Lending Club Case Study

Group Members
1 – Palash Gupta(Group facilitator)
2 – Nithy

Problem Statement

- 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 labelled as 'charged-off' are the 'defaulters'.

Analysis Approach

- If one is able to identify these risky loan applicants, then bad 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.
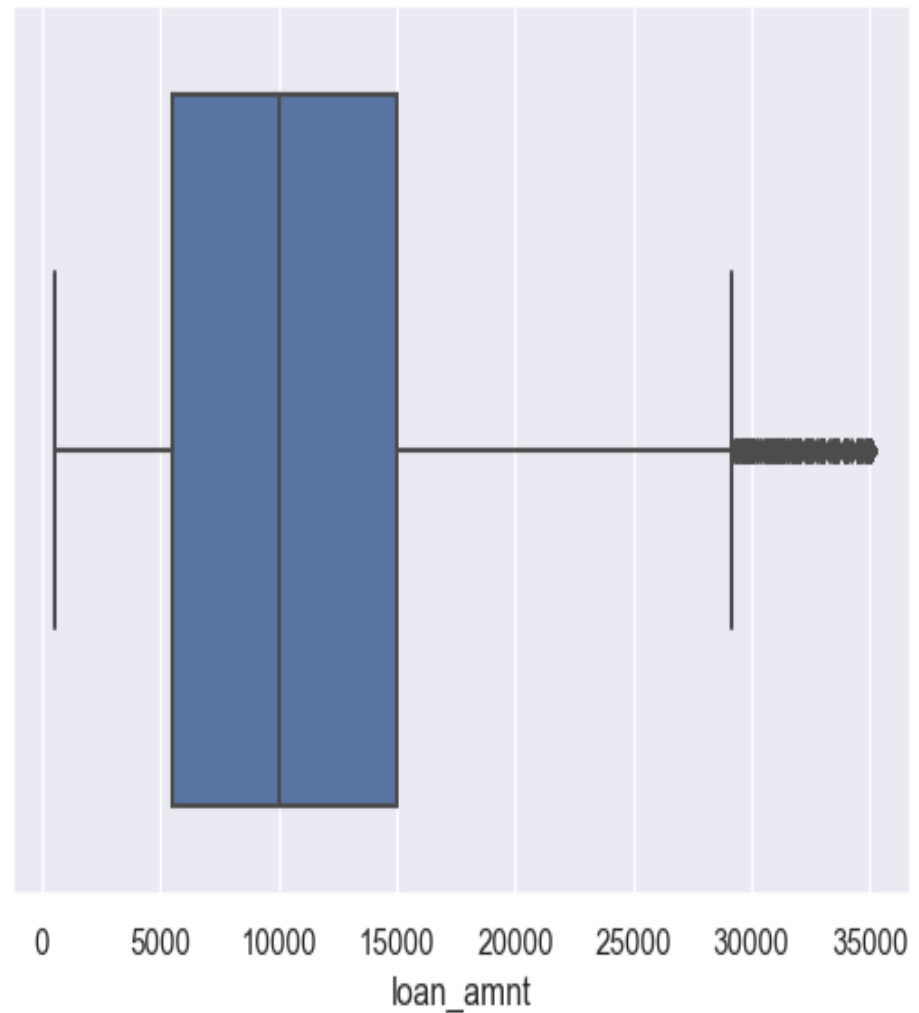
Dataset Used

- The loan dataset contains the complete loan data for all loans issued through the time period 2007 to 2011.

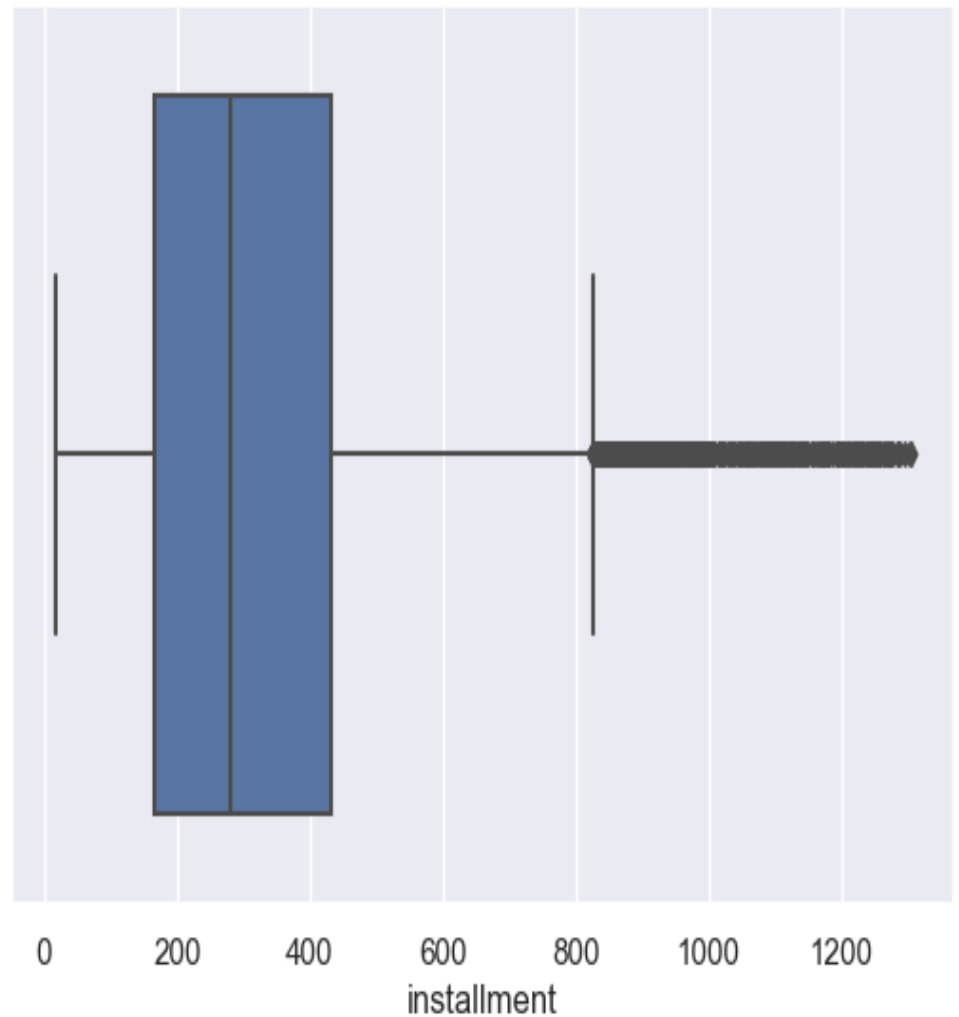# Conclusions from Univariate Analysis of Numerical Columns

Conclusion 1

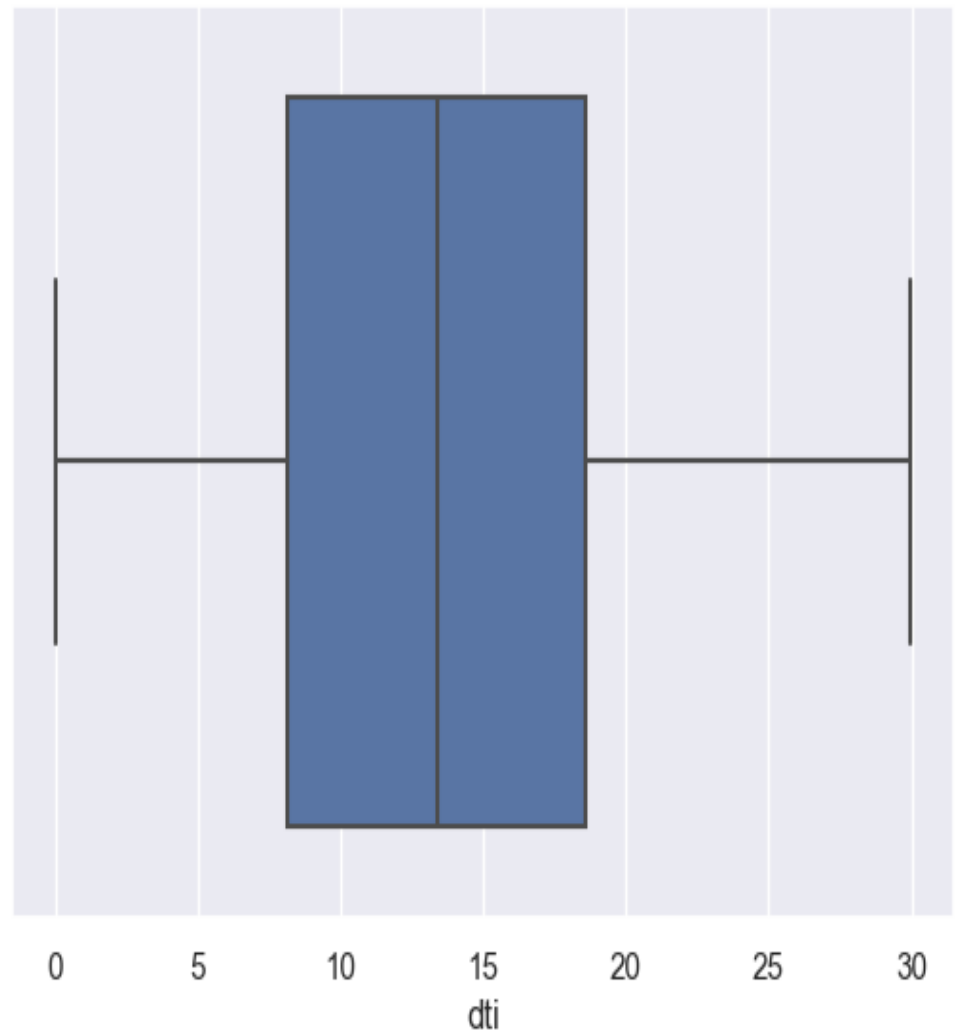Most of the loan
amount lies in range
of 5000-15000

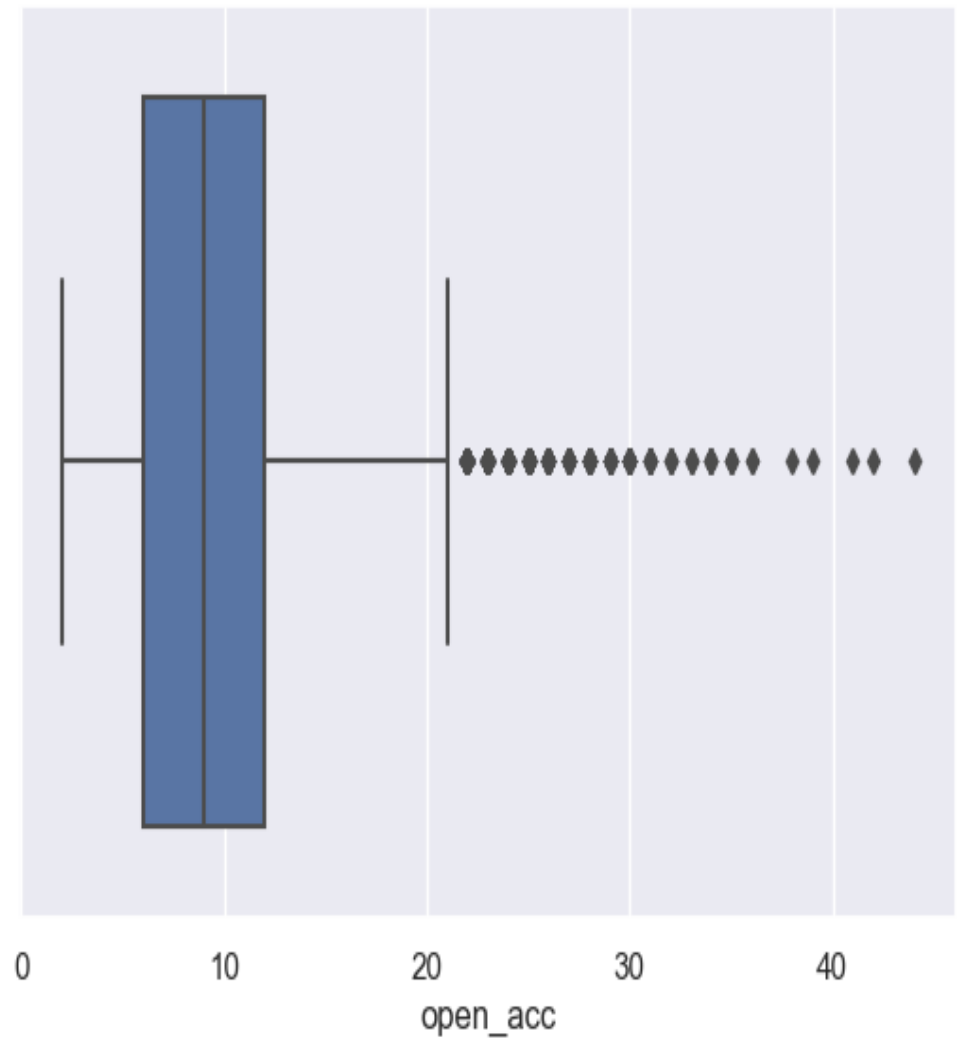**Conclusion 2**

Installments lies in the range of 200-420

Conclusion 3
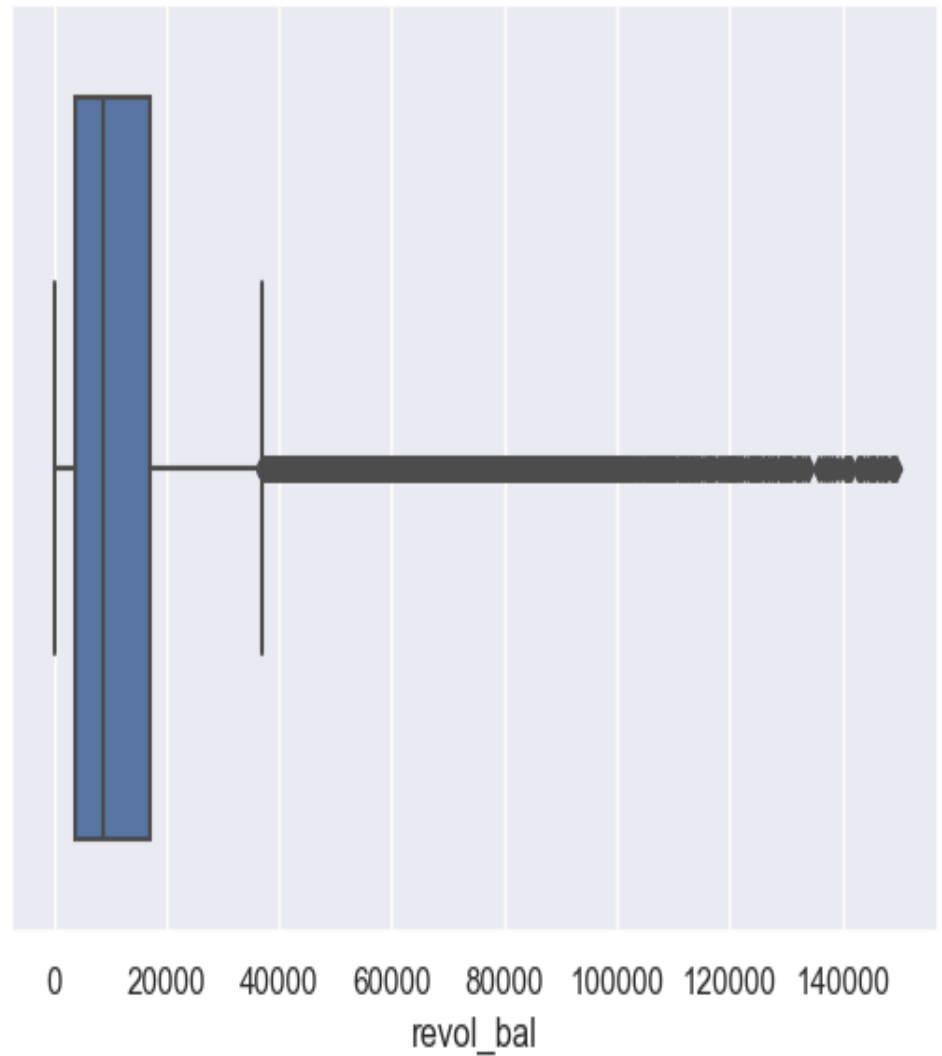
DTI lies in the range of 8-19.

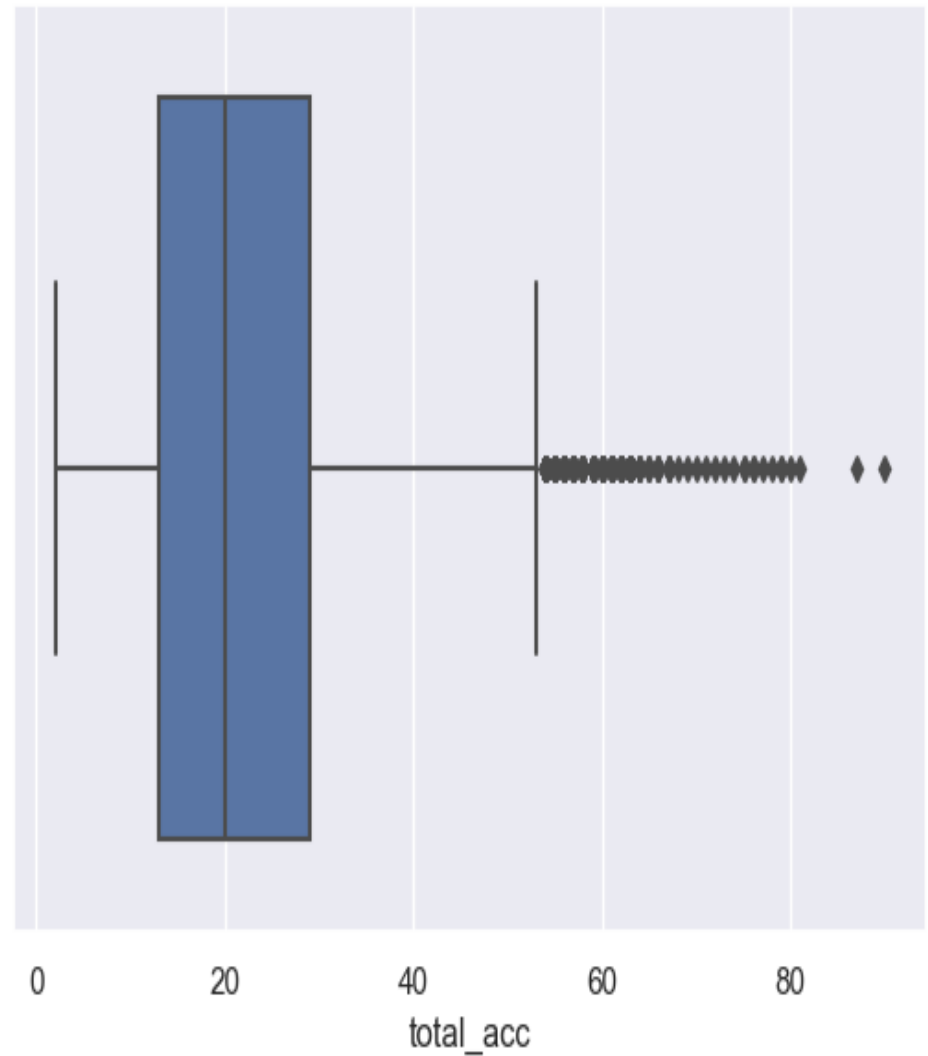**Conclusion 4**

open_acc lies in range
of 6-12.

Conclusion 5

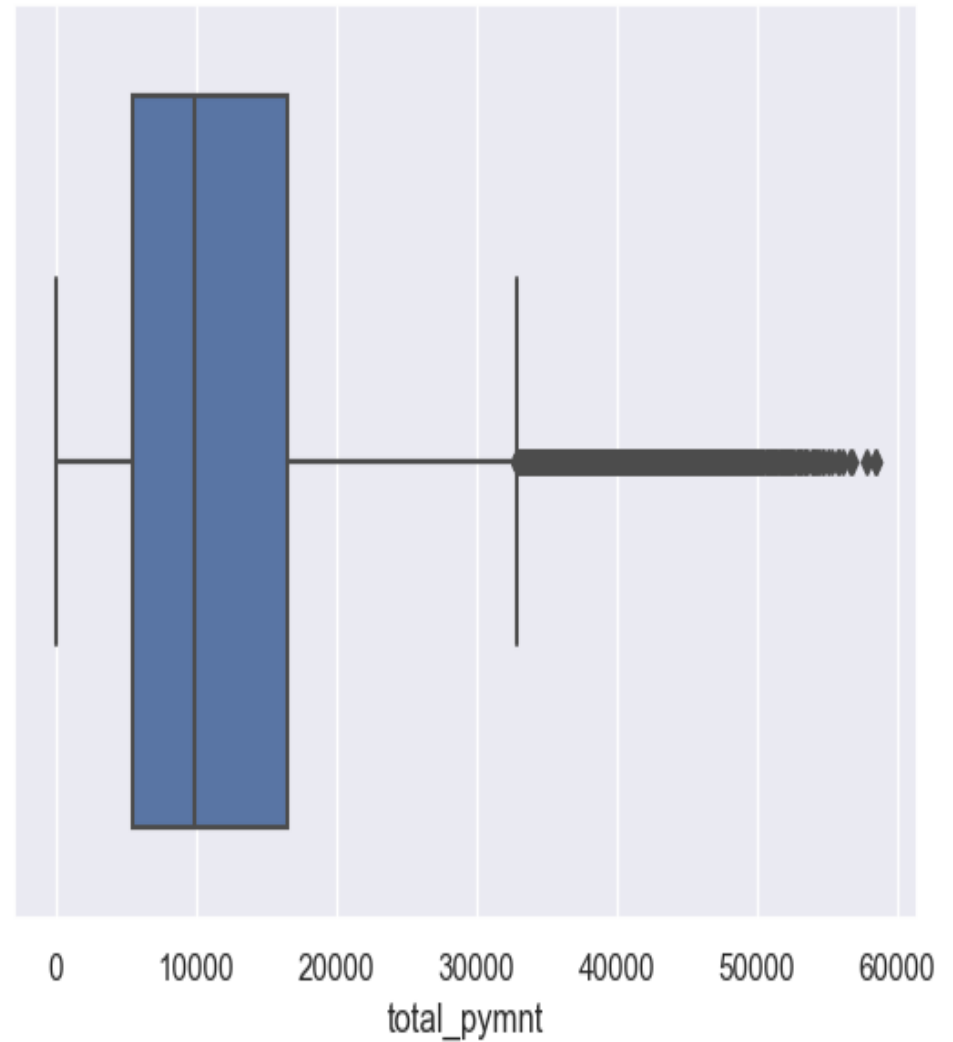Revol_bal lies in the range of 2000-19000.

Conclusion 6

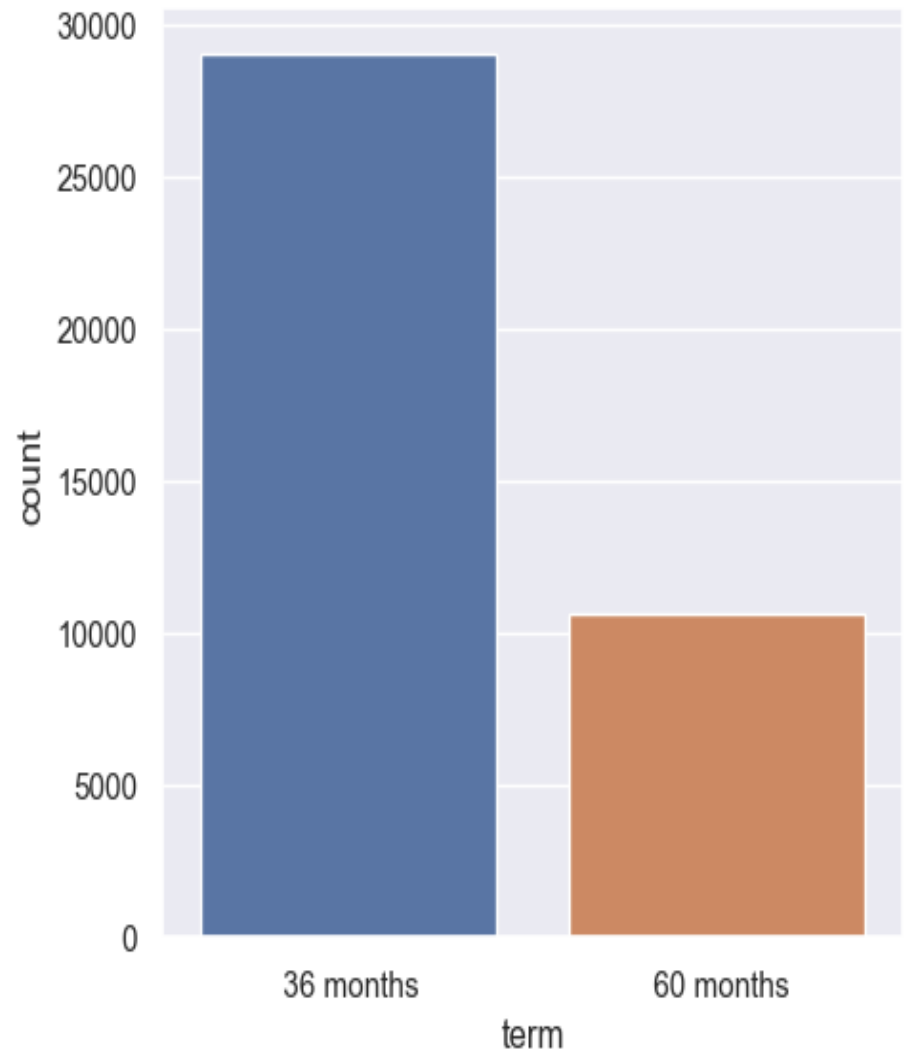total_acc lies in the range of 15-28.

Conclusion 7

total_payment lies in
the range of 7000-
18000.

Conclusions from Univariate Analysis of Categorical columns

Conclusion 1

Most number of loans
are taken for a tenure
of 36 months.

Conclusion 2

Persons belonging to
grade B takes the most
number of loans and
person belonging to
grade G take least
number of loans