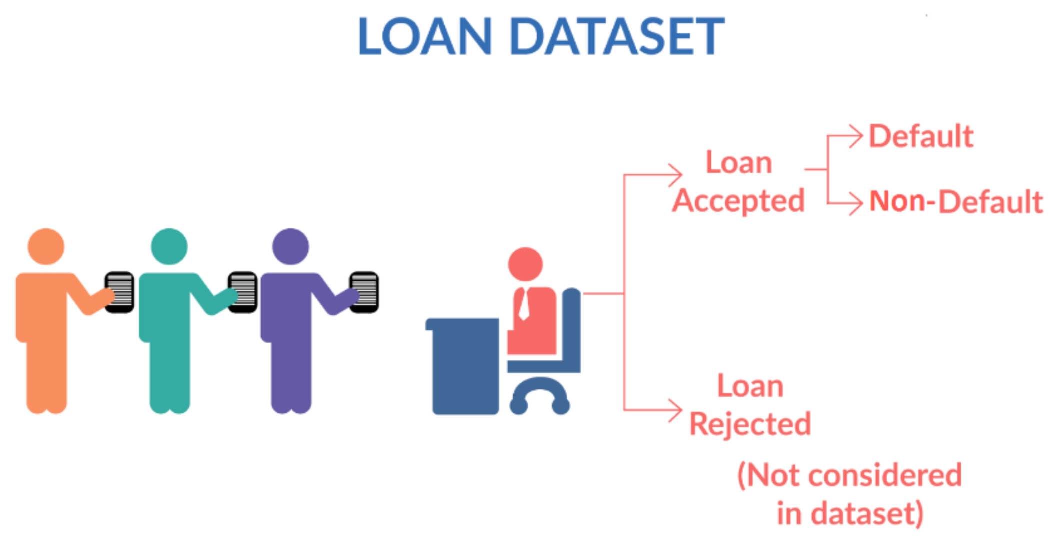
# Lending Club Case Study

Lending company wants to understand the driving factors (or variables) behind loan default, i.e. the variables which are strong indicators of default. The company can utilize this knowledge for its portfolio and risk assessment.



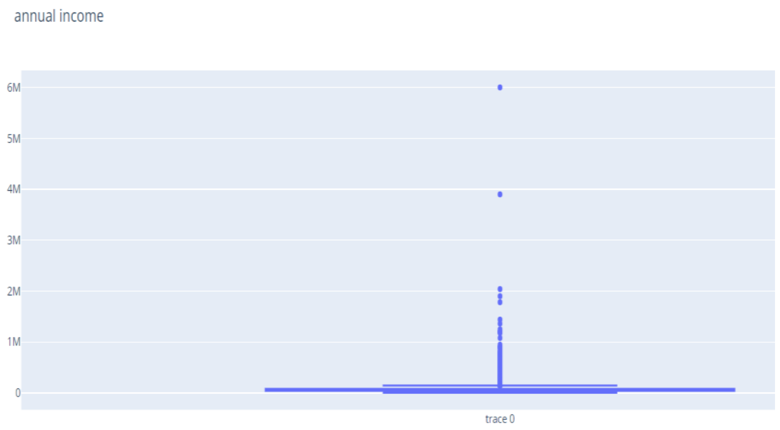
# Exploratory Data Analysis

- Removed Unwanted Column and null columns which are not necessary for this use case study.
- Removed Loan status "current" rows from the dataset as there is uncertainty involved
- Handling missing values
- Standardizing the column for further analysis
- Outlier detection using box plot for better analysis
- Performed Feature engineering to gain better insights

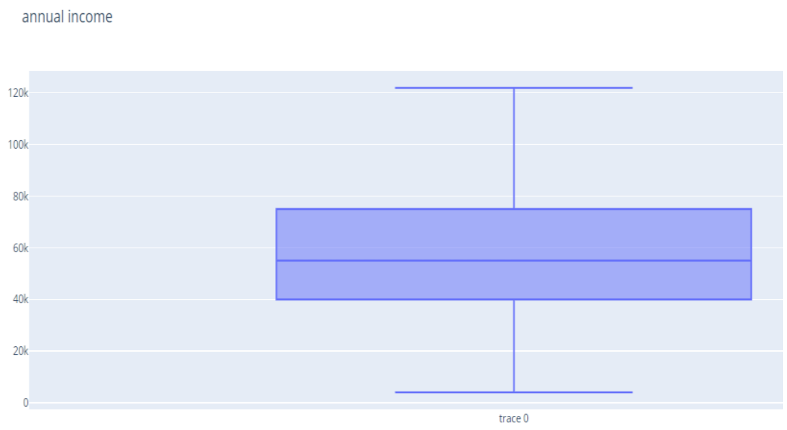
```
[163] loan_df.shape
(39717, 111)
```

Transformation

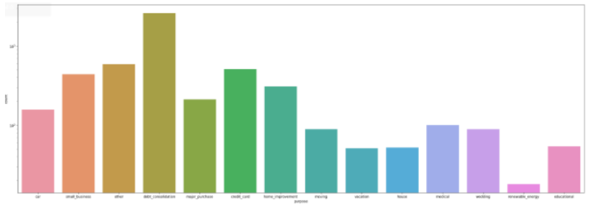
```
[275] loan_df.shape
(35469, 23)
```

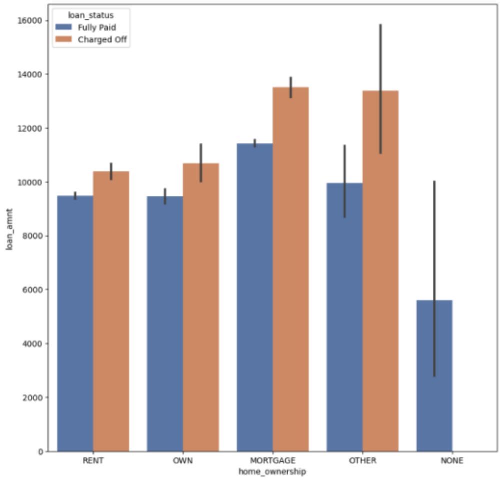



Transformation

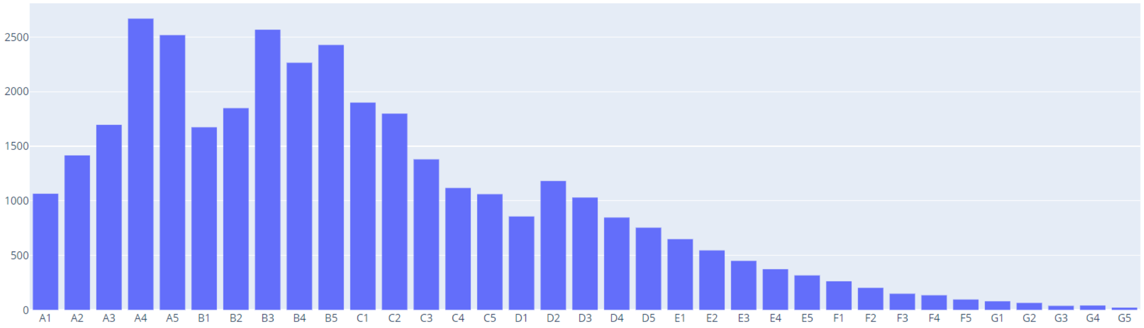


# Exploratory Data Analysis






Sub Grade




- Performed Categorical Data Analysis
- Quantitative Data Analysis
- Segmented Univariate Analysis
- Multivariate Data Analysis for better insights

# Actionable Insights

- Applicant who takes loan for "Debt Consolidation" purpose tends to default more than the other.
- Applicant who takes loan in the month of "Dec" are more prone to default than others.
- Interest rate in higher in all loan amount group where loan status is charged off.
- Applicant with loan amount group between 48K-91K are more prone to default and same goes true for funded amount and funded amount investor group as well
- Applicant with number of installment within range 176-338 tends to charged off more
- Applicant whose annual income group falls between 33K-63K are more prone to charged off more than other
- Applicant whose interest rate group falls between 10%-17% are more prone to charged off more than other
- Applicant whose DTI ratio falls between 11-22 are more prone to charged off than other