

Dataset and Business objective

- Past loan applicants from 2007 to 2011, their metadata, payment behavior and respective loan status 'defaulted' or not.
- Understand the driving factors (or driver variables) behind loan default
- Utilize insights from EDA for portfolio and risk assessment to give loans to safe customers and avoid risky applicants.

id	member_id	loan_amnt	emp_length	home_ownership	annual_inc	verification_status	loan_status
1077501	1296599	5000	10+ years	RENT	24000	Verified	Fully Paid
1077430	1314167	2500	< 1 year	RENT	30000	Source Verified	Charged Off
1077175	1313524	2400	10+ years	RENT	12252	Not Verified	Fully Paid
1076863	1277178	10000	10+ years	RENT	49200	Source Verified	Fully Paid
1075358	1311748	3000	1 year	RENT	80000	Source Verified	Current
1075269	1311441	5000	3 years	RENT	36000	Source Verified	Fully Paid
1069639	1304742	7000	8 years	RENT	47004	Not Verified	Fully Paid
1072053	1288686	3000	9 years	RENT	48000	Source Verified	Fully Paid
1071795	1306957	5600	4 years	OWN	40000	Source Verified	Charged Off

Dataset snapshot

High level approach

Data input

- Read dataset file
- Read Metadata file

Data Cleaning

- Row cleaning
- Column cleaning
- Missing values
- Standardize values
- Fix data types
- Fix data quality issues
- Derive important metrics e.g. "loan_to_income" ratio

Univariate Analysis

- Numerical feature analysis
- Categorical feature analysis

Segmented univariate Analysis

- Variables segmentation on 'loan_status'
- Analysis of numerical and categorical features

Bivariate & Multivariate Analysis

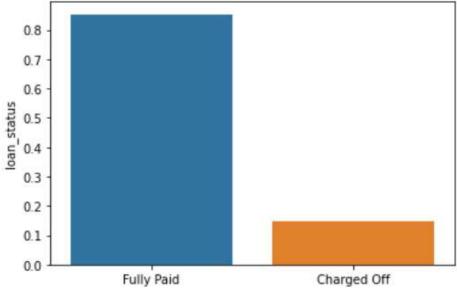
 Analysis of multiple variables together.

Recommendations

 Recommendations using analysis created.

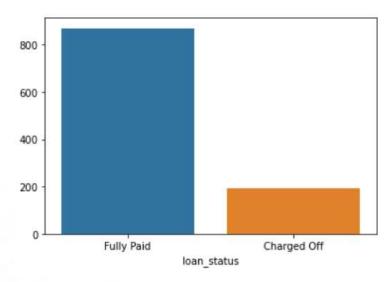
Results (1)

Benchmarking:

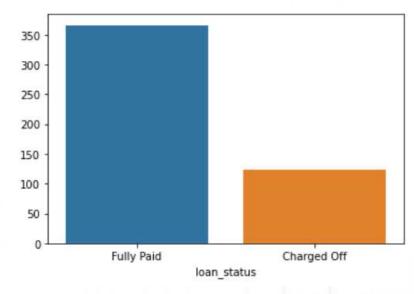


- 14.7% people default in overall dataset.
- This is a benchmark. Any segment of dataset with significant difference in distribution should be considered for further analysis and appropriate actions.

Results (2)



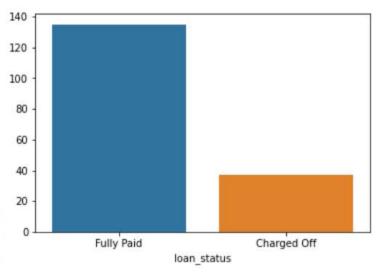
Applicants with credit revolving balance >40000



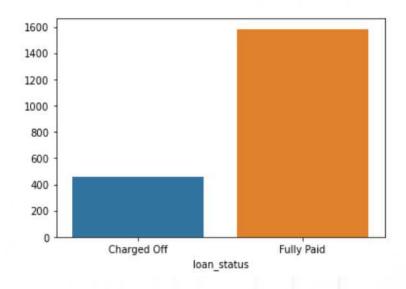
Applicants with credit utilization >98%

- People with credit revolving balance above 40k have higher default rate of 18.17%.
- For credit utilization over 98%, more than 25% customers have defaulted.
- These customers are risky to give loan as they have higher chance of loan default.

Results (3)



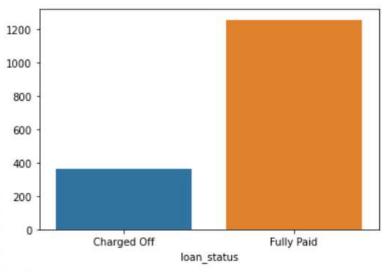
Applicants with credit lines>25



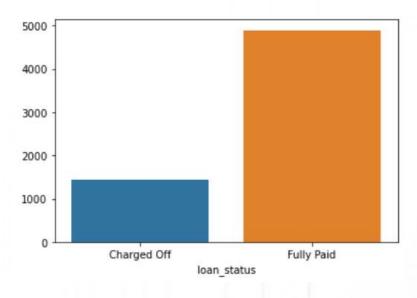
Applicants with public record

- People with credit lines above 25 seem to have high default rate of 21.5%.
- People with public records seem to have high default rate of 22.5%.
- These customers are risky to give loan as they have higher chance of loan default.

Results (4)



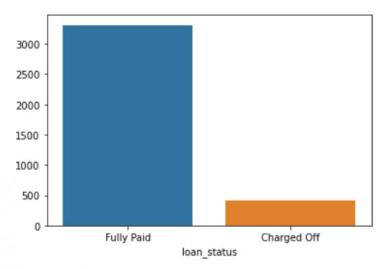
Applicants with record of bankruptcies



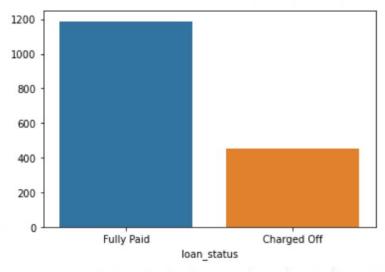
Applicants with loan to income ratio >30

- 22.5% People with public record bankruptcies default out of 1258 people.
- People with high loan_to_income ratio (>30) are risky with 22% default rate.
- These customers are risky to give loan as they have higher chance of loan default.

Results (5)



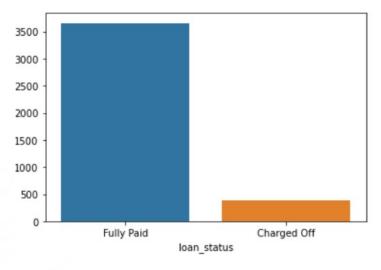
Applicants with loan to income ratio <6



Applicants with loan purpose as small business

- People with low loan_to_income ratio (<6) are safe people to give loan with 10% default rate.
- Loans for small businesses, renewalble energy are of the highest risk with 27.7% people defaulting. People default
 very less for credit card, car and wedding. If these people have low loan to annual income ratio, These seem to be
 very safe loans. These customers are risky to give loan as they have higher chance of loan default.

Results (6)



Applicants with loan purpose as credit card

- People default very less for credit card, car and wedding. If these people have low loan to annual income ratio,
 These seem to be very safe loans.
- These customers are safe to give loan as they have higher chance of loan default.

Recommendations

Safer loans:

Below loan applicants profiles are less risky than most of loan applilcants and may be considered giving loans:

- low loan_to_income ratio (<6)
- 2. Loan purpose: credit card, car and wedding

Risky loans:

Below loan applicants profiles are high risk than most of loan applicants and may be considered giving loans:

- 1. High credit revolving balance (above 40k)
- 2. Credit utilization over 98%
- 3. credit lines above 30
- 4. People with public records
- 5. high loan_to_income ratio (>30)
- 6. Loans for small businesses, renewable energy

Thank you

Learner details:

Indrajeet Jadhav Suchin Chouta

