Lending Club Case Study

ANSHUMAN SINGH
SOURABH MAHESHWARI

Problem Statement & Business Objective – Lending Club Case Study

Problem Statement

Lending club is consumer finance company which specializes in lending various types of loans to urban customers.

There are 2 types of risks associated with company's decision of approving or rejecting a customer's loan application.

- 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.

Business Objective

To identify strong driving factors which are potential indicators of customers defaulting and the loan getting charged off. These factors will be used to identify risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss.

Data Set

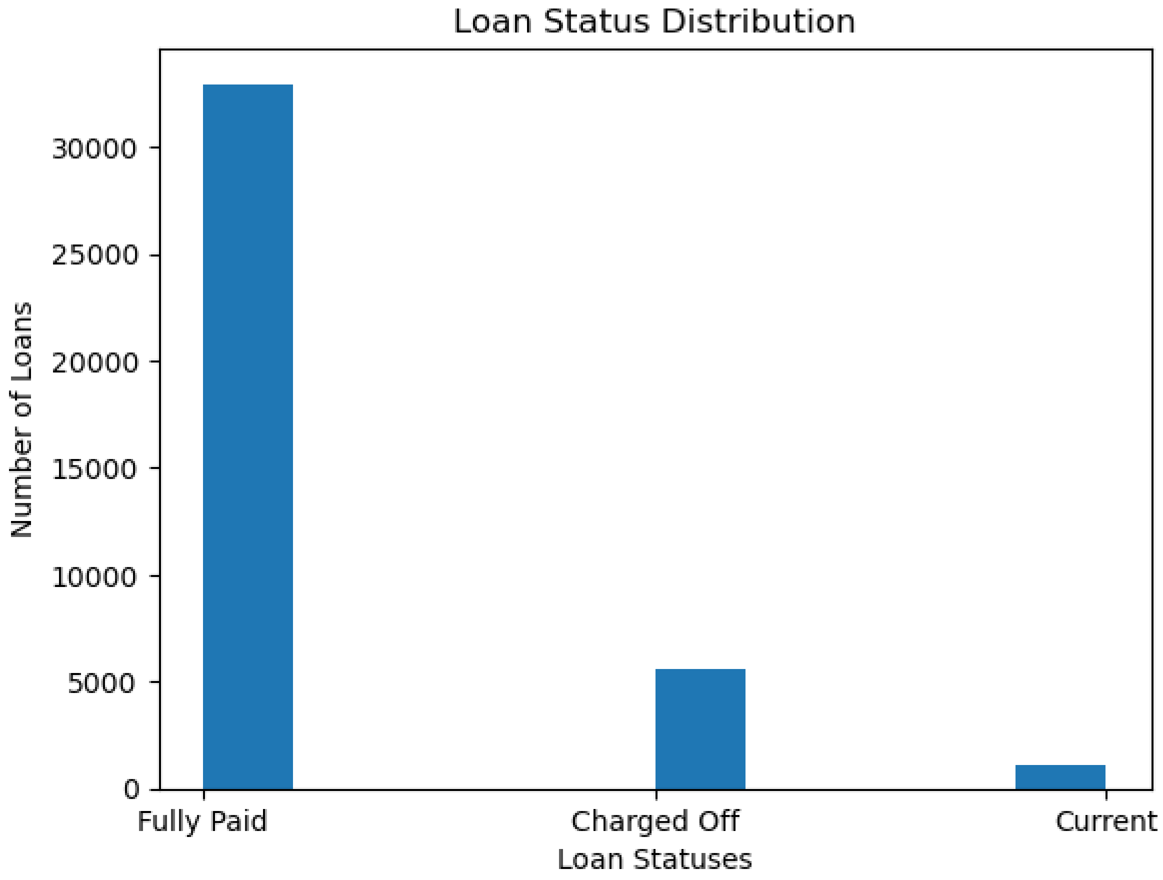
A csv containing data about previous loan applicants have been given (including whether they defaulted or not) along with data dictionary describing the columns.

There are **39,717 records** in the data set with each record containing 111 columns.

The data set also contains information about whether these applications defaulted or not.

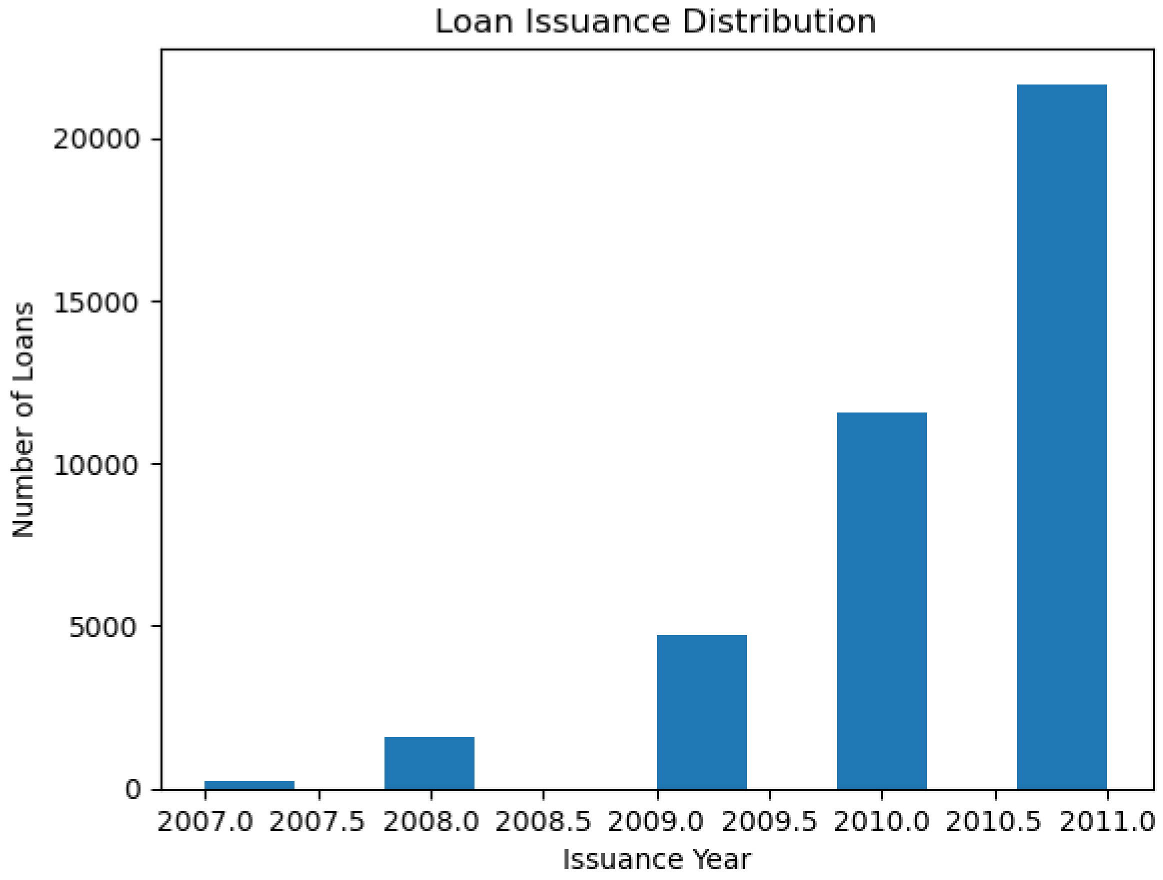
Data Understanding – Overall Default Rate is 14.16 %

- 32,590 *Fully Paid* loans.
- **5,627 loans *charged-off*.**
- 1,140 loans with *current* status.



Data Understanding – Time frame of Data – Jun-2007 – Dec 2011

Issuance Year	No. of Loans
2007	251
2008	1562
2009	4716
2010	11532
2011	21656

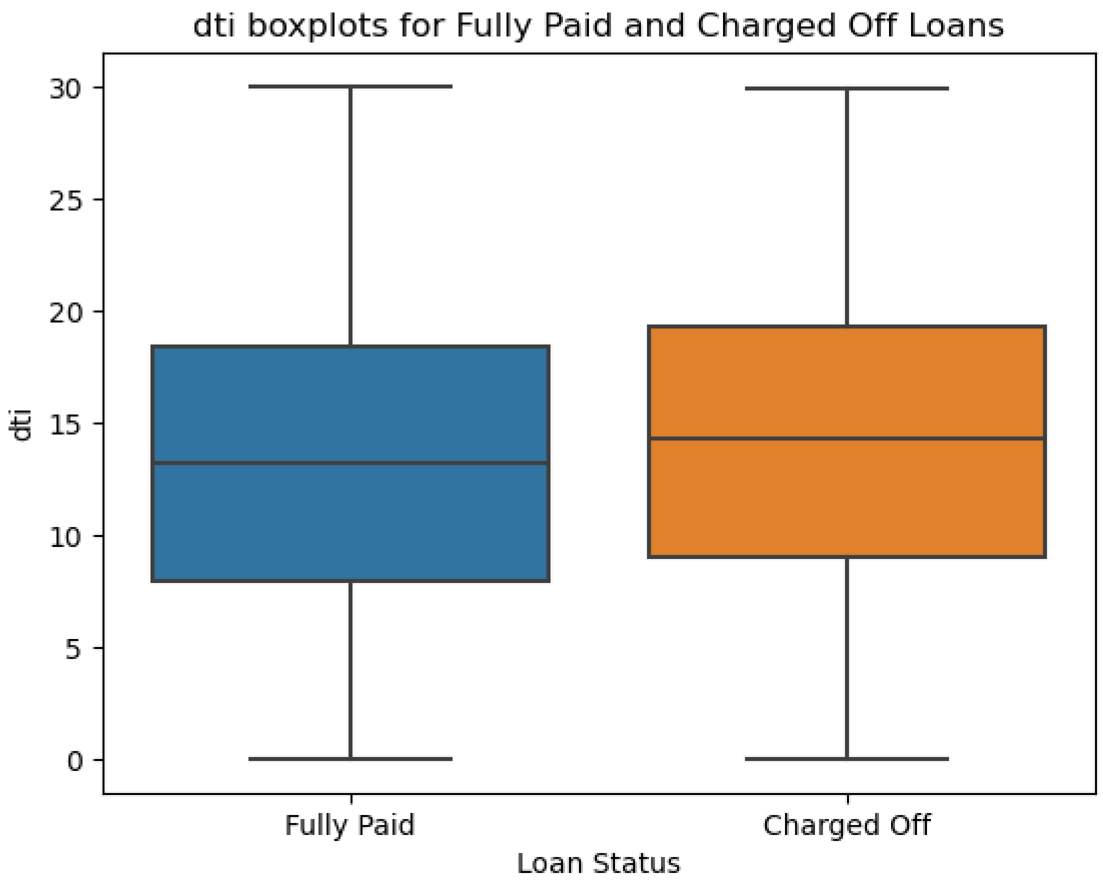


Data Analysis – High **Debt-to-Income (*dti*)** increases the probability of loan getting charged-off

- Median dti value for loans getting charged off is ***14.29***.

- Median dti value for loans getting charged off is ***13.20***.

- Customers with high debt to income (dti) ratio are **more likely to get default on the loan.**

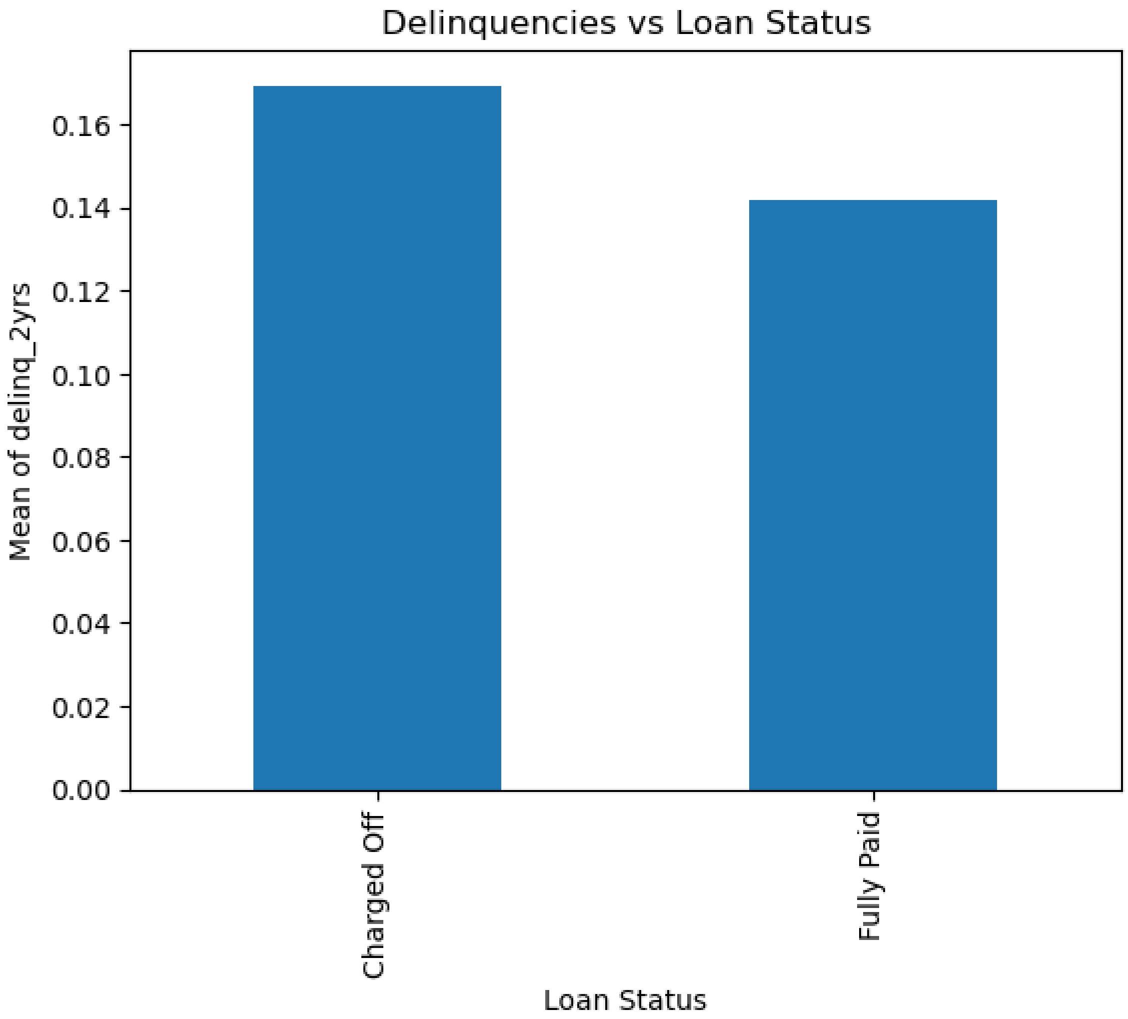


Data Analysis – Higher number of **delinquencies** leads to increased probability of loan default

- Data set has a column titled **delinq_2yrs**. It indicates the number of 30+ days past-due incidences of delinquency in the borrower's credit file for the past 2 years.

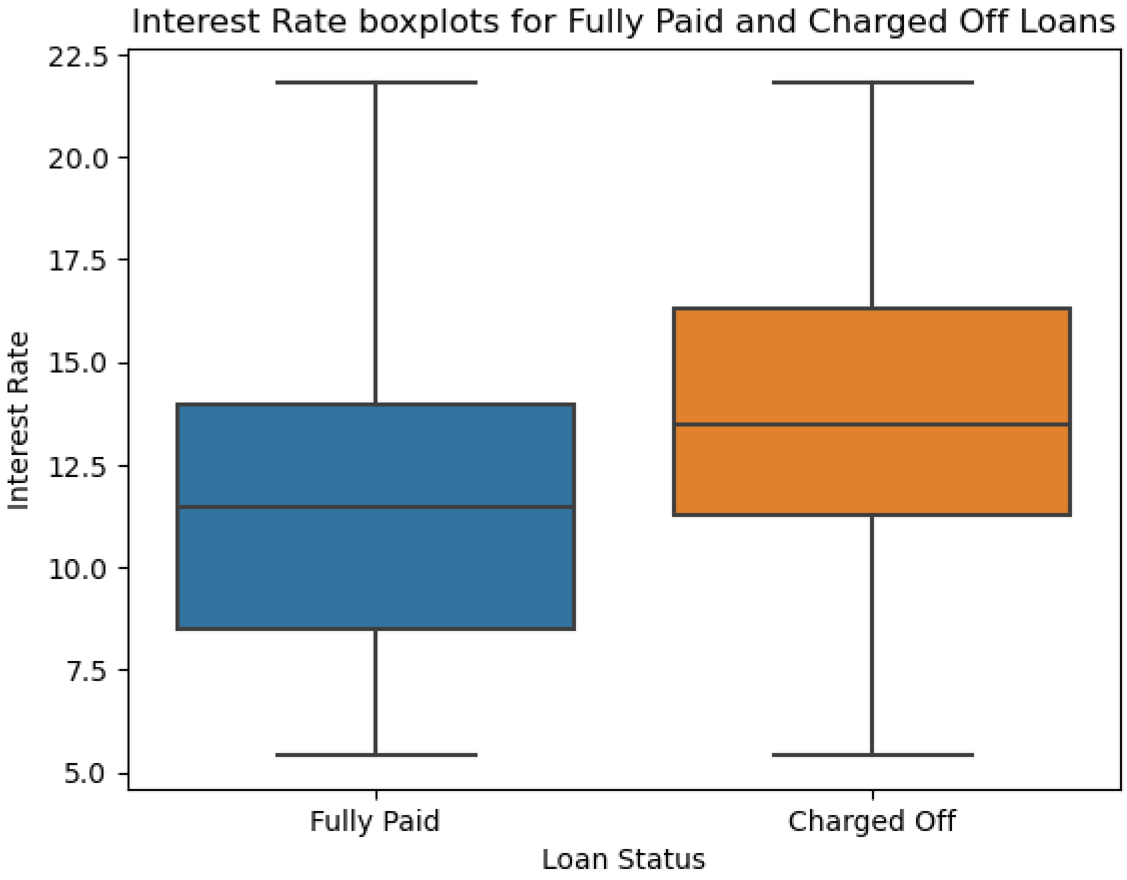
- The average value of delinq_2yrs for charged off loans is ***0.17*** whereas this value is ***0.14*** for all Fully Paid loans.

- This indicates that a **customer with a record of delinquencies in the past is more likely on default** on the loan.



Data Analysis – High *interest_rate* increases the probability of loan getting charged-off

- Median interest_rate value for loans getting charged off is ***13.49***.
- Median interest_rate value for loans getting charged off is ***11.49***.
- Loans where interest is higher are relatively **more likely to get charged-off.**

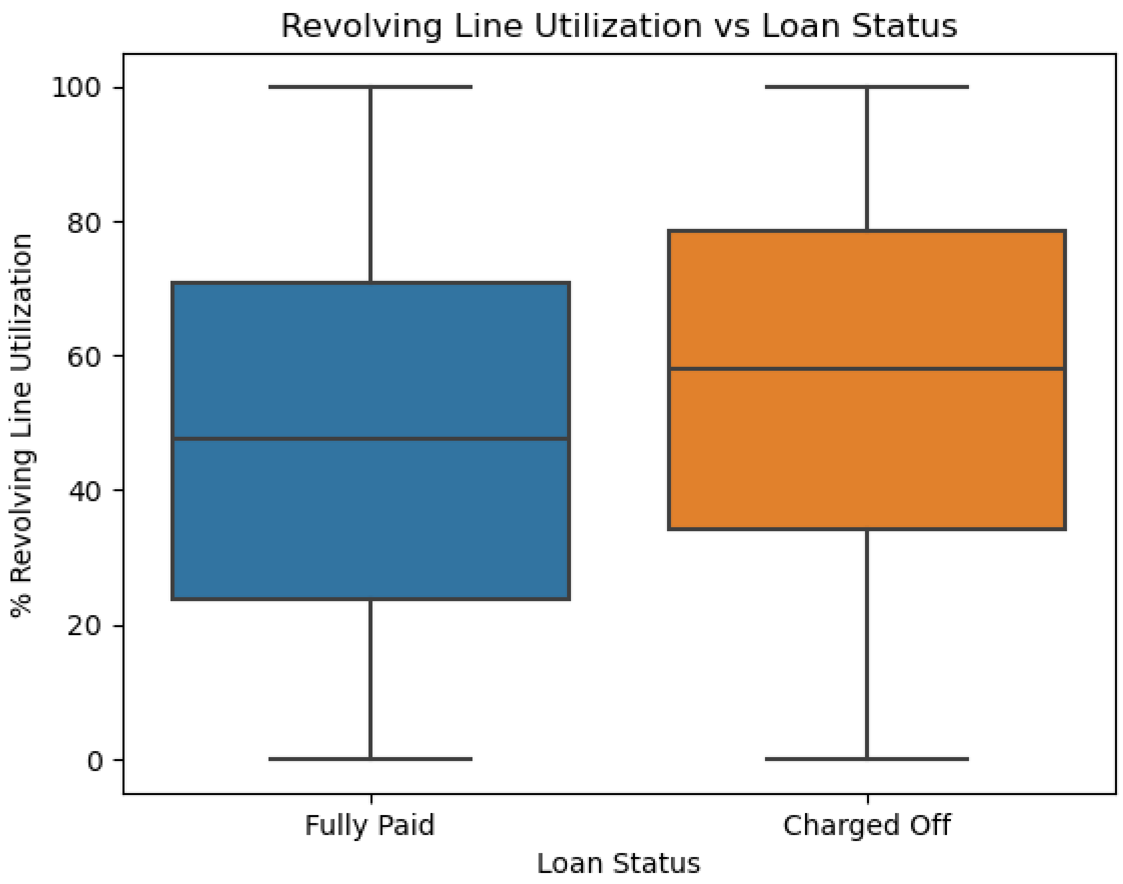


Data Analysis – Higher **revolving line utilization %** indicates higher probability of loan default

- Data set has a column titled **revol_util**. It indicates revolving line utilization rate, or the amount of credit the borrower is using relative to all available revolving credit.

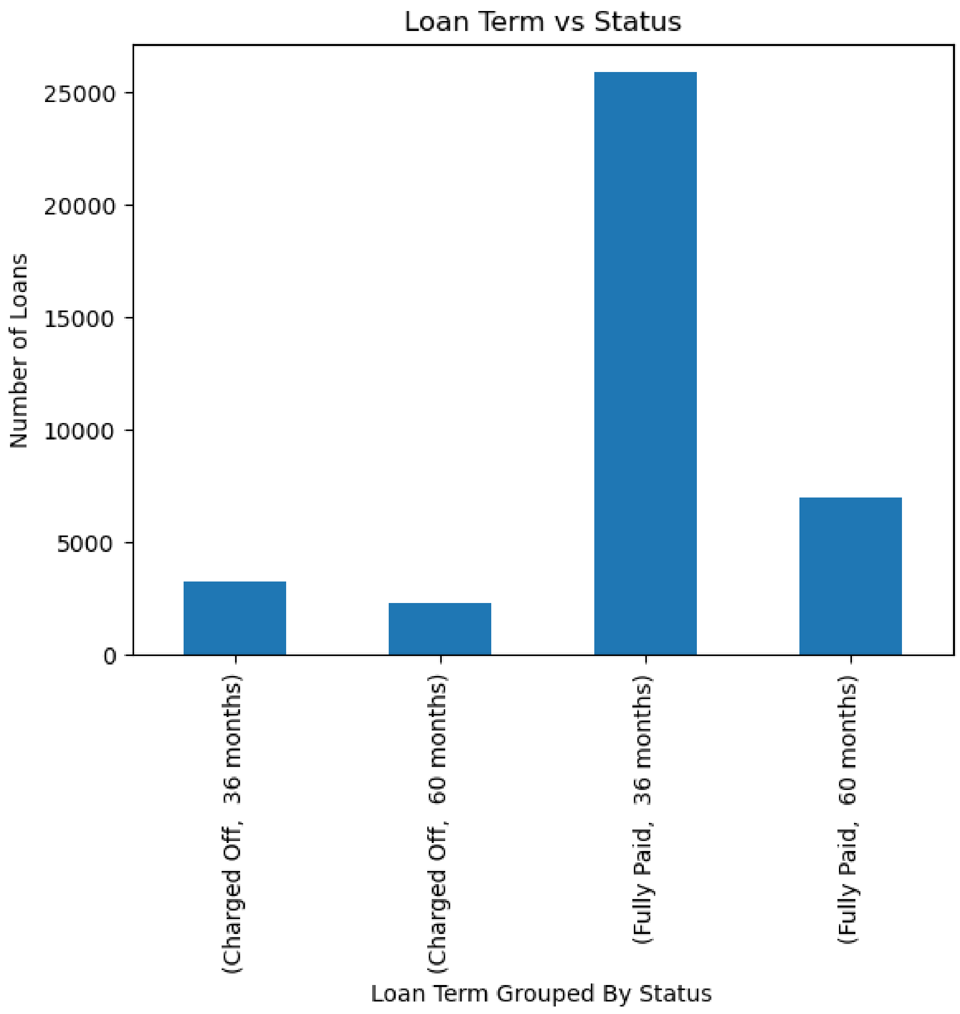
- The median value of revolving line % utilization for charged off loans is ***58%*** whereas this value is **47%** for all Fully Paid loans.

- This indicates that a **customer with higher % utilization of revolving credit line is more likely on default** on the loan.



Data Analysis – % of Loans with 60-month **term** significantly more in charged-off loans (relative to Fully Paid loans)

- Loans have only **2 possible terms** i.e., 36 months and 60 months.
- **Out of all charged-off loans 42% have a 60-month term** whereas out of all Fully Paid loans only 21% loans have 60-month term.
- A 36-month term loan is more likely to get charged off than a 60-month term loan.



Data Analysis Summary - Key Drivers Indicating Loan Default

Variable	Conclusion
Debt to Income Ratio (dti)	Customers with high debt to income (dti) ratio are more likely to get default on the loan.
Delinquencies (delinq_2yrs)	A customer with a record of delinquencies in the past is more likely on default on the loan.
Interest Rate (interest_rate)	Loans where interest is higher are relatively more likely to get charged-off.
Revolving line utilization % (revol_util)	A customer with higher % utilization of revolving credit line is more likely on default on the loan.
Loan Term (term)	A 36-month term loan is more likely to get charged off than a 60-month term loan.