



# Gramener Case Study

AAQIB MUSHTAQ KHAN

NAYAN BHATTACHARJEE

PREETHI DAMODARAN

SHIRIN LOURDES

# Objectives

**A Consumer finance company deals in giving personal 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 data set and hypothesise the multiple cause and determine the metrics for defaulters.
3. Possible Solution to the Issue.

# Problem Analysis

## Understanding Data:

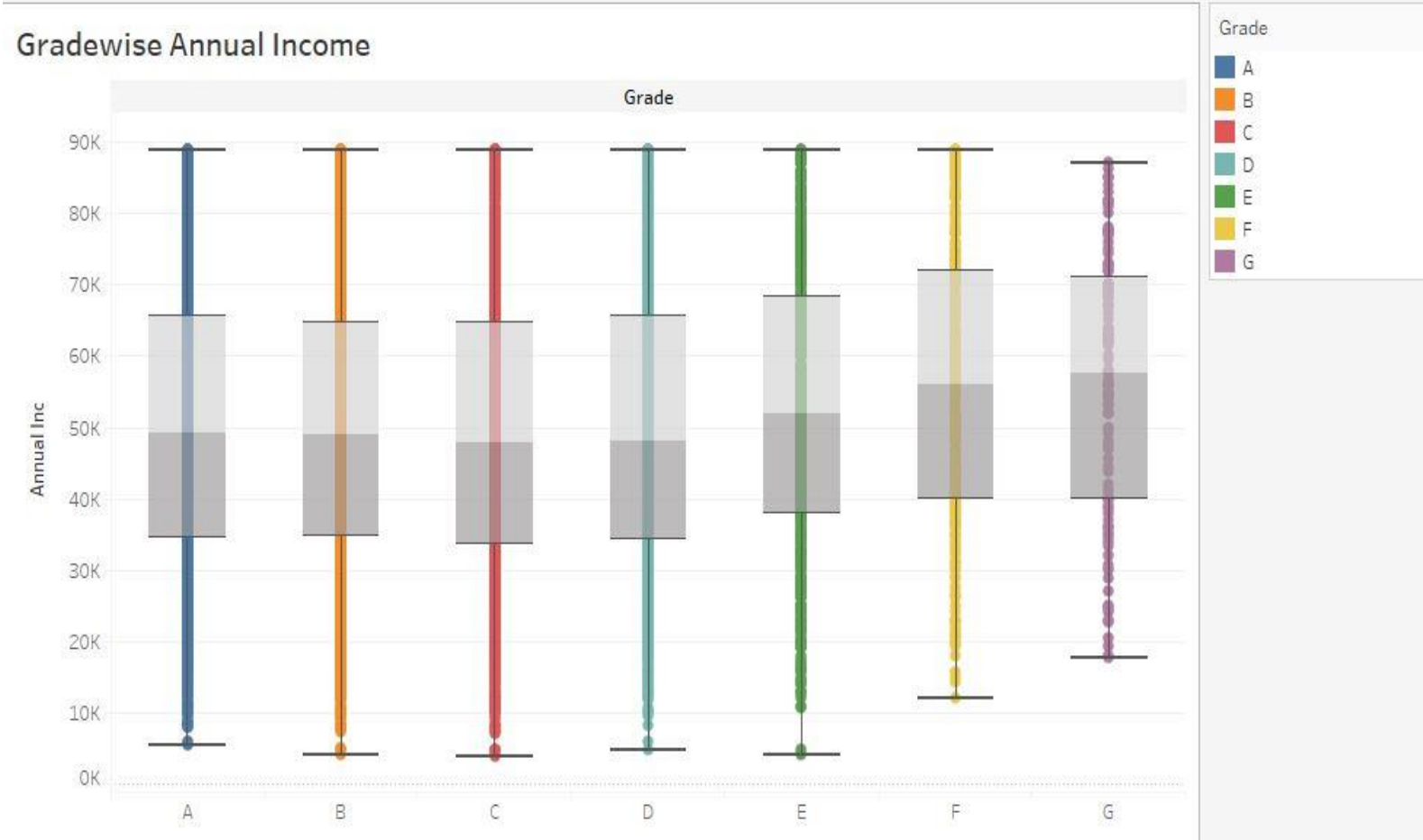
- ▶ The data has information about 4 years of loans that have been issued by the bank. They are mainly of three types *Fully Paid off*, *Current* and *Defaulted*. The data set contains all sorts of information about customer like annual income, employee length, purpose of loan, no. of bankruptcies etc.

## Following are stages in the workflow:

- ▶ Gather data for the analysis.
- ▶ Clean and Format the data for readability.
- ▶ Filter out the outliers based on Annual Income.
- ▶ Filter Data Set on Loan Status.
- ▶ Perform Univariate and Bivariate analysis on the Data.
- ▶ Plot the results of Analysis and Hypnotize the Root of the Issue

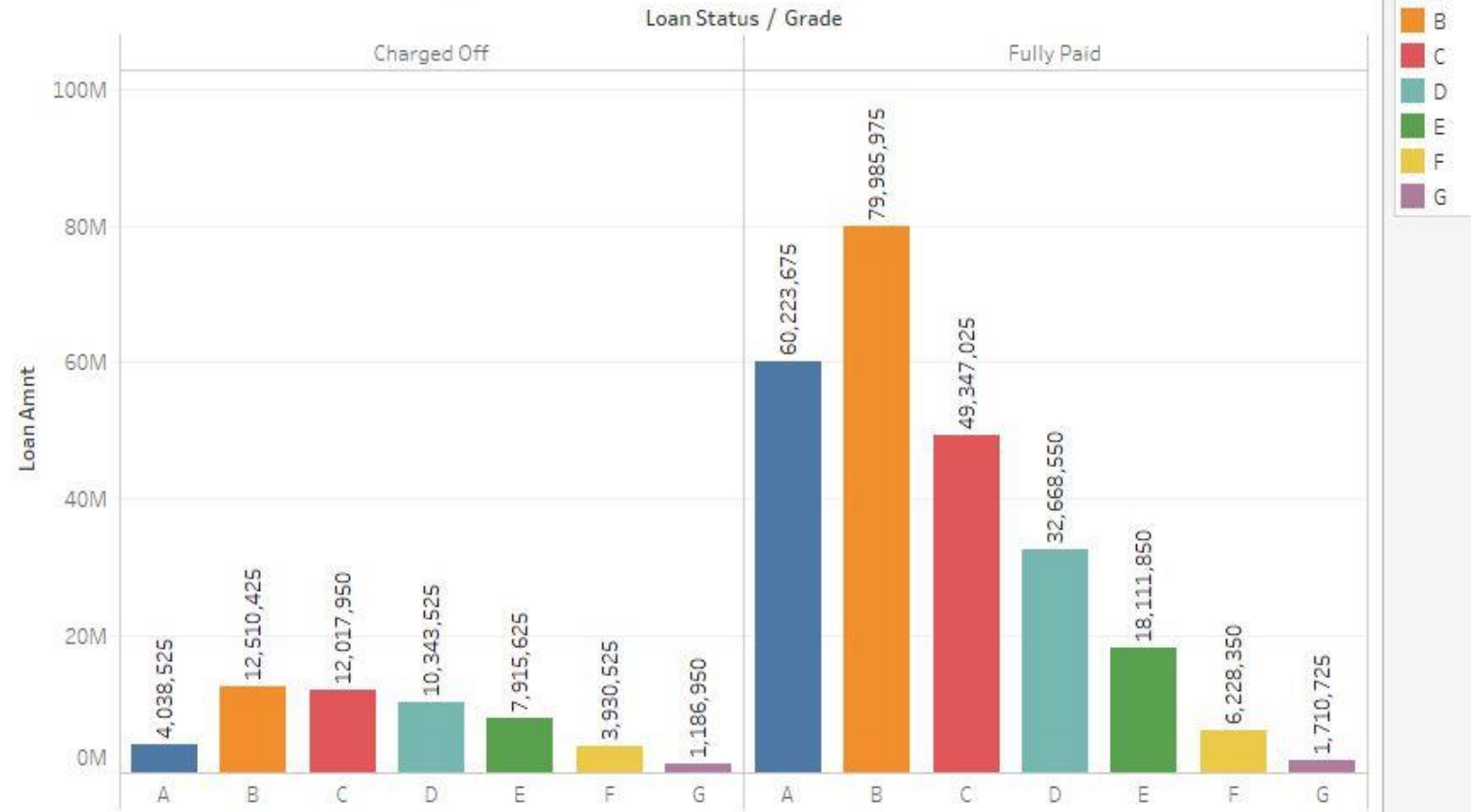
# Plot: Annual Income vs Grade

- The mean 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.



# Plot: Loan amount vs. Loan Status/Grade

Loan Amount vs Loan Status/Grade

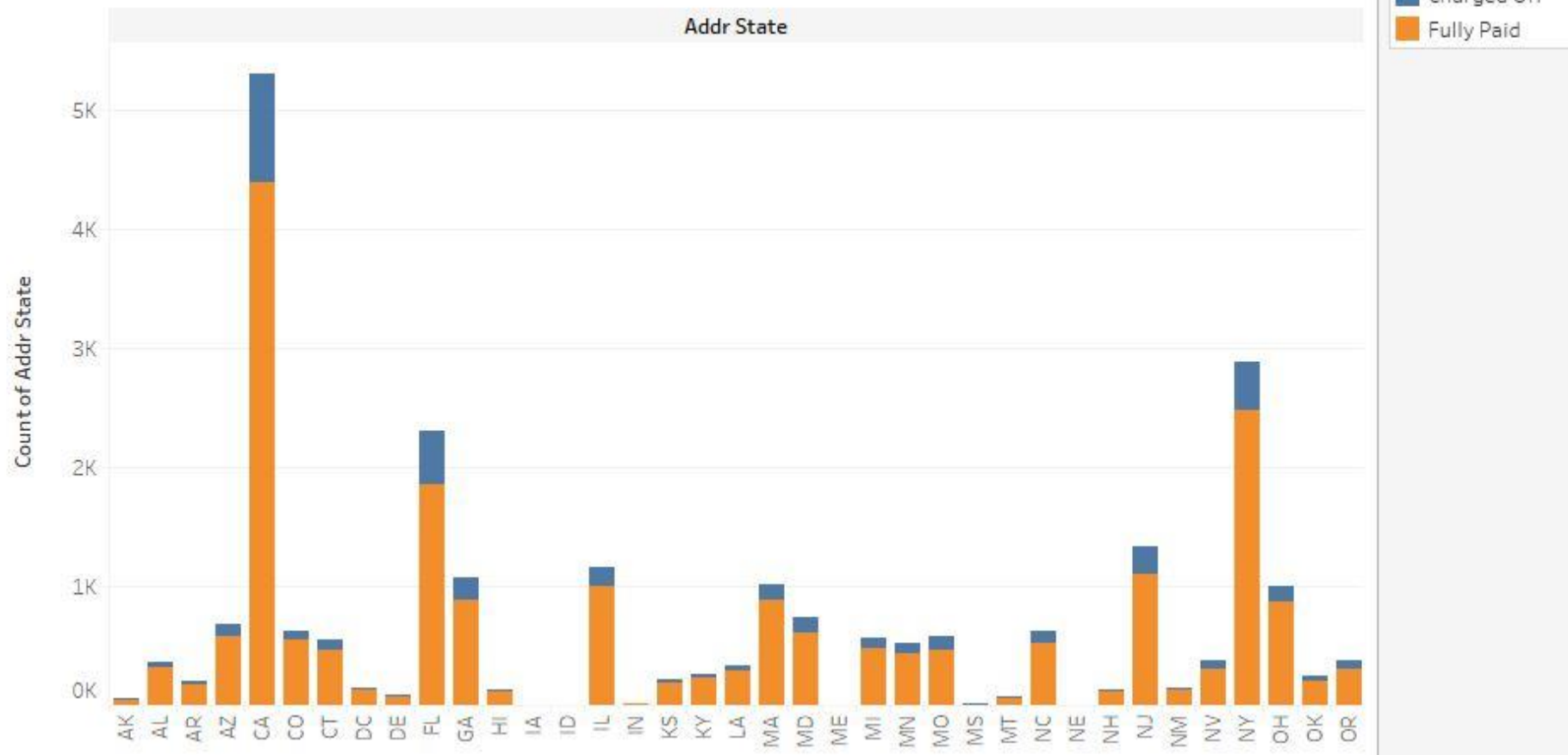


- The graph is right skewed, that shows the grades from D to G follow a decreasing trend for both Default and Fully paid.
- The highest number of defaulters are for grade B.

# Plot: Loan Status vs. State

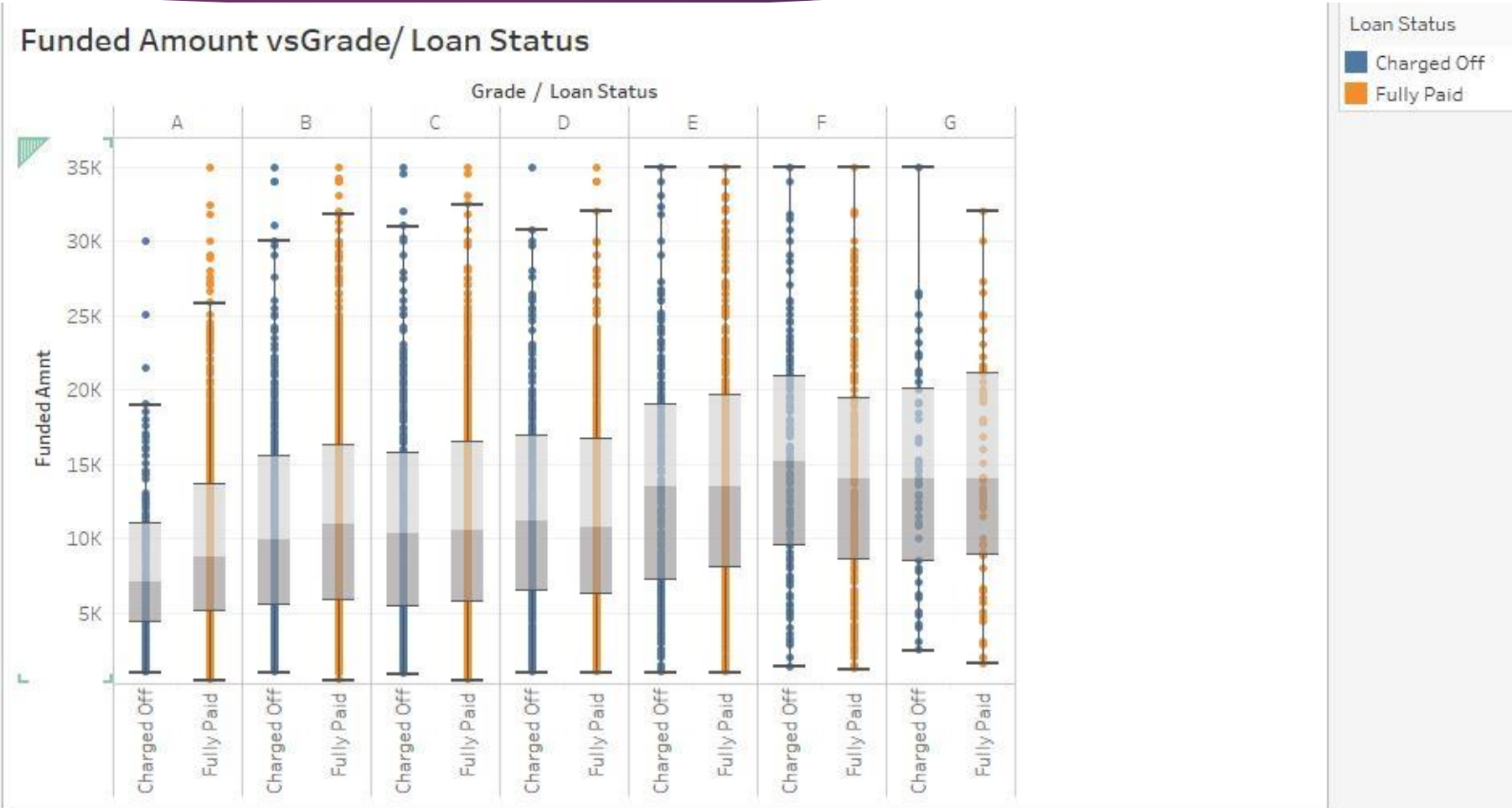
- We can infer that maximum number of loan applicants belong to California(CA). Hence the maximum number of defaulters belong to CA

Loan Status vs State



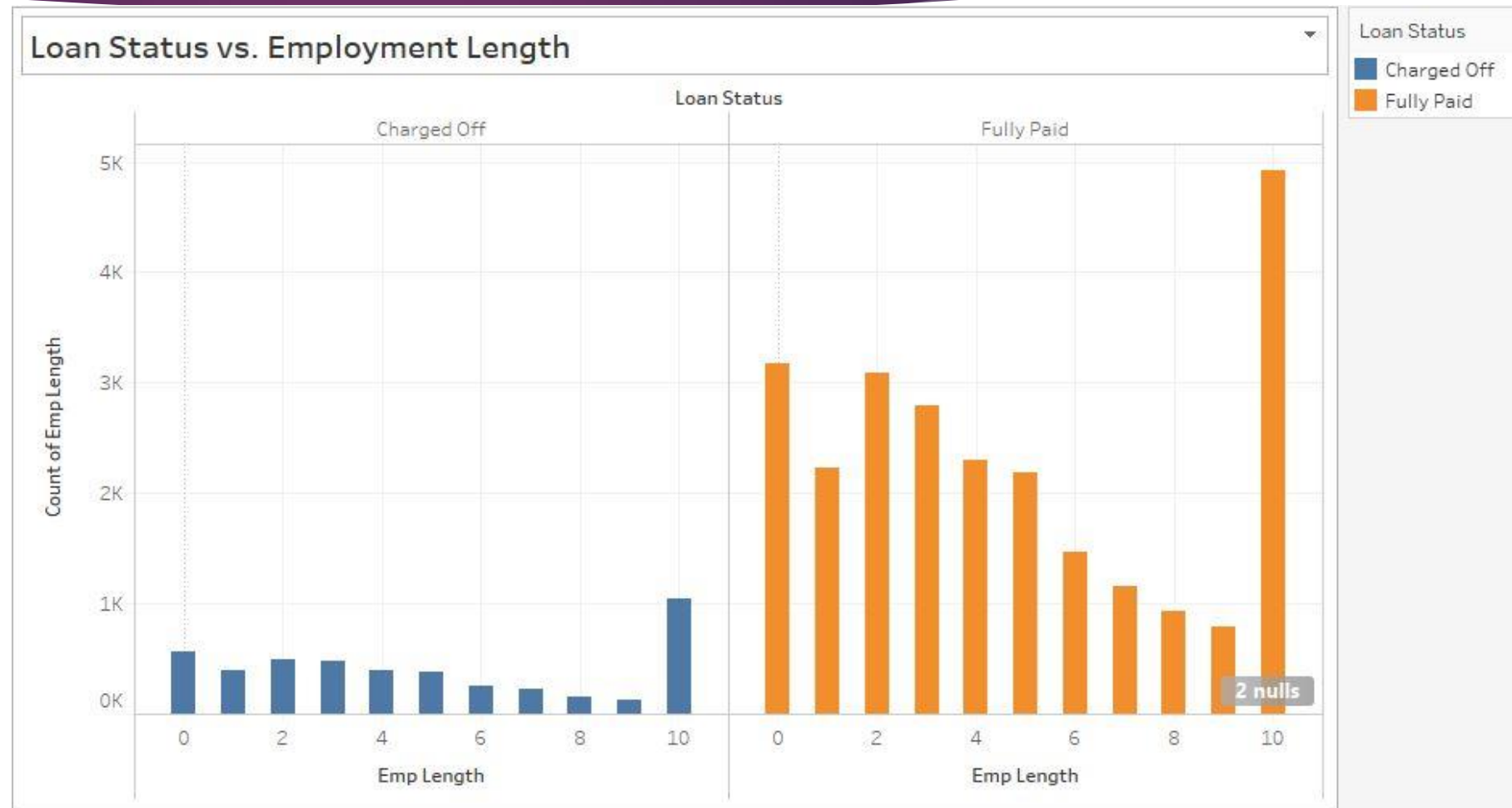
# Plot: Funded amount vs. Grade

- This graph shows that the amount of defaulters are more in the grades from E to G.
- The pattern is very similar where the grades from D to G are less performing grades for the bank.



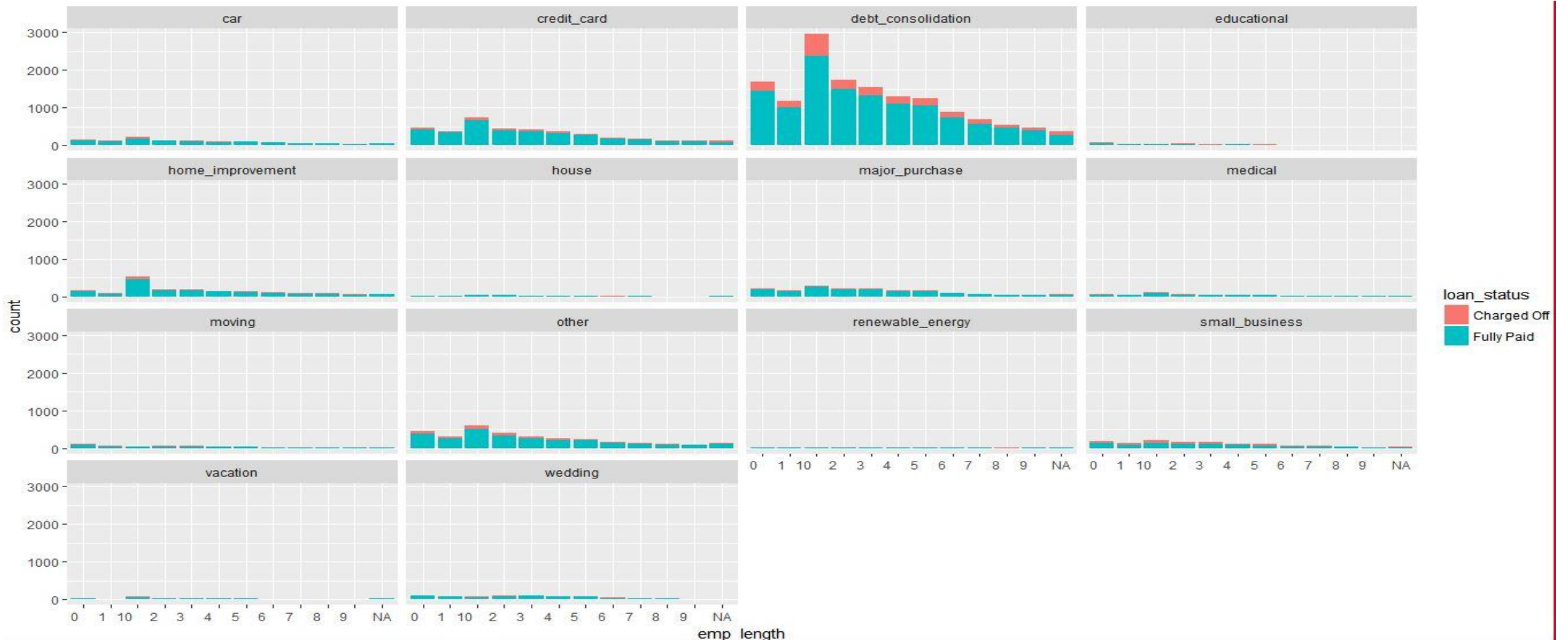
# Plot: Loan Status vs. Employment Length

- We can see here that the number of defaults as well as fully paid is highest for employment length of 10 or more years.





# Plot: Loan Status vs. Employment Length/Purpose



100

- 
- Purpose / Home Ownership**
- Verification Level**
- Home Ownership**
- MORTGAGE
  - OTHER
  - OWN
  - RENT
- Purpose**
- car, credit, debt, education, home, house, major, medical, movie, other, rental, small, vacation, wedding
- Count of Purpose**
- Not Verified
- Source Verified
- Verified
- OTHER, RENT, OTHER, RENT, OTHER, RENT, OWN, OTHER, RENT, OWN, OWN, OWN, OTHER, RENT, OTHER, RENT, OWN, OTHER, RENT, OWN, OWN



# Conclusion

- ▶ After examining annual income and grades, we have found that all the grades have almost similar distribution.
- ▶ *Annual Income, interest rate, Employment length ,Purpose and Verification* have bearing on whether an loan will become fully paid or become charged off. Hence, these parameters need to be critically monitored.
- ▶ In the current disbursal, from grade D to grade G, the percentage of defaults are more. *Grades* also play an important role in loan disbursal and repayment and can be reference from current data to make future decisions.
- ▶ Target based/ last minute loan disbursals should be avoided.

# Recommendations

- ▶ Customer should be classified clearly based on their income, frequency of application, purpose of use, loan duration and intended use based on this interest rate and disbursed amount should be calculated.
- ▶ The process of income verification should be more accurate as it is an important parameter to control defaults.



Thank You!!!