Lending Club Case Study

BINOJ VARGHESE

Table of Contents

1. Overview and Problem Statement
2. Understanding Loan Dataset
3. Approach
4. Analysis
5. Conclusion

Overview and Problem Statement

Company Overview:
This company 'Lending Club' 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.

Business Context:
Lending Club wants to understand the driving factors behind loan default as approving loans to 'risky' applicants is the largest source of financial loss (called credit loss). Identification of these risky loan applicants is crucial to the business to limit the amount of credit loss.

Problem Statement
The aim is to identify patterns using Exploratory Data Analysis (EDA) of the given 'Loan Dataset' which indicates if a person is likely to default, which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc.

Understanding the Loan Dataset

❑ It contains the complete loan data for all loans issued through the time period 2007 to 2011.
❑ Loan rejected details are not considered in this give dataset
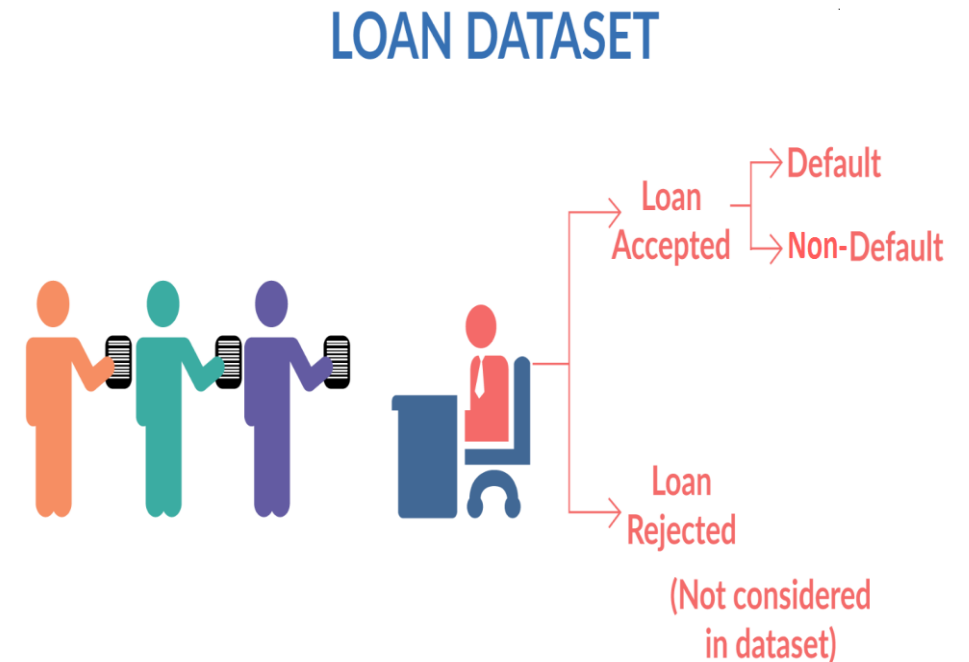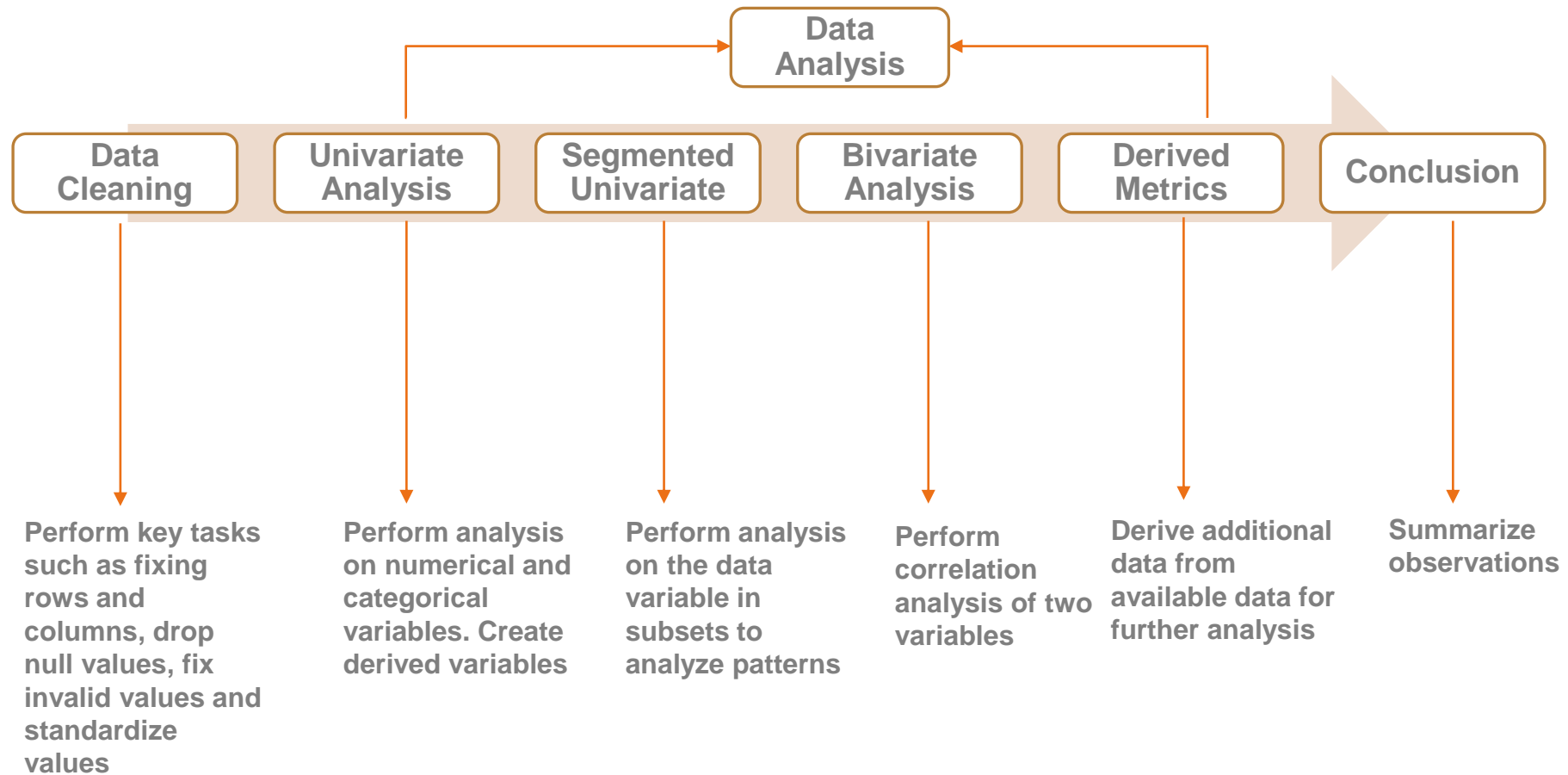❑ Loan dataset contains 39717 rows and 111 columns for review, data cleaning and analysis.

LOAN DATASET

Fig 1.1 Loan Dataset

Analysis Approach

Data Analysis

Data Cleaning | Univariate Analysis | Segmented Univariate | Bivariate Analysis | Derived Metrics | Conclusion

Perform key tasks such as fixing rows and columns, drop null values, fix invalid values and standardize values

Perform analysis on numerical and categorical variables. Create derived variables

Perform analysis on the data variable in subsets to analyze patterns

Perform correlation analysis of two variables

Derive additional data from available data for further analysis

Summarize observations

Analysis – Loan Dataset

❑ Loan Dataset has two categories of loan terms - for a period of 36 and 60 months.
❑ Under 'loan status', loans are categorized as Fully paid, Current (omitted in our analysis) and Charged off.
❑ The ratio of fully paid vs charged off customers is 85 and 15 percent

Observation
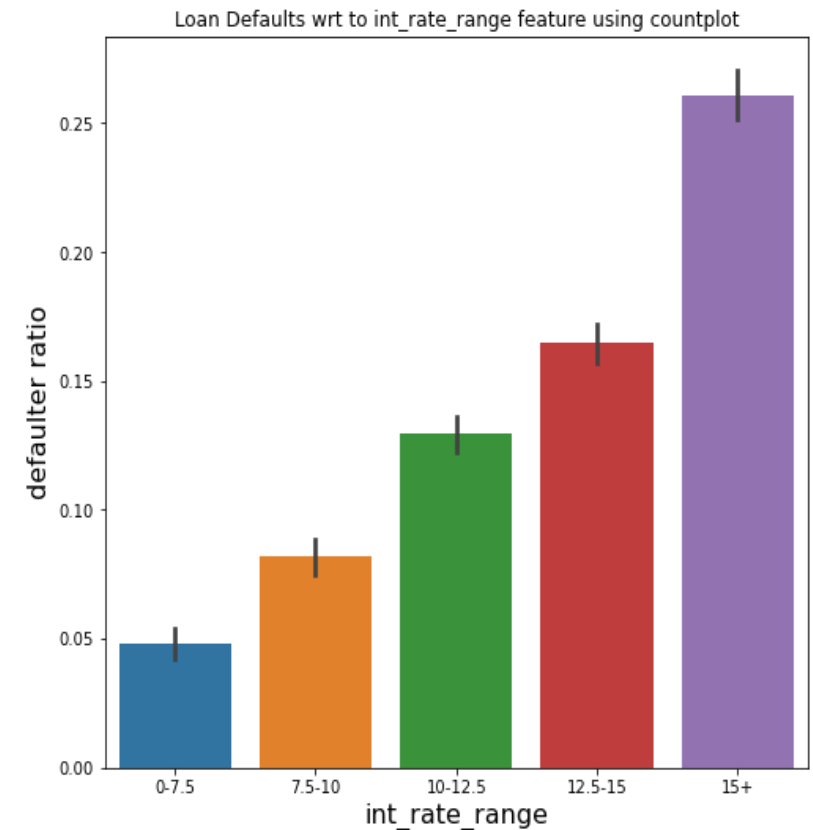❑ Higher number of defaulters for loans with 60 months term

Loan Defaults wrt to term feature using countplot
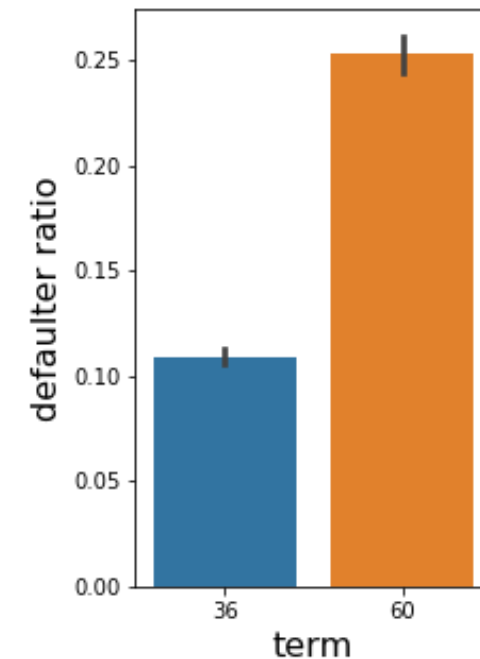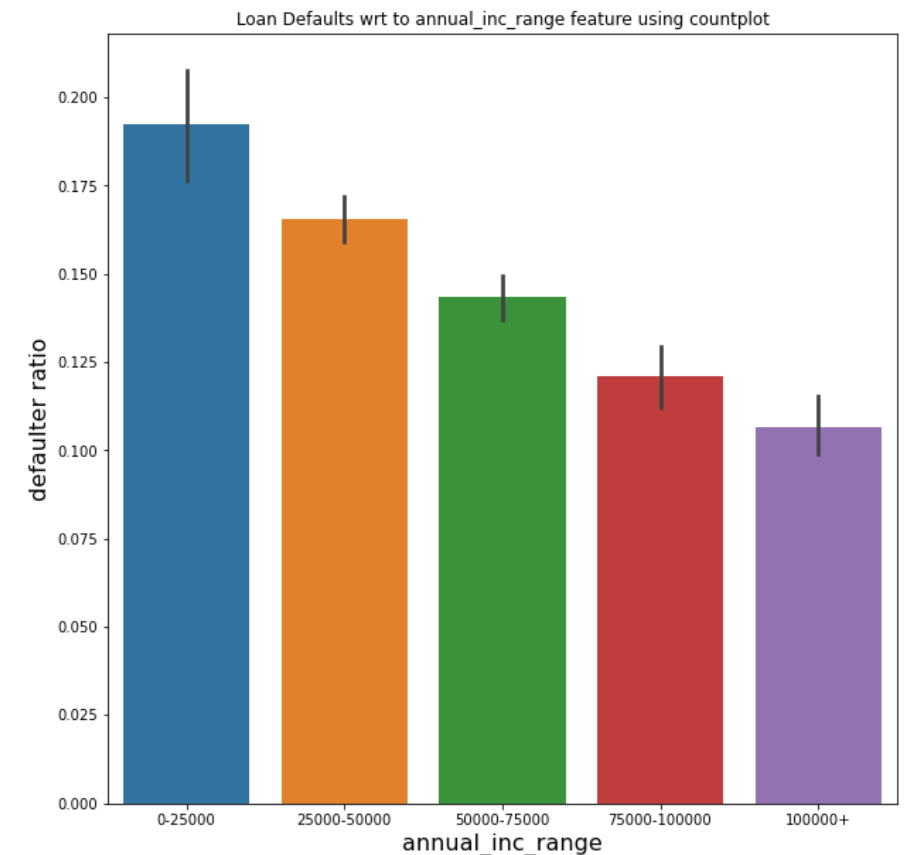
Analysis – Loan Dataset

Observation
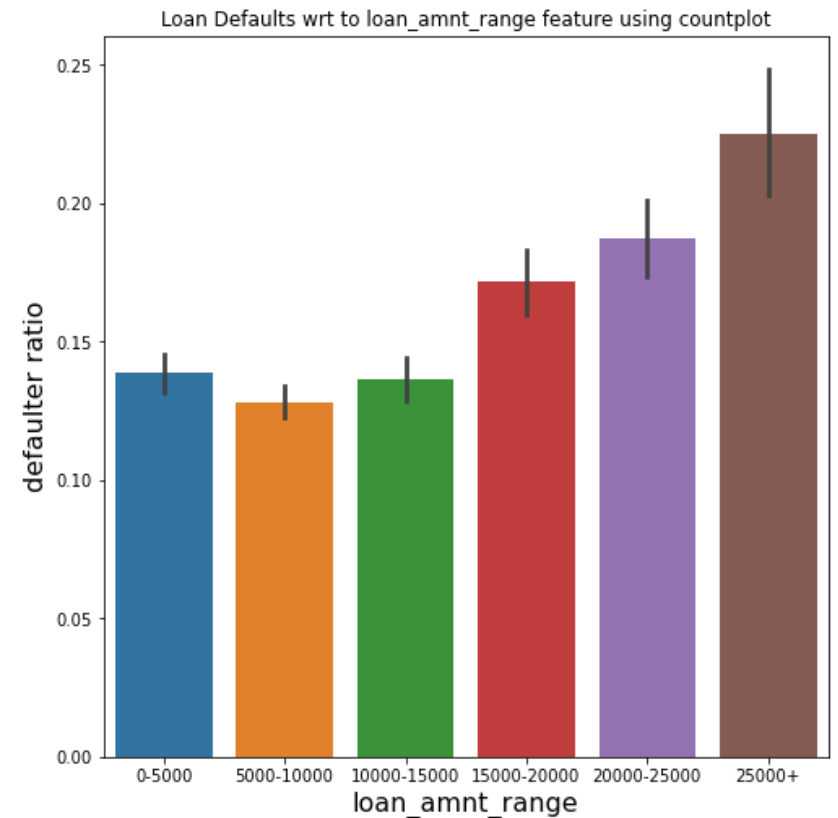❑ Loan defaulters rate increase as the interest increases

Loan Defaults wrt to int_rate_range feature using countplot

Analysis – Loan Dataset

Observation
❑ Loan defaulters rate decrease as the annual income range increases

Loan Defaults wrt to annual_inc_range feature using countplot

Analysis – Loan Dataset

Observation

❑ Loan defaulters rate increases as the loan amount increase. There's a steep increase in defaulters ratio from loan amount > 15000


Loan Defaults wrt to loan_amnt_range feature using countplot

Analysis – Loan Dataset

Observation

❑ Loan defaulters rate increases as the dti* percentage increases.

* dti - A ratio calculated using the borrower's total monthly debt payments on the total debt obligations, excluding mortgage and the requested LC loan, divided by the borrower's self-reported monthly income.
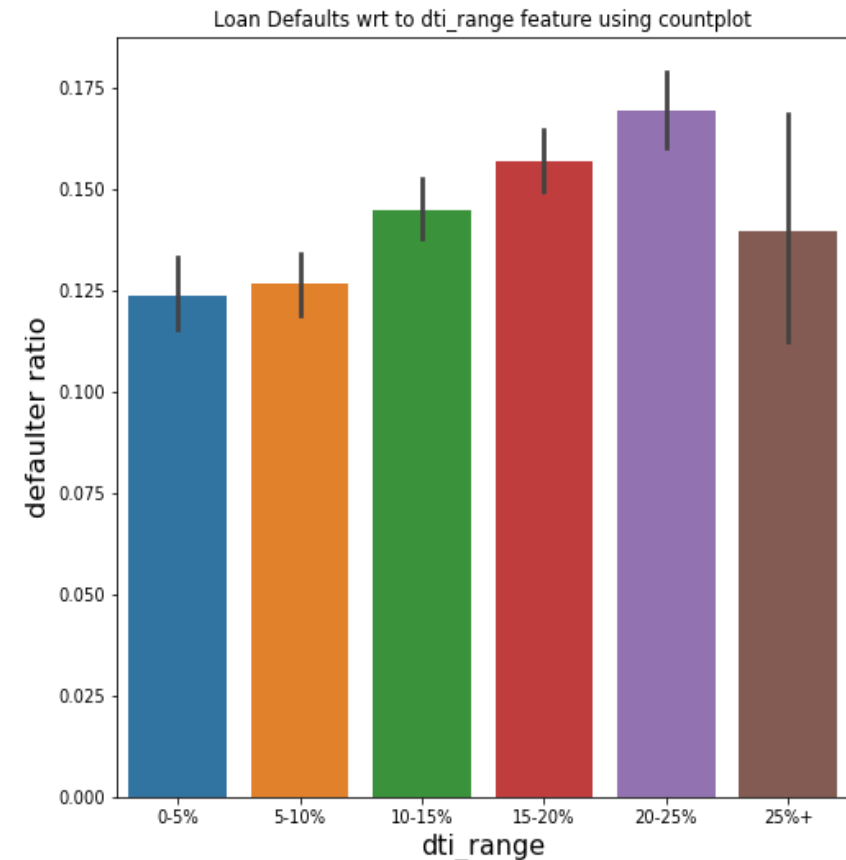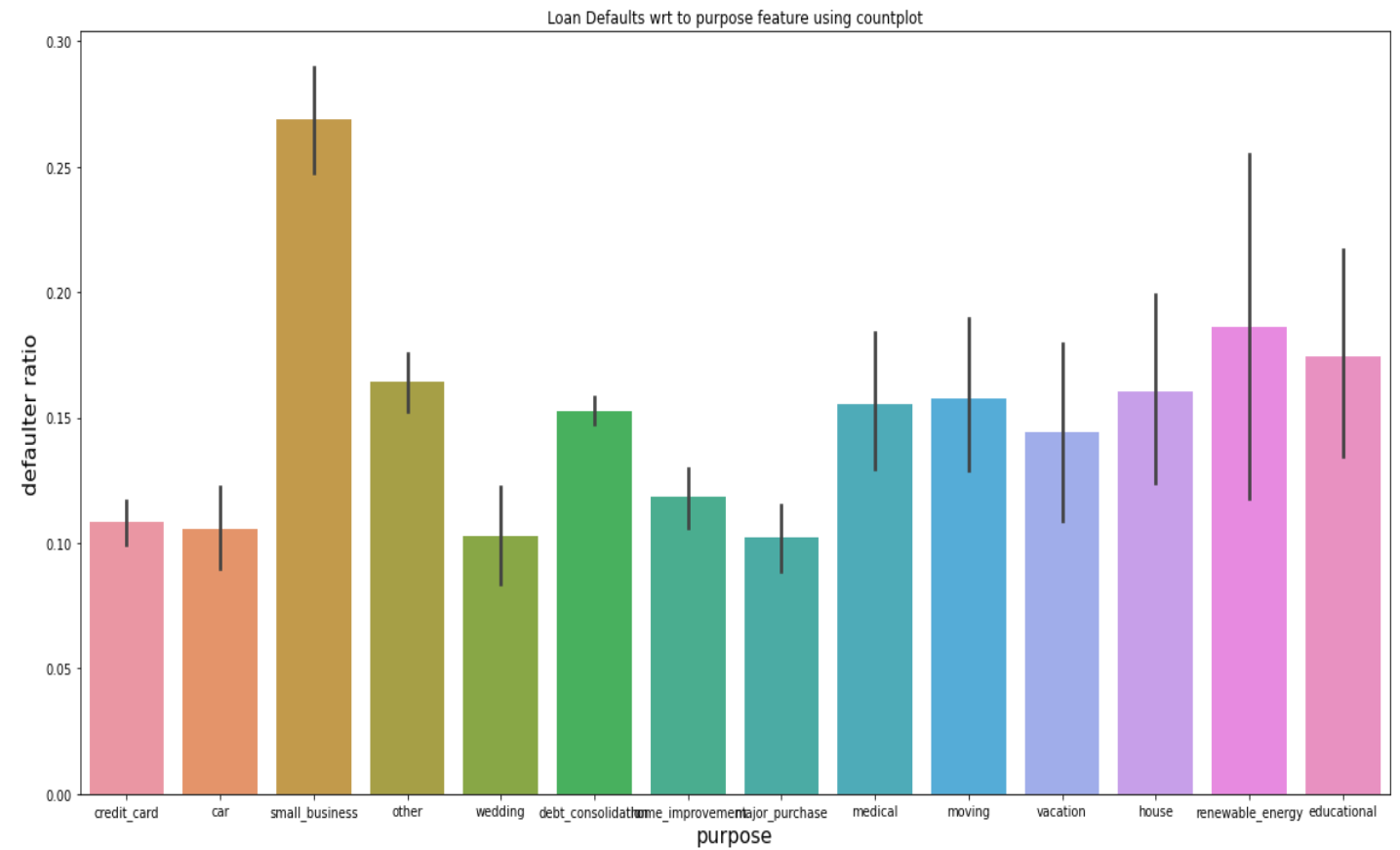
Analysis – Loan Dataset

Observation
❑ Strong correlation between interest rate and grade. Loan defaulters rate increases for higher grades and interest rates.

box plot grade vs int_rate for loan_status
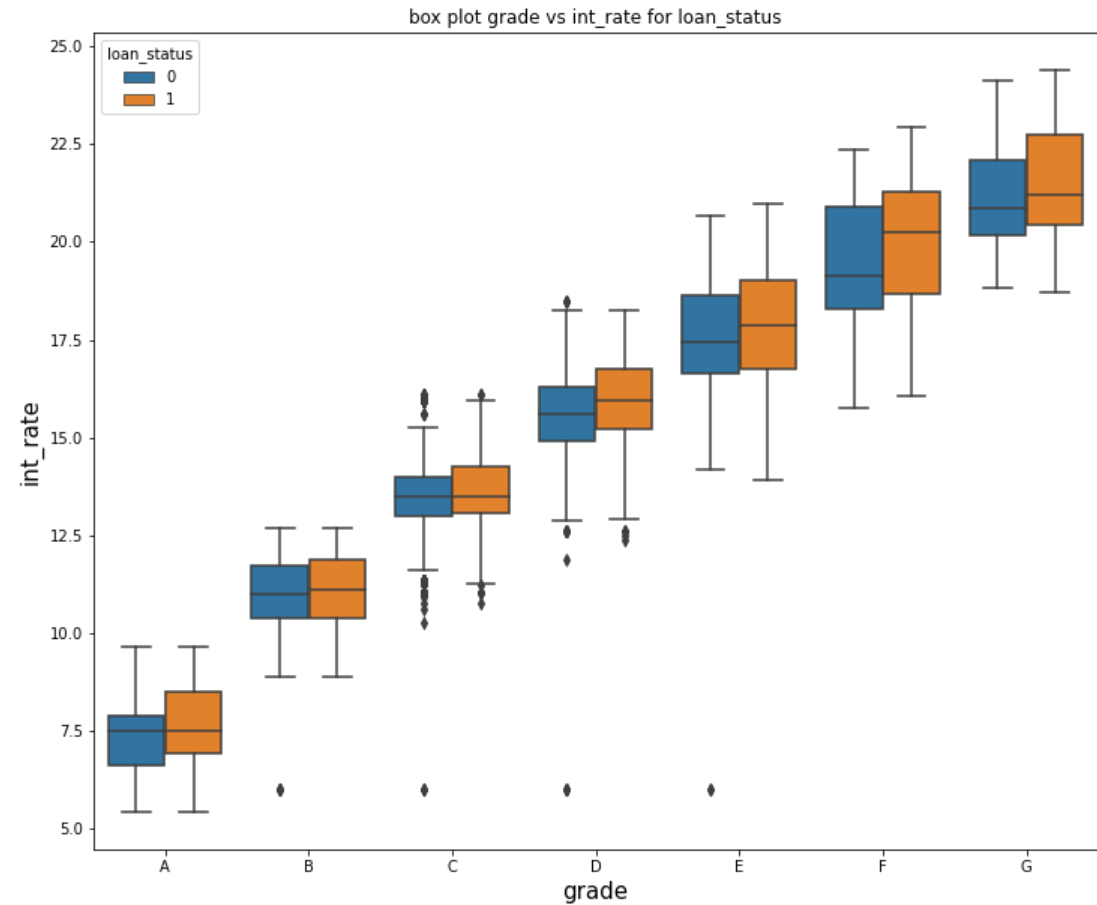
Analysis – Loan Dataset

Observation
❑ Defaulters are highest amongst small business

Loan Defaults wrt to purpose feature using countplot

Conclusion

❑ Loan defaults are high for loans with 60 months term hence focus on shorter term loans would be advisable for risky applications.

❑ Lending Club should restructure interest rates as the defaults increase for higher interest rates.

❑ Loan defaults are the highest amongst small businesses. Lending club should reduce the either the loan amount or the term for such business.

❑ Defaults tend to increase for higher loan amounts, hence Lending Club should focus on providing smaller amounts to reduce risk of defaults.

❑ Lending should also focus on offering loans to customers with higher annual income as defaults are less