

LENDING CLUB CASE STUDY

VARUN M R



OBJECTIVE

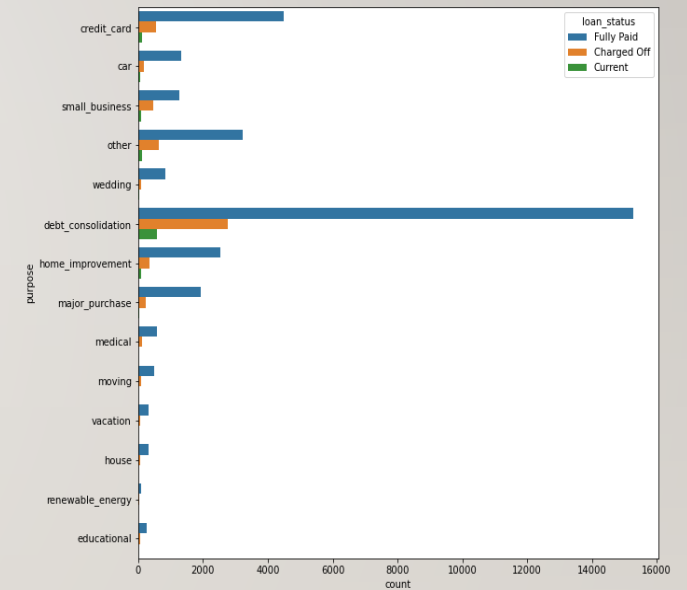
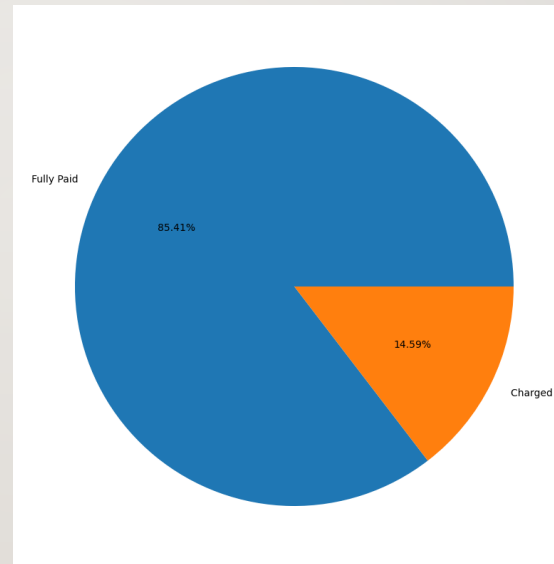
- The client is a financial institution focused on lending using a fast online process system.
- Problem Statement is to predict the risk of defaulting customers using already available customer loan data.
- Exploratory Data Analysis is used to clean, manipulate data followed by Univariate and Bivariate Analysis to find attributes contributing to default of loans.

DATA SET AND PREPARATION OF DATA.

- Data set contains basic characters of a loan like amount, term, interest rate, customer income, status of previous defaults of the customer, and final status of the loan taken, loan taken year etc.
- Data cleaning was done for single value columns, all null value columns, irrelevant columns.
- String and time variables changed into categorical and numerical for better analysis
- New variables created for better analysis of the data.

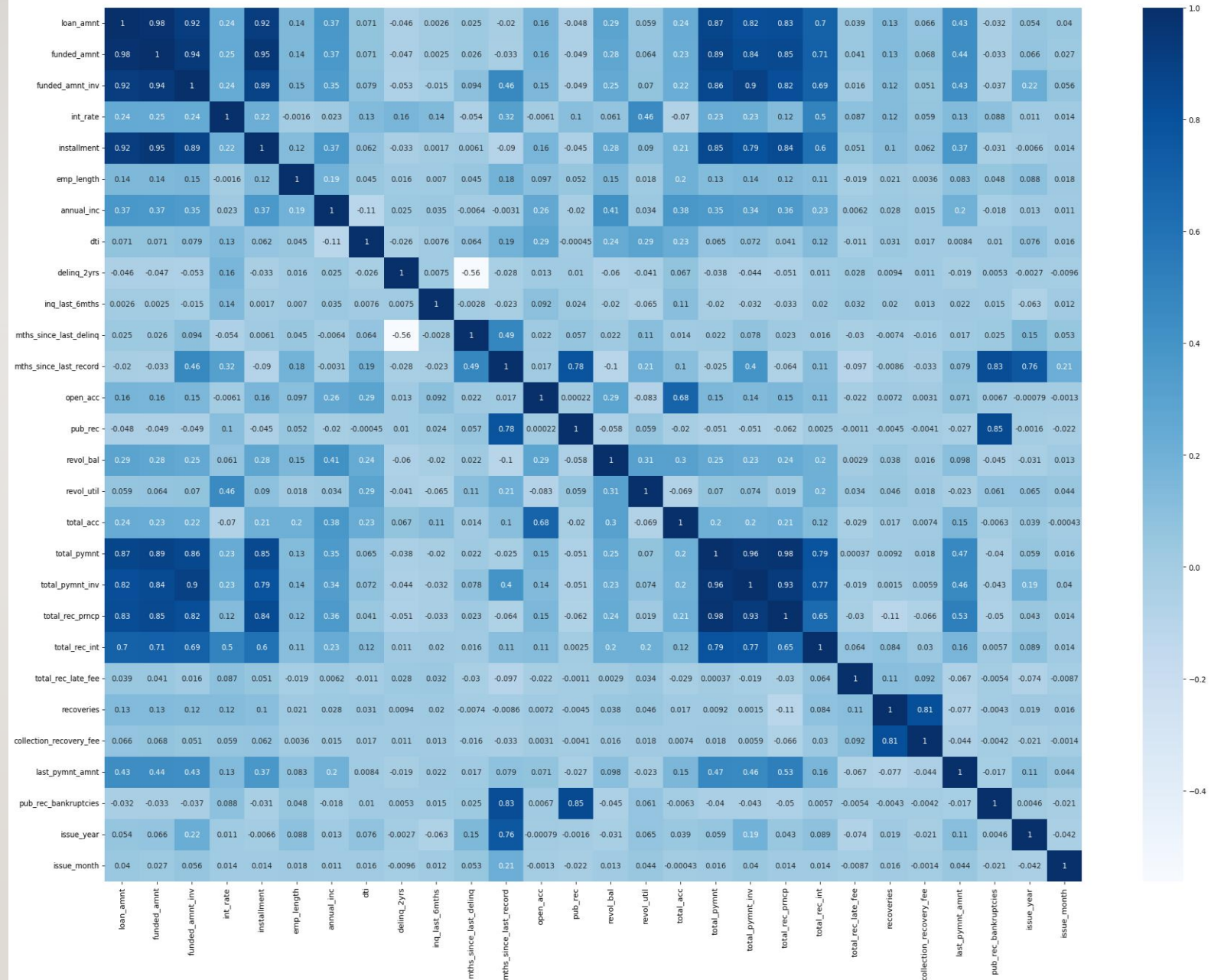
LOAN STATUS & PURPOSE OF LOAN.

- **83% loans are fully paid. Nearly 14% of the loans are charged off.**
- **People took majority of the loans(47%) for debt consolidation. Second main purpose of loans are Credit card payments (13%).**



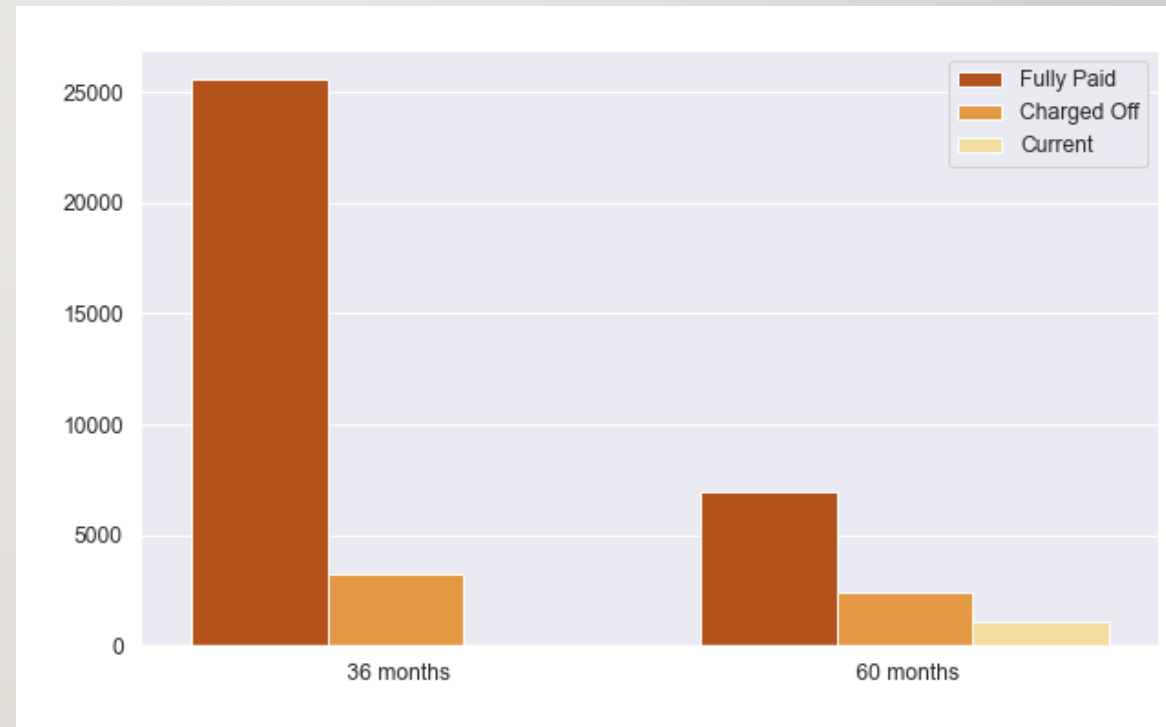
Correlation between Different Variables

- Loan amount, investor amount, funding amount are strongly correlated.
- Annual income with DTI(Debt-to-income ratio) is negatively correlated.
- Positive correlation between annual income and employment years showing income increases with work years.



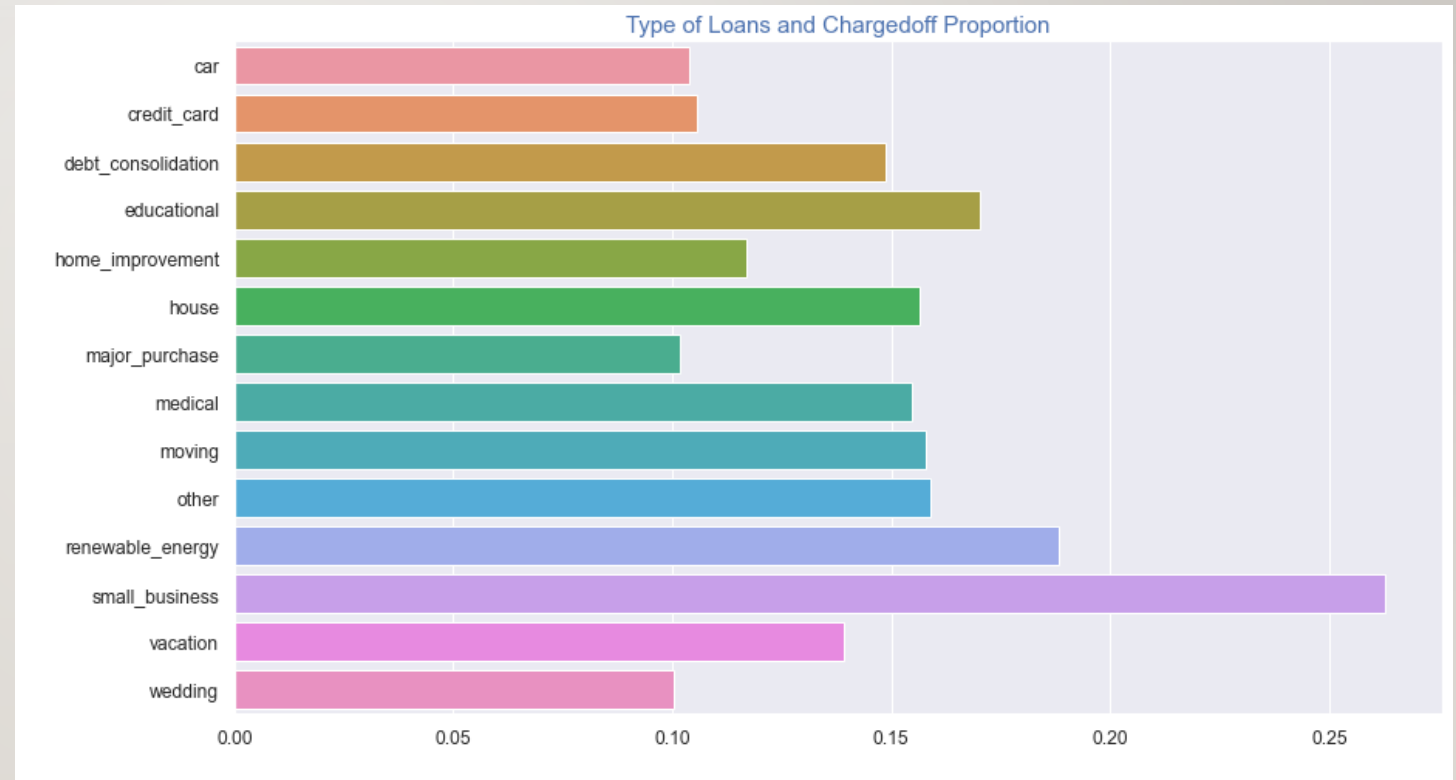
LOAN TERM VS LOAN TIME.

The plot shows proportion of charged off loans increases with increase in loan term.



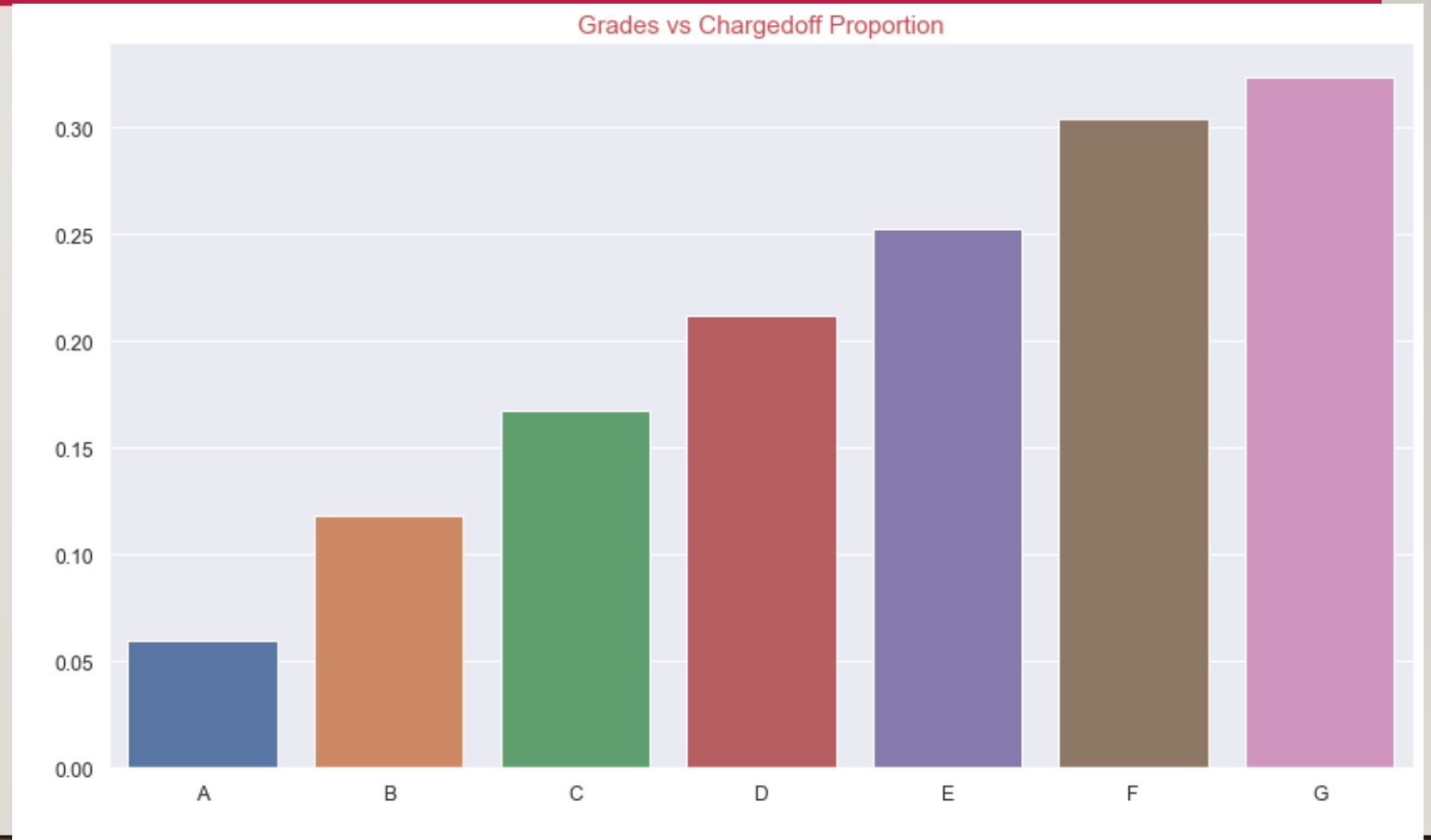
TYPE OF LOAN GIVEN AND DEFAULT.

- Small business, renewable energy, educational loans shown highest chargedoff proportion.
- Wedding, vacation, home improvement loans seen less default. May be also due to small amounts and lower loan term.



GRADE OF THE LOAN VS CHARGED OFF PROPORTION

- Change of the Grades from A to G shows steady increase in charged off loans.



RECOMMENDATIONS

- Income range 80000+ has less chances of charged off but 0-20000 has high chances of charged off.
- Debt Income Ratio, Grade of the Loan, Annual Income, Previous Record of Bankruptcy are strong indicator loan default.
- Small scale Industries , Educational Loans need thorough risk assessment.
- Renewable energy showing good completion of loans.
- Grades F, G have higher risk of default compared to others.