



# Lending Club Case Study

---

SHIVASHANKAR & SHIPRA

# Objective

---

- Identification of applicants using EDA is the aim of this case study.
- Analyzing the driving factors (or driver 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.

# Data cleaning

---

The first few steps involve making sure that there are no **missing values** or **incorrect data** types before we proceed to the analysis stage.

**For Missing Values: Some common techniques to treat this issue are:**

- Dropping the rows containing the missing values
- Deleting the duplicated entries
- Imputing the missing values
- Keep the missing values if they don't affect the analysis

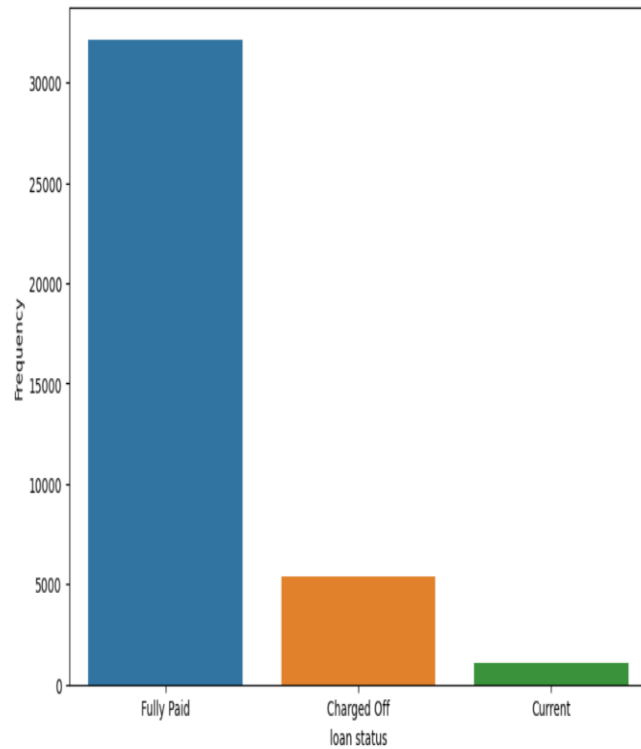
**Incorrect Data Types:**

- Clean certain values
- Clean and convert an entire column

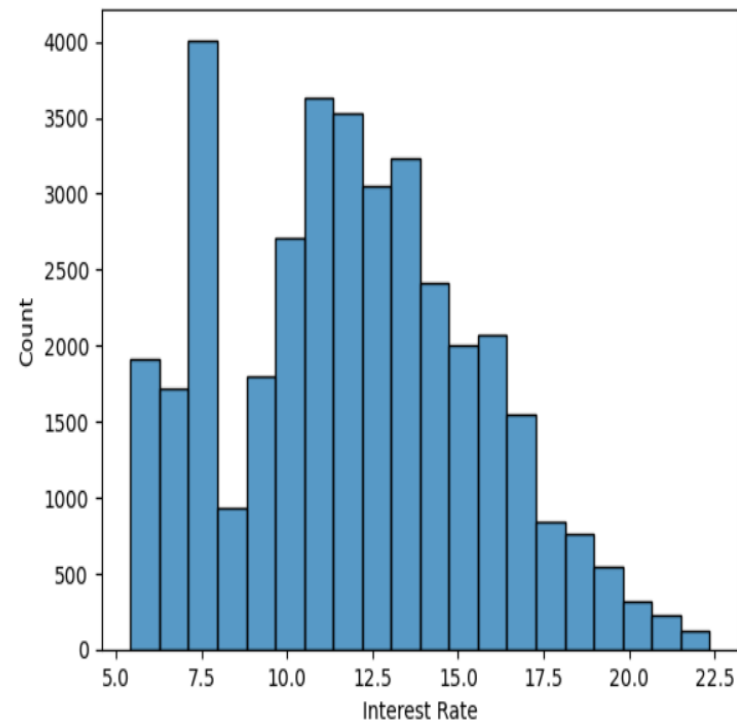
# Univariate Analysis

---

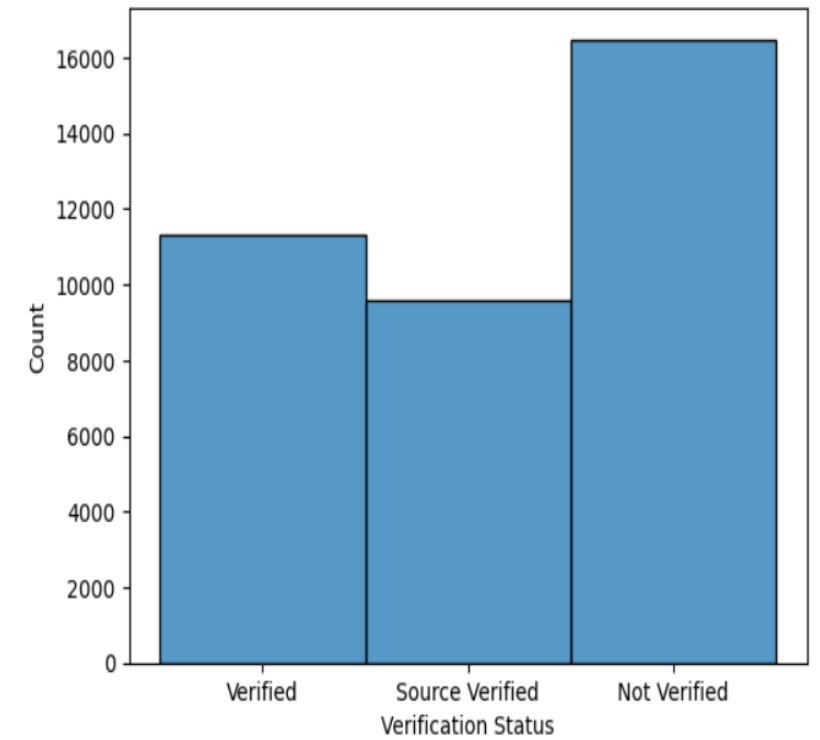
Bar Chart of loan Status



Hist Plot of Interest rate

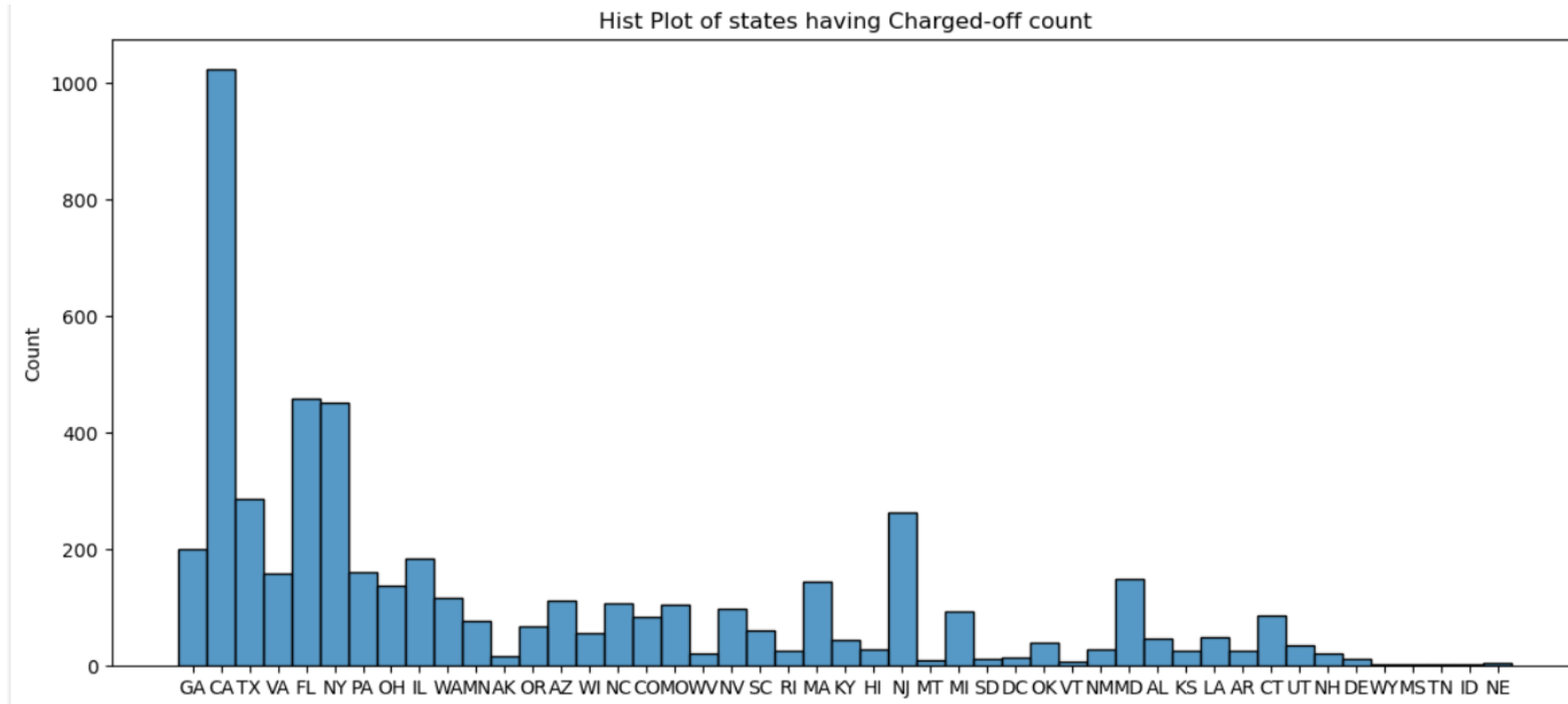


Hist Plot of Verification status of loans

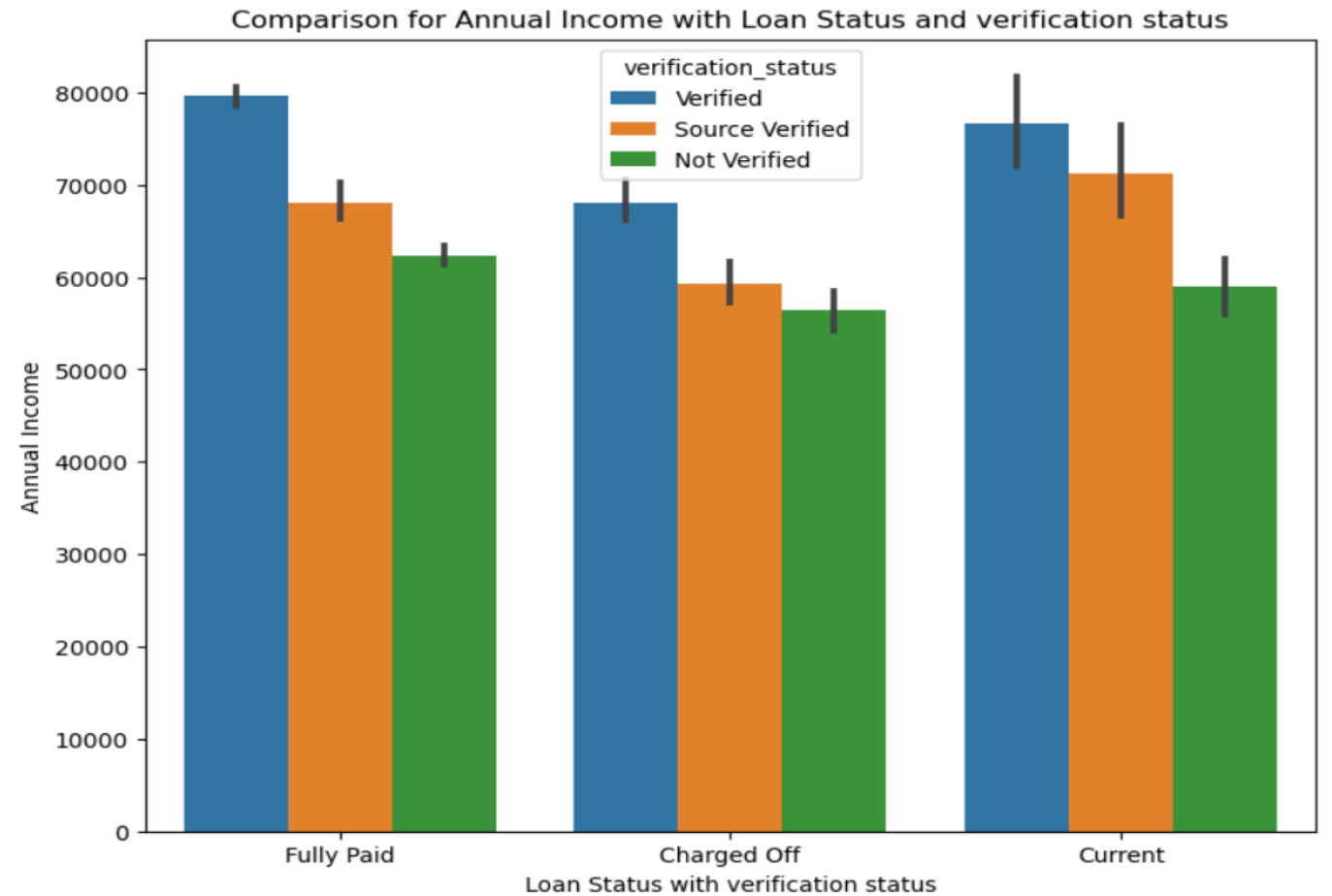
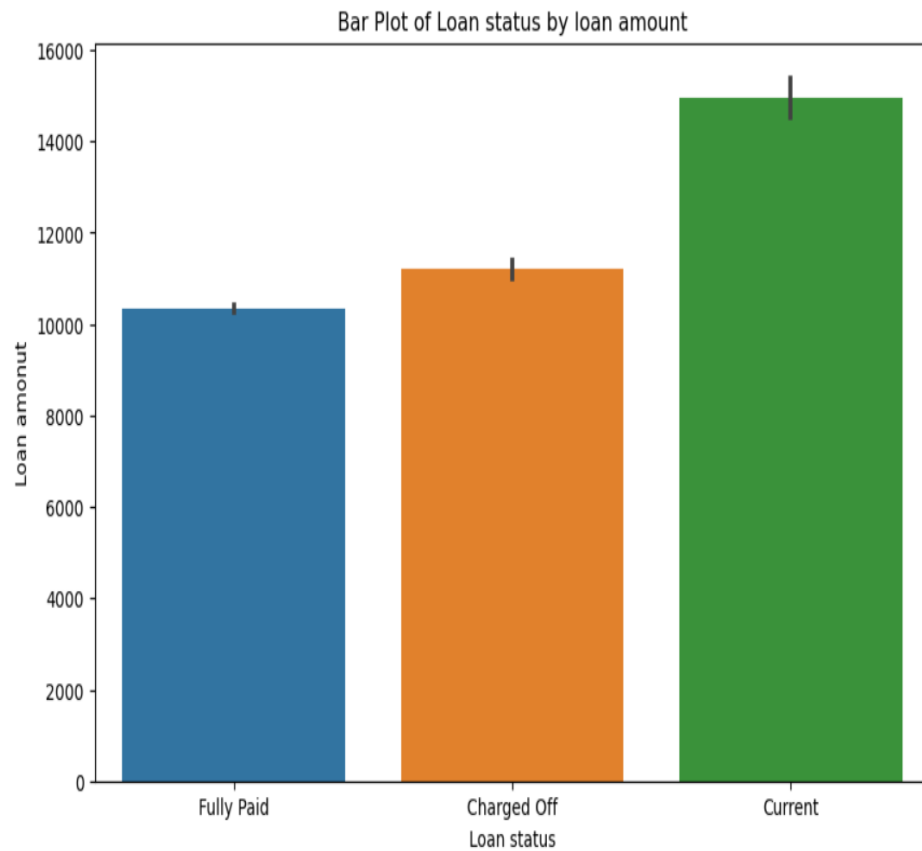


# Univariate Analysis for Charged off

---

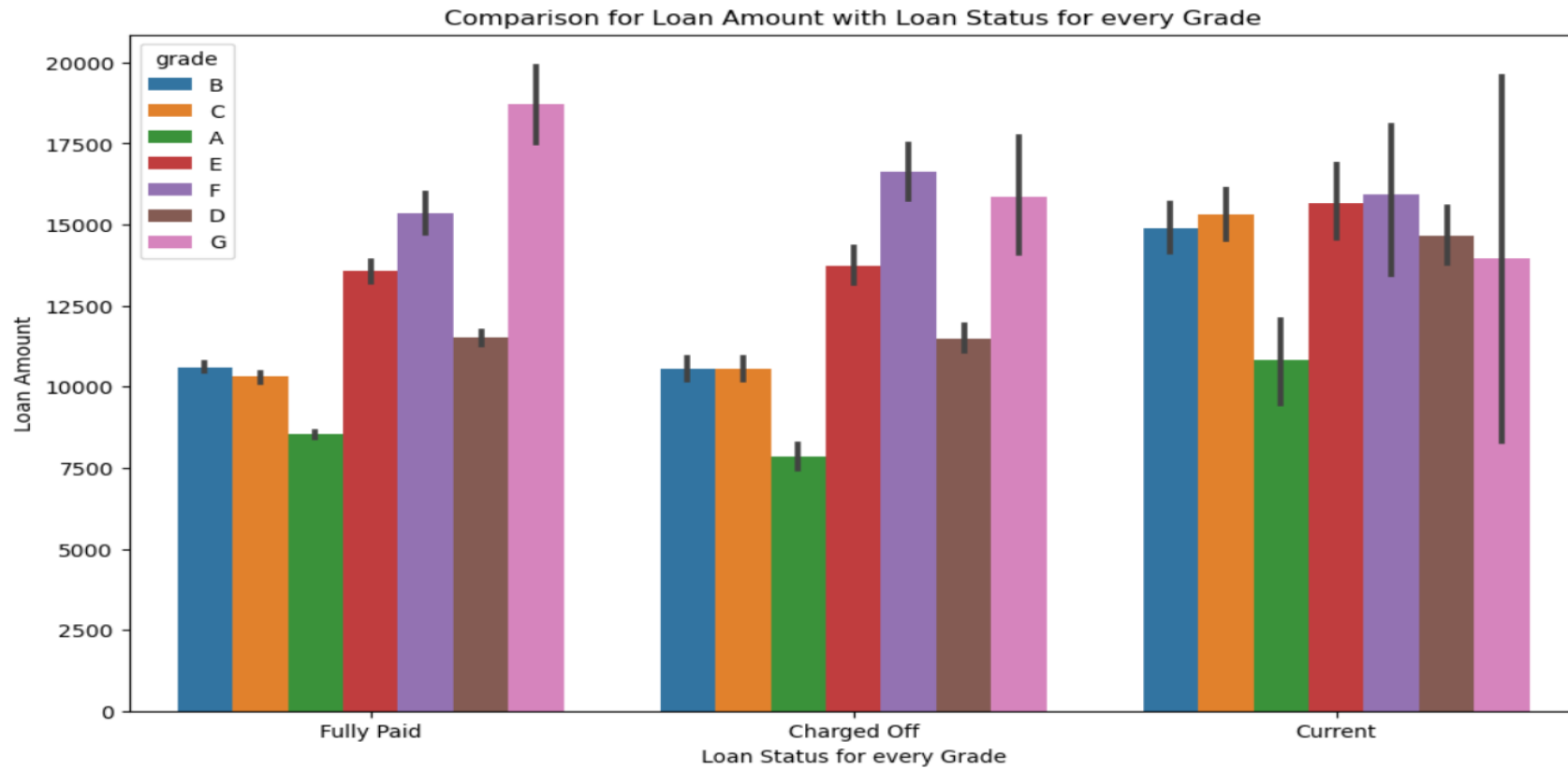


# Bivariate Analysis

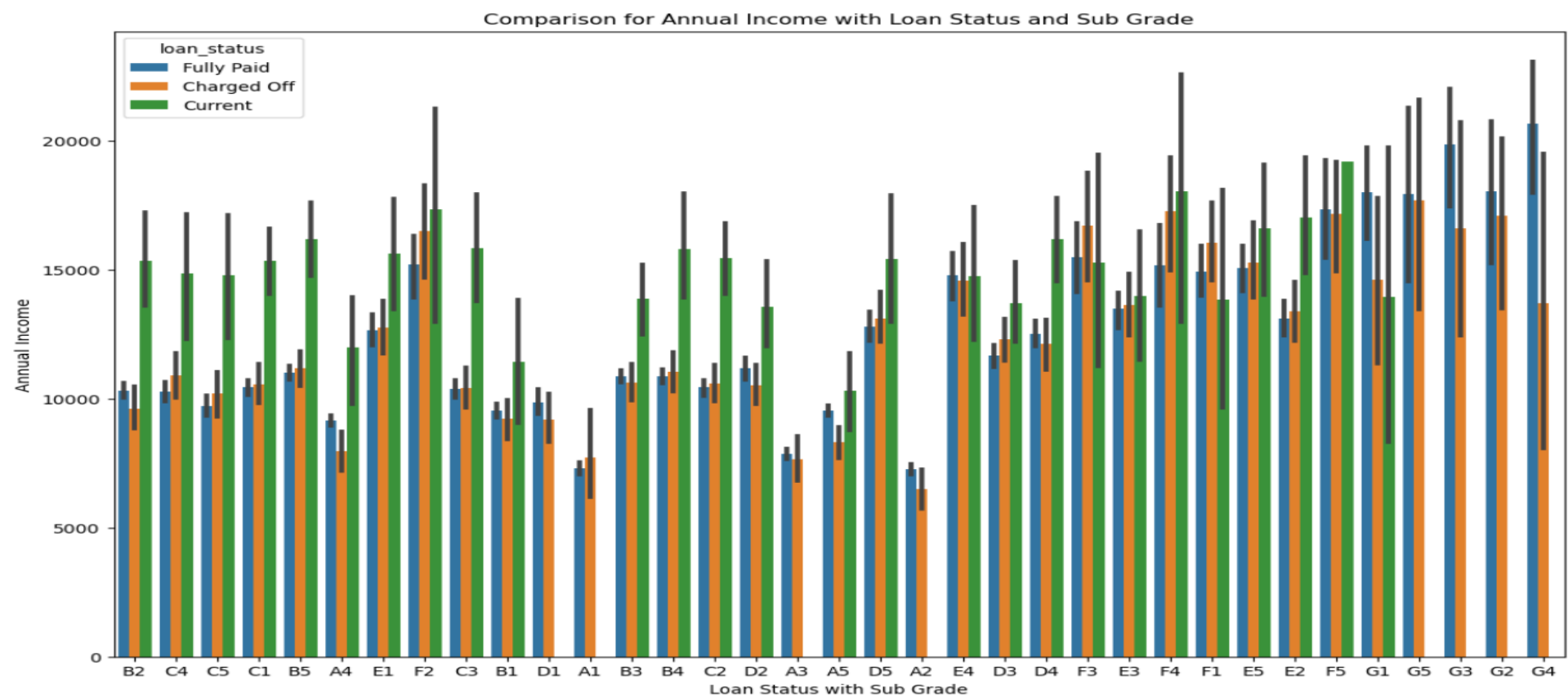


# Bivariate Analysis

---

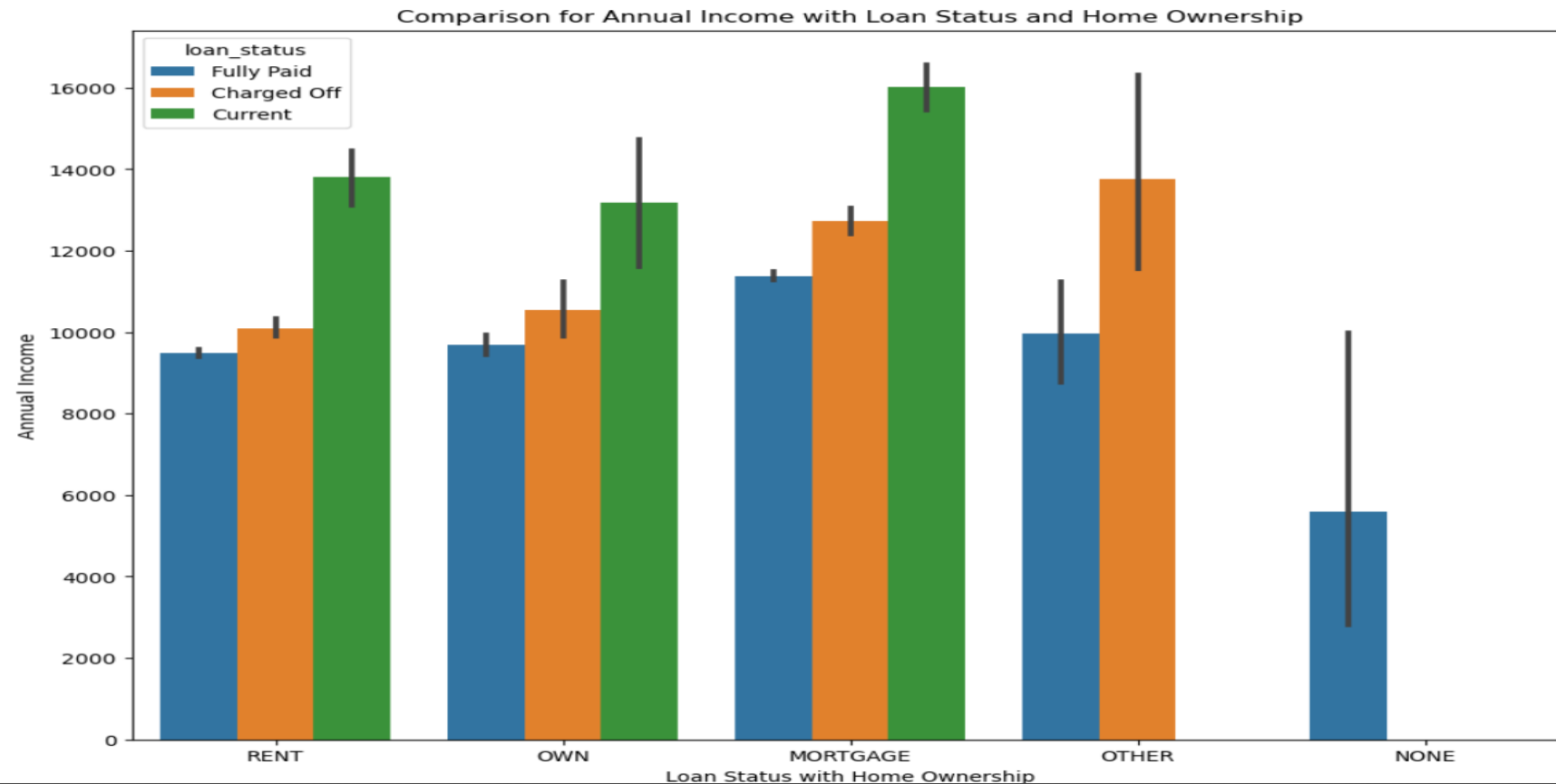


# Bivariate Analysis

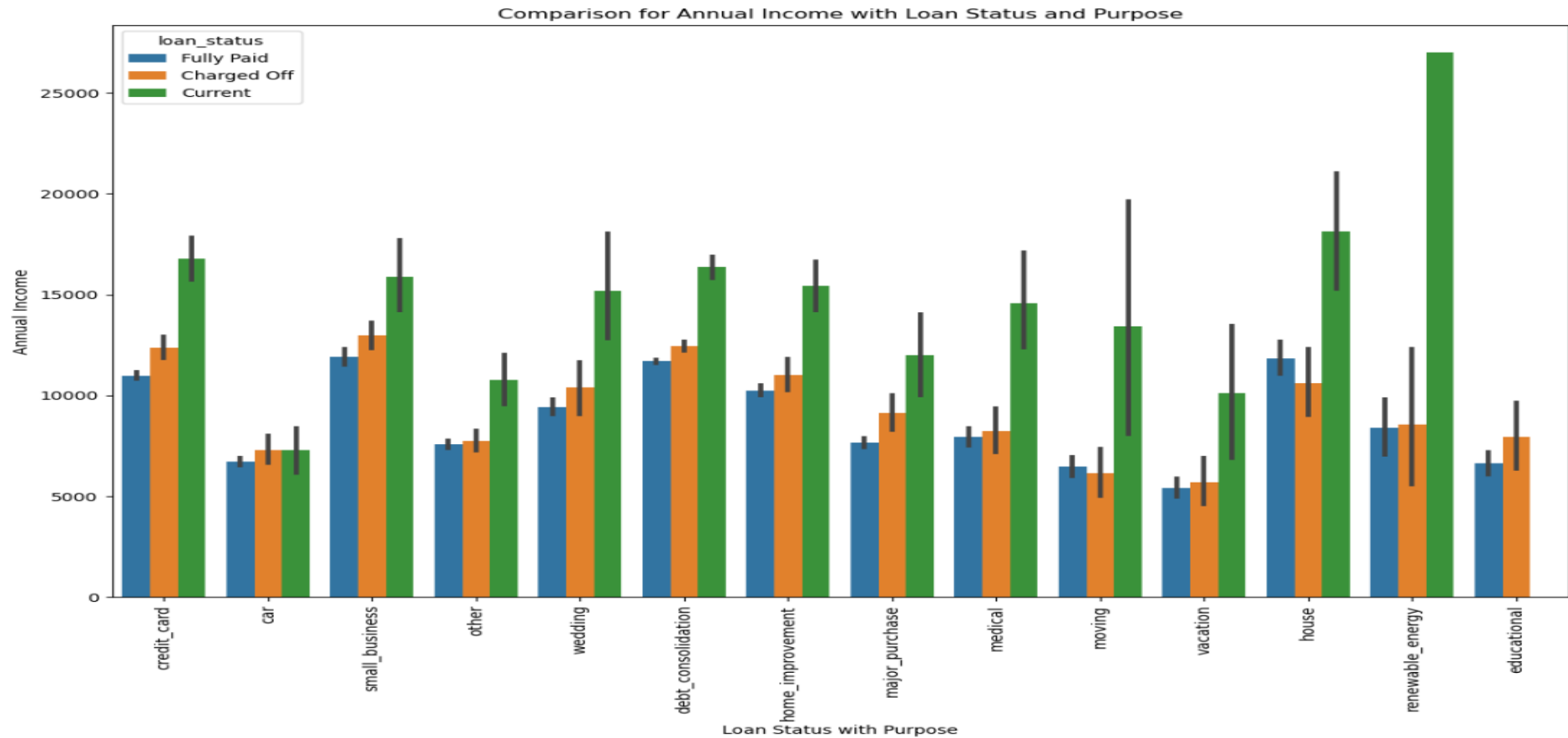




# Bivariate Analysis



# Bivariate Analysis



# Conclusions

---

- More risk with loan applications with E, F and G grade employees.
- More risk with loan applications with higher year (8years, 9years, 10+ years) of experience employees.
- More risk with loan applications with higher income employees.
- More risk with loan applications with the employees who have declared home ownership as Mortgage or other.
- More risk with loan applications with the employees who have taking loan for credit card, small business and debit consolidation.
- More risk with loan applications from California employees.

Thank you

---