Lending CLUB Case Study - Submission

Members:

1. Deepak Hiremath
2. Ebino Reji

Problem Statement

- Lending Club 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). Borrowers who **default** cause the largest amount of loss to the lenders. In this case, the customers labelled as 'charged-off' are the 'defaulters’.
- Goal is 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.
- The company wants to understand the **driving factors** behind loan default

Objectives

- Analyze the data of past loan applicants
- Do exploratory data analysis and **consumer attributes** and **loan attributes** influence the tendency of default
- Identify the key factors behind loan defaulters, which will help company take better decisions on loan applications

Steps

- Steps Followed

Steps: Data Cleaning

Performed below steps in data cleaning:

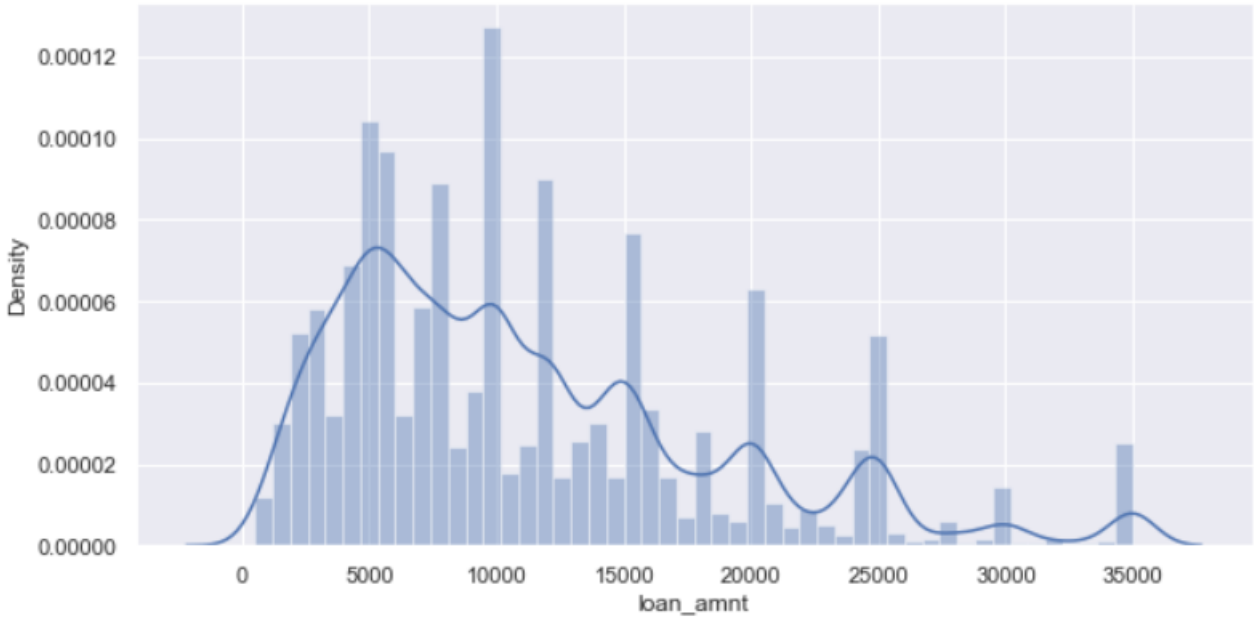
- Dropped fields which had null values
- Dropped fields which had single category
- Dropped fields which have identical data
- Dropped rows which had identical data
- Data Standardization:
  - Treated the missing values
  - Converted the data types of few columns to standard type
- Drop fields which are not required for current study:
  - Some fields which are not adding value to current study are removed
- Deriving new fields
  - Derived new fields based on existing data
  - Example: DTI_Bad_Good , Loan_year, Loan Month

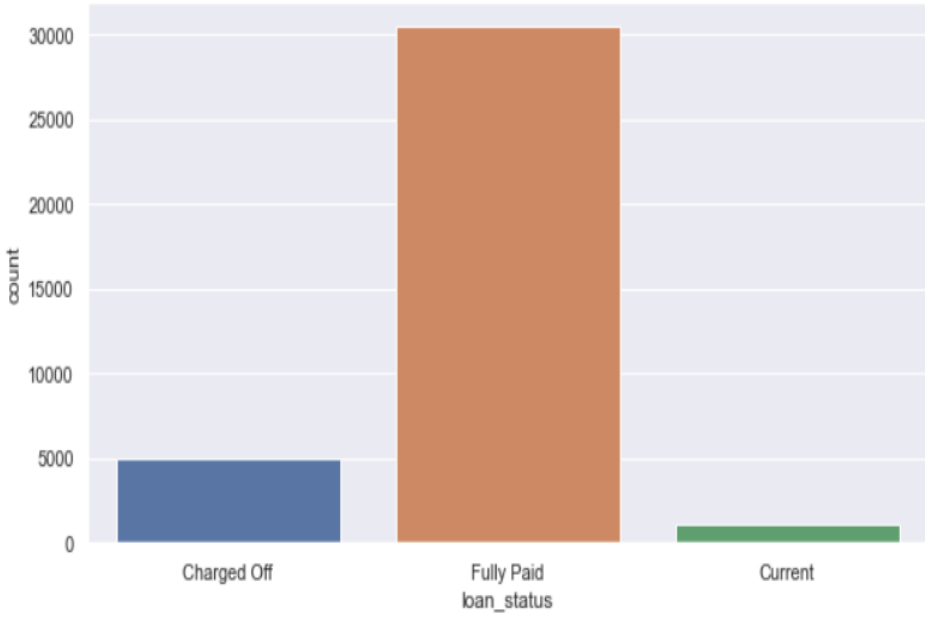
Steps: EDA

Performed below steps in EDA:

- Univariate Analysis
- Segmented Univariate Analysis
- Bivariate Analysis
- Prepared summary List

Analysis: Overall Loan Status

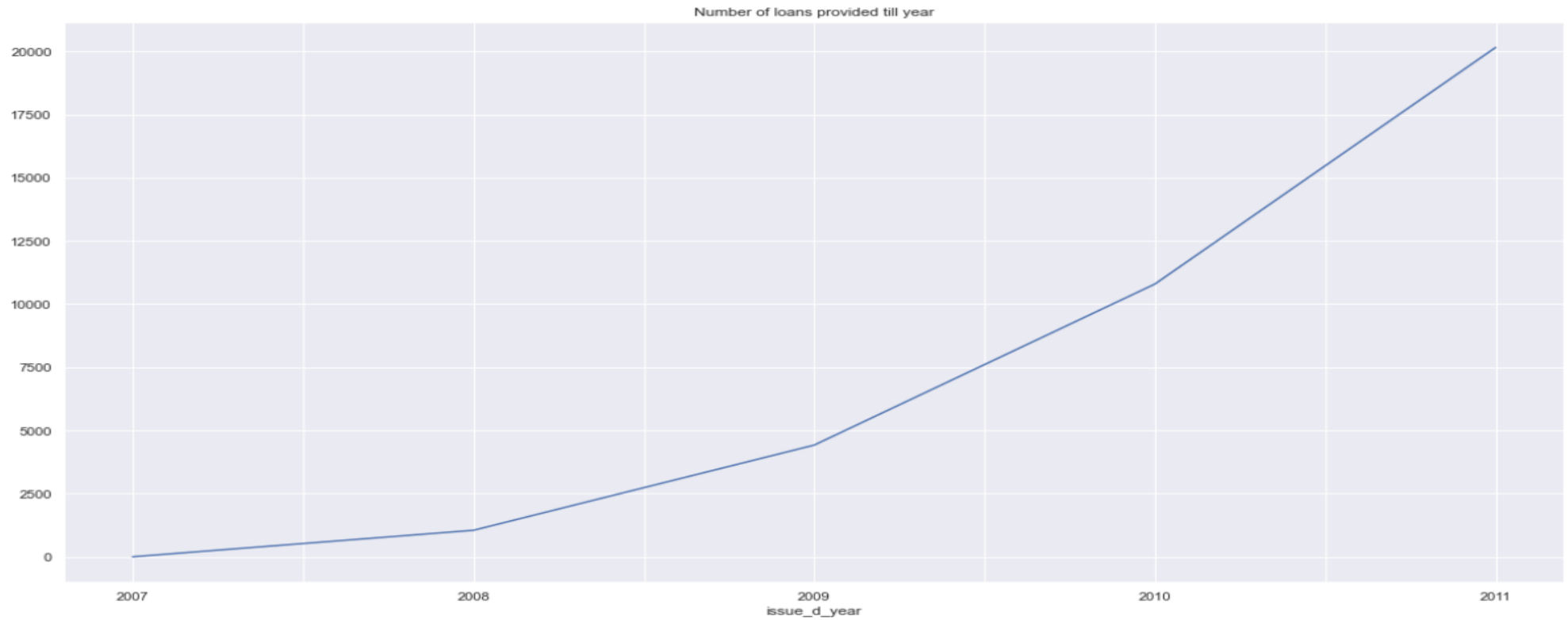





```
{'percecntage_charged_off': 13.567370241264786,
 'percentage_fully_paid': 83.50671094886503}
```

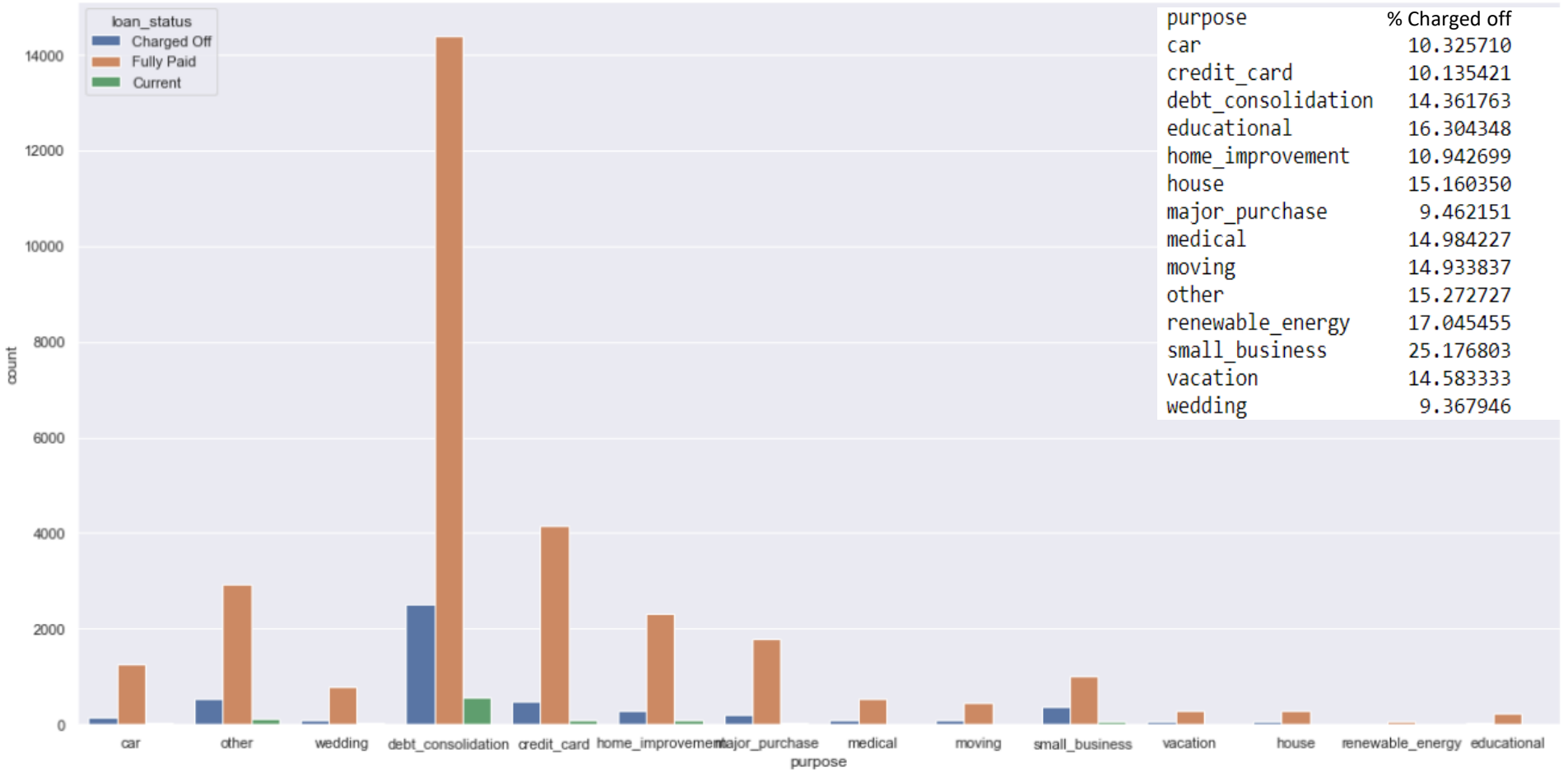
- Lending Club gave more loans of lesser amount
- The total defaulters are 13.57%
- The total fully paid loans are 83.5%

Analysis: Overall Loan Status



- Lending Club has given more loans year over year

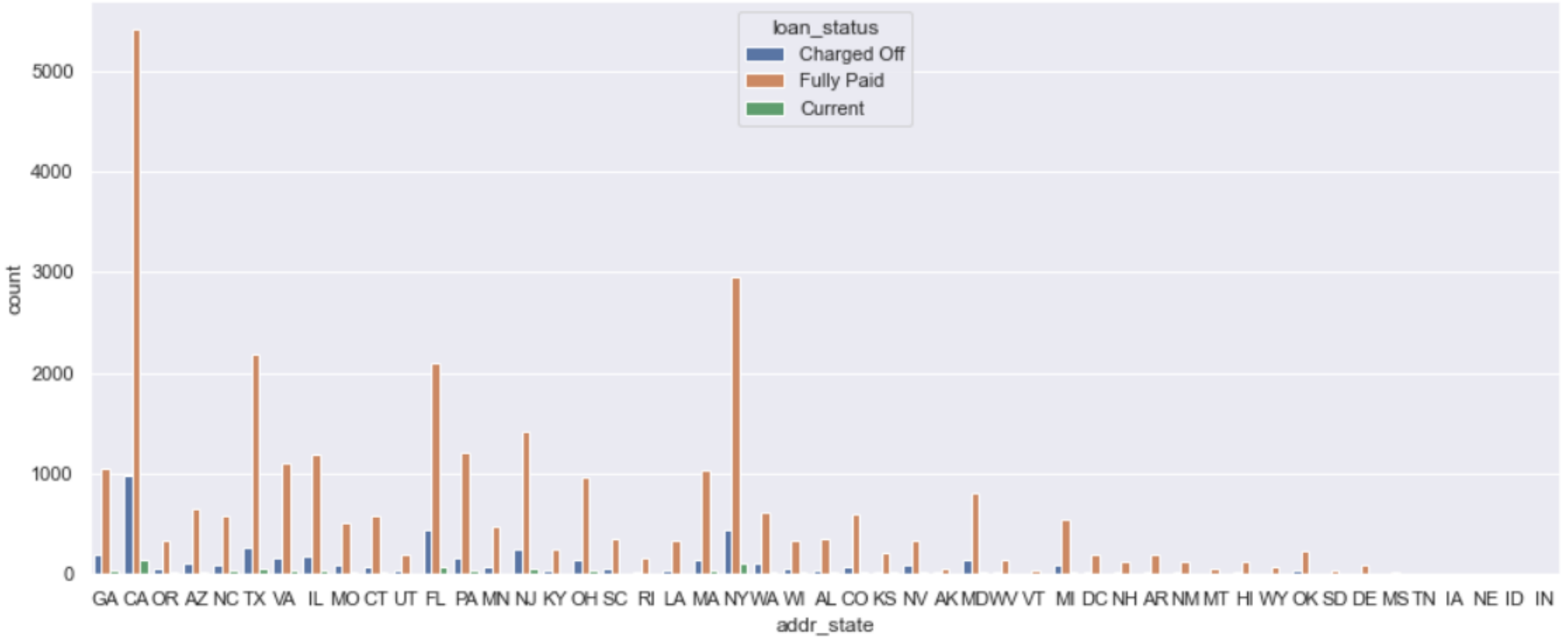
Analysis: Loan Status



purpose	% Charged off
car	10.325710
credit_card	10.135421
debt_consolidation	14.361763
educational	16.304348
home_improvement	10.942699
house	15.160350
major_purchase	9.462151
medical	14.984227
moving	14.933837
other	15.272727
renewable_energy	17.045455
small_business	25.176803
vacation	14.583333
wedding	9.367946

- **Around 47% loans are provided for debt_consolidation, where we found only 14% defaulters**
- **25% defaulters are under small_business loans.**
- **The company should be very cautious while giving loans in small_business sector**

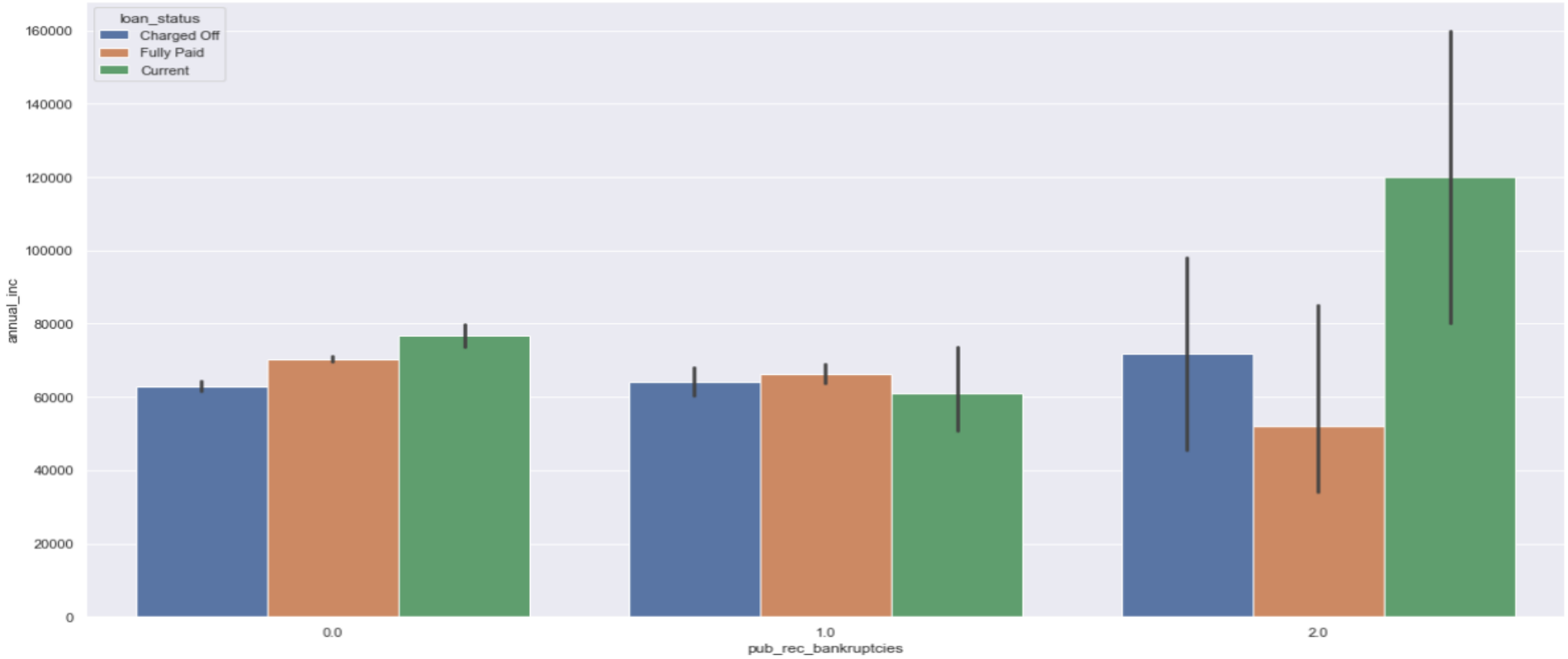
Analysis: Loan & US States



State	% Charged off
AK	20.000000
AL	11.380145
AR	11.111111
AZ	13.537676
CA	14.932681
CO	11.315417
CT	11.504425
DC	6.796117
DE	10.377358
FL	16.704893
GA	15.027111
HI	16.149068
IL	13.113604
KS	9.638554
KY	14.052288
LA	11.306533
MA	11.374795
MD	14.373717
MI	13.567073
MN	12.788632
MO	15.224359
MS	5.555556
MT	12.328767
NC	13.920455
NH	14.569536
NJ	14.857806
NM	16.766467
NV	21.086957
NY	12.535776
OH	12.126538
OK	13.074205
OR	15.291262
PA	11.291461
RI	13.227513
SC	13.793103
SD	17.543860

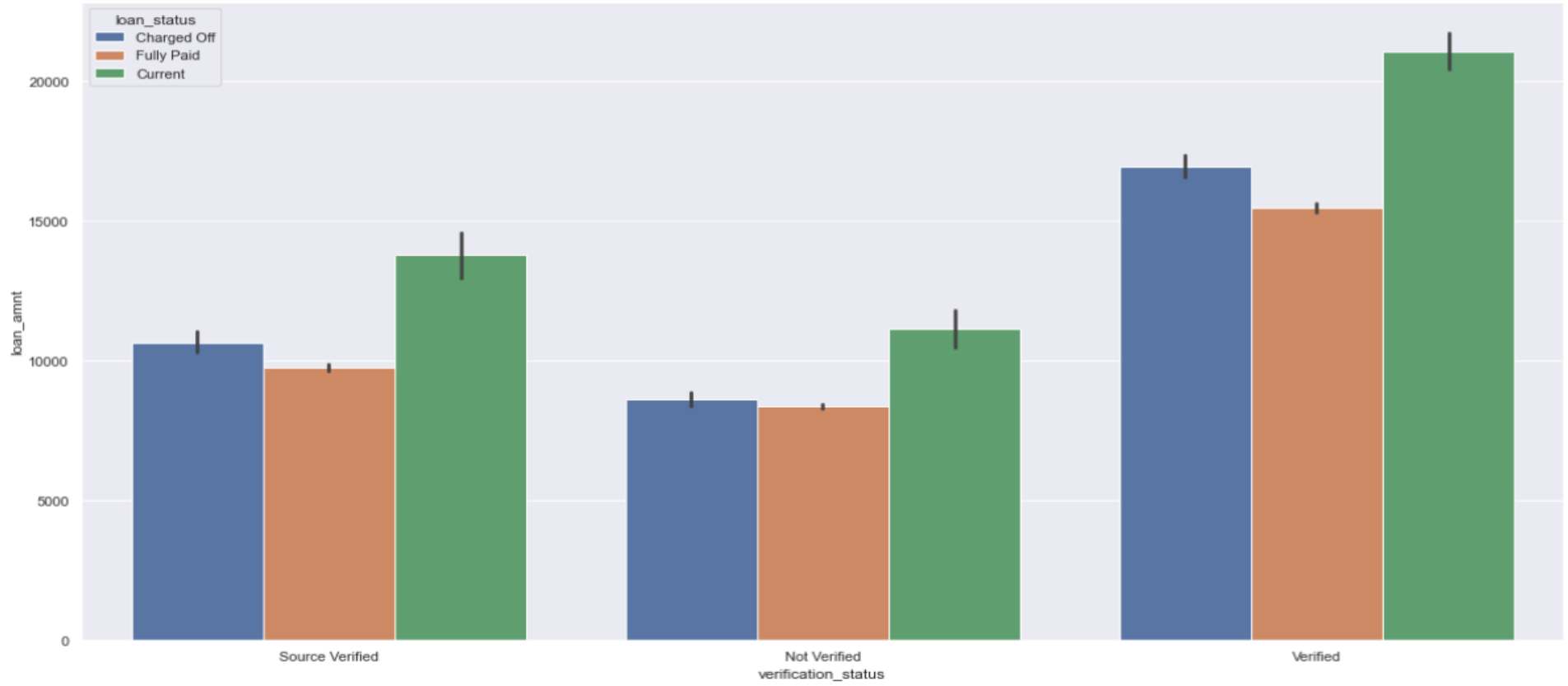
- Maximum loans are provided in California state
- Defaulters are high in Nevada 21%, Alaska 20%, Tenessee 20%. Company should be giving less loans in these states.

Analysis: Public Record of Bankruptcies



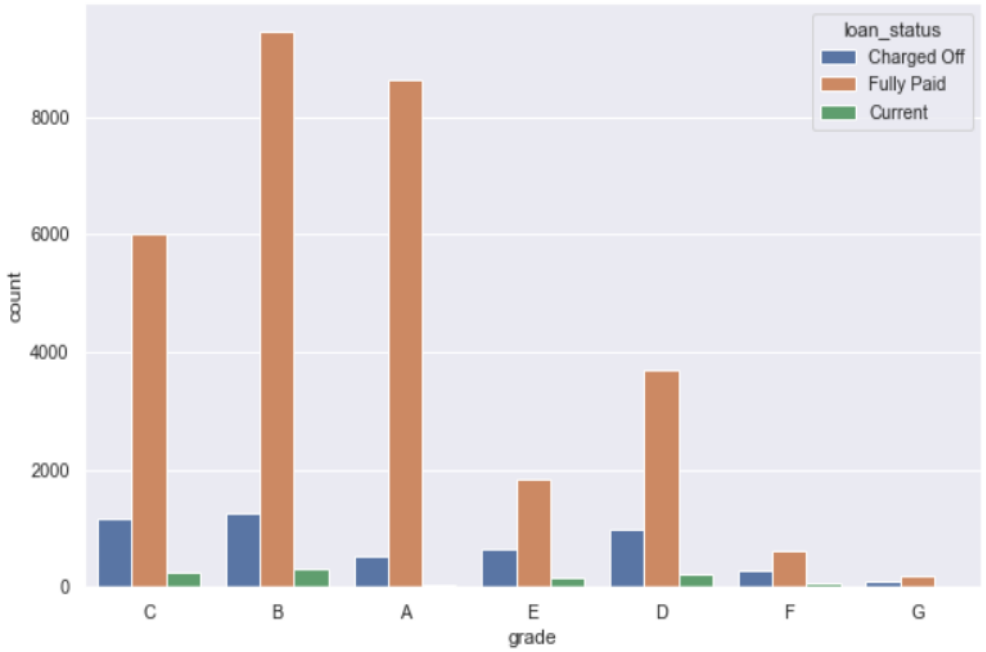
- Lending Club has given more loans/high value loans to people with higher public record of bankruptcies
- Higher the public record of bankruptcies higher is the risk of recovery

Analysis: loan amount vs loan verified vs loan status



- In the above univariate analysis of verification status we found that charged off loans are more under not verified.
- Looks like if amount borrowed is higher then defaulters are also higher.
- Company should be cautious while given higher value loans

Analysis: Grades

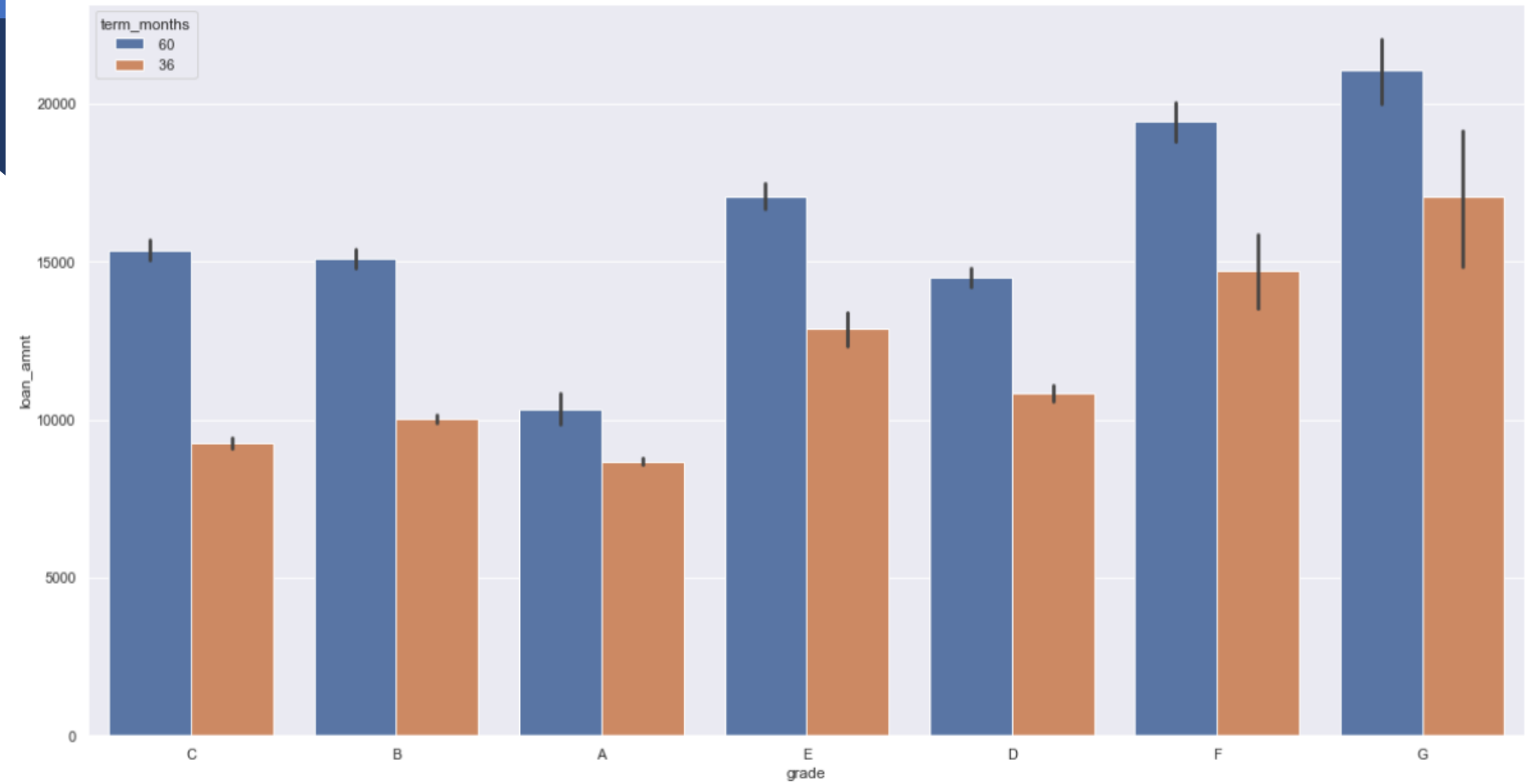


grade	% Charged off
A	5.591281
B	11.354780
C	15.744853
D	20.102564
E	24.452830
F	29.514964
G	31.208054

Name: grade, dtype: float64

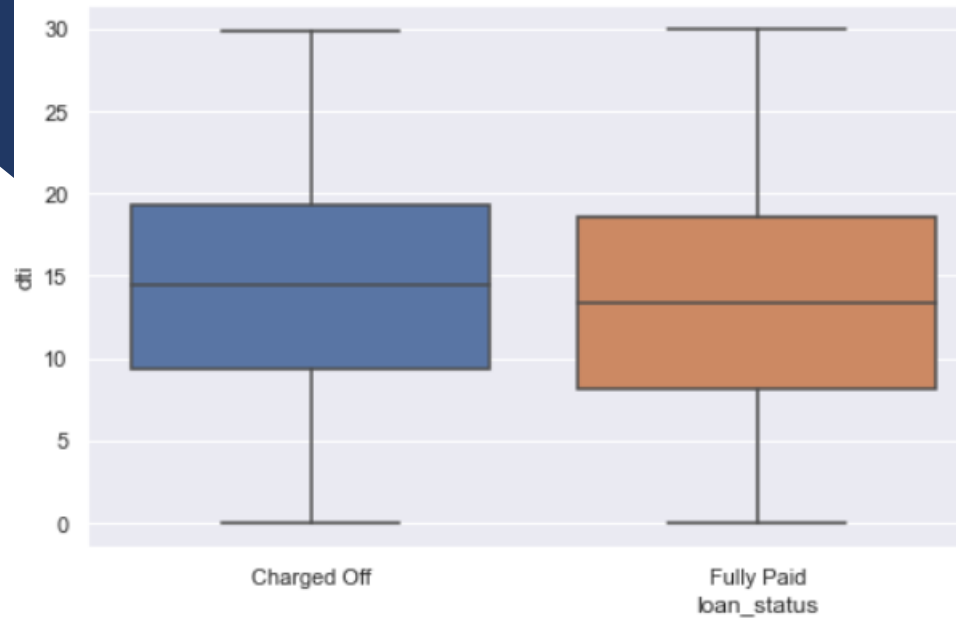
- More loans are given to grade A, B, C category customers. It is good for company as the probability of returning the loan is high.
- The above plot also tells more defaulters are in B,C,D category
- But the percentage of defaulters are more in G, F, E, D with 31%, 29%, 25%, 20% respectively.
- So company should decrease the loans to high risky borrowers

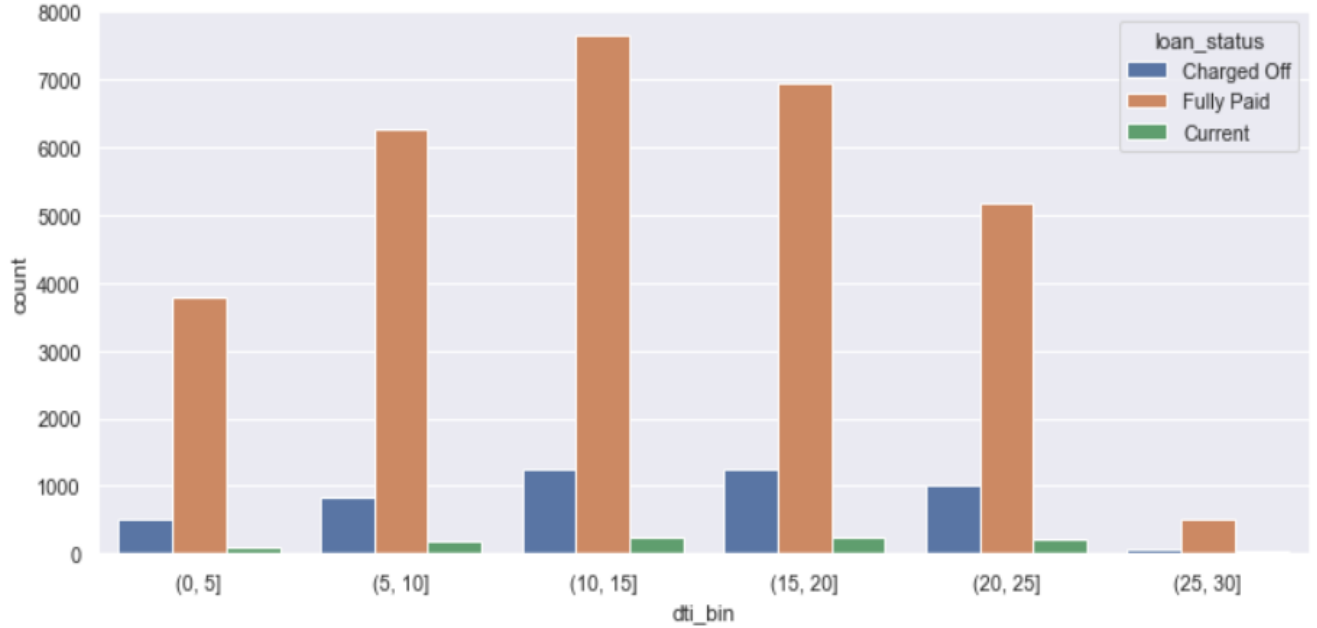
Analysis: Grades vs Loan amount vs Term months



- In previous slide we saw **higher loans are at high risk, may be its because of loan grades.**
- **Higher loans are provided at higher grades, which is of higher interest**
- Company should be cautious about interest rates

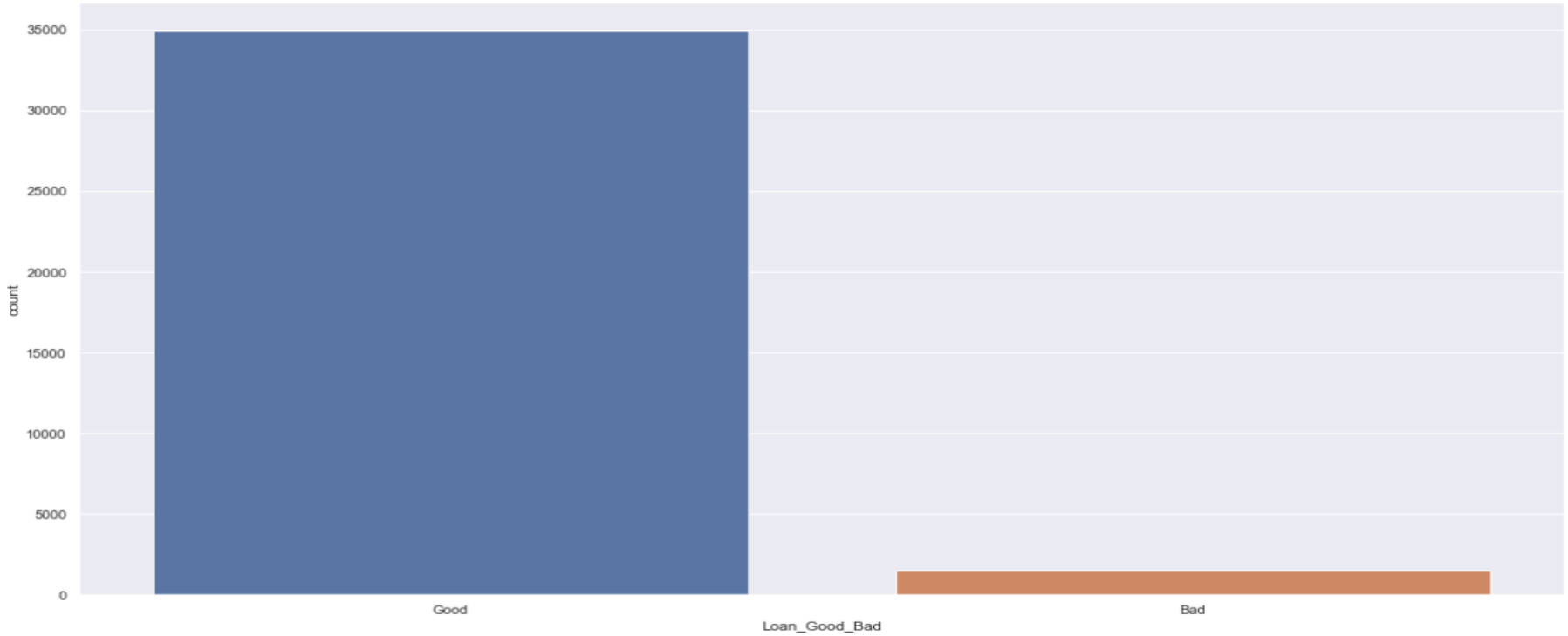
Analysis: DTI





- The analysis on dti shows that, higher the dti more chances of loan being charged off
- Companies can check if the asked loan amount vs dti ratio is crossing 60% then lending loan should be stopped
- In the next slide I have explained more about DTI

Analysis: DTI (derived)



- Found that there are some loans which were bad loans
- Company can use this derived field to analyze for further loans

- In the below formula I am trying to calculate if the given loans are good or bad based on 'dti'.
- If loan amount given was greater than (annual income * (60% - %dti)) then the loan was a Bad loan else its a good loan.

Conclusion

- Make sure the loan amount is not exceeding the threshold of dti (derived in previous slide)
- Higher value loans are at high risk(may be its because of loan grades, which is of higher interest), so decrease the interest of higher value loans, increase it for lower value loans
- The company should be very cautious while giving loans in small_business sector, better to decrease the loans in this sector by 50%. Increase in Car, Home development sectors
- Higher the public record of bankruptcies higher is the risk of recovery, stop giving loans to people with higher bankruptcies