LENDING CLUB CASE STUDY ASSIGNMENT

SUBMISSION

Name: RAJAGOPAL SHENOY G

ANINDYA SUDHIR

Problem Statement – Business Objective

- Identify the **risky loan applicants** using EDA, so that such loans can be reduced to cut down the amount of credit loss

- Understand the **driving factors (or driver variables)** behind loan default, i.e. the variables which are strong indicators of default.

Business Understanding

- 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.

- 2 risks associated with banks decision to approve loans.
  - 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

Approach

Data Cleaning

1. Deleting NA and single valued columns
2. Duplicate Check
3. Formatting the data in correct format. Like remove % from interest rate.
4. Converting data to correct data types.

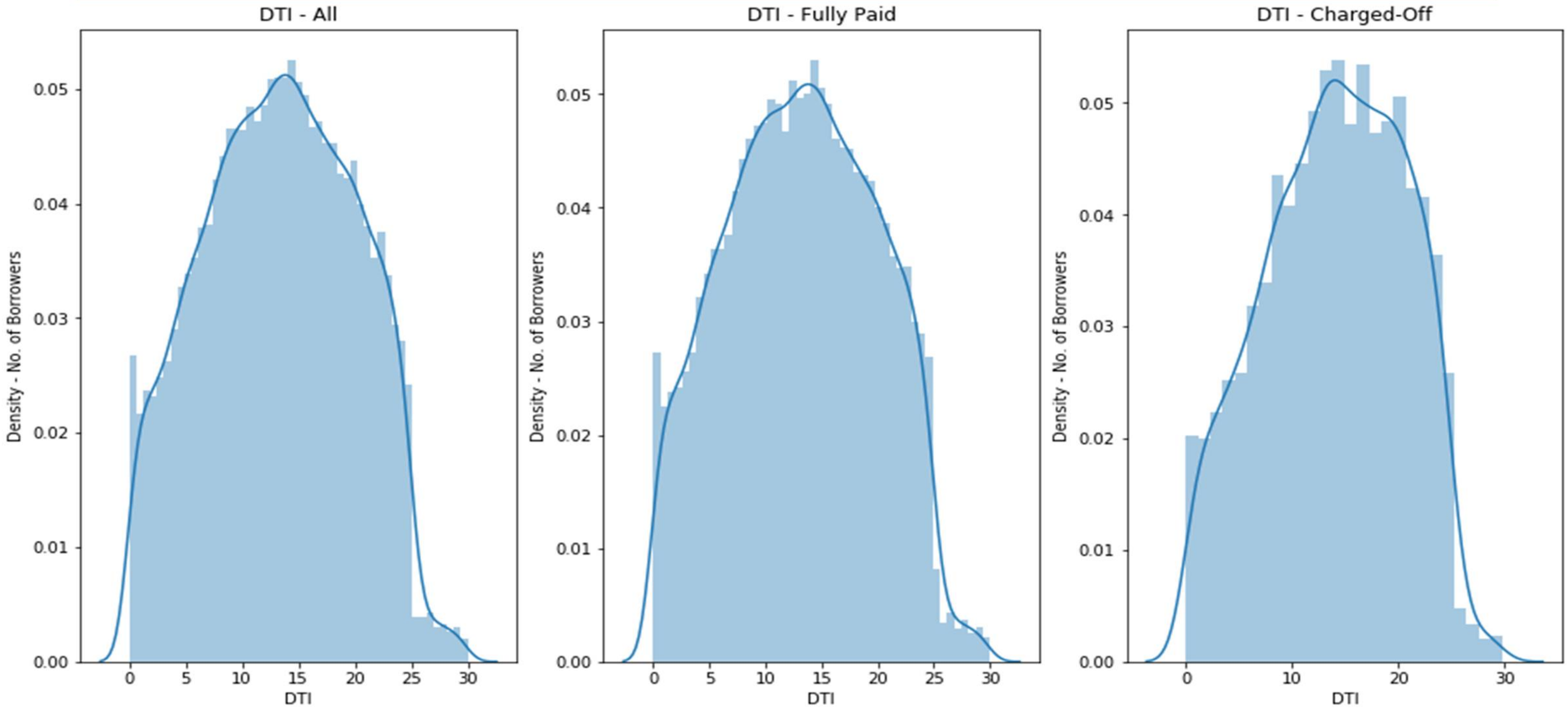
Analysis

1. Taking domain knowledge by understanding Lending Club related loan and financial terms and processes.
2. Exploratory Data Analysis performed:
   - Univariate and Segmented Univariate
   - Bivariate Analysis on various continuous and categorical variables

Data Plot

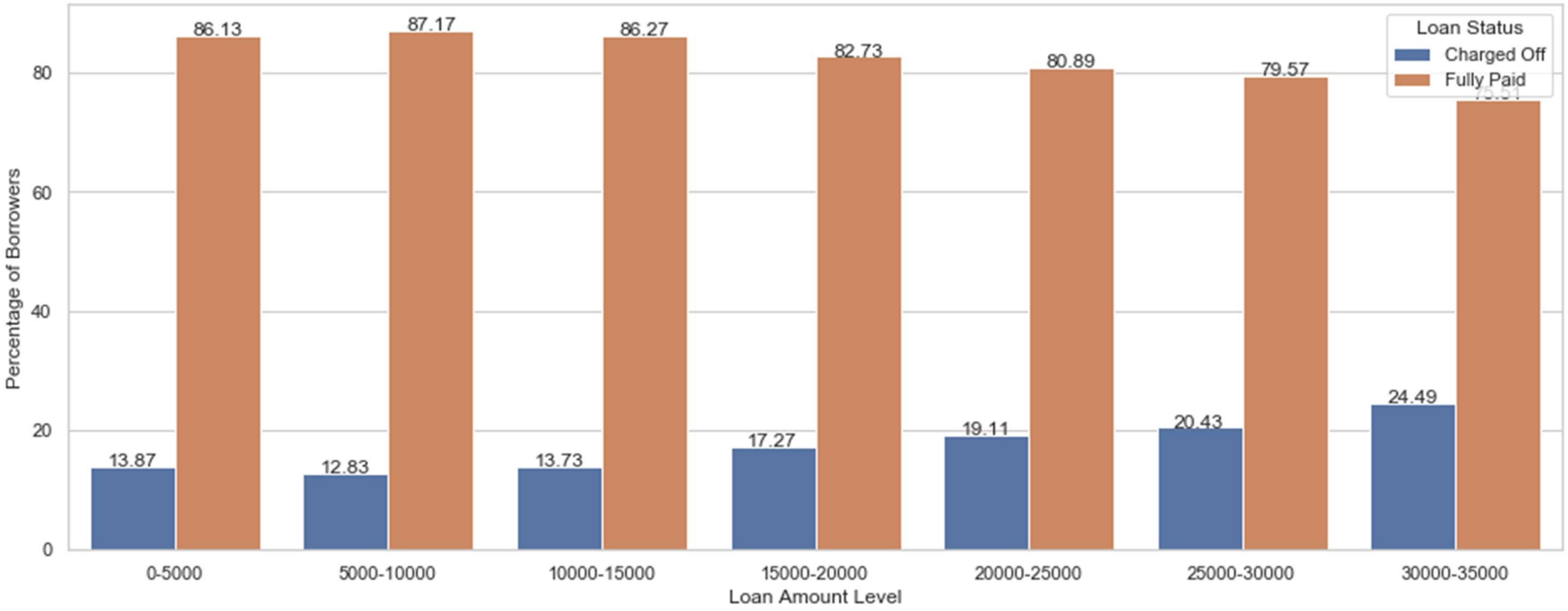
Visualizing results from the analysis in the form of bar charts, distplot etc to have a clear understanding of it.

Plot – DTI vs Loan Status



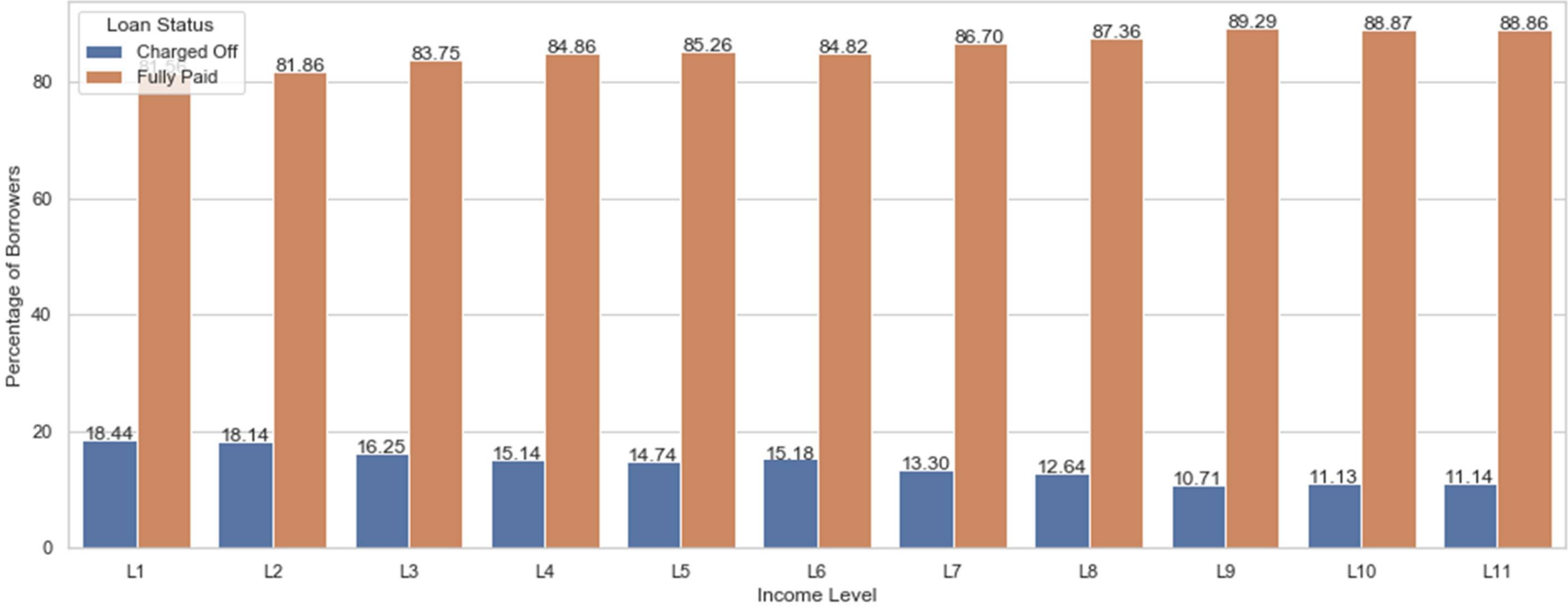
- The Upper Critical Value for Fully Paid normal curve: 13.21 considering 5% α
- The Lower Critical Value for Charged Off normal curve: 13.86 considering 5% α
- Borrowers with DTI lesser than 13.21 are more likely to pay off their loans. Borrowers with DTI higher than 13.86 are more likely to default.

Plot – Loan Amount vs Loan Status



- Borrowers with loan amount > 15,000 USD tend to default, those with < 15,000 USD tend to pay
- Also, borrowers who apply for bigger amounts of loan have relatively higher chances of defaulting

Plot – Annual Income vs Loan Status

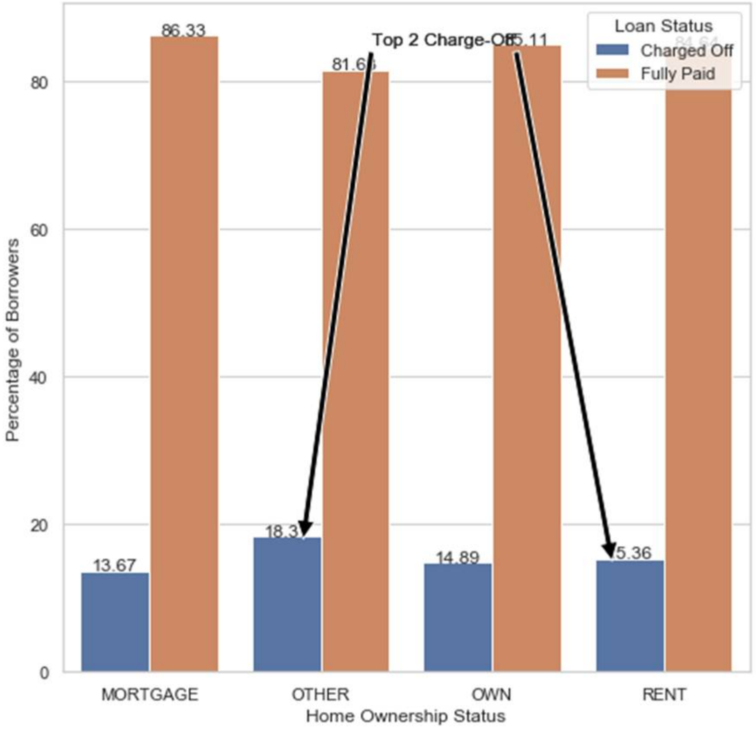


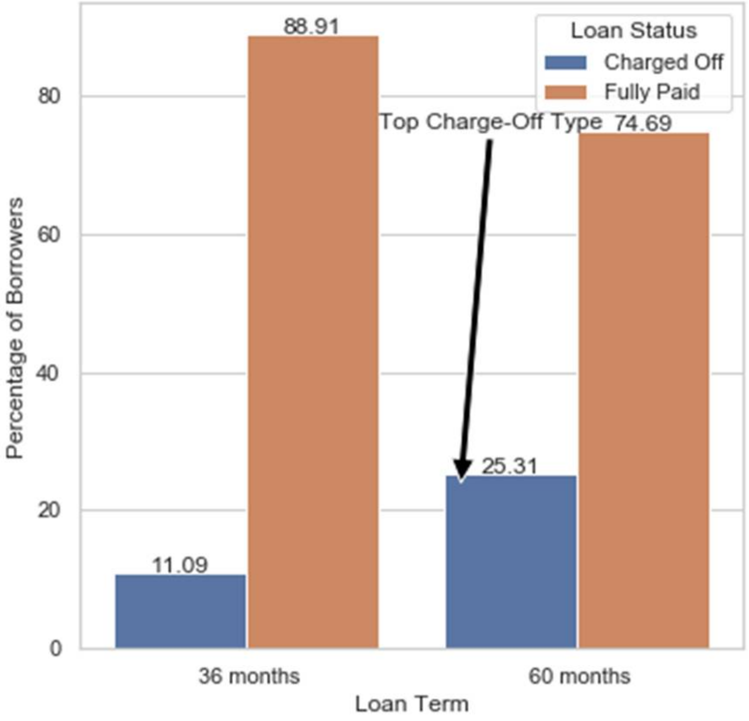
- Borrowers with annual income < 55,000 USD tend to default and those with > 65,000 USD tend to pay
- Also, borrowers with higher annual income have relatively lesser chances of defaulting.

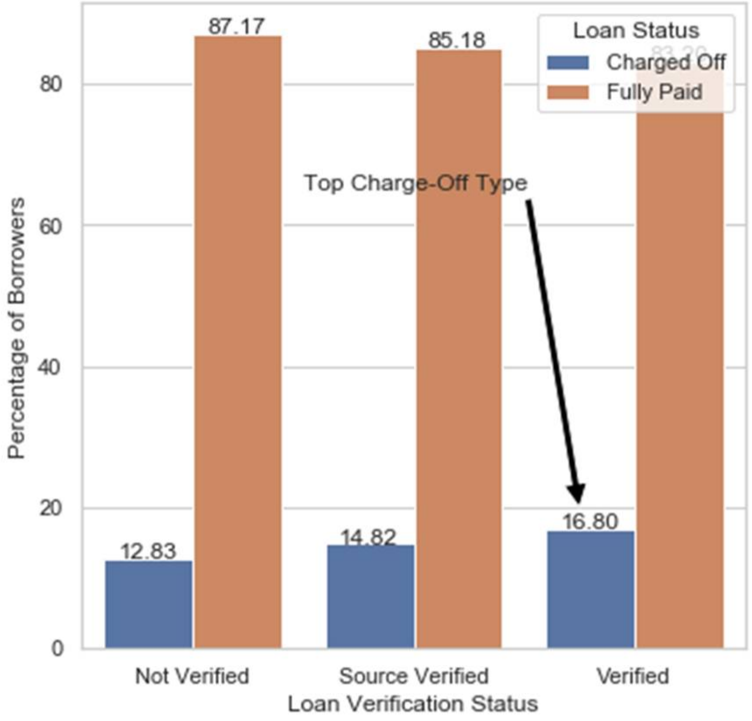
Note - L1 to L11 represent [4000, 30000.0, 37196.16, 44500.0, 50004.0, 58868.0, 65004.0, 75000.0, 90000.0, 115000.0, 234144, 6000000] USD.

Plot – Home Ownership, Loan Term, Verification Status vs Loan Status

UpGrad

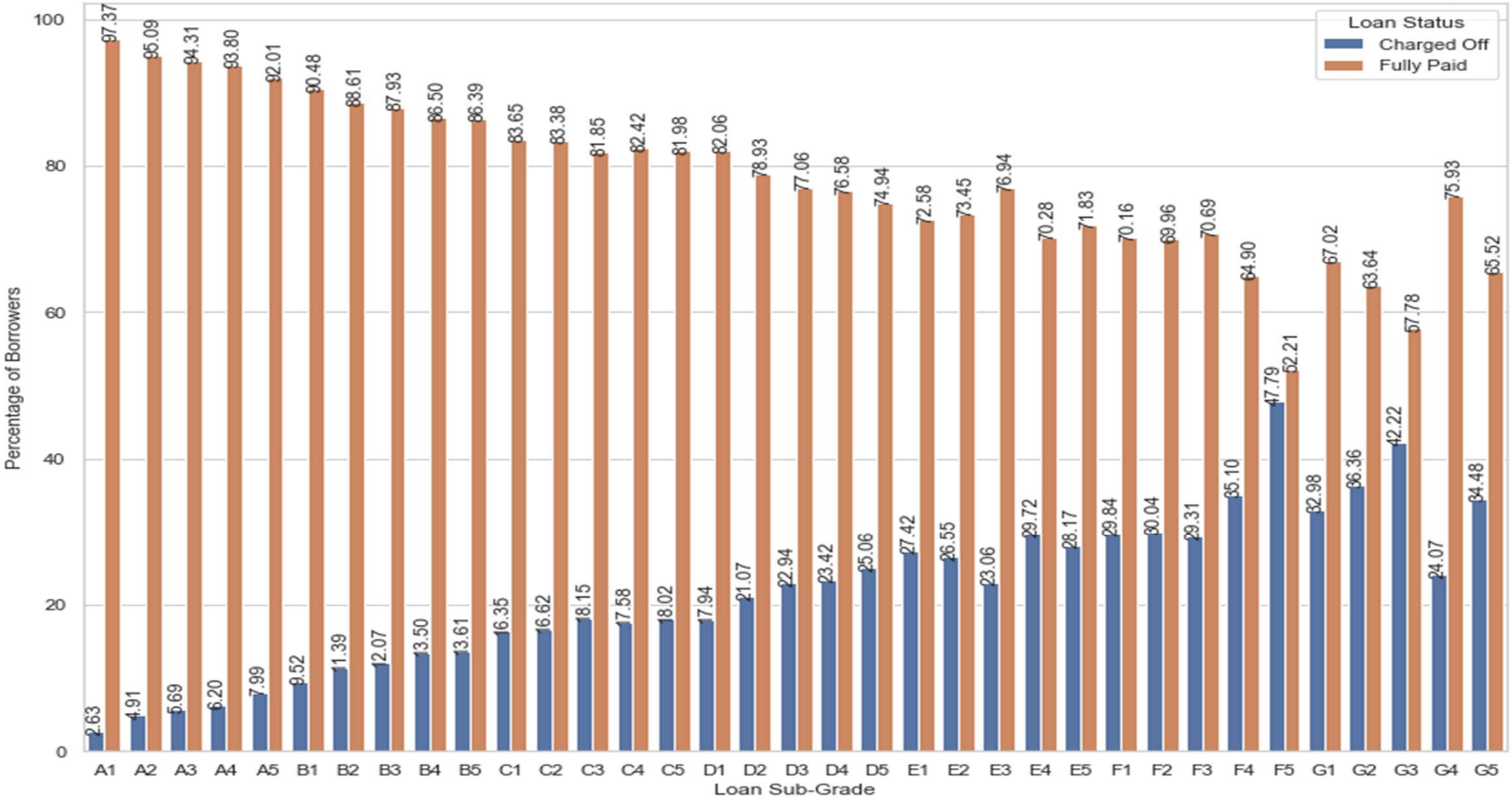






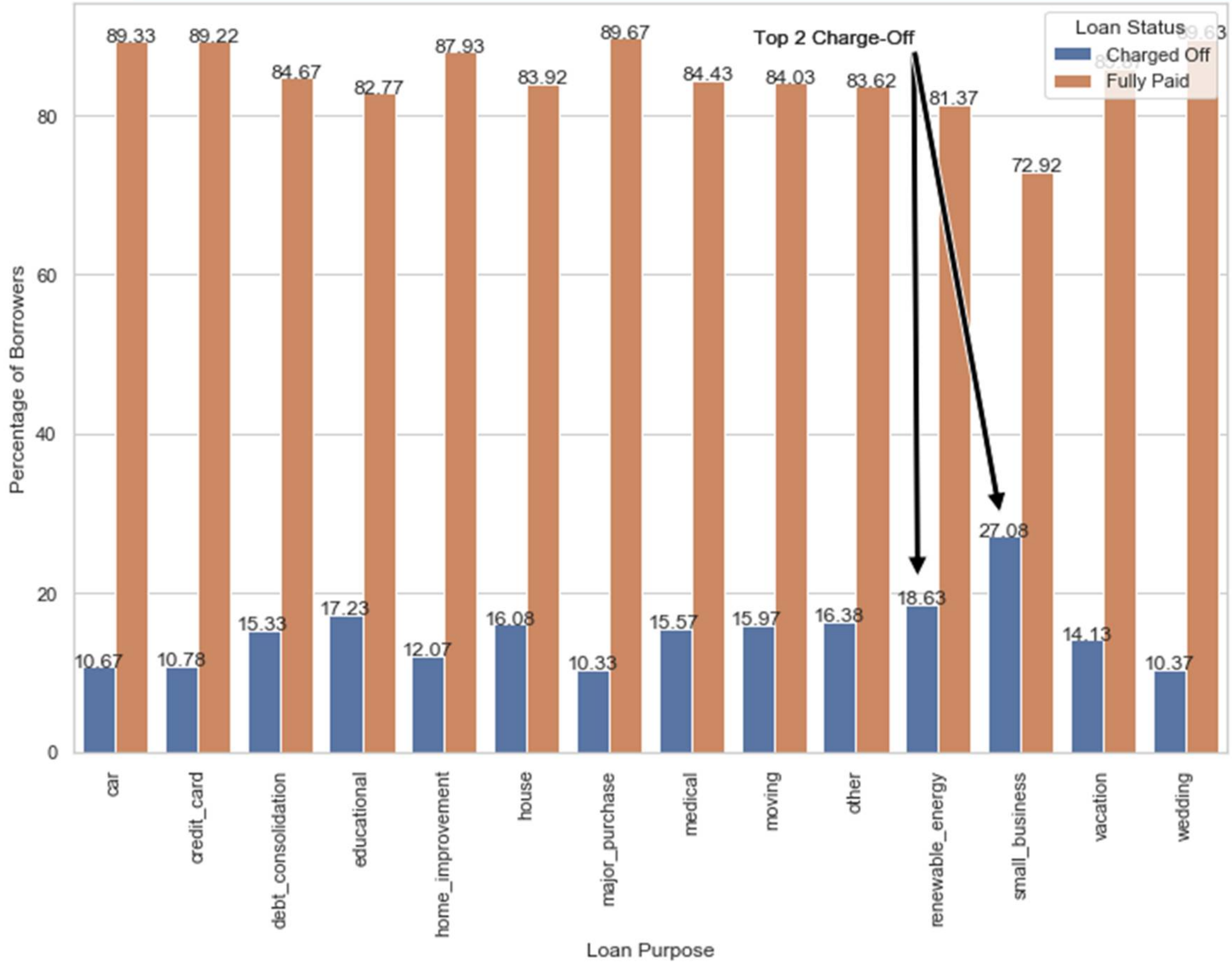
- Borrowers living in rented houses or with home ownership status as Other are more likely to default
- Borrowers who apply for 5 years term loan are defaulting more compared to borrowers who apply for 3 years term loan
- Verified Borrowers tend to default more than Not Verified Borrowers

Plot – Sub-Grade vs Loan Status



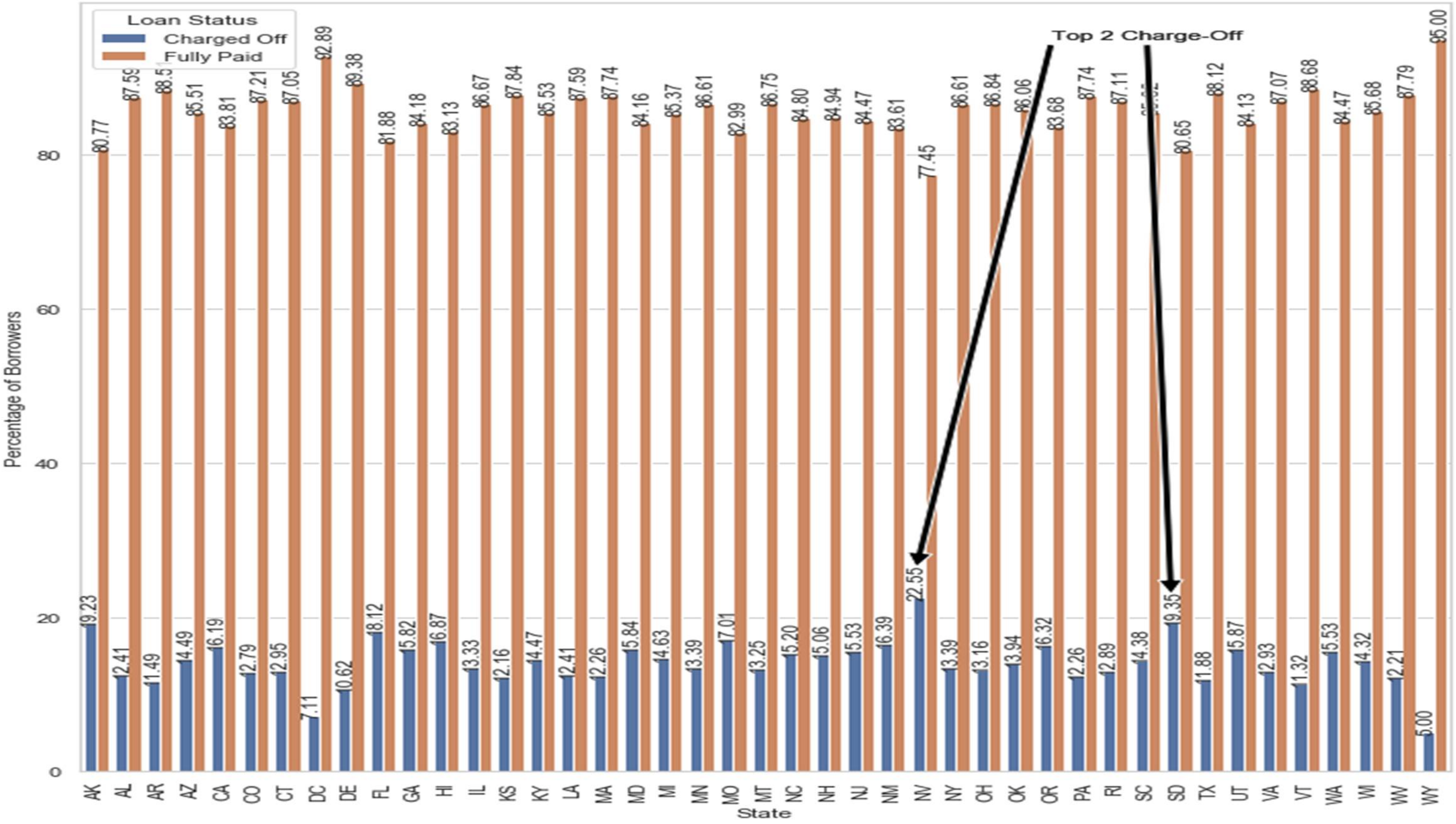
- Borrowers who have taken loan with Grade beyond D2 tend do default more than others
- Borrowers who have taken loan with Grade A or B tend to pay off the loan more than others
- A has the least default rate and the default rate keeps on increasing from A to G

Plot – Purpose vs Loan Status



- Borrowers applying for loans for purposes of “Small business” and “renewable energy” are more likely to default.
- Borrowers applying for loans for purposes of “major purchase” and “wedding” are less likely to default.

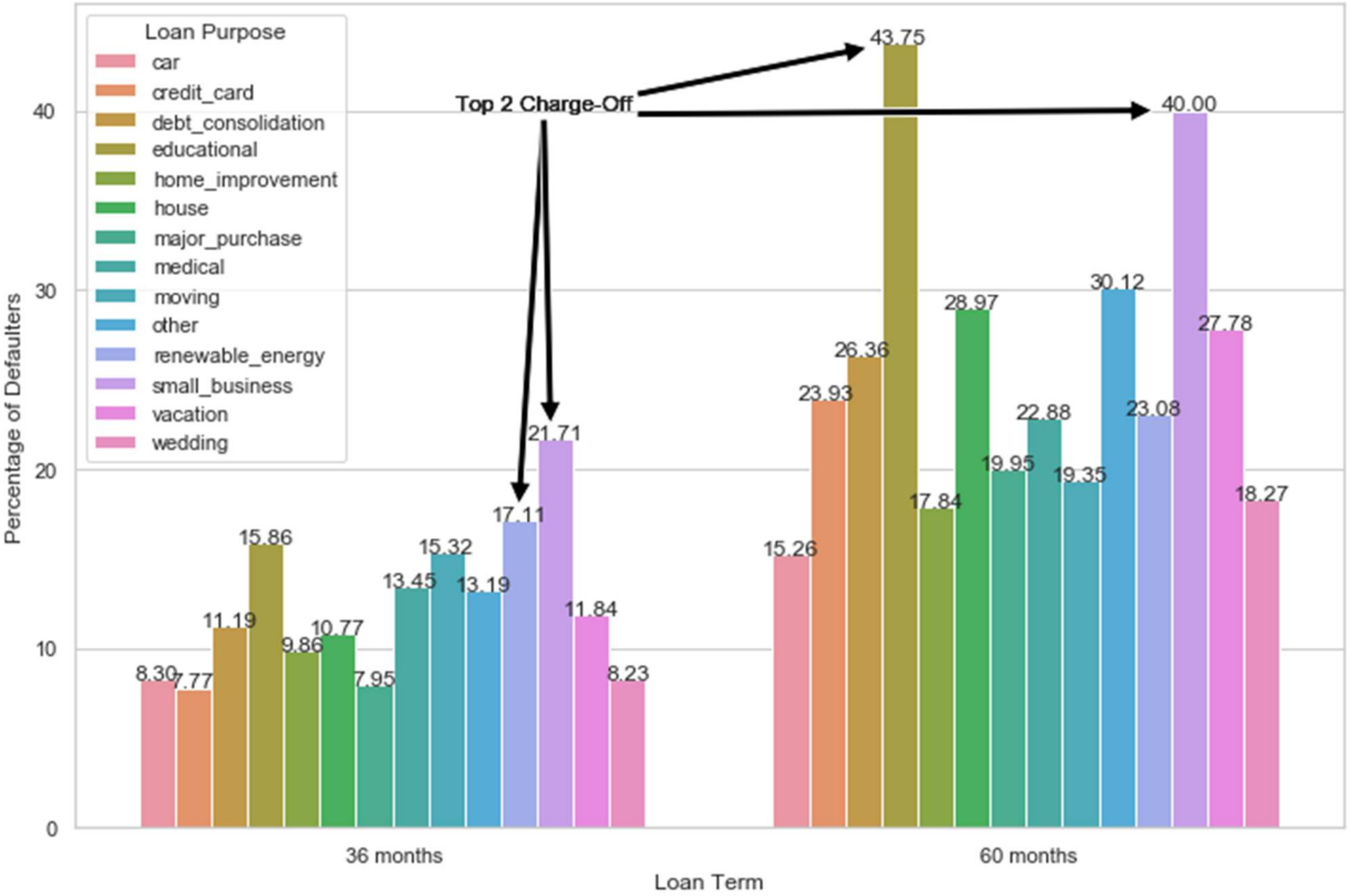
Plot – State vs Loan Status



- Borrowers living in US states of Nevada (NV) and South Dakota(SD) are more likely to default

- Borrowers living in US states of Wyoming (WY) and Washington DC (DC) are least likely to default

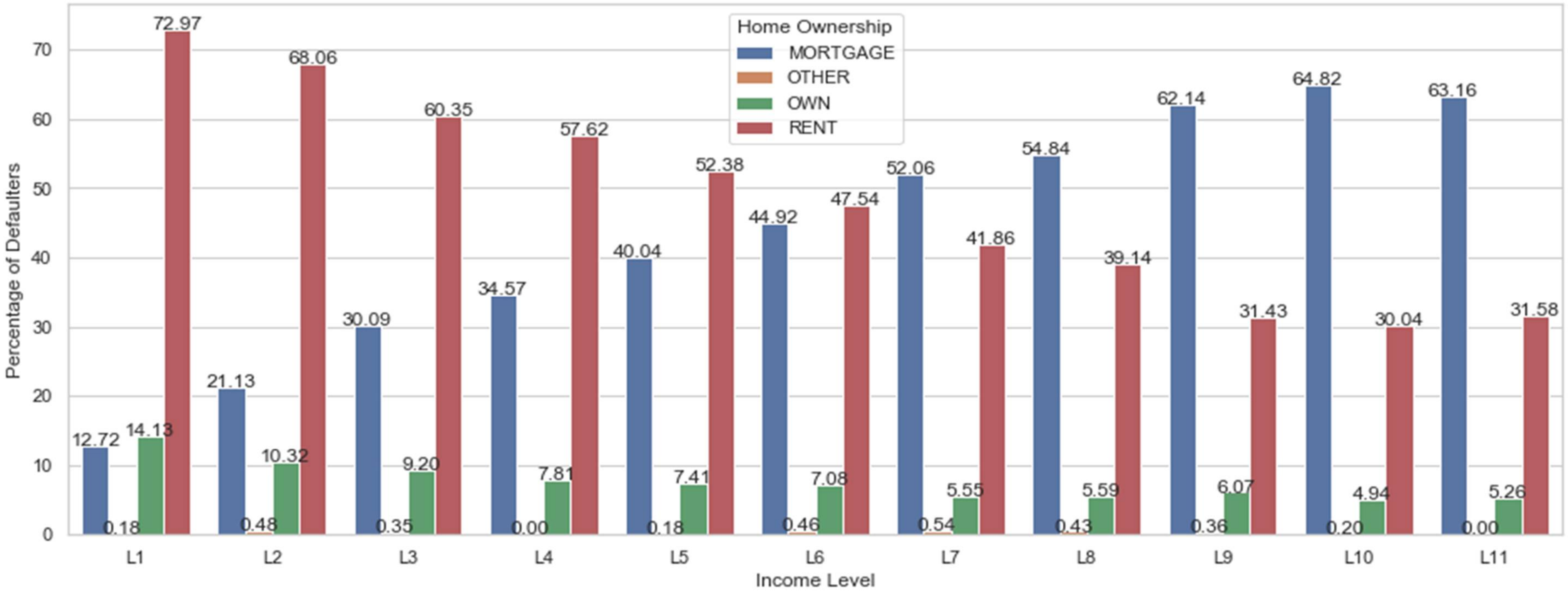
Plot – Loan Term and Purpose vs Loan Status



- Borrowers who have taken loan for the purpose renewable energy and small business with 36 months tenure are more likely to default

- Borrowers who have taken loan for the purpose educational and small business with 60 months tenure are more likely to default

Plot – Home and Annual Income vs Loan Status



- As the income level decreases, the borrowers living in Rented houses tend to default the loan more.
- As the income level increases, the borrowers living in Mortgaged houses tend to default the loan more

Note - L1 to L11 represent [4000, 30000.0, 37196.16, 44500.0, 50004.0, 58868.0, 65004.0, 75000.0, 90000.0, 115000.0, 234144, 6000000] USD.

Summary

The driving factors (or driver variables) behind loan default are :

DTI

- Default: More than 13.86
- Fully Paid: Less than 13.21
- Higher the DTI, higher is the default rate

Loan Amount

- Default: more than 15,000 USD
- Fully Paid: Less than 15000 USD
- Higher the loan amount, higher is the default rate

Annual Income

- Default: Less than 55,000 USD
- Fully Paid: More than 65,000 USD
- Higher the annual income, lower is the default rate

Purpose

- Default: Small business and renewable energy
- Fully Paid: Major purchase and wedding

State

- Default: Nevada (NV) and South Dakota(SD)
- Fully Paid: Wyoming (WY) and Washington DC (DC)

Term and Purpose

- Default:
1. Term: 60 months, Purpose: Education and Small Business
2. Term: 36 months, Purpose: Renewable Energy and Small Business

Home and Annual Income

- Default:
1. Borrowers in Rented houses as income level decreases
2. Borrowers in Mortgaged houses as income level increases

Home Ownership

- Default:- Rent and Other
- Fully Paid:- Own and Mortgage

Loan Sub Grade

- Default: Grade beyond D2
- Fully Paid: In Grade A or B

Loan Term

- Default: 5 years
- Fully Paid: 3 years

Verification Status

- Default: Verified Borrowers
- Fully Paid: Not Verified Borrowers.