Lending Club Case Study

Sachin G

Promblem Statement

A consumer finance firm, offering various loan products to urban clients, faces two major risks when processing loan applications: the risk of "business loss" (stemming from denying loans that would have been repaid) and the risk of "credit loss" (resulting from approving loans that eventually default). The core challenge is to strike a balance between these risks when making lending decisions. This analysis aims to mitigate "credit loss," which occurs when a borrower defaults, causing financial damage to the company. The main objective of this study is to identify the factors contributing to loan defaults through Exploratory Data Analysis (EDA). By uncovering these factors, the company can develop a robust risk assessment model to minimise the approval of loans to high-risk applicants—those prone to default. The focus of the analysis is on differentiating borrowers who have "fully repaid" their loans, those with "ongoing" loans, and "defaulters" (i.e., loans that have been "charged-off"). Understanding the indicators of default will allow the company to reduce "credit loss" and enhance its "loan approval" process, ultimately leading to better risk management and more effective portfolio oversight.

Objectives

1. Leverage Exploratory Data Analysis (EDA) to identify the key variables influencing loan defaults.
2. The objective is to detect patterns that distinguish high-risk applicants who are more likely to default on their loans.
3. Based on these insights, design strategies to minimize credit loss by reducing the approval of loans to risky borrowers.
4. Ensure that future loan approval decisions strike an optimal balance between seizing business opportunities and mitigating financial risks from defaults, improving overall portfolio performance.

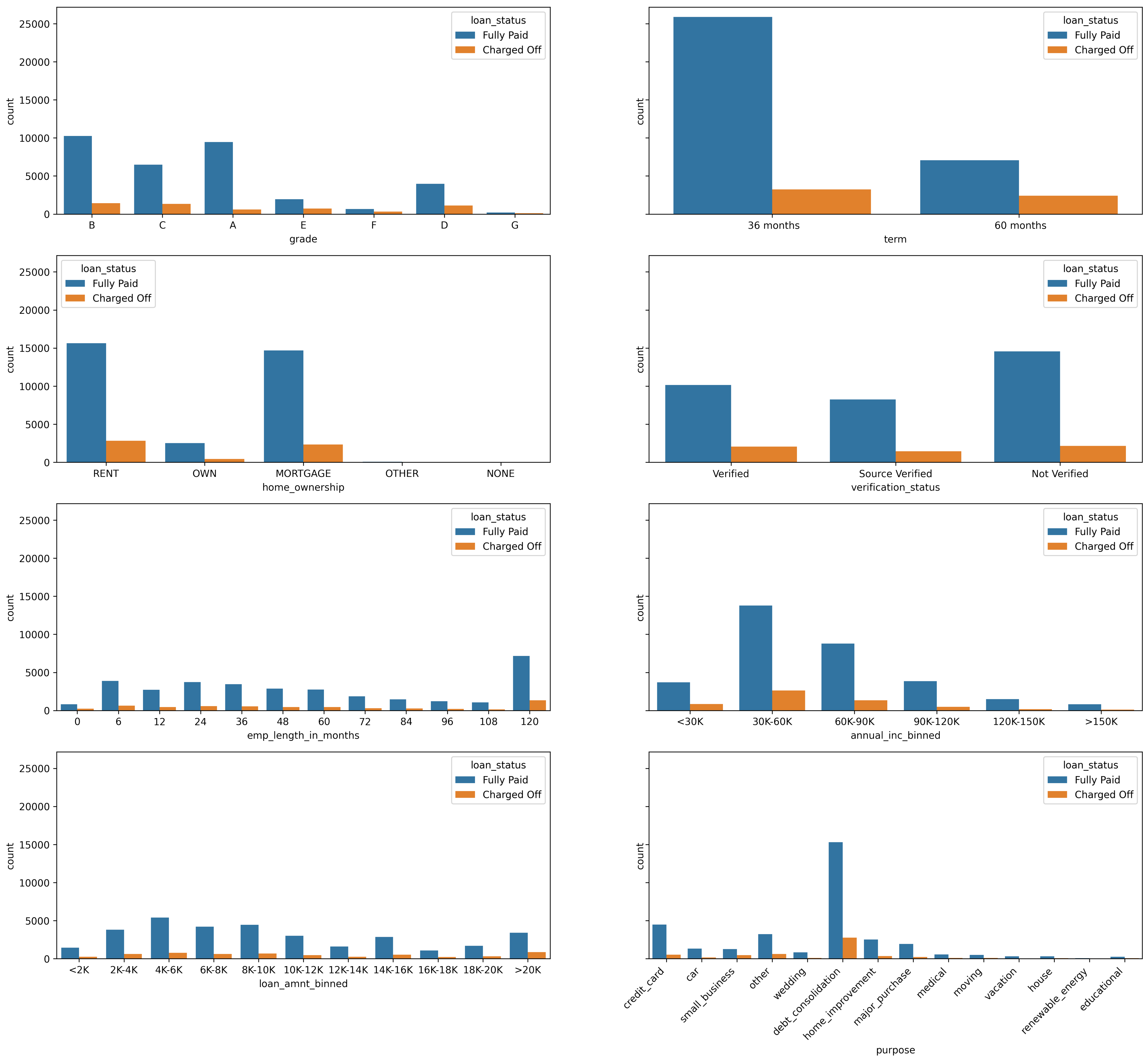
Methodology Used

- Data Cleaning
- Data Understanding
- Data Analysis
- Recommendations

Loan Status Evaluation across Categorical Variables

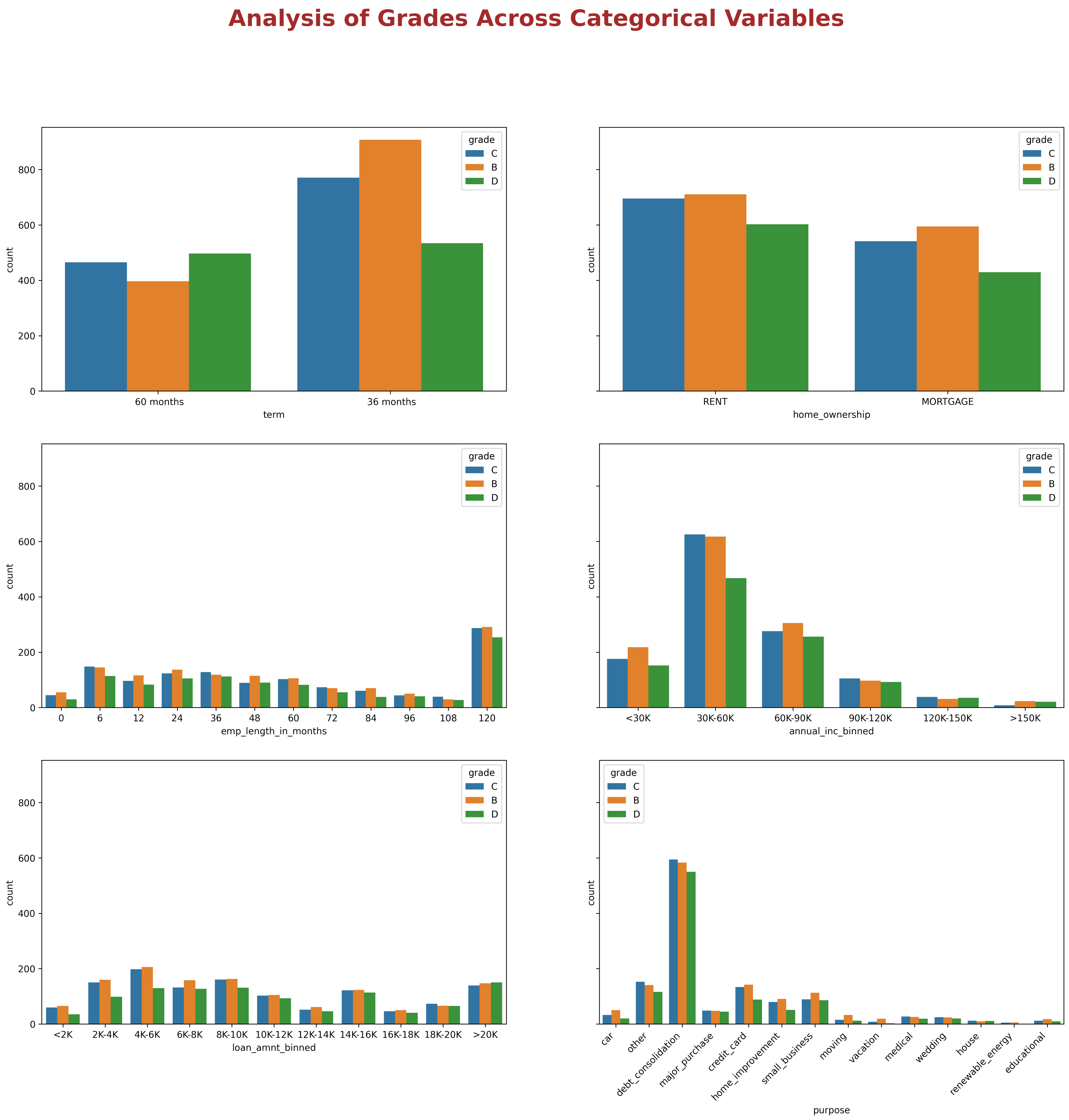
- Grade: Grades B, C, and D exhibit a higher proportion of Charged Off loans compared to Grades A, E, F, and G. This implies that loans with lower grades are more prone to default.
- Term: Loans with a 36-month term have a slightly higher incidence of Charged Off loans in comparison to those with a 60-month term.
- Home Ownership: Borrowers classified as RENTERS and MORTGAGE holders show a slightly greater proportion of Charged Off loans compared to those categorized as OWNERS or OTHER.
- Verification Status: No distinct patterns were identified in this category.
- Employment Length: Borrowers with over 10 years of employment display a slightly higher proportion of Charged Off loans. However, those with less than 1 year, as well as 2, 3, and 4 years, also exhibit some irregular trends.
- Annual Income: The likelihood of Charged Off loans decreases as annual income rises, indicating that higher-income borrowers are less likely to default. Borrowers earning between \$30K and \$60K appear to be at a higher risk of default.
- Purpose: Loans for debt consolidation seem to have the highest default rates (charged off), while categories such as credit card, small business, other, and home improvement also show increased risk levels.

Loan Status Evaluation in Relation to Categorical Variables



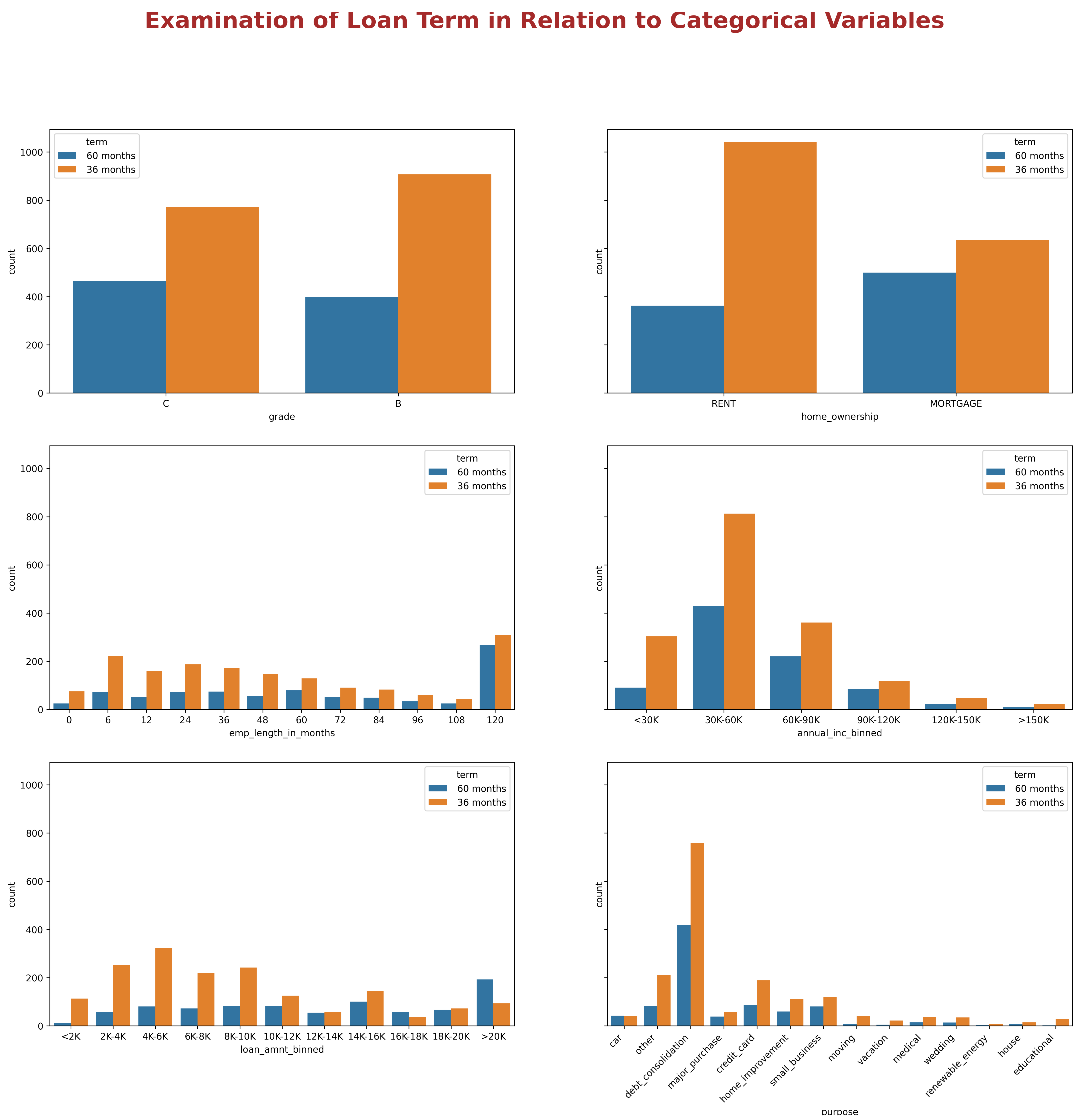
Grade Evaluation across Categorical Variables

- Grade B and C demonstrate a strong correlation with all other categories.



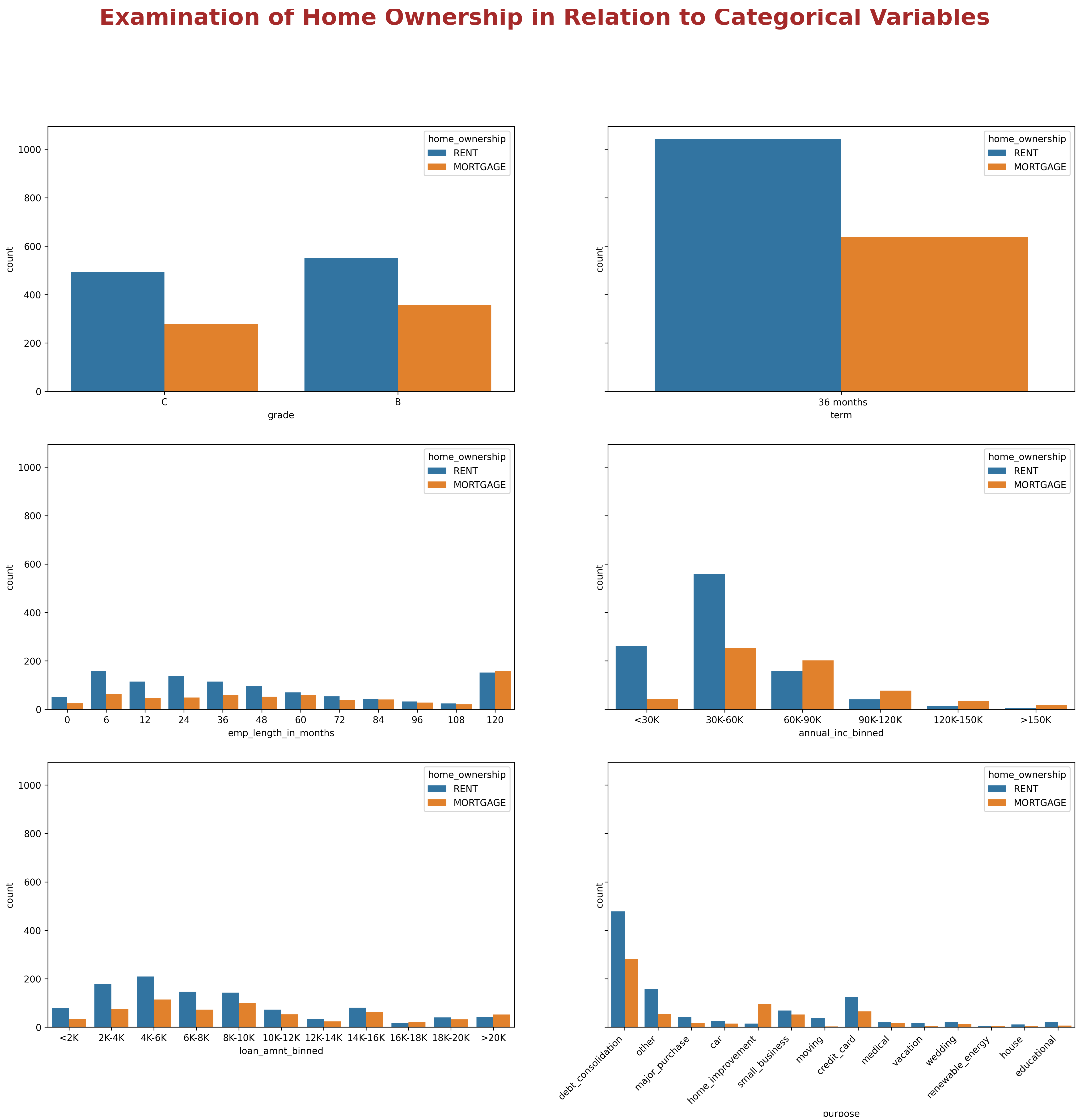
Loan Term Evaluation across Categorical Variables

- The 36-month duration exhibits a strong correlation with all other categories.



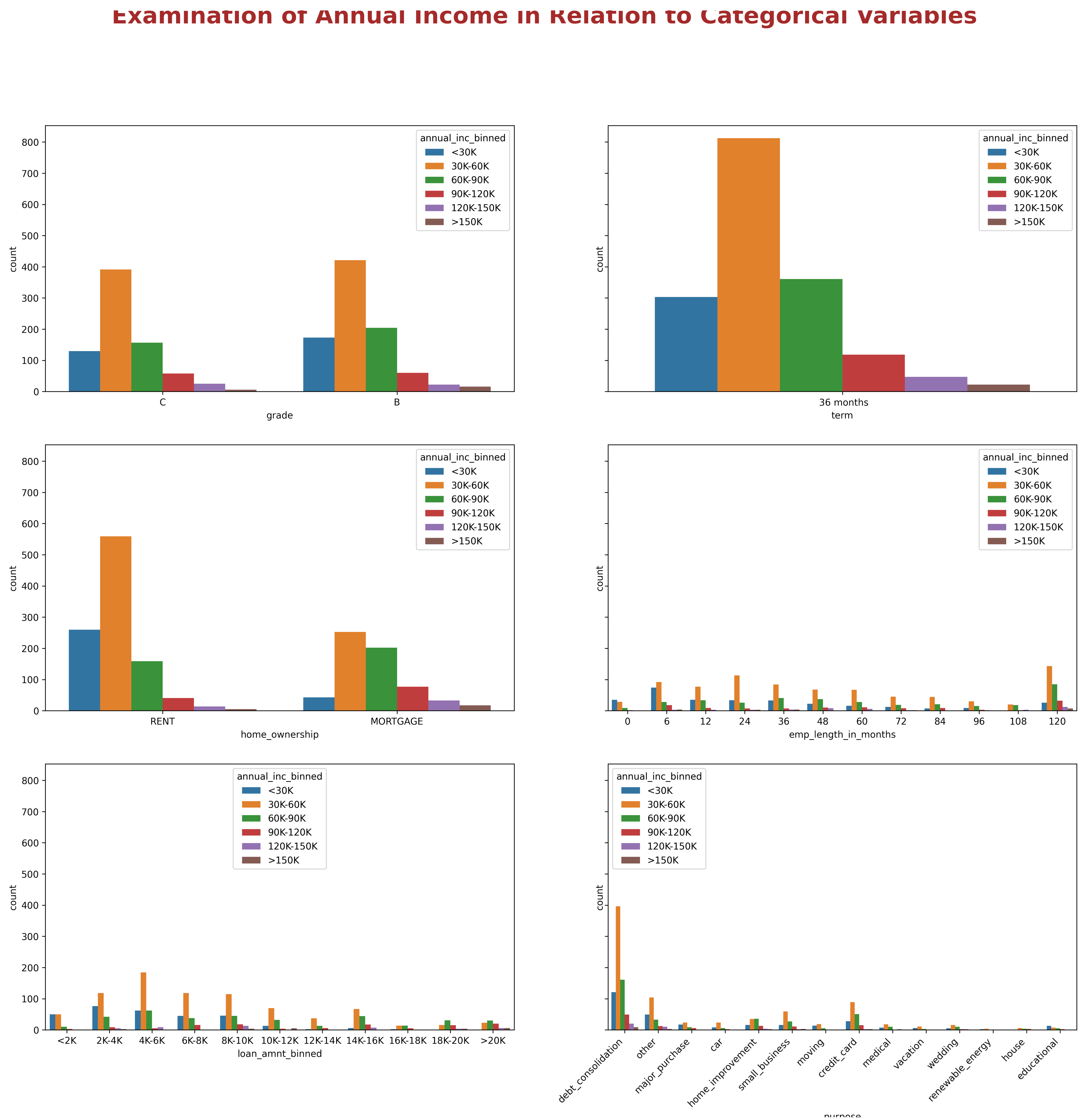
Home Ownership Evaluation across Categorical Variables

- The RENT category demonstrates a strong correlation with all other categories, except for employment length.



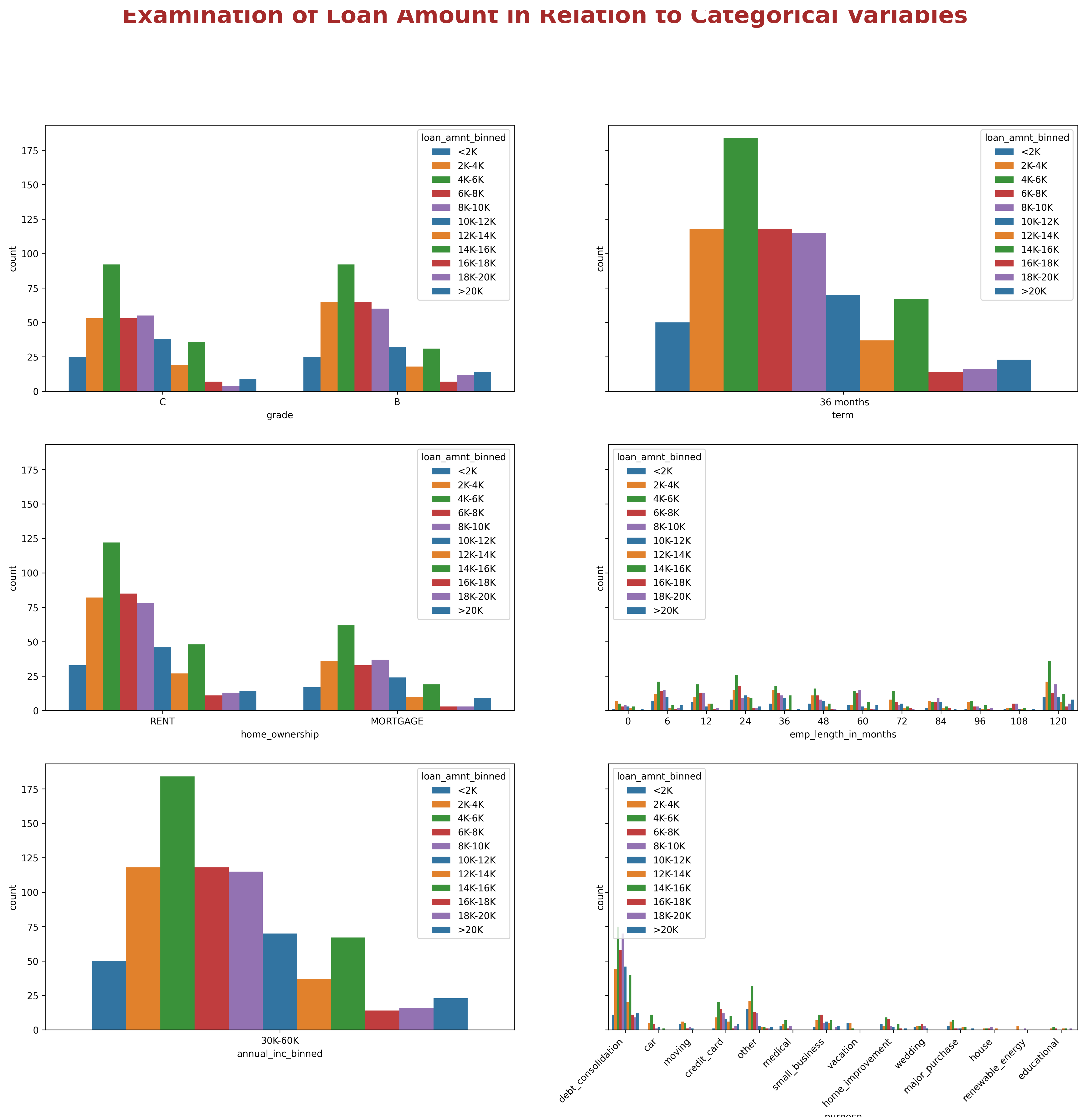
Annual Income Examination across Categorical Variable

- An annual income of $30K to $60k exhibits a strong correlation with all the other categories



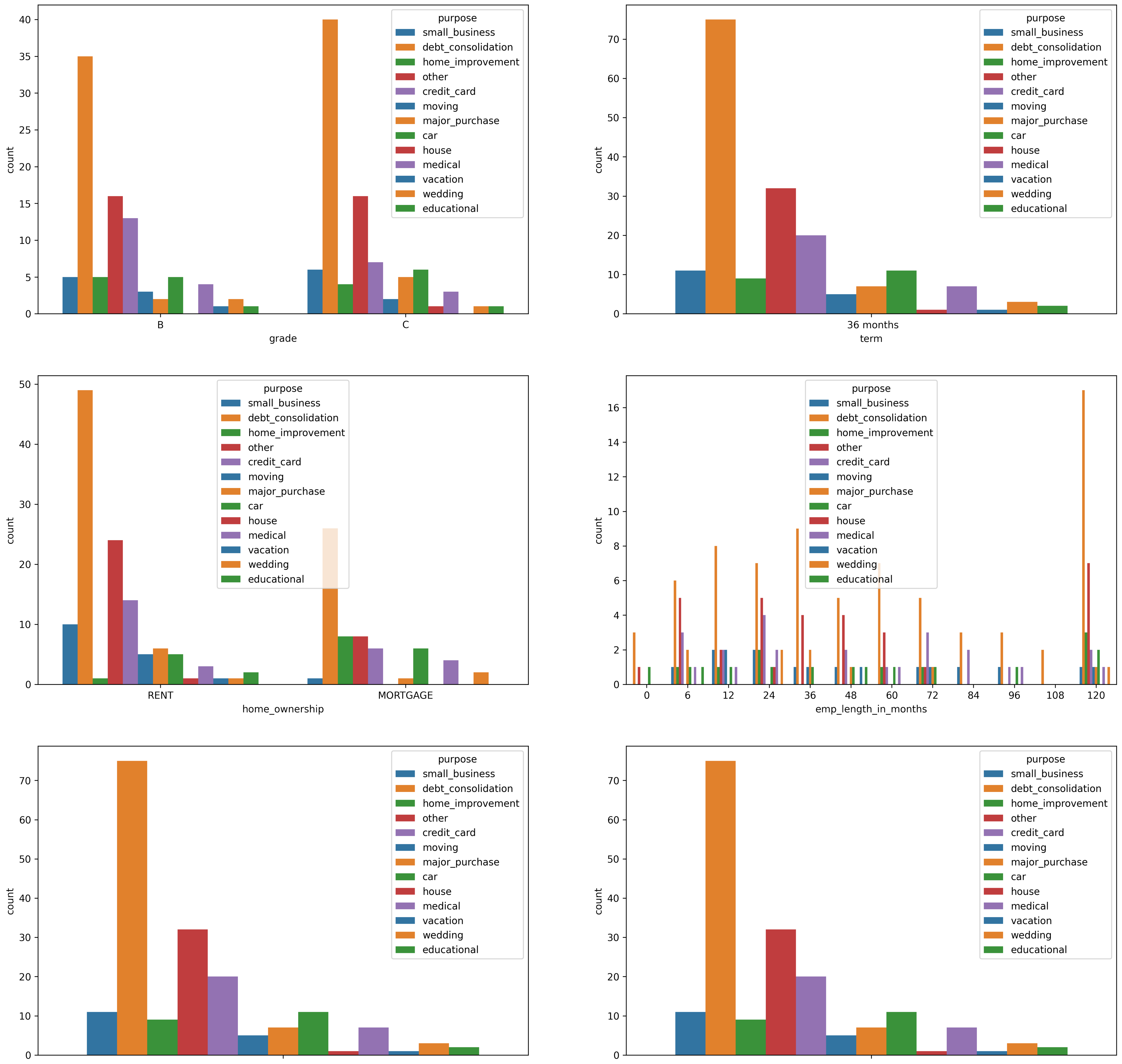
Loan Amount Examination across Categorical Variables

- The loan amount ranging from $4k to $6k demonstrates a strong correlation with all other categories.



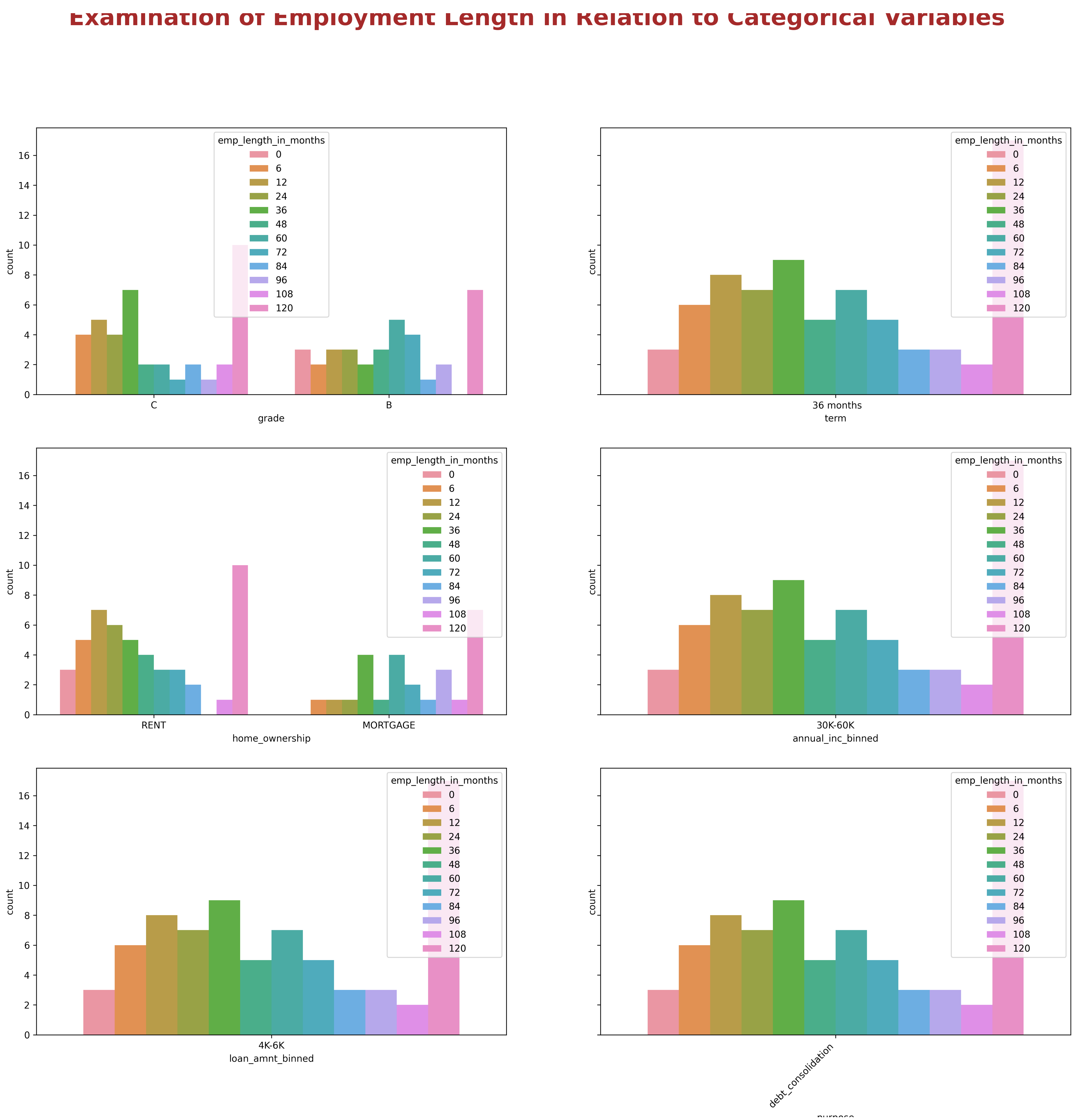
Purpose Examination across Categorical Variable

- Debt consolidation demonstrates a strong correlation with all other categories.



Employment Length Examination across Categorical Variables

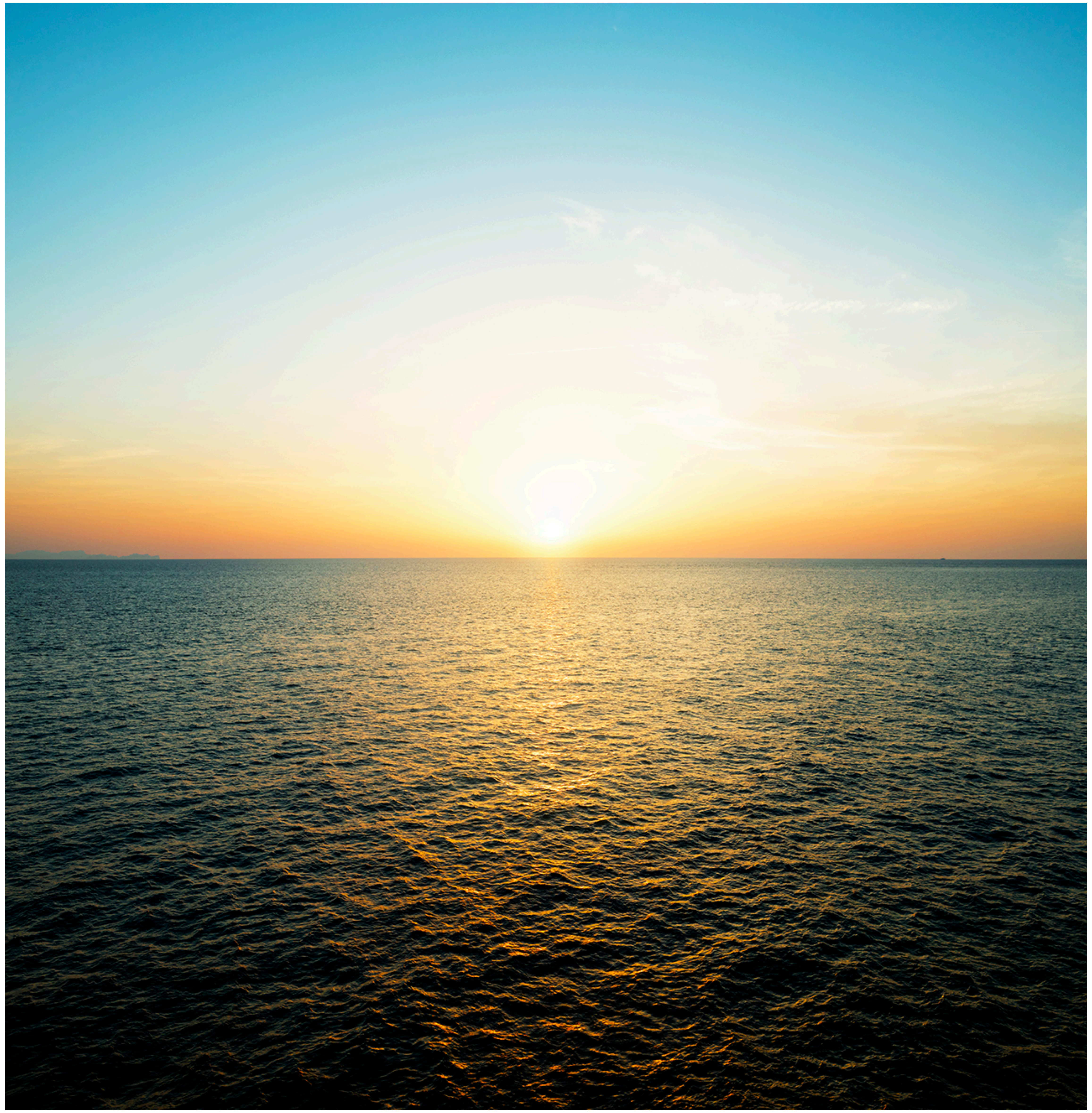
- An employment length of 10 or more years exhibits a strong correlation with all other categories.



Observation

Following combinations infers probability of loan default

- Grade: B, C
- Term: 36 month
- Home Ownership: Rent
- Employment Length: 10+ years
- Annual Income: $30-$60k
- Loan Amount: $4k-$6k
- Purpose: Debt consolidation



Conclusion

Key Factors to Forecast Loan Default Probability and Mitigate Credit Loss

- Loan Grade: Borrowers classified with a loan grade B or C demonstrate a greater tendency to default.
- Loan Term: Loans with a term of 36 months exhibit a slightly higher default rate in comparison to those with a 60-month term.
- Home Ownership Status: Renters armoire susceptible to default than homeowners, highlighting "RENT" as a significant risk factor.
- Employment Length: Borrowers with over 10 years of employment experience show a marginally higher default risk that those with shorter employment durations.
- Annual Income: Individuals earning between $30k to $60k, especially for debt consolidation purposes, have a heightened likelihood of default.
- Loan Purpose: Loans aimed at debt consolidation and credit card refinancing exhibit higher default rates. Other categories, including small business loans, general loans and home improvement loans, also present significant risk.
- Credit Inqurries in last 8 months: Borrowers with a greater number of credit inquiries in the preceding six months are more likely to default, suggesting a correlation between recent credit activity and default risk.