

Gramener Case Study



Nikhil Tiwari
DDA1710184

Objective

- ❖ XYZ Bank deals in giving personal Financial Loans of all sorts. Business Objectives:
 1. Analyse the Loan data set of 4 years of current, Fully paid and Defaulted Loans.
 2. Analyse the dataset and hypothesize the multiple cause of Defaults.
 3. Possible Solution to the Issue

Problem analysis

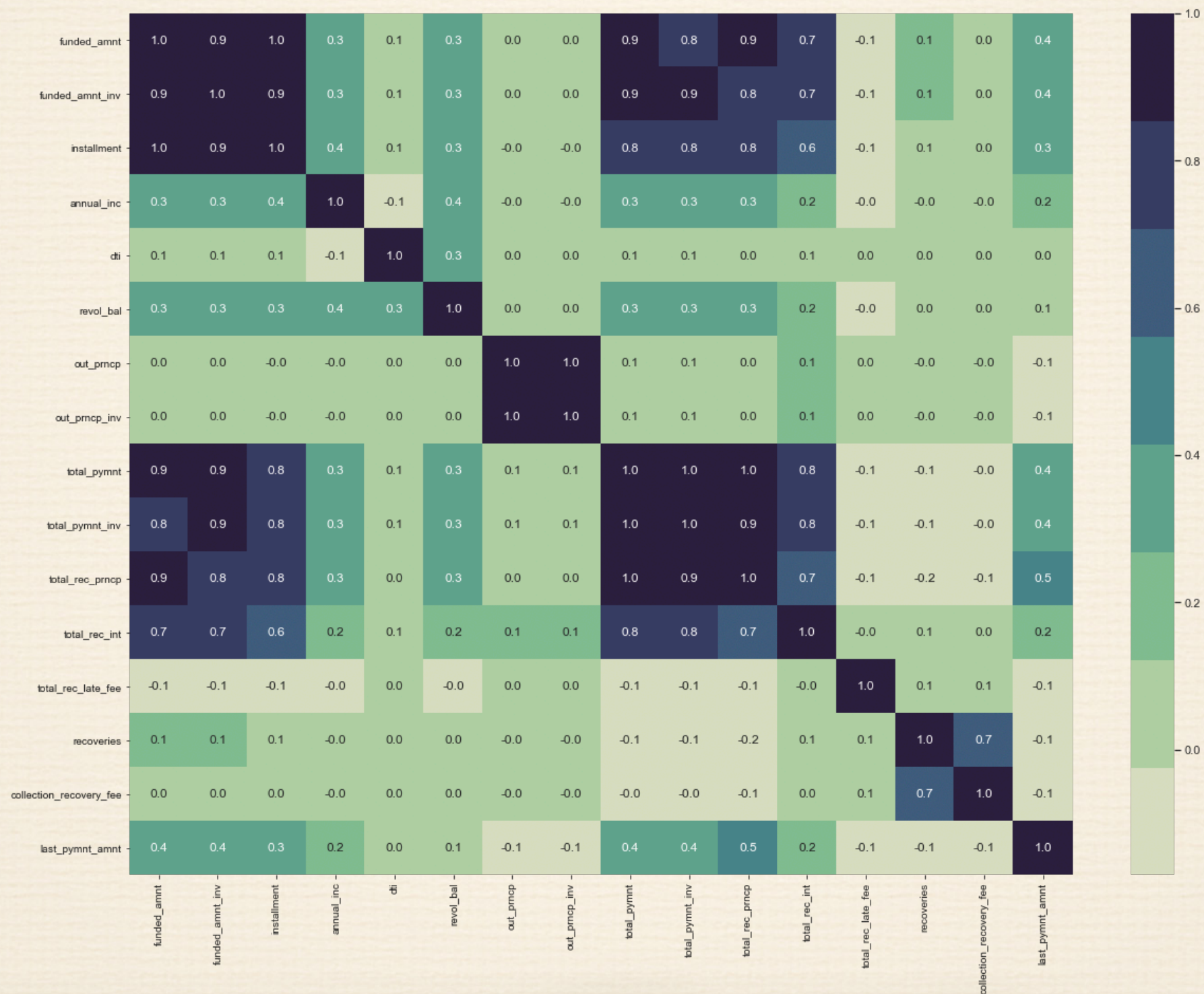
- ❖ Understanding Data: This data has information about 4 years of loans that have been issued by the bank. They are mainly of three types Fully Paid off, Current and Defaulted. The data set contains all sorts of information about customer like annual income, employee length, purpose of loan, no. of bankruptcies etc.
- ❖ On high level following are stages in the workflow:
 1. Gather data for the analysis.
 2. Clean and Format the data for readability.
 3. Extrapolate the Derived Metrics.
 4. Filter out the outliers based on Annual Income.
 5. Filter Data Set on Loan Status.
 6. Perform Univariate and Bivariate analysis on the Data.
 7. Plot the results of Analysis and Hypothesize the Root of the Issue.

Problem Analysis

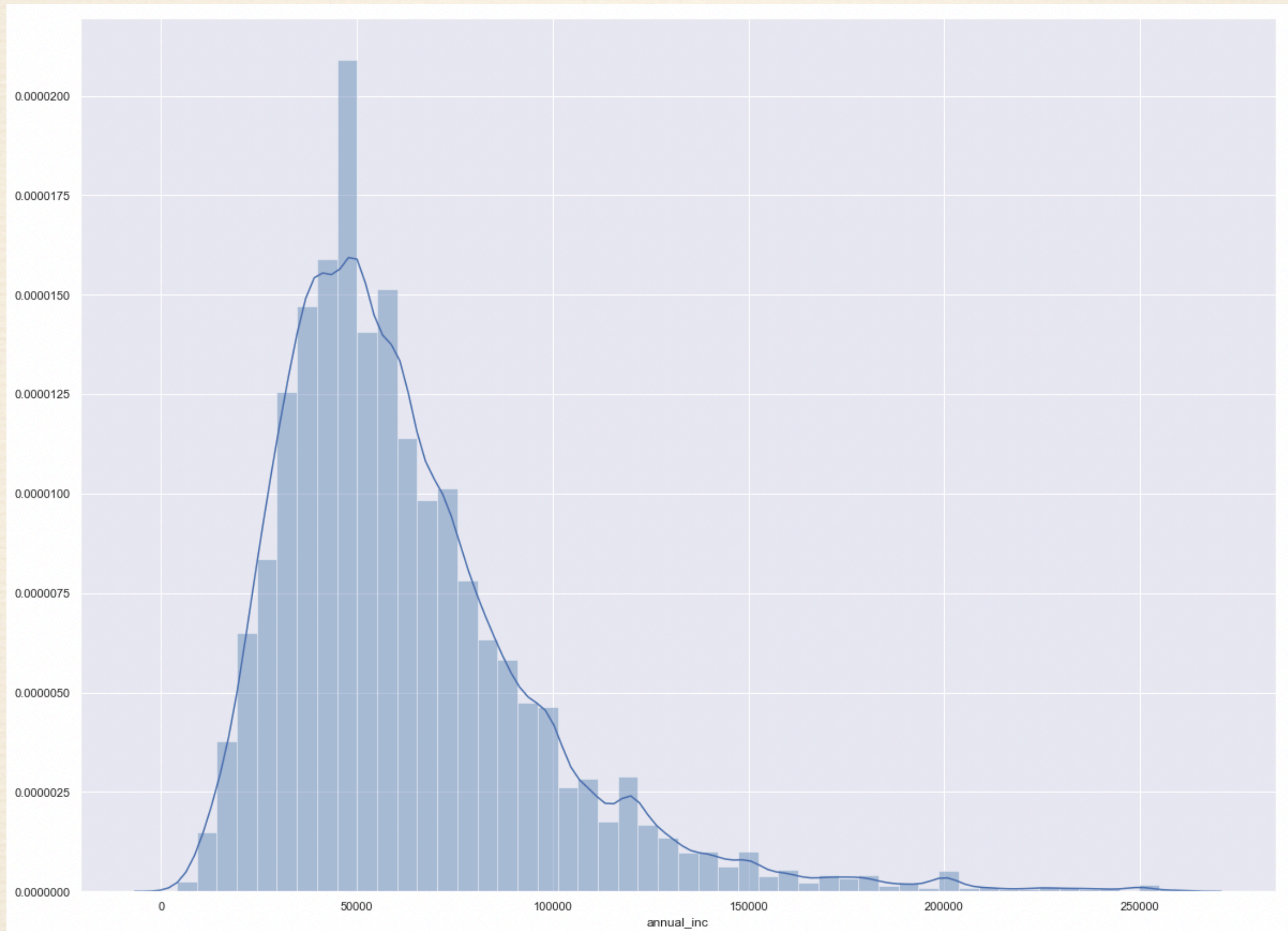
❖ Assumptions:

1. As current loans have not yet been completed, we decided to ignore them during analysis.
2. Outliers (Based on Z-Score) have been considered and addressed during analysis.
3. Tools: Python 3.6 & Jupyter Notebook

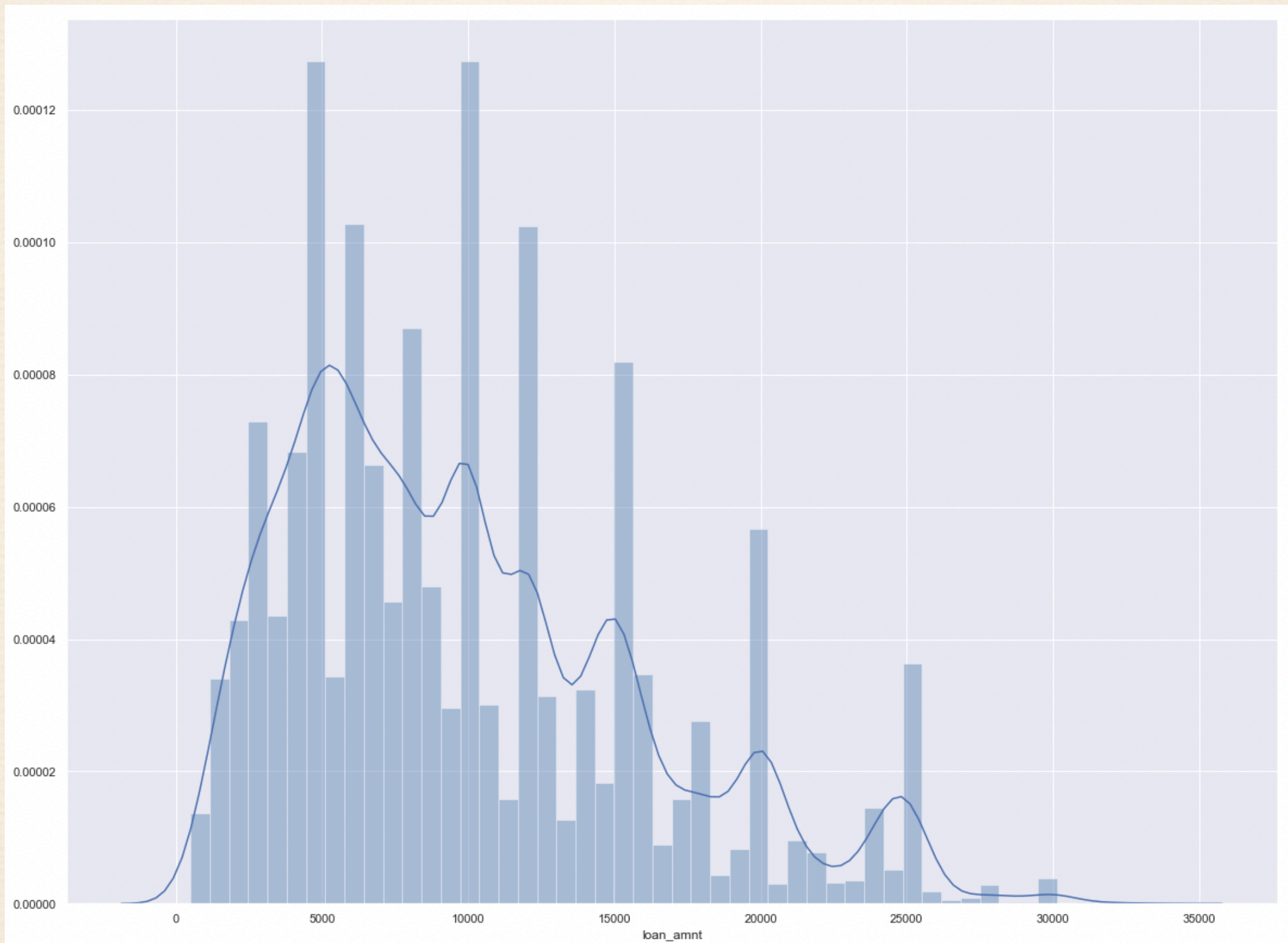
Correlation among variable (Cleaned Dataset)



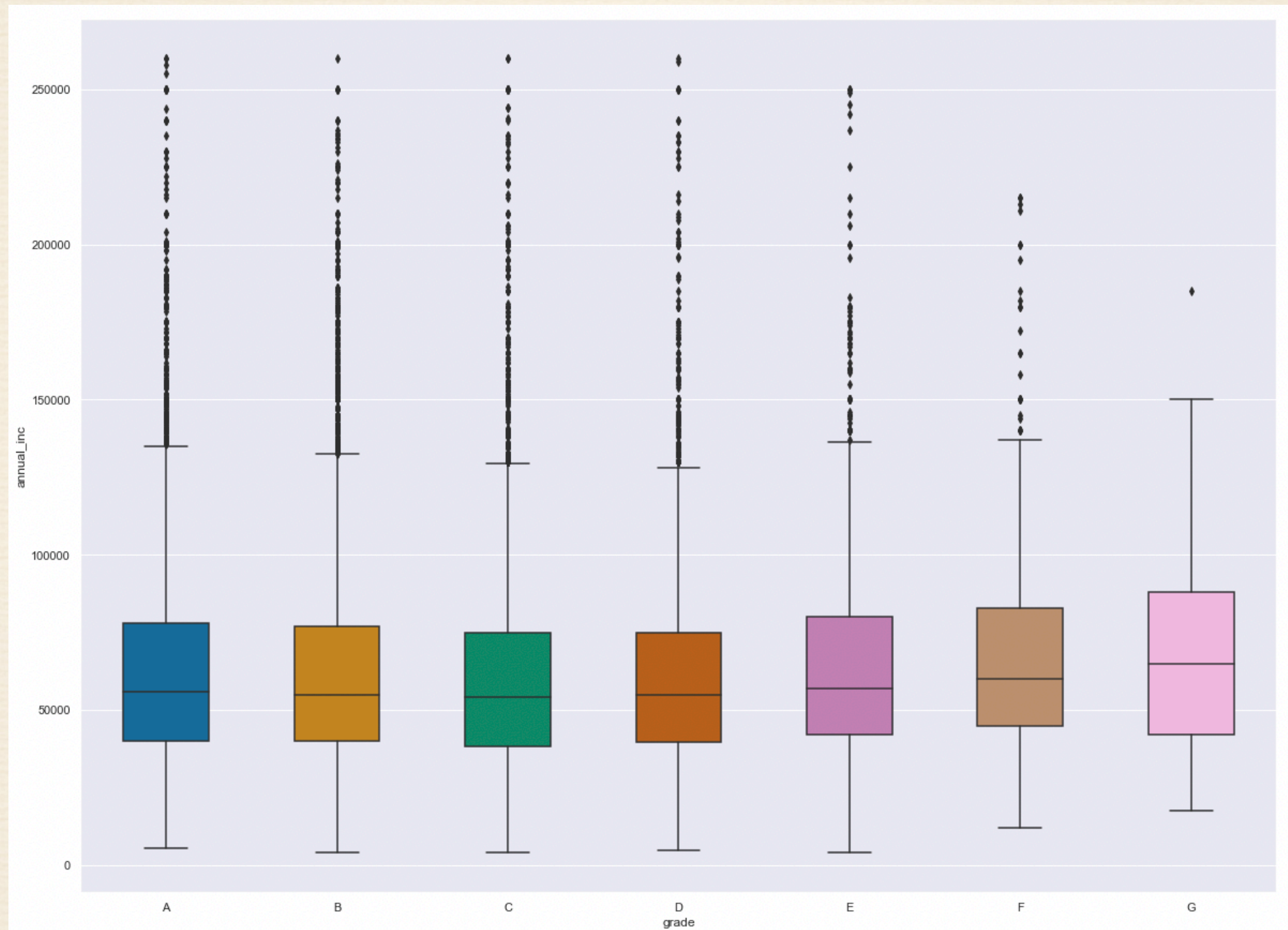
Income Distribution



Loan Distribution

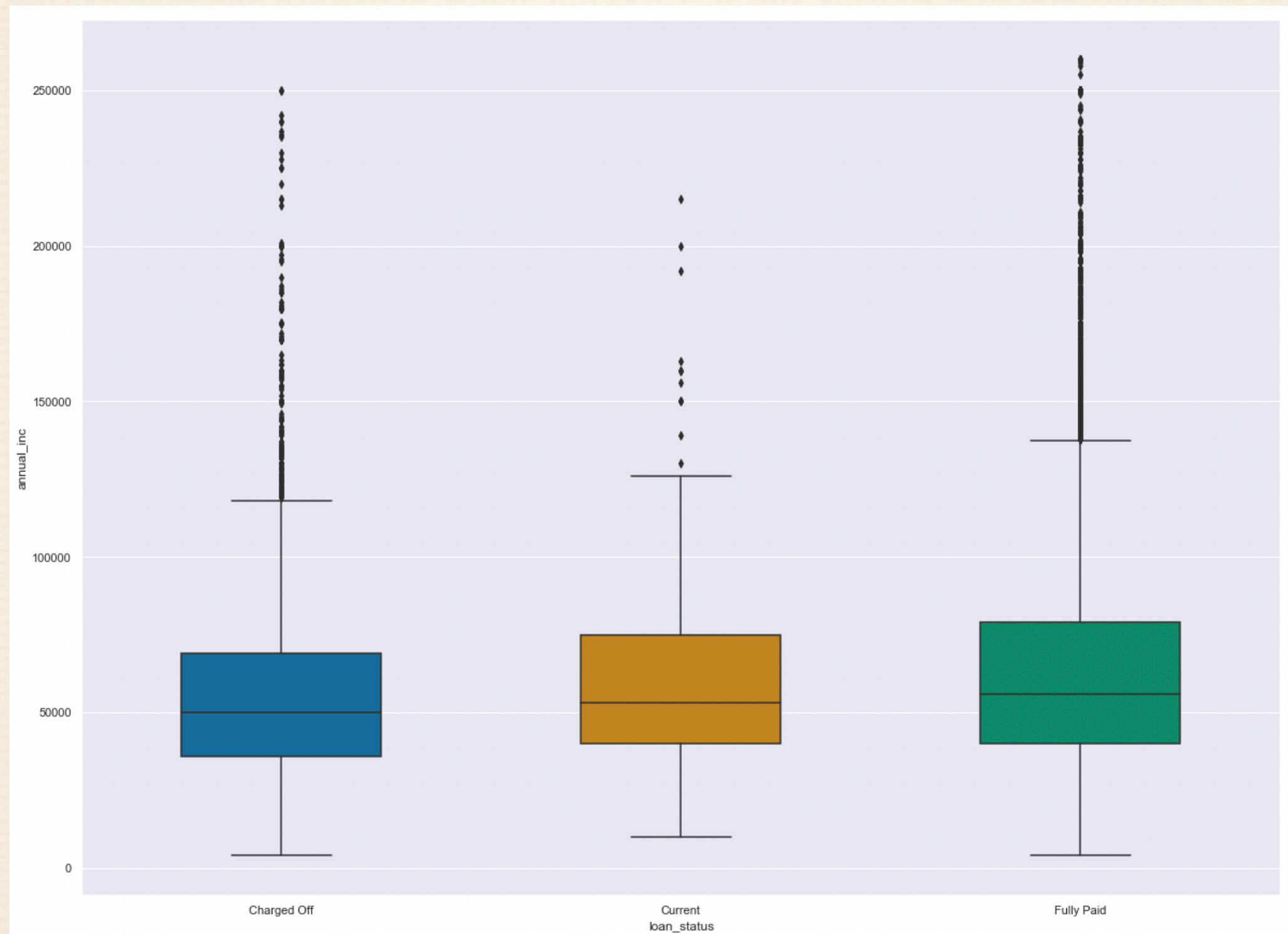


Annual income vs Grade

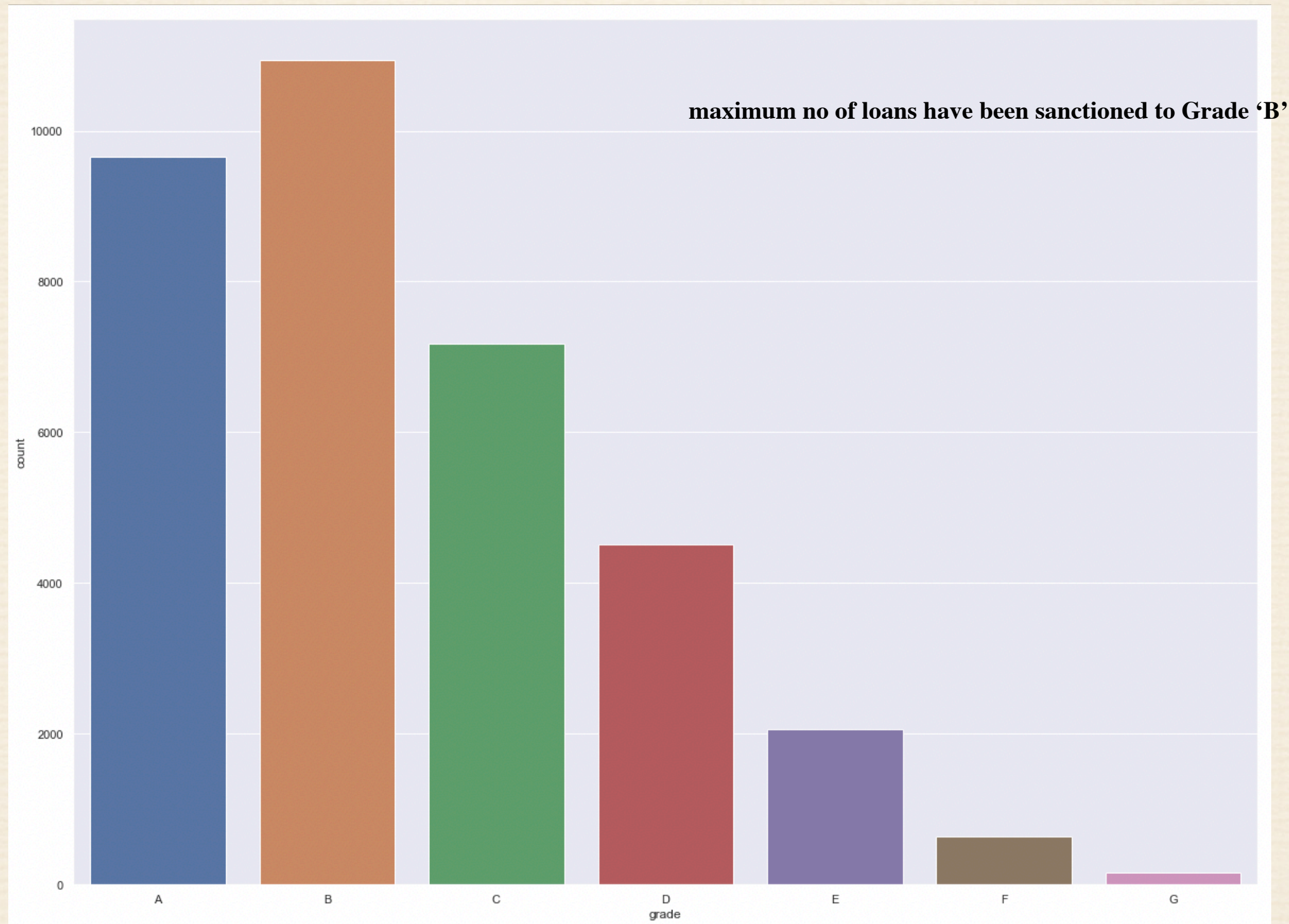


The median income as well as the range of the incomes is almost same across all the grades. It is a little more for grades E,F and G.

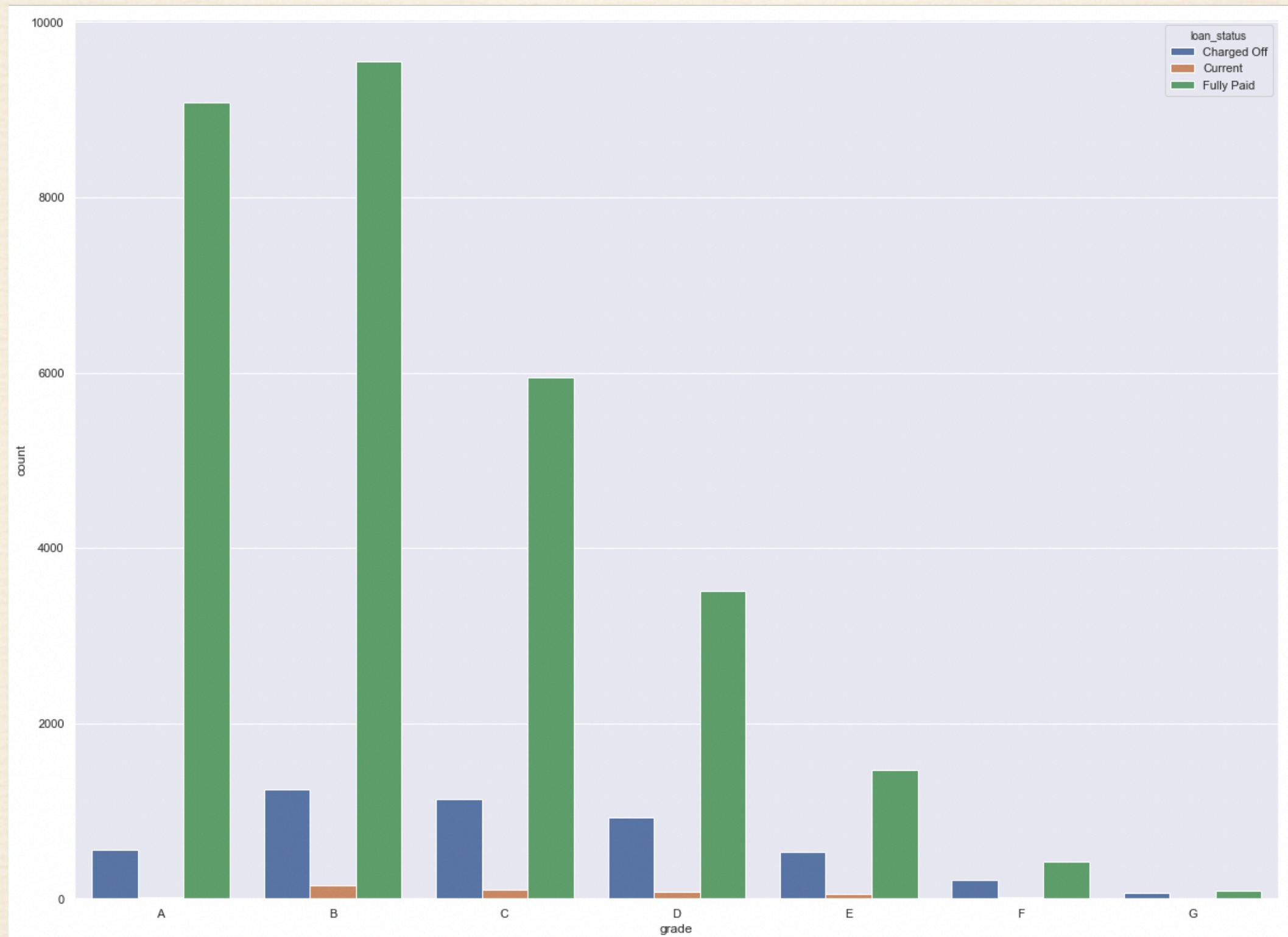
Annual income vs Loan Status



Count of Loans vs Grade

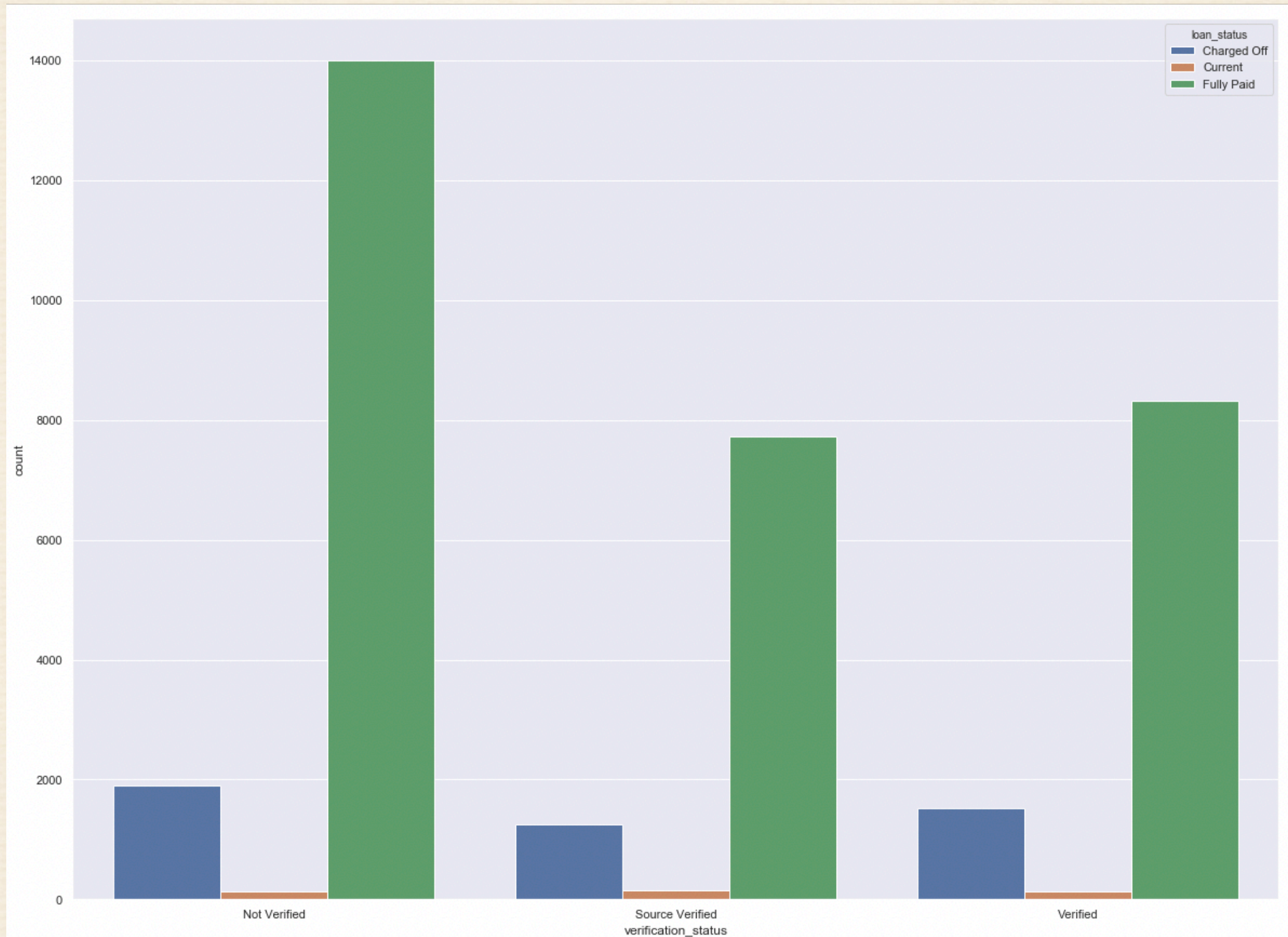


Gradewise count of loans vs their status



Grade B has maximum loans sanctioned and has maximum charged off loans, agency should revisit their sanction criteria.

Loan status vs verification



as default rate is relatively higher in case of not verified loans extra precautions should be exercised before sanctioning (although its a high risk and high reward scenario). a complete verification should be done.

Conclusion

- ❖ After examining annual income and grades, we have found that all the grades have almost similar distribution.
- ❖ Annual Income, interest rate and verification have bearing on whether a loan will become fully paid or become charged off. Hence, these parameters need to be critically monitored.
- ❖ In the current disbursement, from grade D to grade G, the percentage of defaults are more. Grades also play an important role in loan disbursement and repayment and can be reference from current data to make future decisions.
- ❖ Target based/ last minute loan disbursements should be avoided as default rate in last month is very high.