

Lending Club Case Study

Group Members

Shivakumar
Gurupreet

Introduction:

The Lending Club Loan Approval Analysis case study involves the examination of a dataset named "loan.csv" to determine whether a particular person's loan application should be approved or denied. Lending Club is a peer-to-peer lending platform that connects borrowers with investors. The goal of this analysis is to perform an Exploratory Data Analysis (EDA) to gain insights into the dataset.

Dataset Overview:

The dataset "loan.csv" contains a comprehensive set of information about past loan applicants, including various attributes such as credit scores, employment details, loan amounts, interest rates, loan status, and more. Each row in the dataset represents a unique loan application.

Key Objectives:

1. **Data Preprocessing:** Cleanse and prepare the dataset by handling missing values, outliers, and data formatting issues.
2. **Exploratory Data Analysis (EDA):** Perform an in-depth analysis of the dataset to understand the patterns, relationships, and trends within the data.
3. **Data Visualization:** Create visualizations such as histograms, scatter plots, and correlation matrices to visualize the data and gain insights.

Data Understanding

```
data = pd.read_csv("drive/MyDrive/Colab Notebooks/Lending Club Case Study/loan.csv")
print("Data Shape: {}".format(data.shape))
print("Data Size:".format(data.size))
print("INFO: ")
print(data.info())
```

```
Data Shape: (39717, 111)
Data Size:
INFO:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 39717 entries, 0 to 39716
Columns: 111 entries, id to total_il_high_credit_limit
dtypes: float64(74), int64(13), object(24)
memory usage: 33.6+ MB
None
```

We got a dataset from consumer finance company which specialises in lending various types of loans to urban customers.

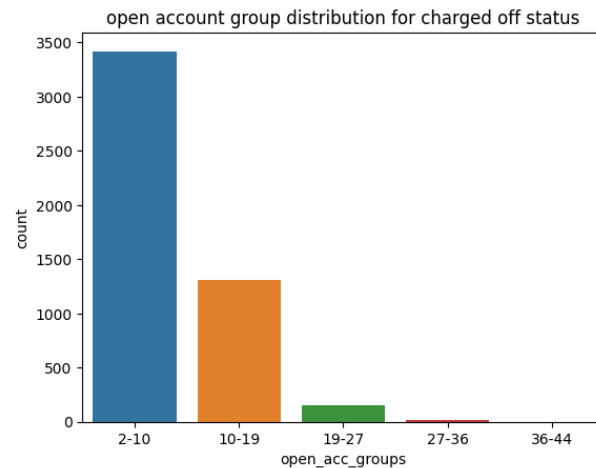
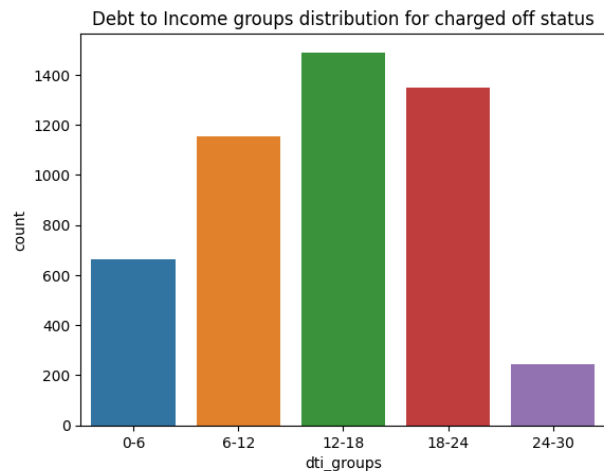
It contains the complete loan data for all loans issued through the time period 2007 to 2011.

Total No of records:

Rows: 39717

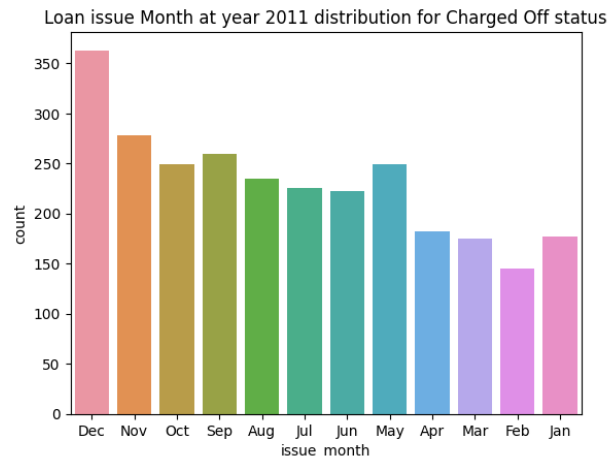
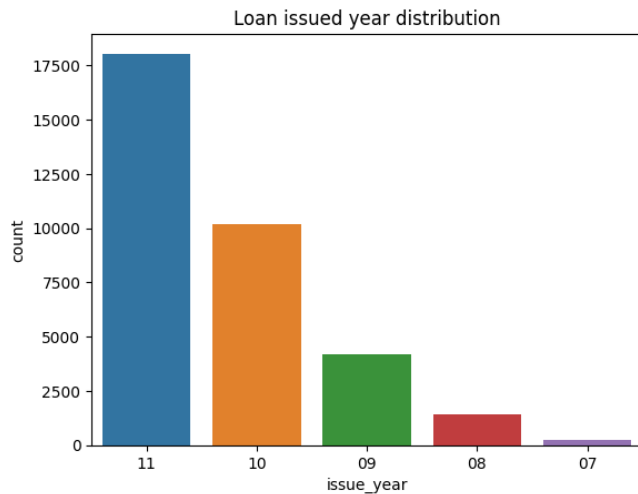
Columns: 111

Analysis



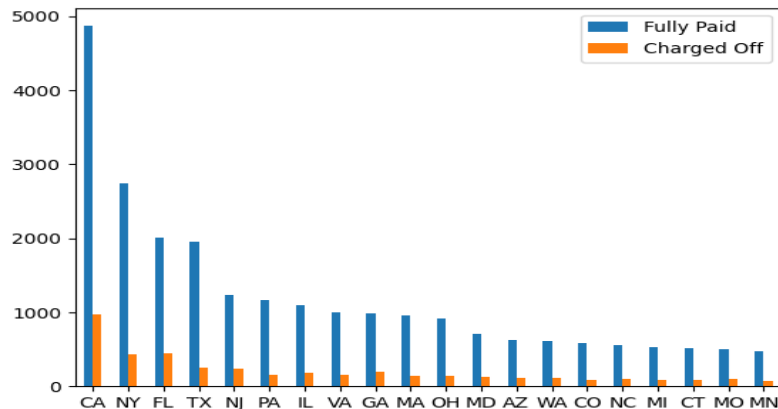
- From the above graph, we observed that, those who have the the dti score between 12-18 are mostly like a defaulters.
- Similarly, we observed most defaulters are having the open account credit number is credit file is between 2-10.

Analysis



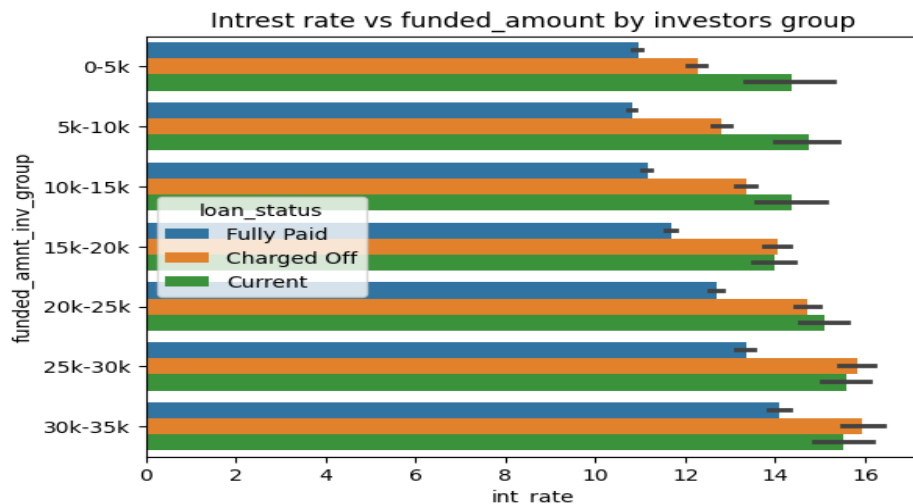
- From this financial dataset, mostly defaulters are falls in the year of 2011 and also, they brought the loan from the company at end of year (December).

Analysis



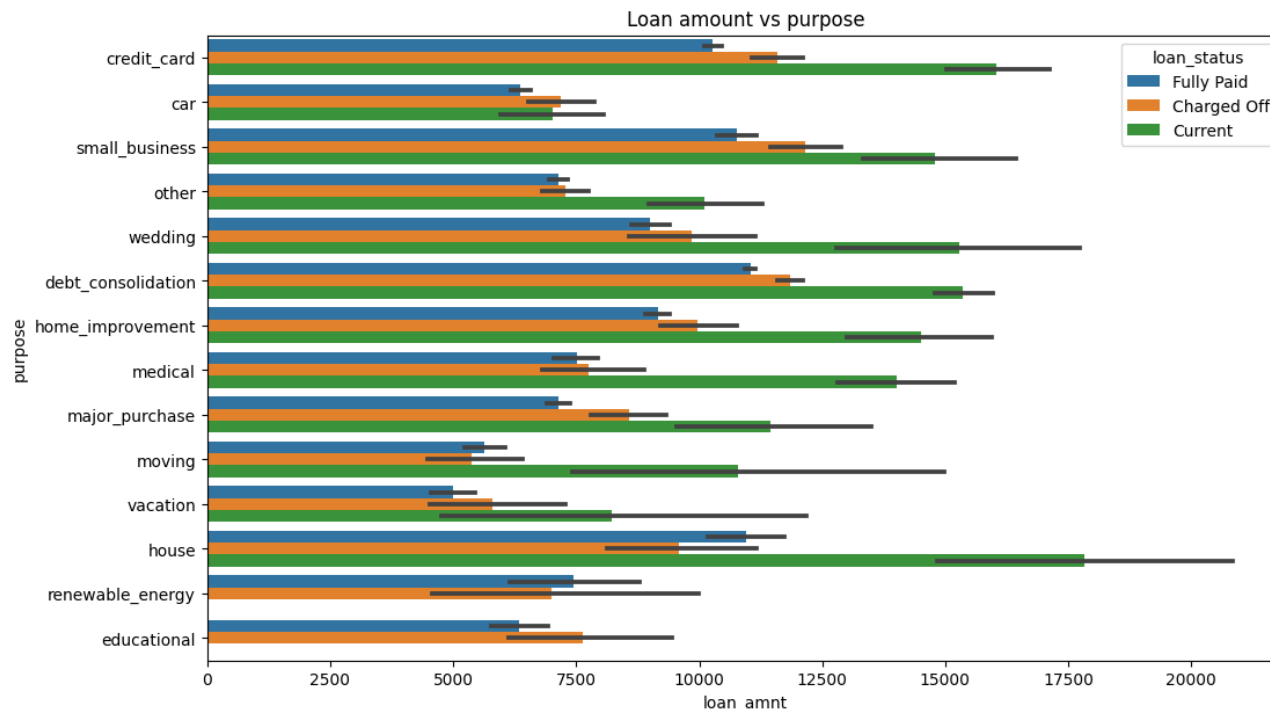
If we see the dataset with state wise, we observed some interesting fact, those who lives in CA are likely getting more defaulters.

Analysis



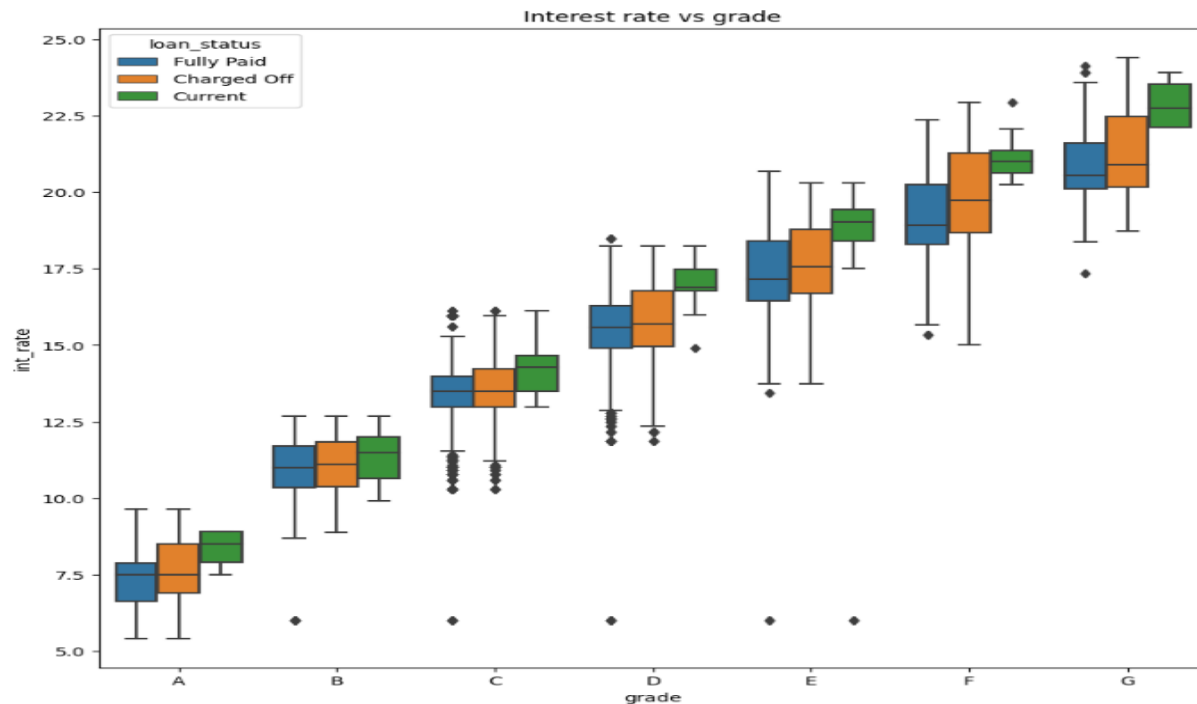
People who opted for low loan amount with high interest rate are mostly marked as defaulters.

Analysis



- Those who mentioning the debt_consolidation are mostly “defaulters” and those who opted for credit card, small business are “likely defaulters”.
- For House, moving, renewable energy are mostly paid their amount back to company.

Analysis



- Here we have the clean observation, when the grade increases then the interest rate also getting increase.
- Those who falls under the grade “E,F,G” category has the more defaulters.

Conclusion

- When the loan amount is low or mid-range and taken the time period of 60 months and with high interest rate are mostly prone to defaulters.
- Those who fall under the grade "E,F,G" category have the more defaulters.
- Lending clubs have to consider giving loans for the purpose "debt_consolidation" during end of the year.
- If DTI (Debt To Income) is in the range of "24-30" are the defaulters.
- If Open Acc (Pre approved loan based on current credit file) is below 15 are the defaulters.

FEATURES	DISAPPROVE	REDUCE LOAN AMOUNT	FULLY PAID
PURPOSE	Debt consolidation	Credit card, small business	House, moving, renewable energy
TIME	Year end	JAN - MAR	MAY-AUG
DTI	24-30 range	6-18	0-6
OPEN ACC	Below 15	20-35	35 and above
INTEREST RATE	Above 13	Below 16	Below 18