Lending Club Case Study

Group Members:

- Sunny Prakash
- Sudhir Wani

Case Study Brief*:*

You work for a consumer finance company which specializes in lending various types of loans to urban customers. When the company receives a loan application, the company has to make a decision for loan approval based on the applicant's profile. Two types of risks are associated with the bank's decision:

- If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company
- If the applicant is not likely to repay the loan, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company.

When a person applies for a loan, there are two types of decisions that could be taken by the company:

1**.** **Loan accepted**: If the company approves the loan, there are 3 possible scenarios described below:

- **Fully paid**: Applicant has fully paid the loan (the principal and the interest rate)
- **Current**: Applicant is in the process of paying the installments, i.e. the tenure of the loan is not yet completed. These candidates are not labeled as 'defaulted'.
- **Charged-off**: Applicant has not paid the installments in due time for a long period of time, i.e. he/she has defaulted on the loan

2. **Loan rejected**: The company had rejected the loan (because the candidate does not meet their requirements etc.). Since the loan was rejected, there is no transactional history of those applicants with the company and so this data is not available with the company (and thus in this dataset)

Business Objectives:

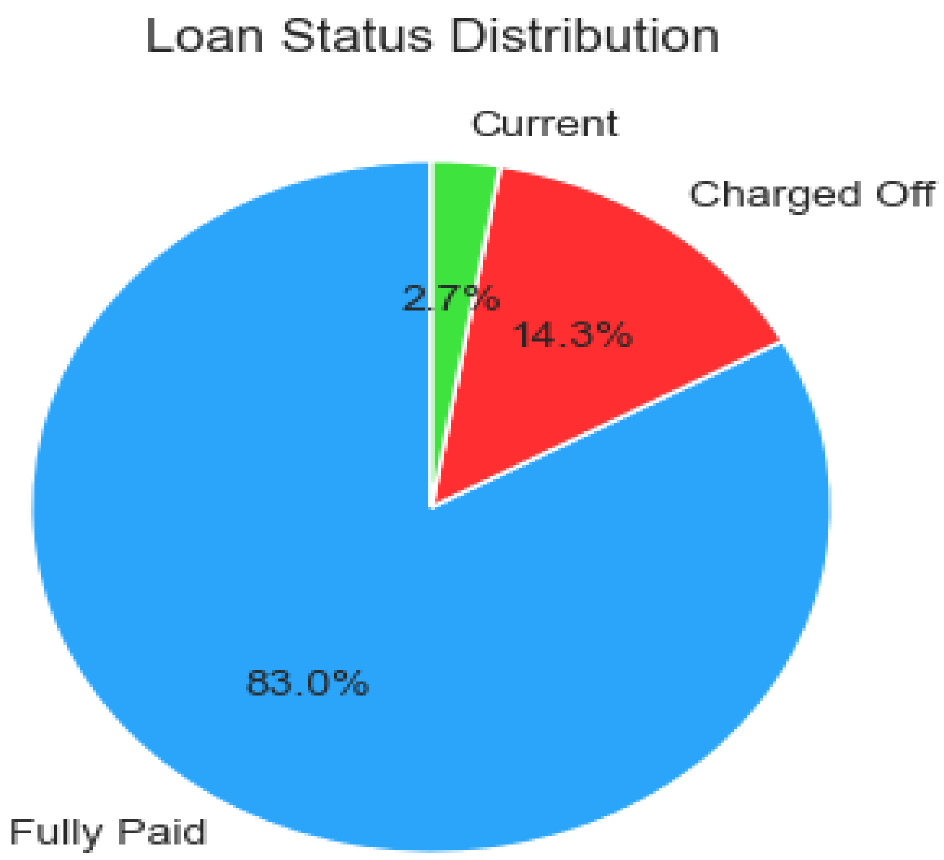
This company is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures. Borrowers can easily access lower interest rate loans through a fast online interface.

Like most other lending companies, lending loans to ‘risky’ applicants is the largest source of financial loss (called credit loss). Credit loss is the amount of money lost by the lender when the borrower refuses to pay or runs away with the money owed. In other words, borrowers who **default** cause the largest amount of loss to the lenders. In this case, the customers labeled as 'charged-off' are the 'defaulters'.

If one is able to identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss. Identification of such applicants using EDA is the aim of this case study.

In other words, the company wants to understand the **driving factors (or driver variables)** behind loan default, i.e. the variables which are strong indicators of default. The company can utilize this knowledge for its portfolio and risk assessment.

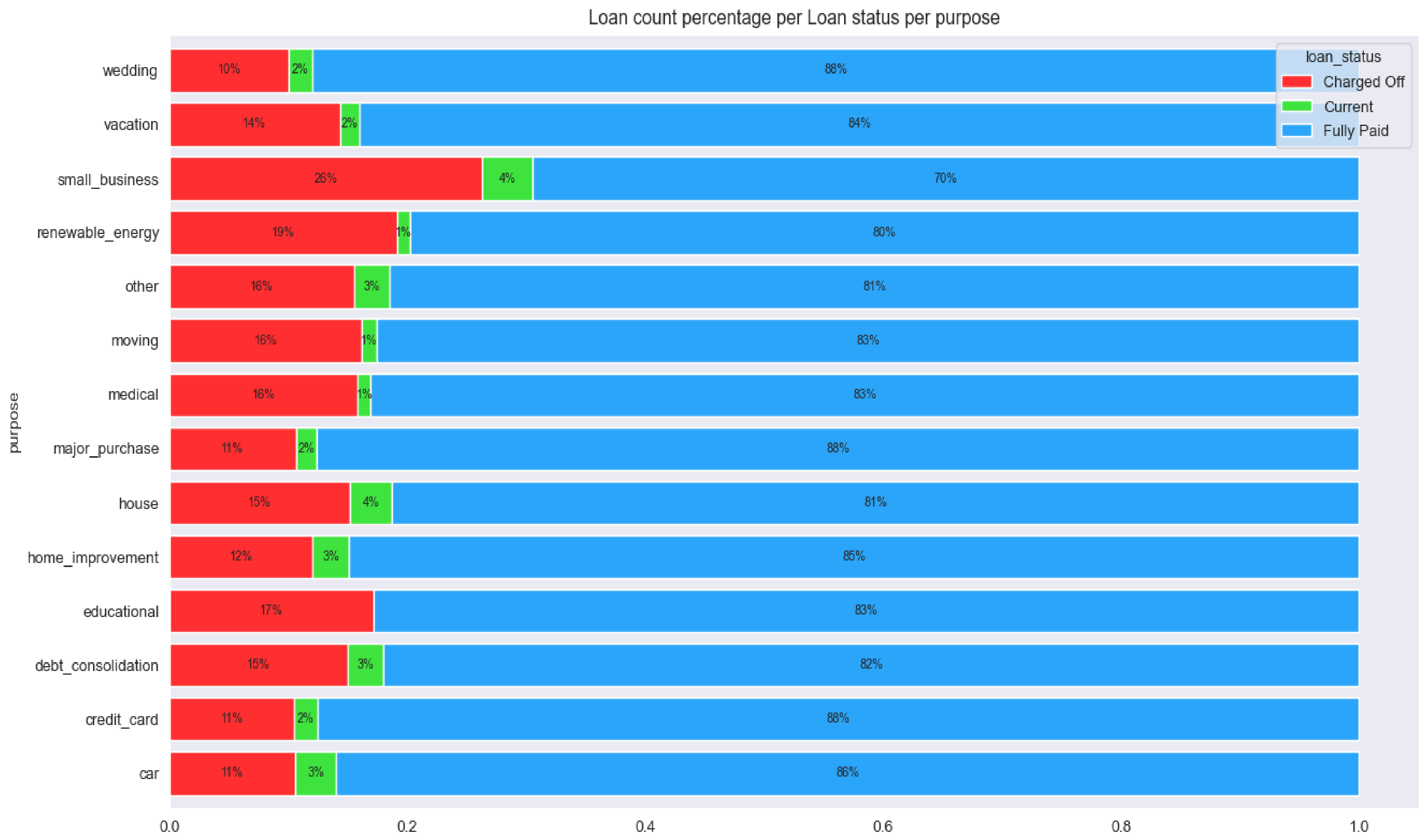
Overall Loan Status



Key Messages:

- *Approximately **14%** of the loans were Charged off out of total loans.*

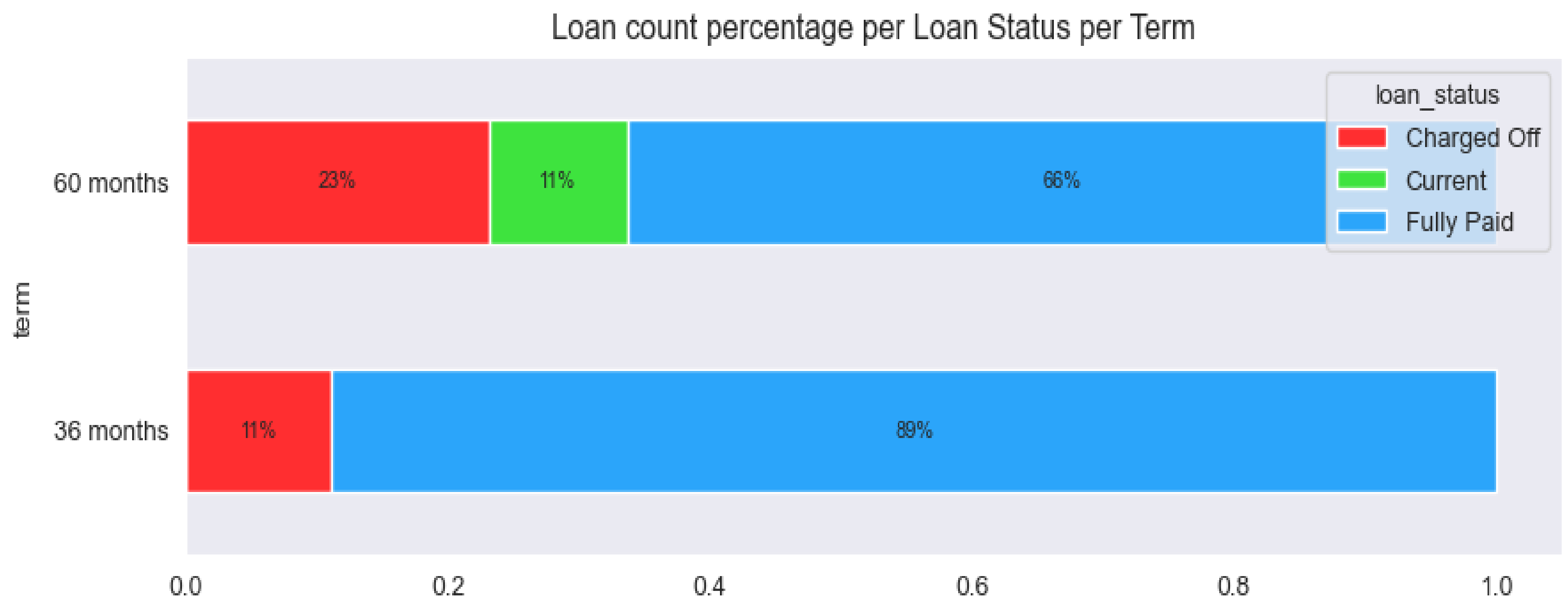
Loan Status Distribution by Loan Purpose: Charged Off vs Fully Paid vs Current



Key Messages:

- ***Small Business*** *loans have the highest* ***Charged Off*** *rate (26%), indicating higher risk in this category.*
- *Only 70% of small business loans were* ***Fully Paid****, which is the lowest among all categories*

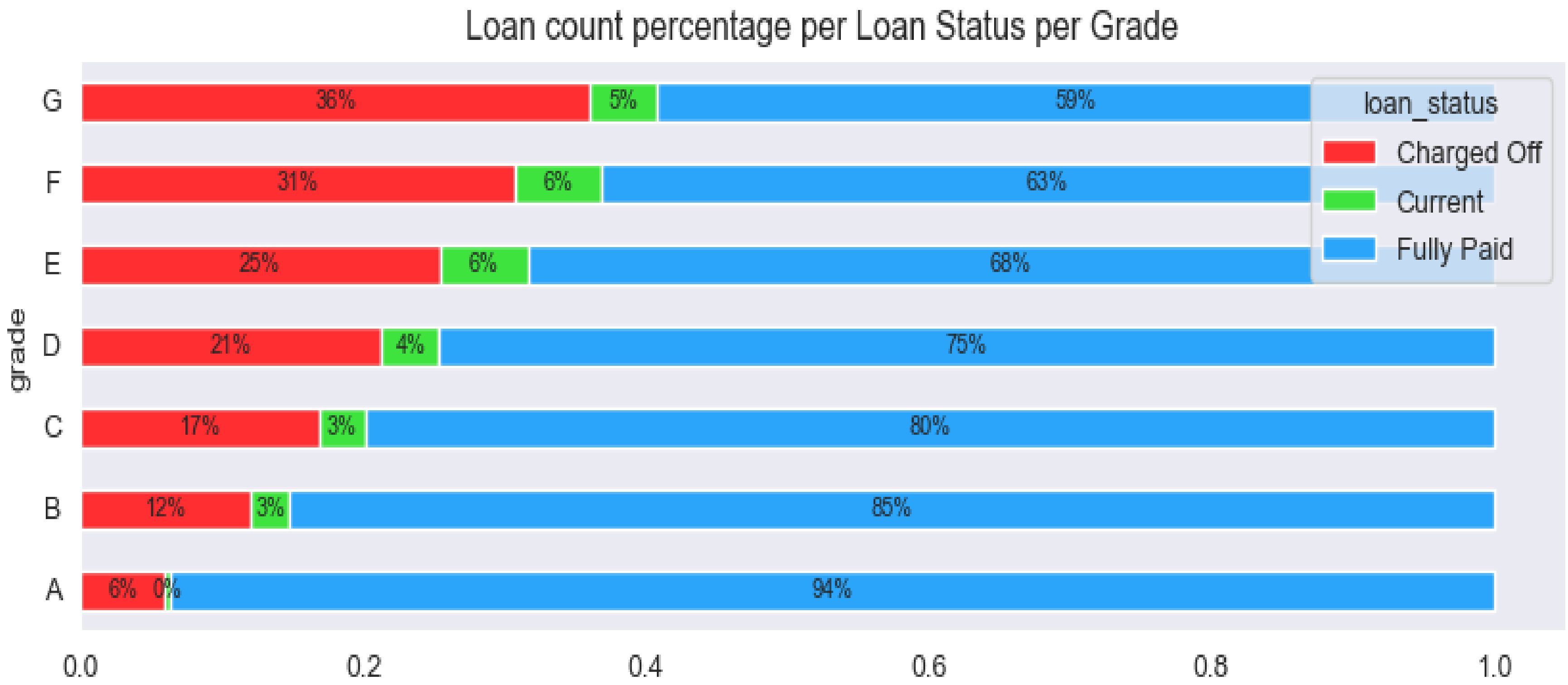
Loan Status Distribution by Loan Term: 36 vs 60 Months



Key Messages:

- ***60 month loans*** *have a much higher* ***Charged Off*** *rate (23%) compared to* ***36 month loans*** *(11%), indicating they are riskier.*
- *Only* ***66%*** *of 60 month loans were* ***Fully Paid****, while* ***89%*** *of 36 month loans were successfully repaid*

Loan Status Distribution by Grade

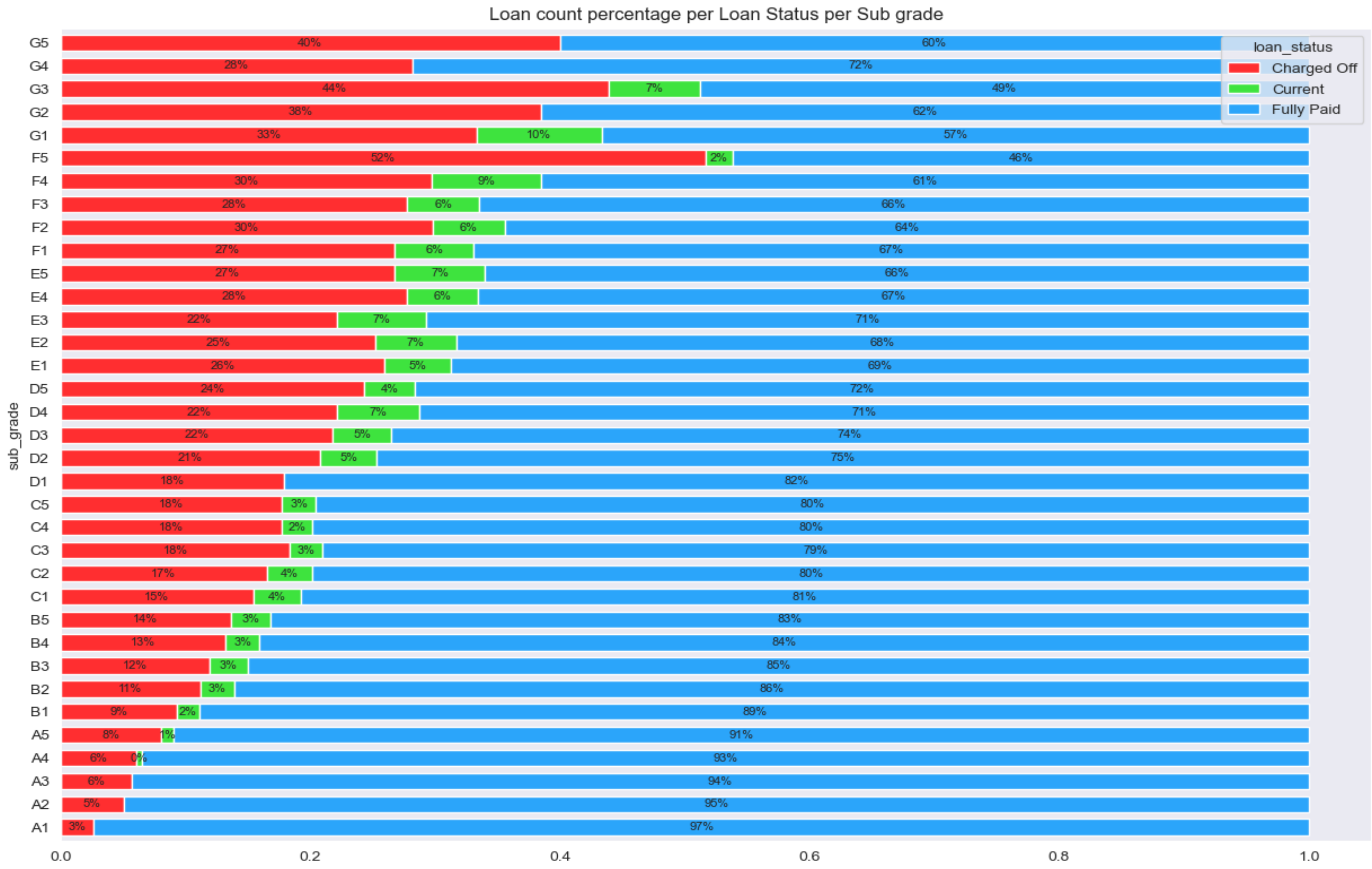


Key Messages:

- *Lower grade loans (F and G) have significantly higher default rates compared to higher grade loans, with* ***Grade G*** *having a* ***38%*** *default rate.*

- *Careful monitoring and assessment of lower-grade loans are crucial to mitigate financial losses.*

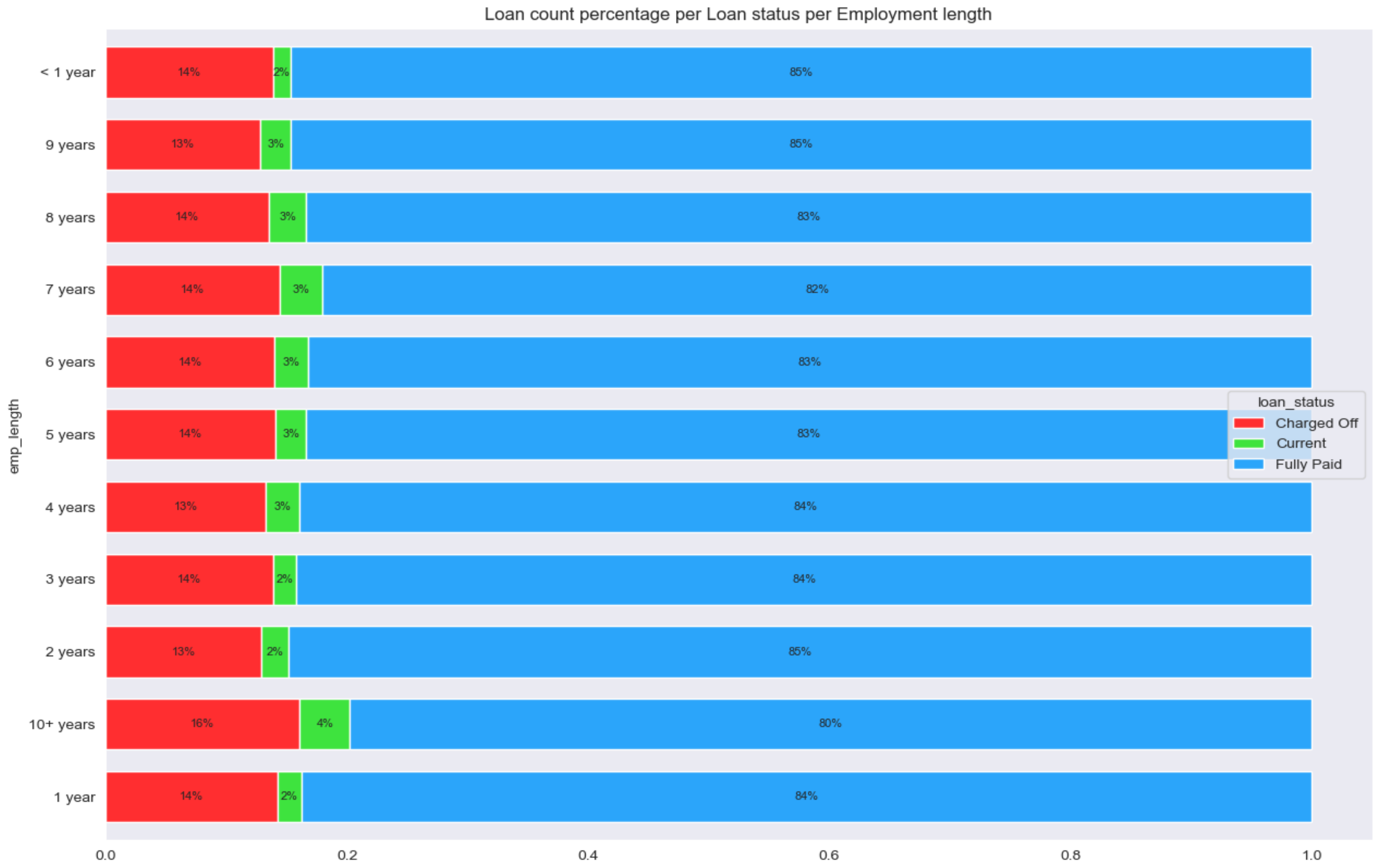
Loan Status Distribution by Sub-grade



Key Messages:

- *Sub-grades like* ***G5*** *exhibit the highest default rate, with* ***40%*** *of loans charged off.*
- *Higher sub-grades like A1 are significantly less risky compared to lower ones like G5*

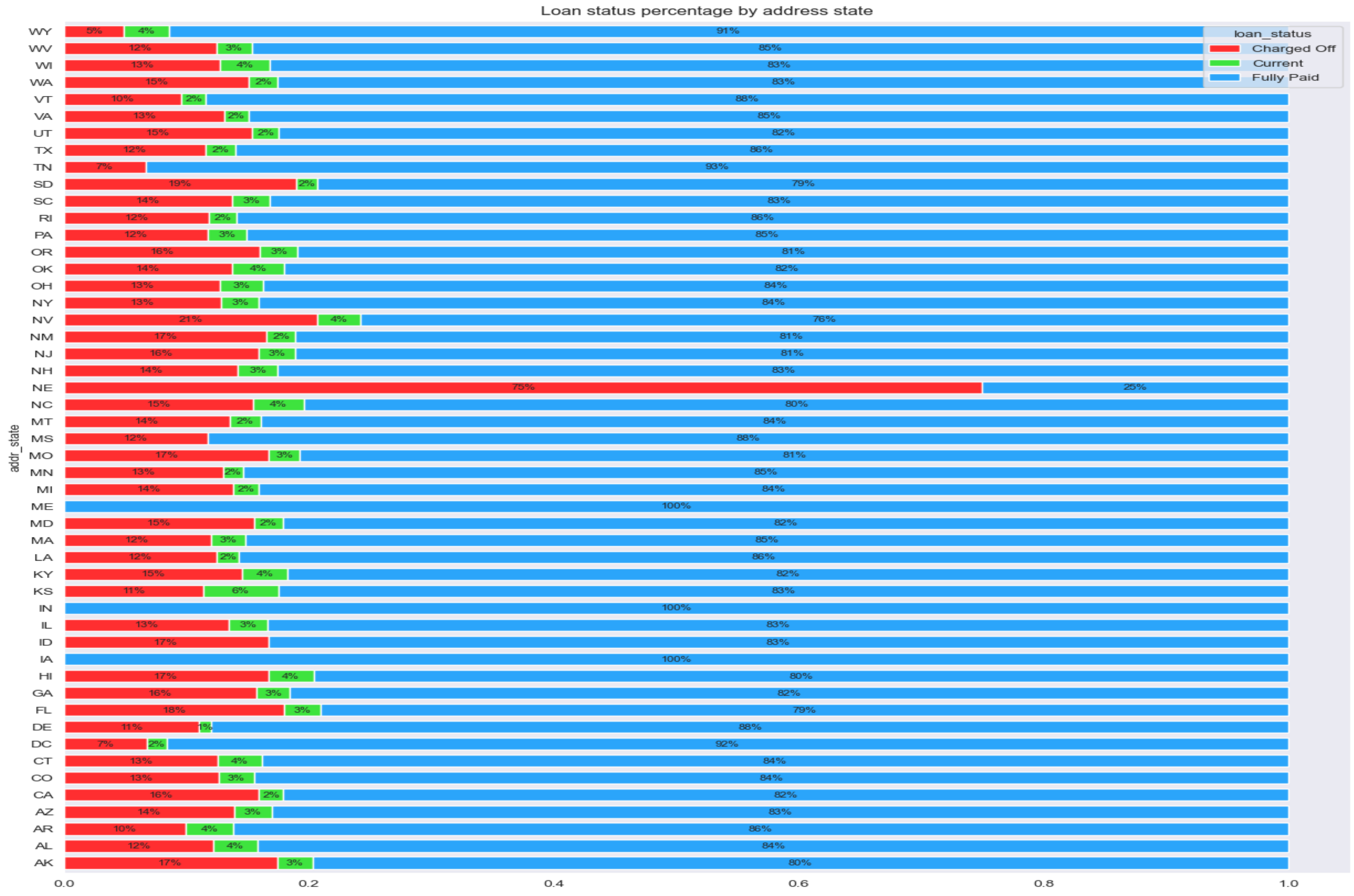
Loan Status Distribution by Employment Length



Key Messages:

- *Employment length **doesn't drastically impact** loan status distribution*
- *Slightly higher percentage of charged-off loans for those with **10+ years** of employment **(16%)** compared to other employment lengths **(13%-14%)**.*

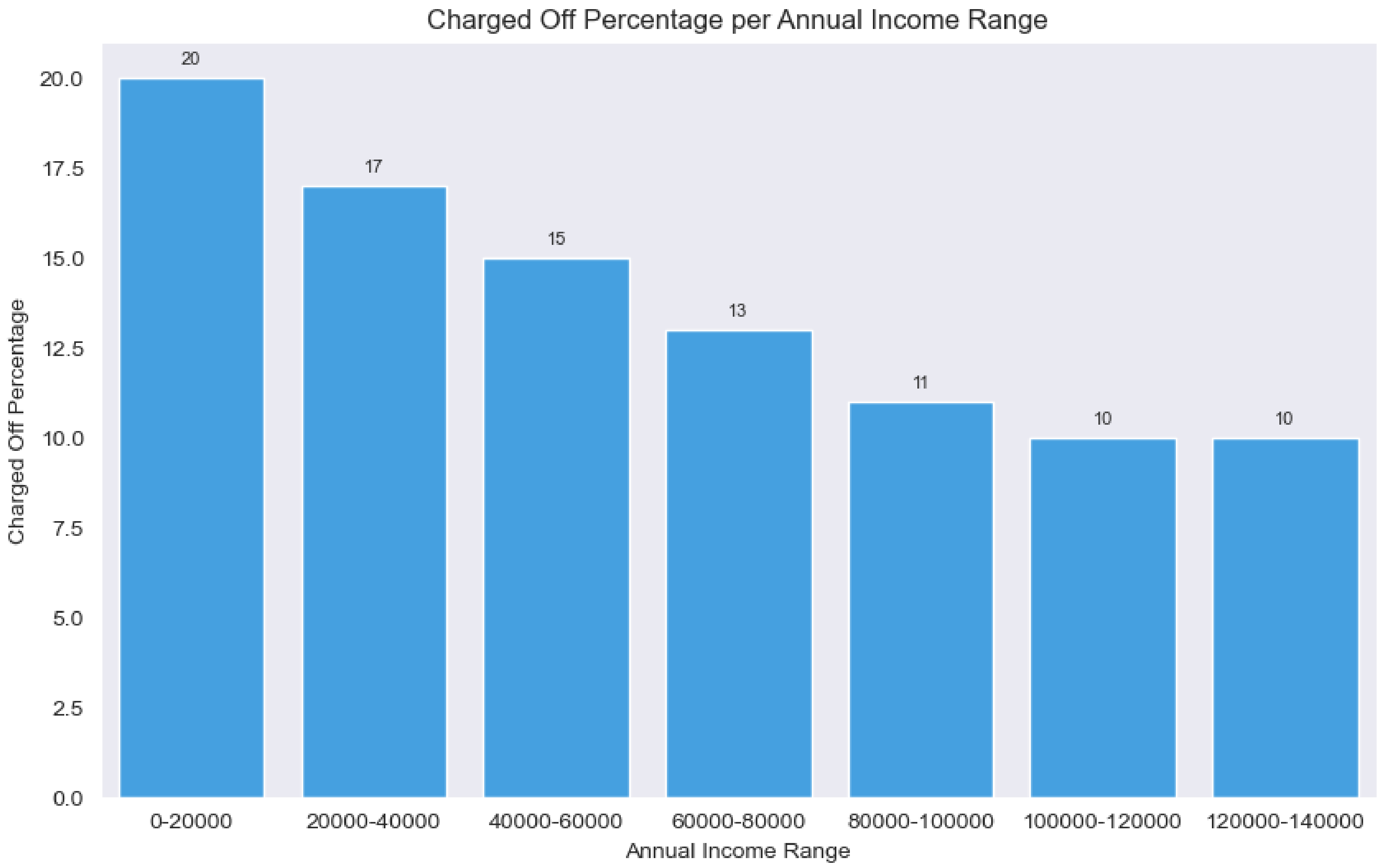
Loan Status Distribution by State



Key Messages:

- *70% of loans from NE state are the most likely to be charged off, although their overall numbers are quite low*

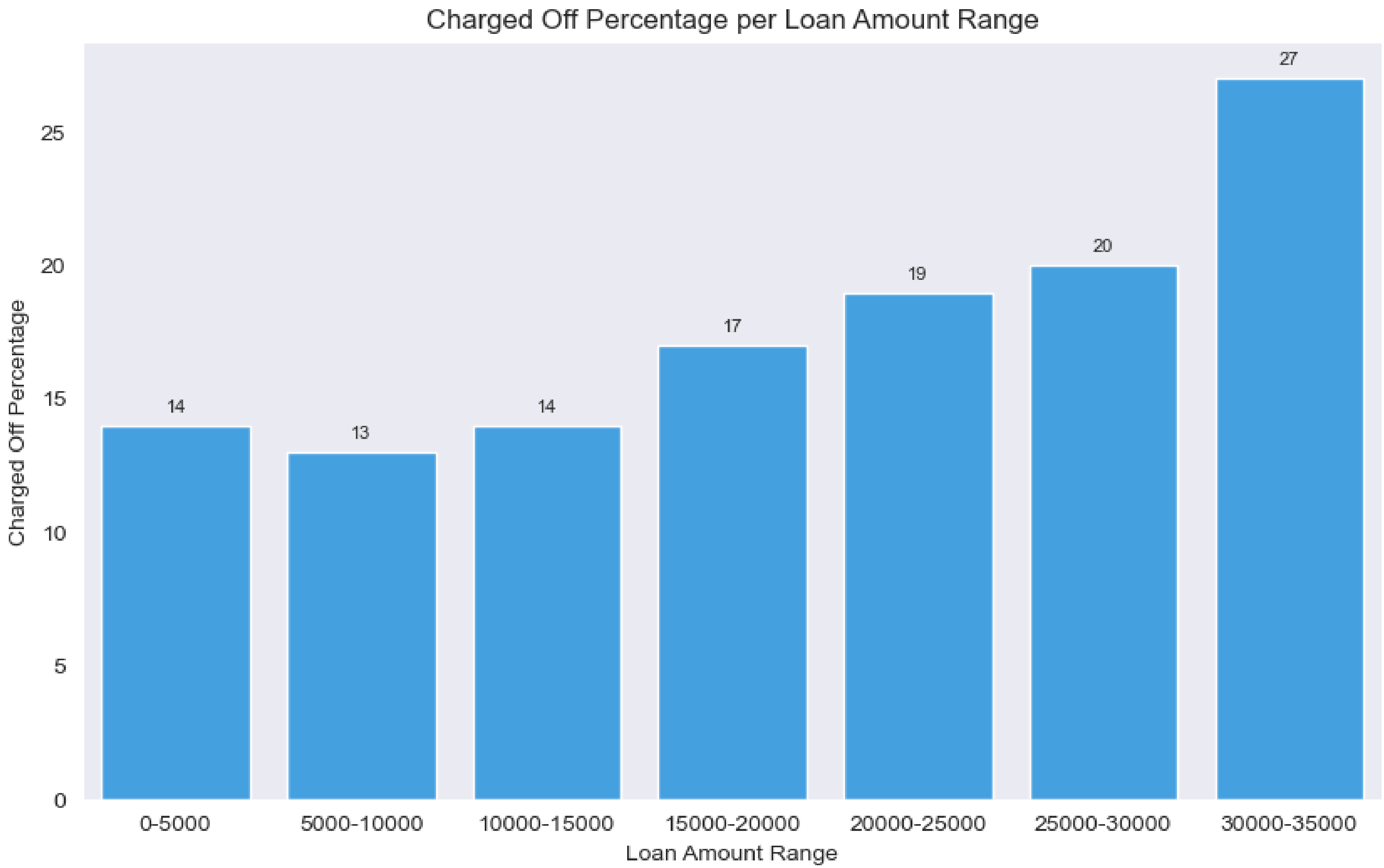
Impact of Annual Income on Loan Default Rates



Key Messages:

- *The charged-off percentage decreases as the annual income increases, highlighting that individuals with higher incomes are less likely to default on loans.*
- *Significant drop in default rates as income increases from 0-20K to 100-140K, indicating financial stability and better loan repayment capability in higher income brackets.*

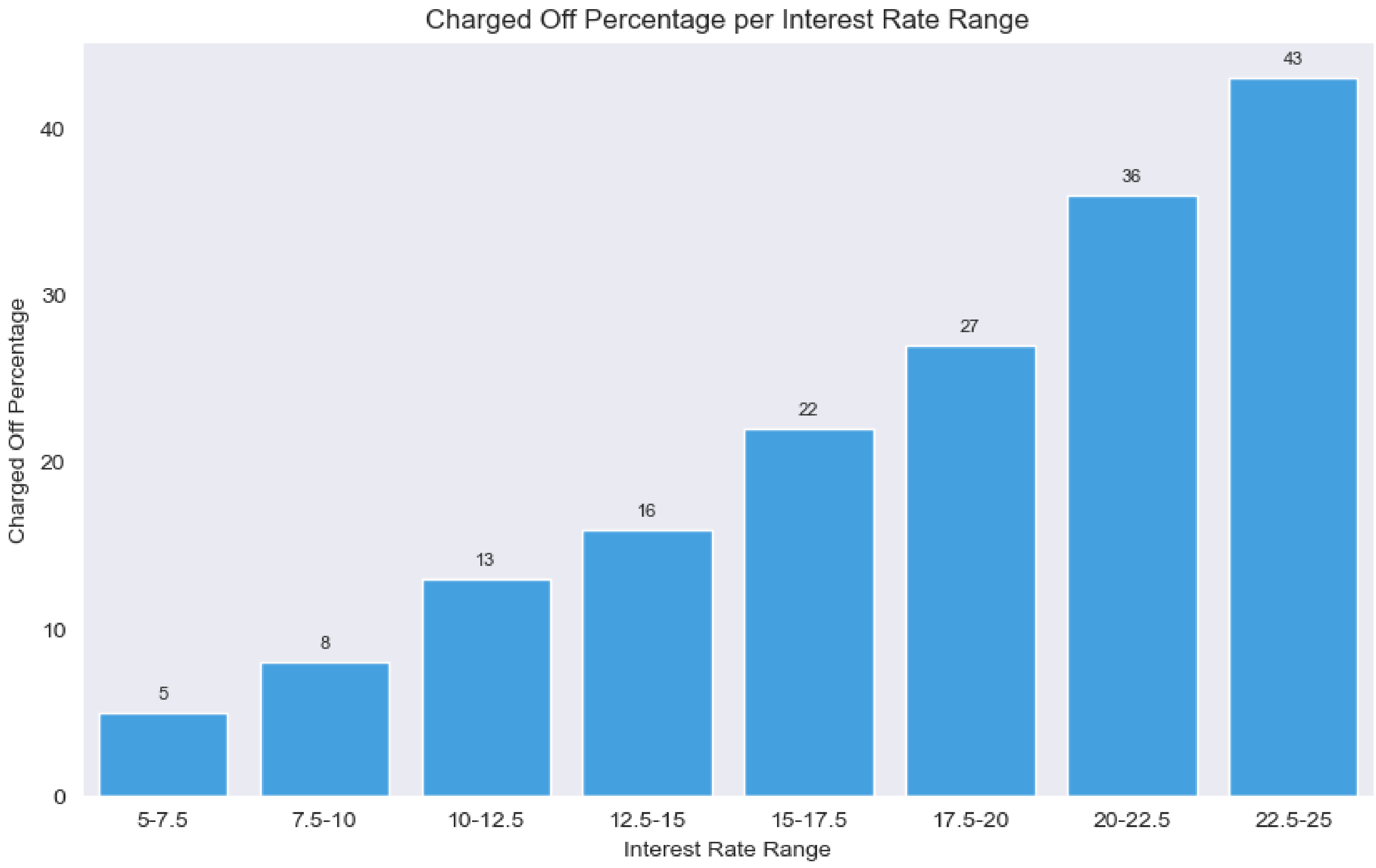
Impact of Loan Amount on Loan Default Rates



Key Messages:

- *As loan amounts increase, so does the charged-off percentage, indicating higher risk with larger loans.*
- *The **30,000-35,000** range has the highest charged-off rate at **27%**.*

Impact of Interest Rate on Loan Default Rates

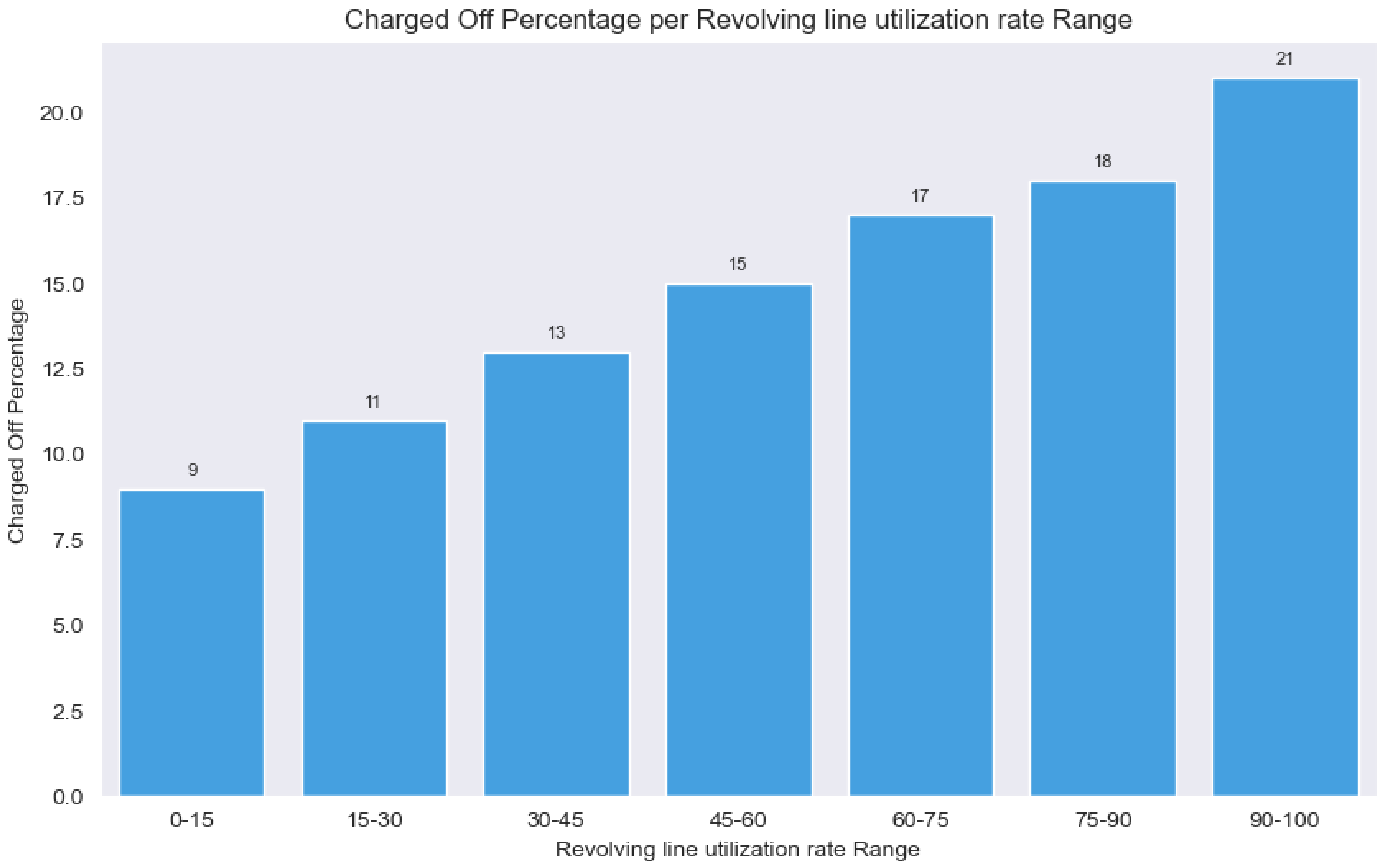


Key Messages:

- *Charged off percentages clearly rise as interest rate ranges increase, showing higher interest rates are linked to more charged off loans.*

- ***43%** Charged off in the 22.5-25% interest rate range.*

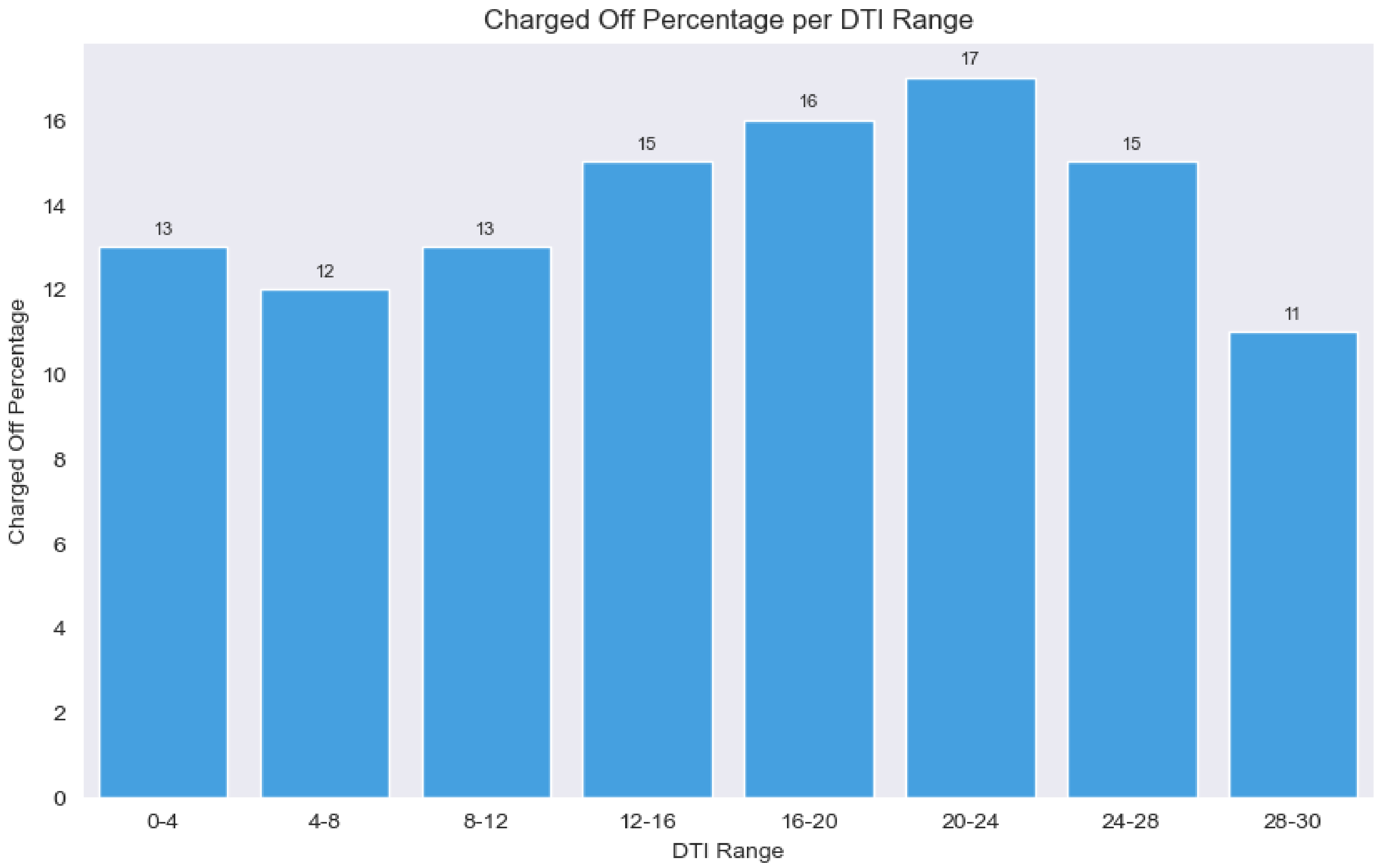
Impact of Revolving Line Utilization Rate on Loan Default Rates



Key Messages:

- *As revolving line utilization rate increases, so does the charged-off percentage.*
- *Highest Charged Off is Observed in the* ***90-100%*** *utilization range at* ***21%***

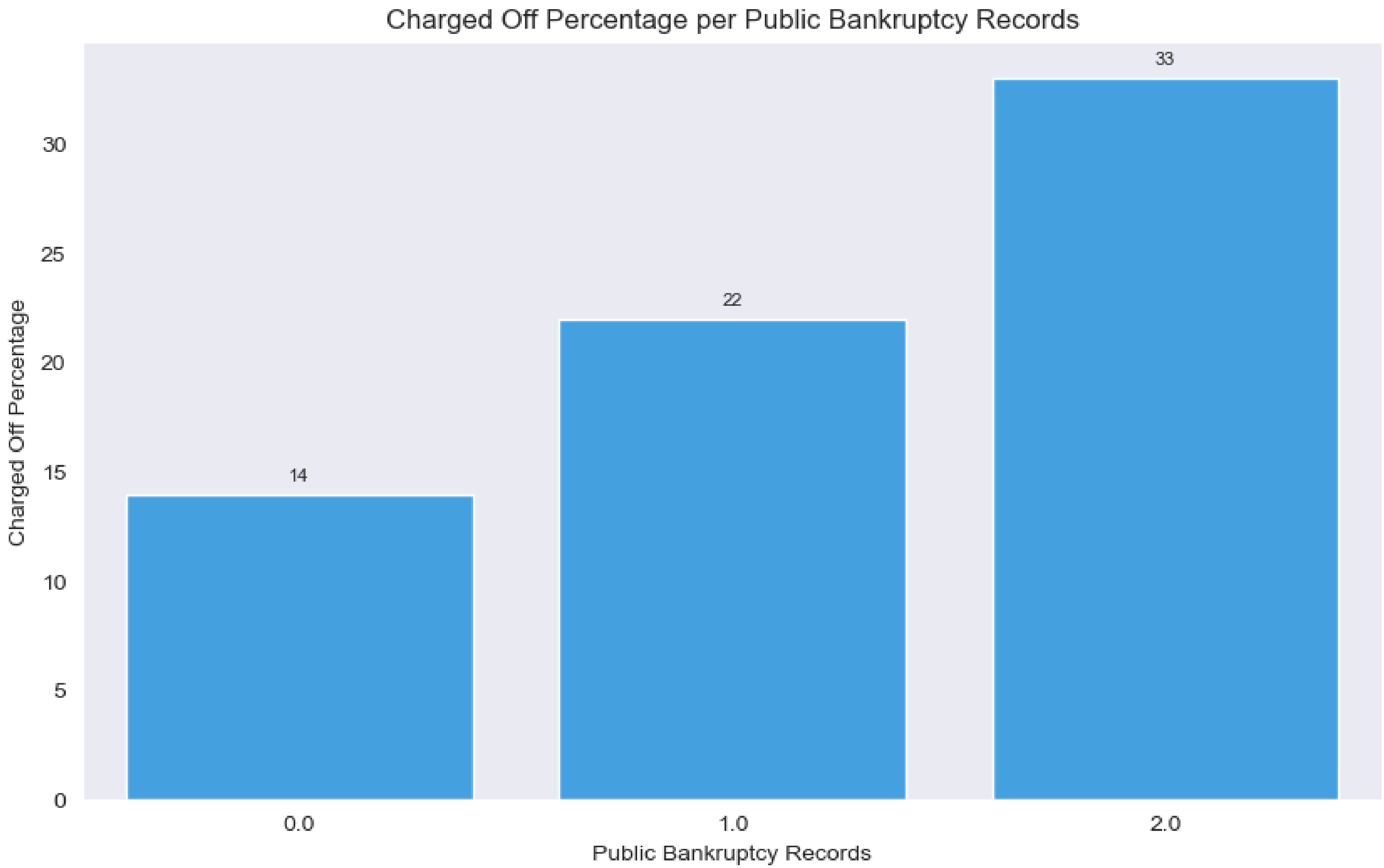
Impact of DTI on Loan Default Rates



Key Messages:

- *The **20-24** DTI range has the highest charged-off rate at **17%.***
- *Moderate DTI indicate higher default risk, while extremely high DTIs might have other influencing factors that reduce risk.*

Impact of Public Bankruptcy Records on Loan Default Rates



Key Messages:

- *The charged-off percentage increases as the number of public bankruptcy records rises.*
- *Borrowers with more public bankruptcy records are at a significantly higher risk of having their loans charged off.*

Key Drivers for Loan Default

- **High Interest Rates:** The highest default rates correlate strongly with higher interest rates.
- **High Revolving Line Utilization Rates:** Increased utilization rates on revolving lines are linked to higher default percentages.
- **Lower Loan Grades:** Loans with lower grades (e.g., G and F) have significantly higher default rates.
- **Public Bankruptcy Records:** More public bankruptcy records indicate a higher risk of loan default.
- **Higher Loan Amounts:** Larger loan amounts generally have a higher default risk.
- **Loan Purpose:** Specific purposes, such as small business loans, show higher default rates.
- **Loan Term:** Longer loan terms, like 60 months, have a higher likelihood of default compared to shorter terms.
- **Annual Income:** Lower annual income levels are associated with higher default rates.