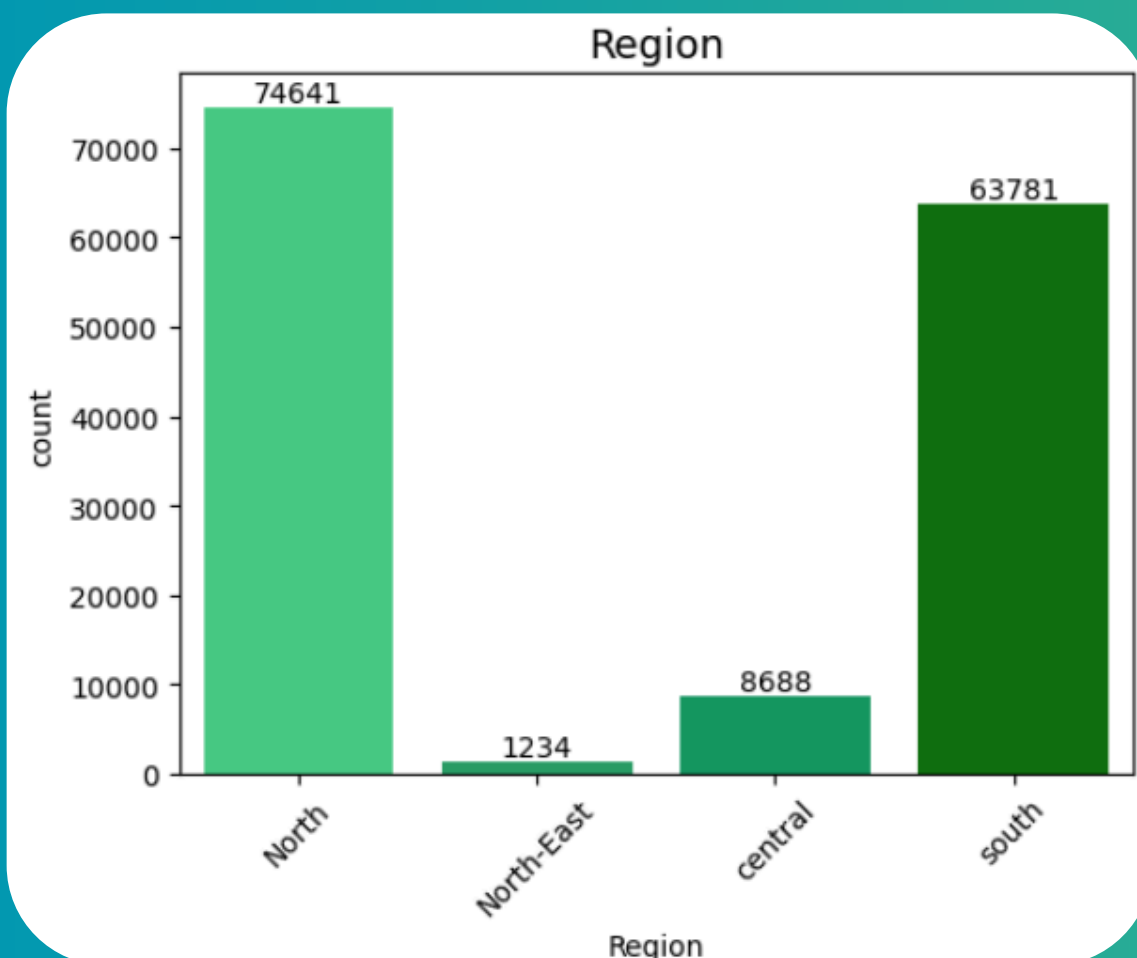


## AGE ANALYSIS

- **Most applicants** fall within the **middle-age range**, particularly between **35 and 64**. Younger and older age groups are underrepresented.
- The **default risk seems to increase with age beyond 55**, especially for the **65-74 and >74 age** groups, where **defaulters** make up a significant portion of the total. The **younger age groups**, though fewer in number, show higher sensitivity to economic stress, reflected in their **default numbers**.

### Recommendation:

**Develop tailored financial education** and support programs for **middle-aged (45-54) and near-retirement (55-64) borrowers**, as these groups have both the highest loan participation and default risks.



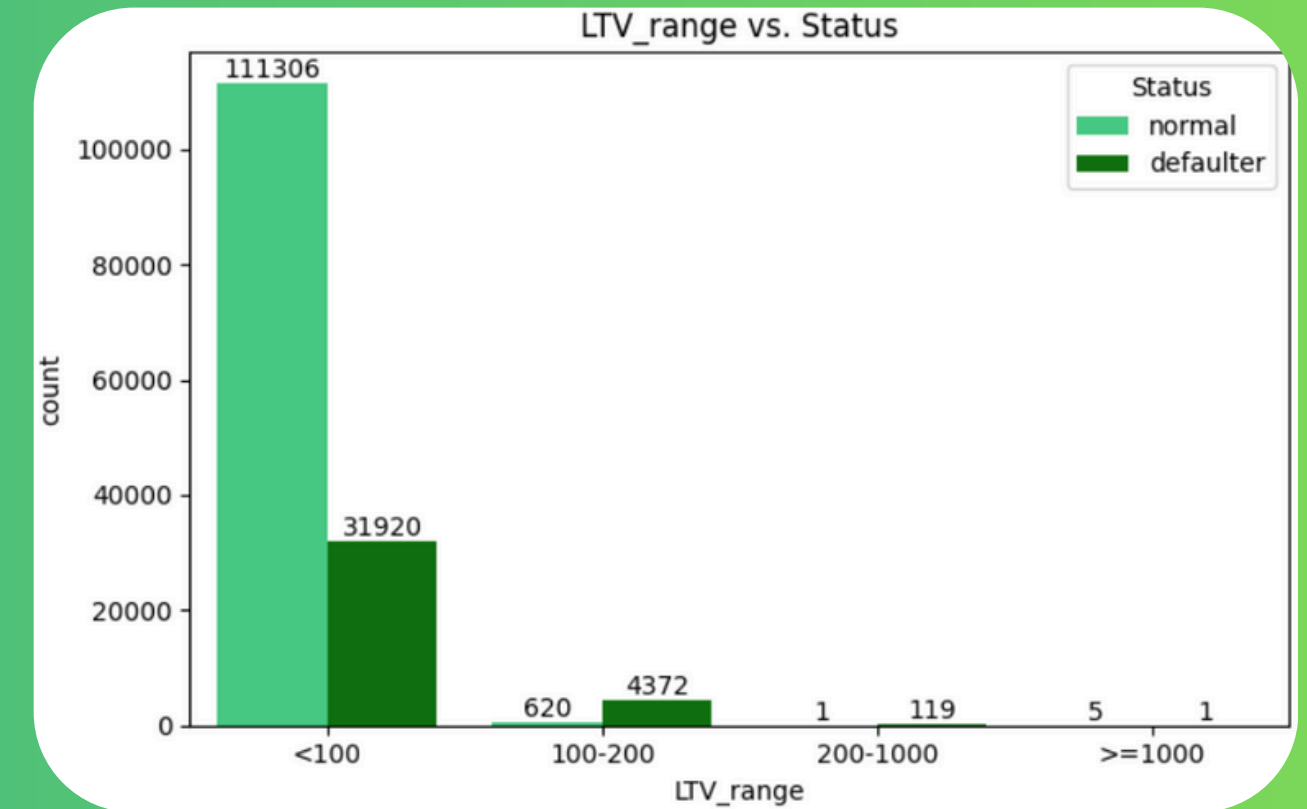
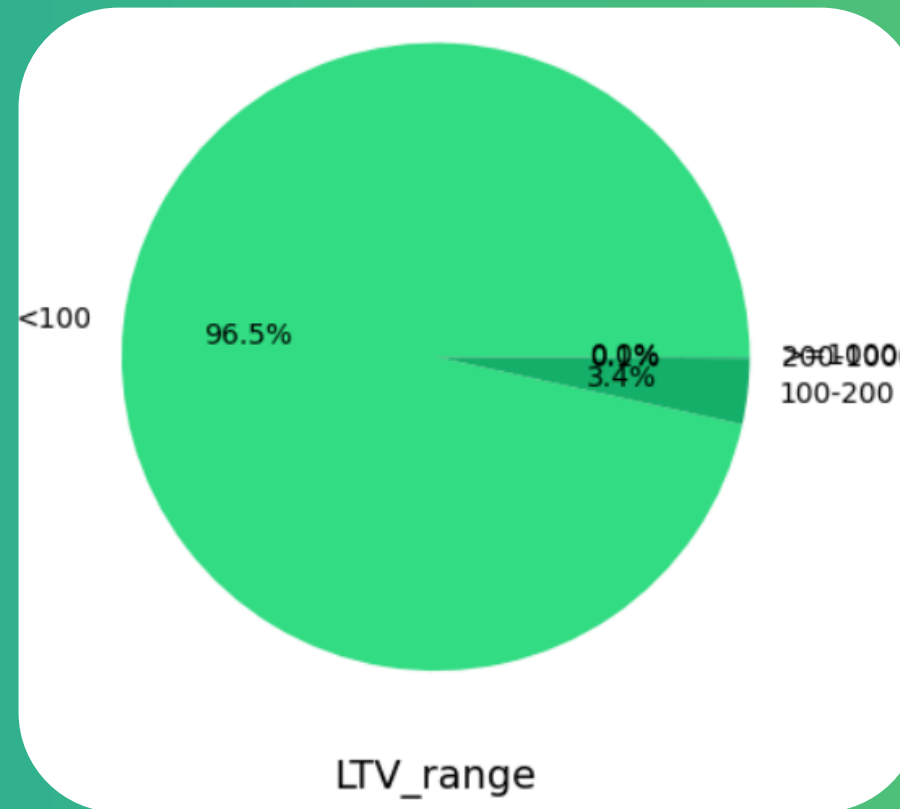
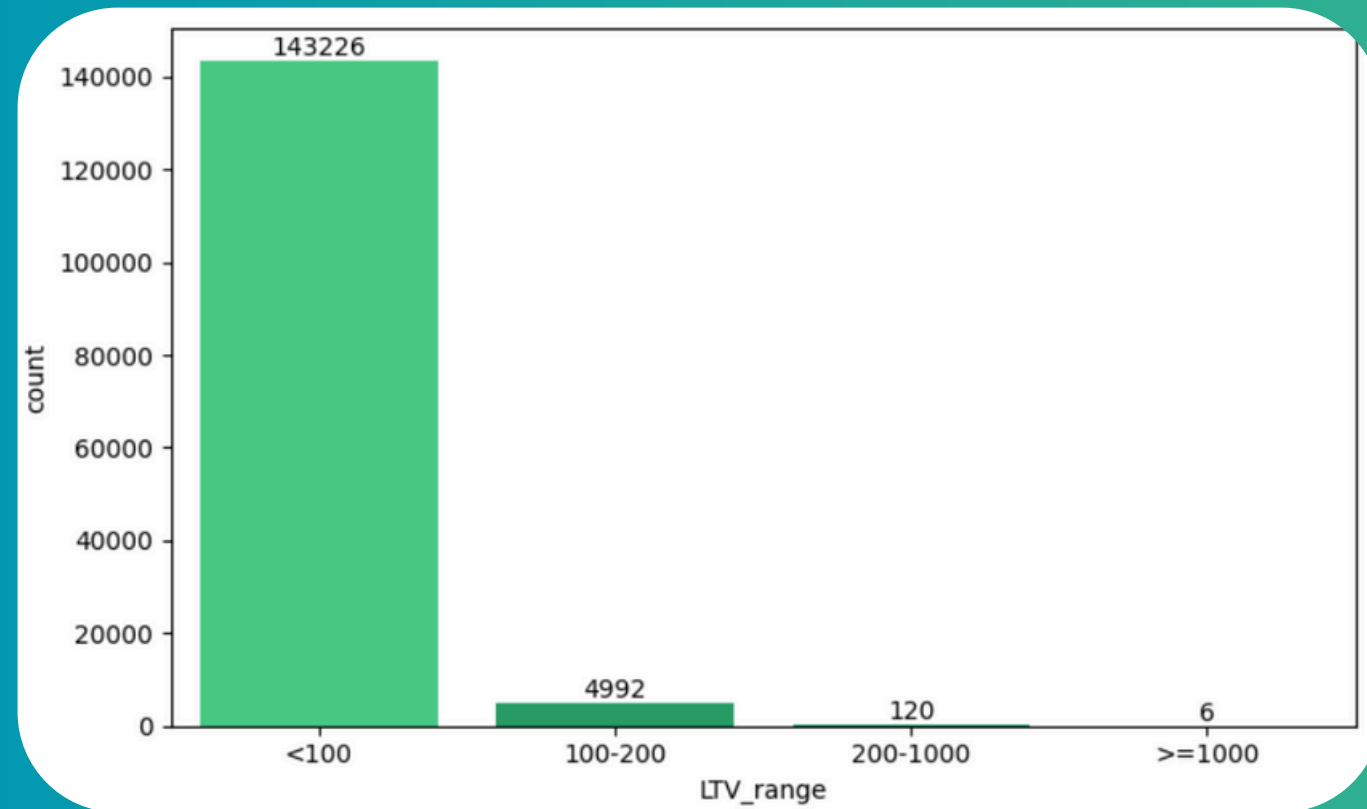
## REGION ANALYSIS

- The **majority of applicants** are from the **North and South regions**, with very few from the Central and North-East regions.
- **North Region** shows a strong performance with the **highest** number of **normal loans** and a relatively lower default rate.
- **North-East Region** is the **highest-risk area**, with nearly **a third of loans in default** despite having the smallest number of loans.

### Recommendation:

Since the North-East and Central regions show significantly higher default rates, the business should consider **implementing stricter loan approval criteria or higher collateral requirements in these areas.**





## LTV RANGE ANALYSIS

- (**<100**): The **majority of loans** have an LTV ratio of less than 100%, indicating that borrowers generally have **more equity in the property than the loan amount**.
- (**>=1000**): Very few loans (6 loans) have an LTV ratio greater than 1000%, which is a **rare and extremely high-risk situation**.
- Almost all loans with **LTV between 200 and 1000%** are in **default (99.17%)**, showing that excessively high LTV ratios lead to an **extremely high risk of default**.

### Recommendation:

**Lenders should be cautious in approving loans where LTV exceeds 100**, as these loans are prone to default. Special attention should be given to loans with  $LTV > 200$ , as they are almost guaranteed to default.

# INSIGHTS AND RECOMMENDATIONS BASED ON HYPOTHESIS TESTING

## Loan amount Vs. Status, Mann-Whitney U test (Right tailed):

### Insight:

Borrowers in the **Normal group** tend to **have higher loan amounts** compared to those in the **Defaulter group**, indicating that **larger loans are associated with more responsible repayment behavior**.

### Recommendation:

**Focus on strengthening risk assessment criteria for smaller loan amounts**, as borrowers with smaller loans show a higher likelihood of default. Offering financial literacy programs or additional credit counseling for smaller loan applicants may help mitigate this risk.

## Rate of Interest Vs. Status, Mann-Whitney U test (Right tailed):

### Insight:

The **Normal group** tends to have loans with a **higher rate of interest** compared to the **Defaulter group**, suggesting that borrowers who **manage to pay back their loans** successfully are more likely to do so even when faced with **higher interest rates**.

### Recommendation:

Increase vigilance for borrowers with lower interest rates, as they might pose a higher risk of default. **Consider reviewing the loan approval criteria for lower-interest loans** and enhancing support mechanisms, such as offering financial counseling, to reduce the risk of default in this segment.

## Upfront\_charges Vs. Status, Mann-Whitney U test (Left tailed):

### Insight:

Borrowers in the **Defaulter group** tend to have **higher upfront charges** compared to those in the **Normal group**, indicating that **higher initial costs may be associated with an increased likelihood of default**.

### Recommendation:

**Consider offering loans with lower or flexible upfront charges to reduce the financial burden on borrowers**. This could improve repayment rates and lower the default risk, particularly for those who might struggle with high initial loan costs.

# INSIGHTS AND RECOMMENDATIONS BASED ON HYPOTHESIS TESTING

## Property Value Vs. Status, Mann-Whitney U test (Right tailed):

### Insight:

The **Normal group** tends to have a **higher property value** compared to the **Defaulter group**, suggesting that borrowers with **more valuable properties are less likely to default on their loans**.

### Recommendation:

Implement a policy to conduct thorough evaluations of property values during the loan application process, **prioritizing loans secured by higher-value properties**. This could enhance risk assessment strategies and potentially reduce default rates by focusing on more stable asset backing.

## Income Vs. Status, Mann-Whitney U test (Right tailed):

### Insight:

The **Normal group** has a **higher average income** compared to the **Defaulter group**, indicating that borrowers with **greater income levels are less likely to default** on their loans.

### Recommendation:

Consider **implementing income-based lending criteria that favor borrowers with higher income levels**, as this could improve loan performance and reduce default rates. Additionally, targeted financial education programs could be developed for lower-income borrowers to enhance their financial literacy and repayment capabilities.

## LTV Vs. Status, Mann-Whitney U test (Left tailed):

### Insight:

The **Normal group** has a **lower loan-to-value (LTV)** ratio compared to the **Defaulter group**, suggesting that **borrowers who take on less debt relative to the value of the collateral (property) are less likely to default**.

### Recommendation:

**Implement stricter LTV limits for loan approvals to mitigate risk.** Encourage borrowers to maintain a lower LTV ratio, which could improve loan stability and reduce the likelihood of defaults. Consider providing additional support or incentives for borrowers who achieve lower LTV ratios.

**THANK YOU**