

# Gramener Case Study

## SUBMISSION

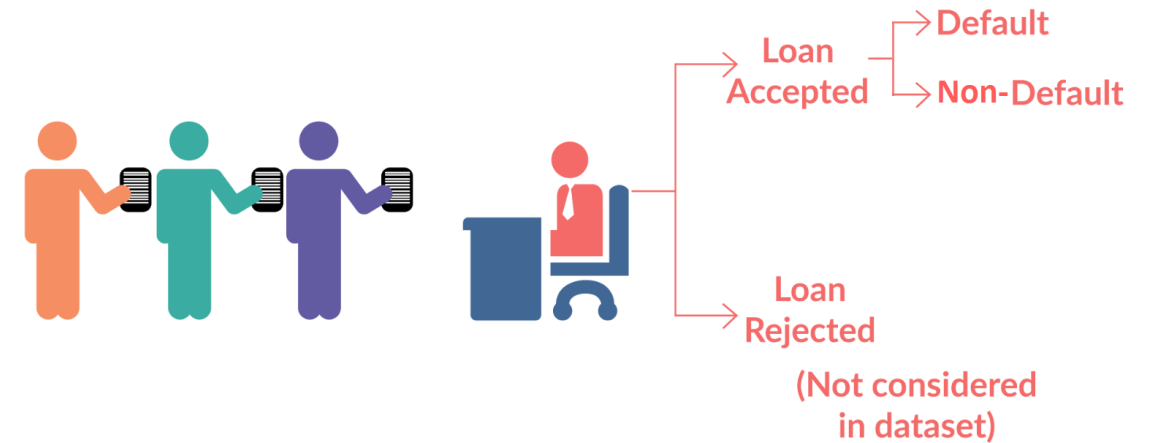
30<sup>th</sup> December 2018

### Group Members:

1. Snehal Bhosle
2. Vineesha Reddy
3. Vipul Shrivastava
4. Yatin Kanchan

Lending Club is a **consumer finance company** which specialises in lending various types of loans to urban customers. Give a set of data they want to understand how they can take right decision of approving the loan and minimize the risk of losing money while lending to those customers.

## LOAN DATASET

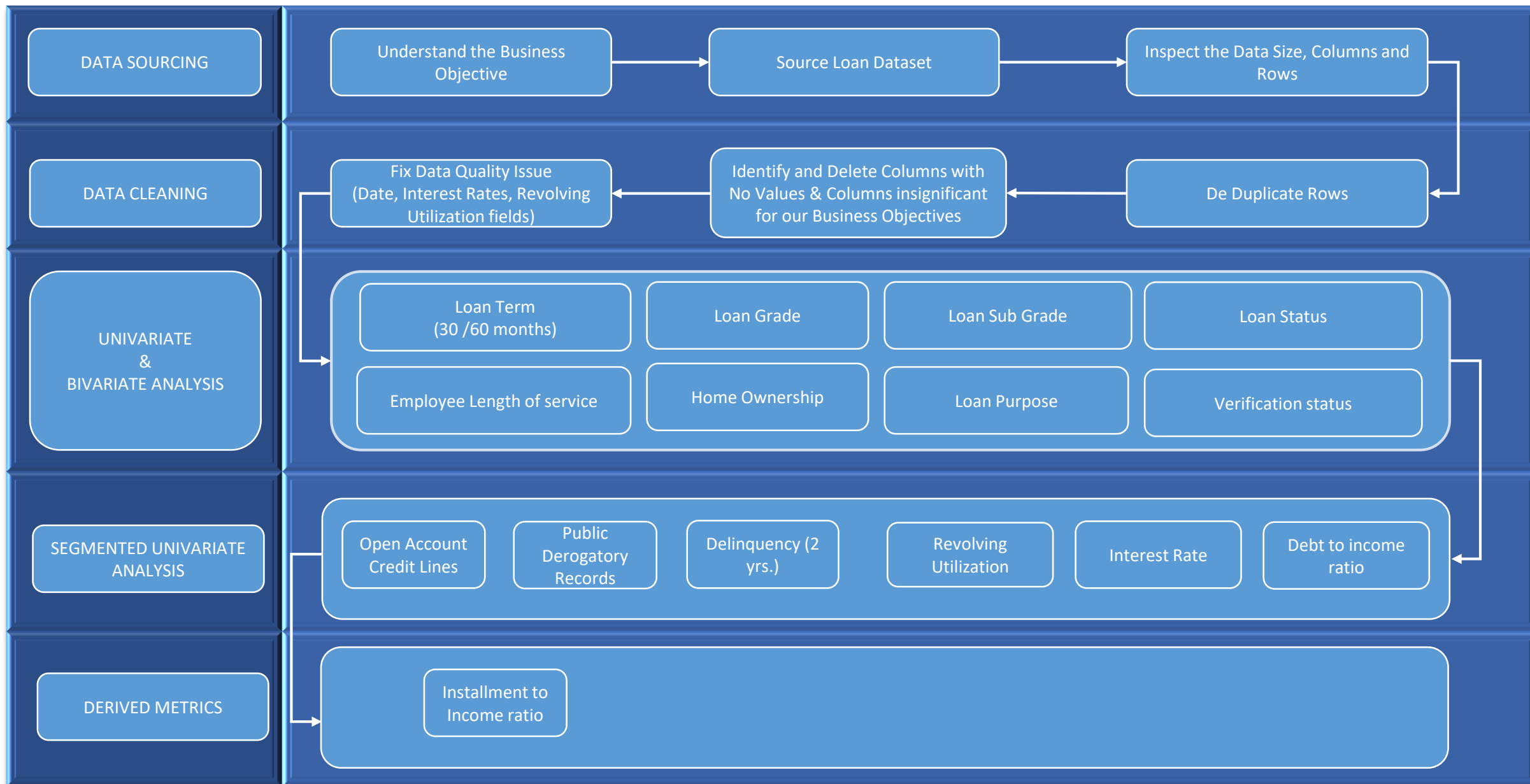


### Business Objectives of Data Analysis

Company wants to understand the **driving factors (or driver variables)** behind loan default, i.e. the variables which are strong indicators of default

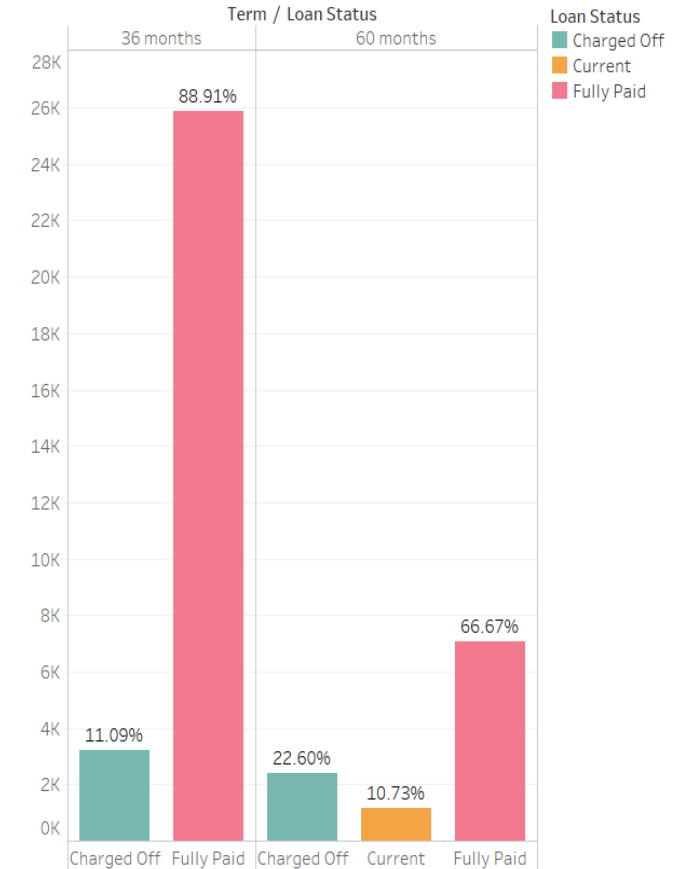
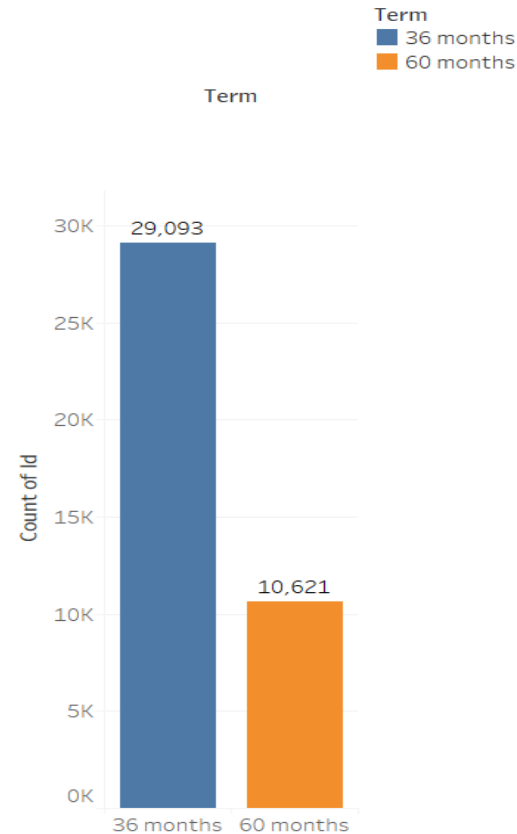
### Goals of Data Analysis

The **aim** is to identify patterns which indicate 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



## How the **Loan Term Analysis** was done ?

- Univariate and Bivariate Analysis
- Consider Loan Status = 'Charged Off' & 'Fully Paid'.
- Identify the count of Loans for 36 months and 60 months Term.
- Determine the Loan performance by Loan Status split.



**Figure 2**

## Findings

- **Figure 1** : indicates Applicants mostly prefer short Term Loan of 36 months.
- **Figure 2** : indicates, most of the Applicants have gone Default on Long Term loans.
- **Loan Term** could be one of the **Key Driving Factors**. This can confirmed on further analysis.

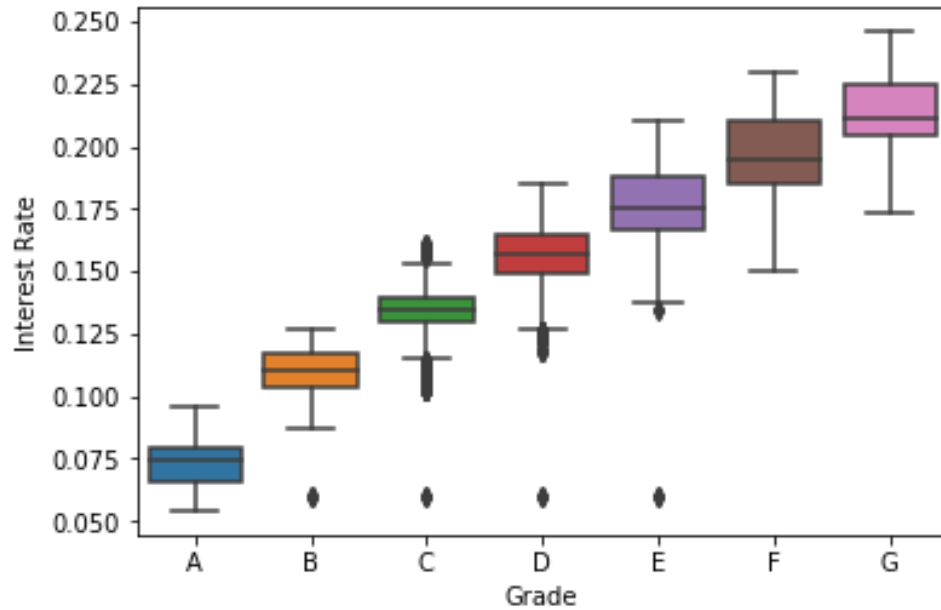


Figure 1.

## Findings

- **Figure 1** : Lower the Grade\*\*, Higher is the Interest Rate loans given.
- **Figure 2a** : Grade B loans are more preferred Loan provided to Applicants.
- **Figure 2b** : **Lower Grade loans** are at **higher Risk of Default**

\*\* Higher to Lower Grade here is from A to G

## How the **Loan Grade Analysis** was done ?

- Univariate and Bivariate Analysis.
- Consider All Loan Status to verify Range of Interest Rates (Figure 1) for each Grade.
- Compare Charge off loan percentage vs each grade.

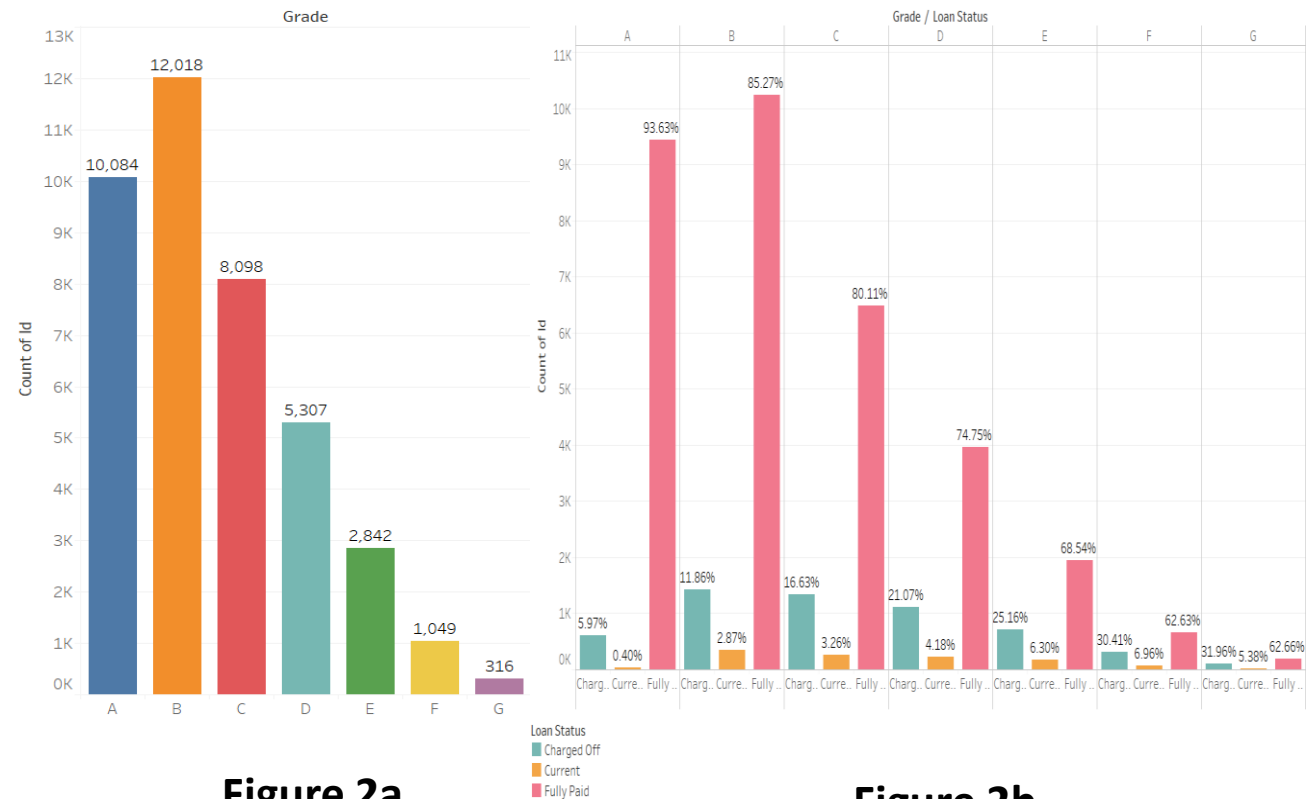


Figure 2a.

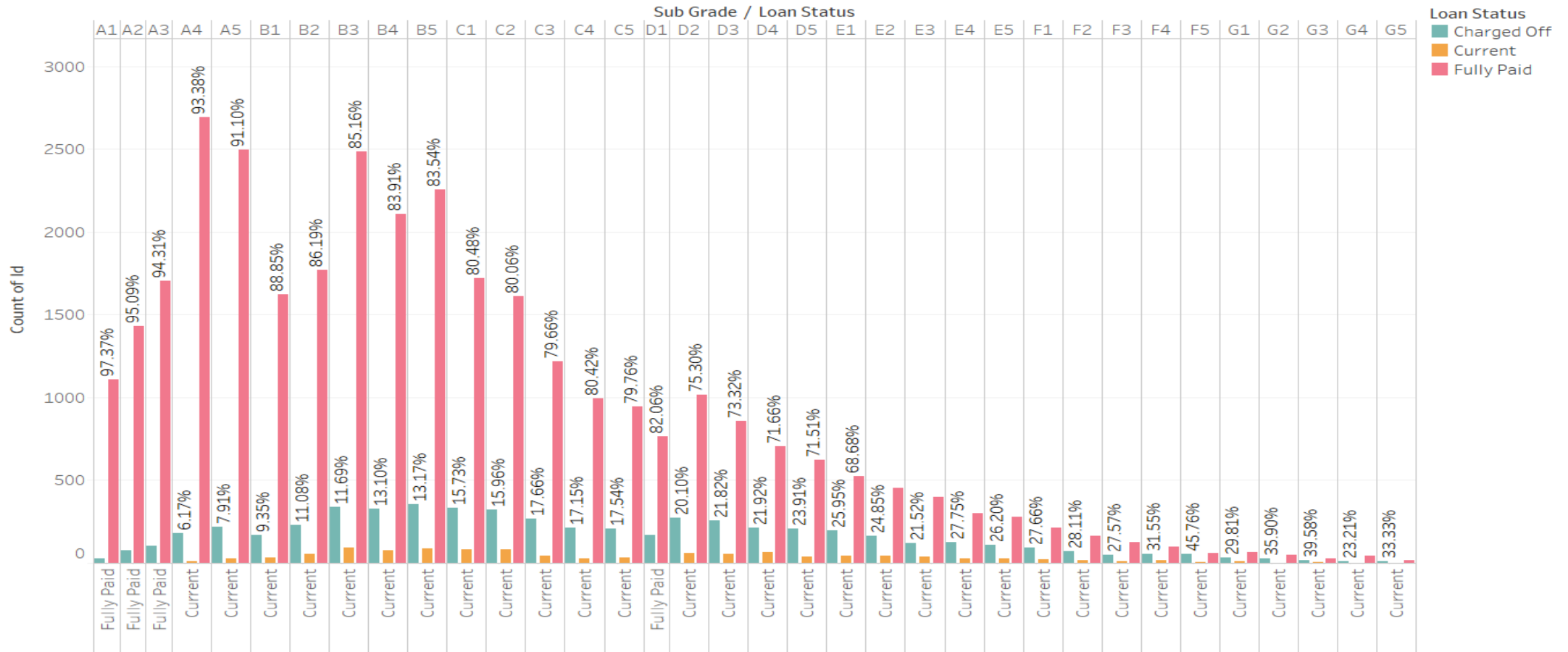
Figure 2b.

## How the **Loan Sub-Grade Analysis** was done ?

- Univariate and Bivariate Analysis.
- Consider All Loan Status to verify Range of Interest Rates (Figure) for each Sub-Grade inside Grade.
- Compare Charge off loan count vs each sub-grade.

## Findings

- There are high charge off cases in 'E' , 'F' & 'G' grades and 'F5' , 'G3' & 'G2' subgrades
- **Lower Sub-Grade loans** are **at higher Risk of Default**

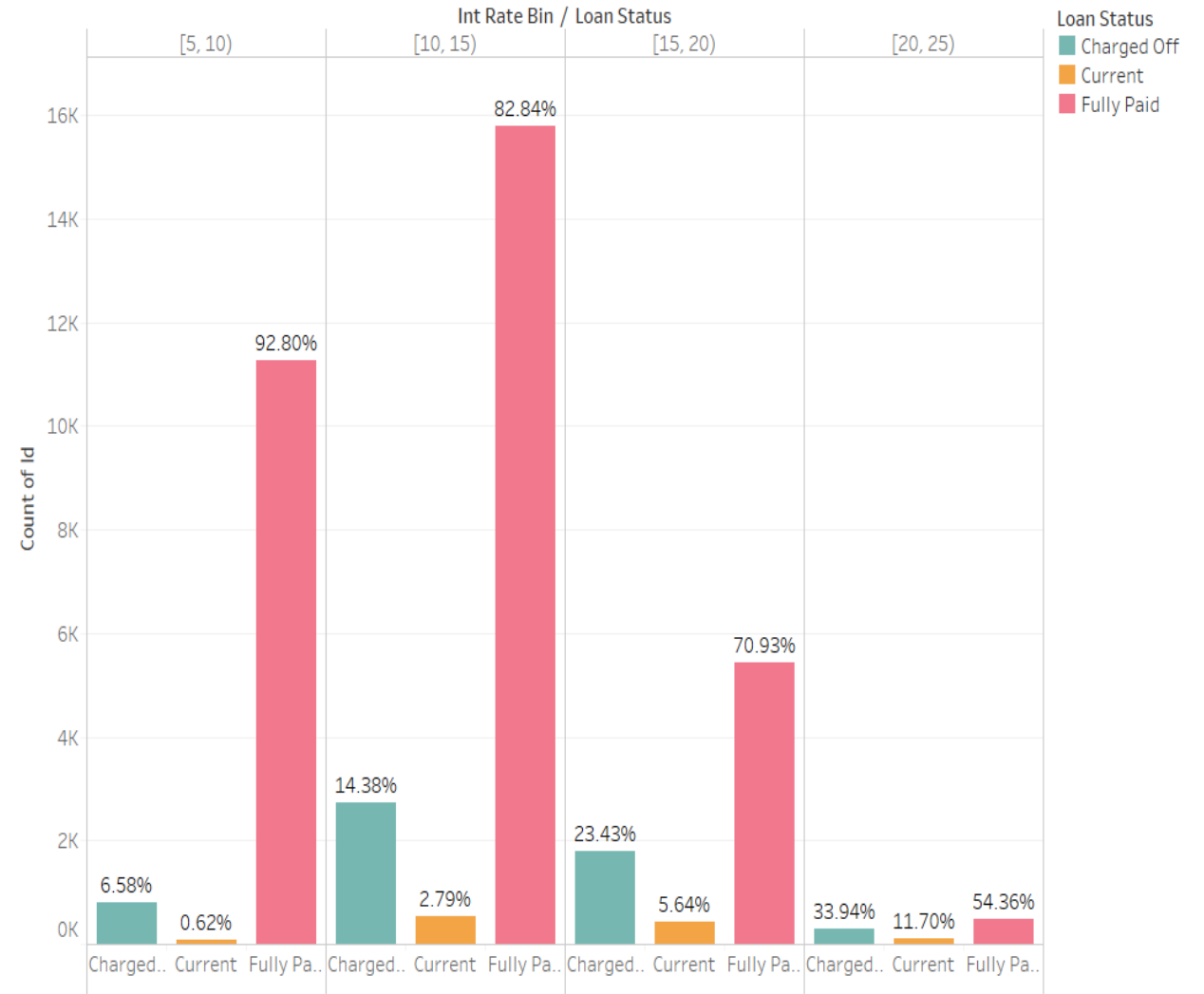


## How the **Loan Interest Rate Analysis** was done ?

- Segmented Univariate Analysis.
- Consider All Loan Status to verify Range of Interest Rates.
- Compare Charge off loan percentage vs interest-rate range.

## Findings

- Higher the interest rates are the higher the percentage of charged off loans



## How the **Verification Status Analysis** was done ?

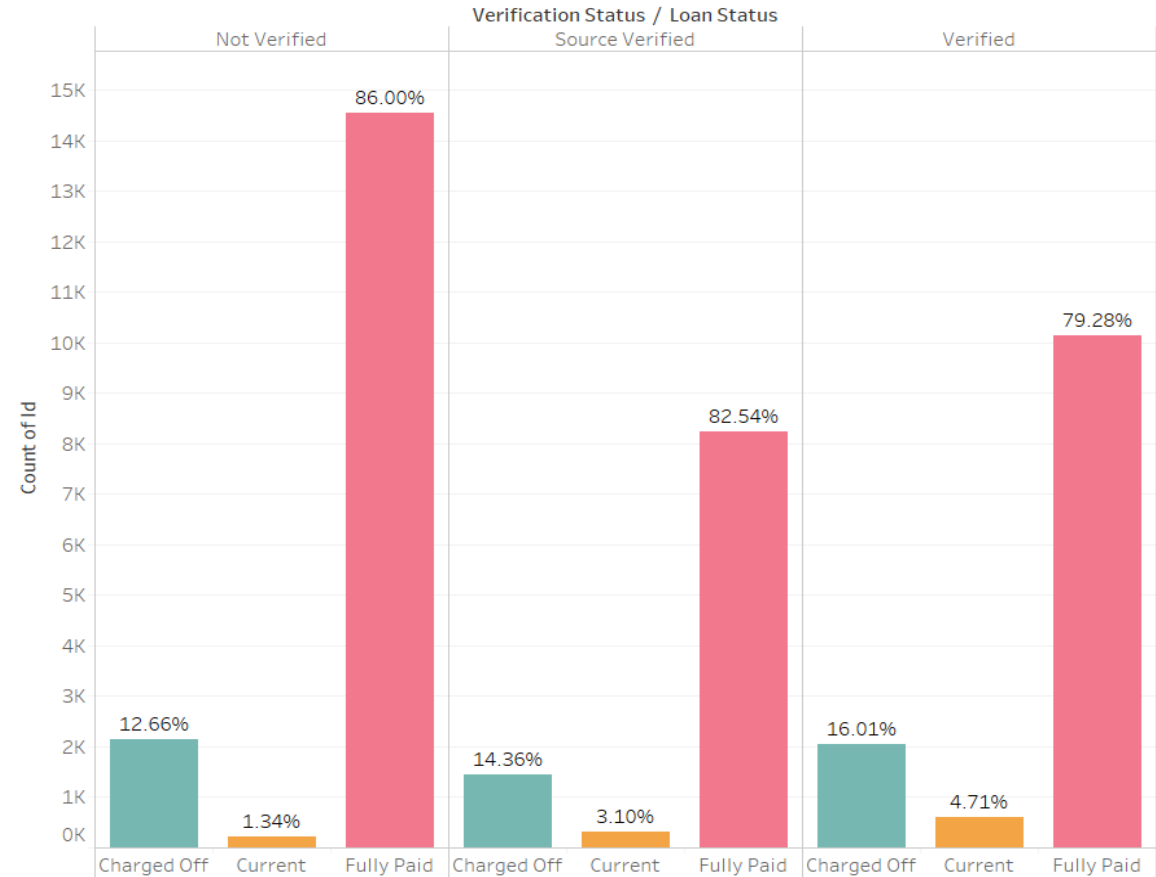
- Univariate and Bivariate Analysis.
- Percentage of Loans that had Applicants Income not verified.
- Compare Charge off loan percentage vs verification status.

## Findings

- **Figure 1** : The Defaulted Loans spread appears higher for Verified than Non - verified Applicant's source of income.
- **Income Source Verification** could also be one of the **Driving Factors**.

## Recommendation

- The Consumer Finance company must make the **Applicant Income source verification process more stringent**.
- In spite of Applicant's Source of Income Verified, the Loans are still at higher Risk of Defaulting.



**Figure 1.**

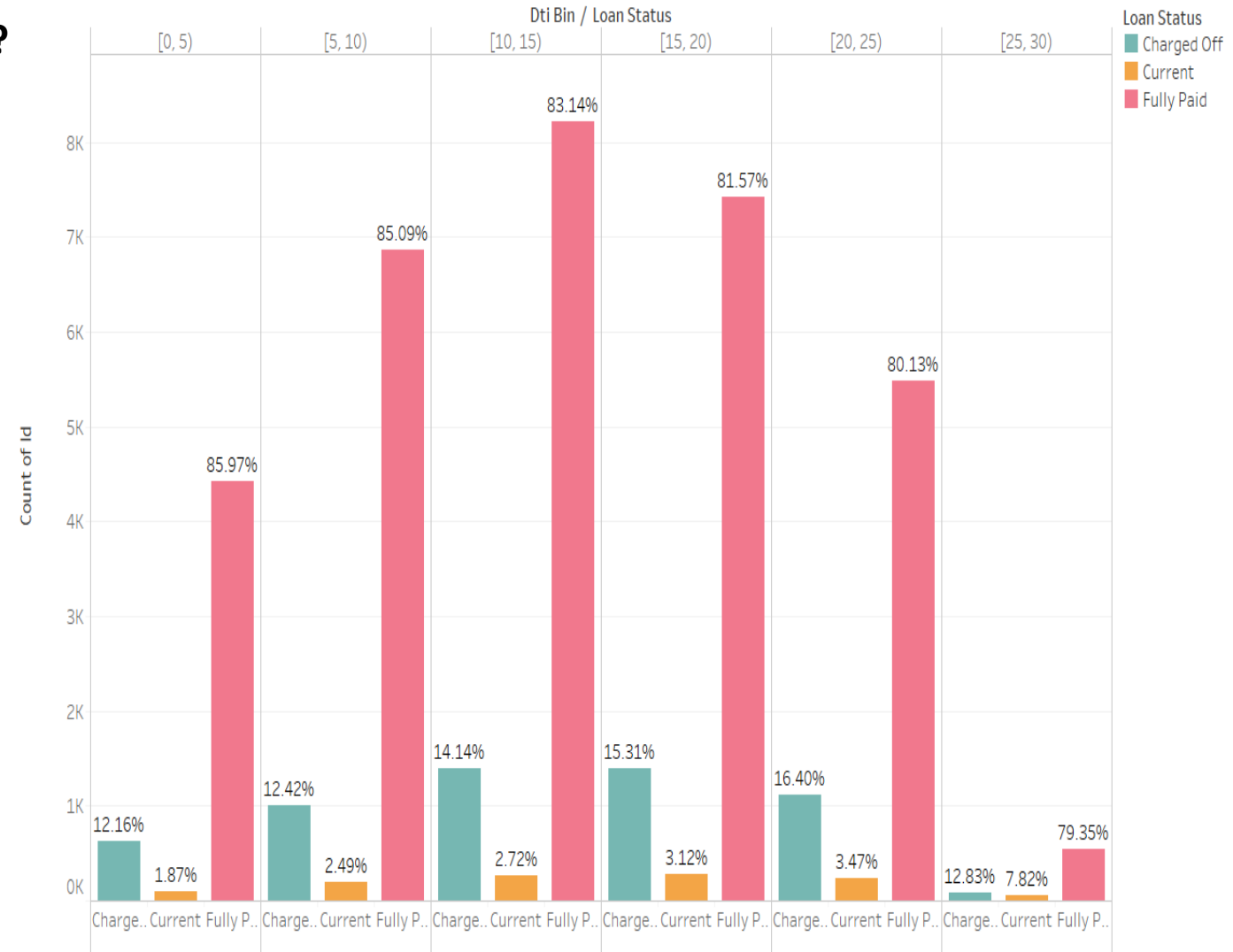


## How the Dti (Debt to income) analysis was done ?

- Univariate Segmented Analysis.
- Percentage of Loans comparison with Dti (debt to income) ratio.
- Compare Charge off loan percentage vs Dti.

## Findings

- As Debt to Income(dti) ratio increases the Charged Off loans increases
- The Charged off loans drops in the DTI range between 25 to 30



# Open Account (Credit Line) & Delinquency Analysis

## How the **Open Account Analysis** was done ?

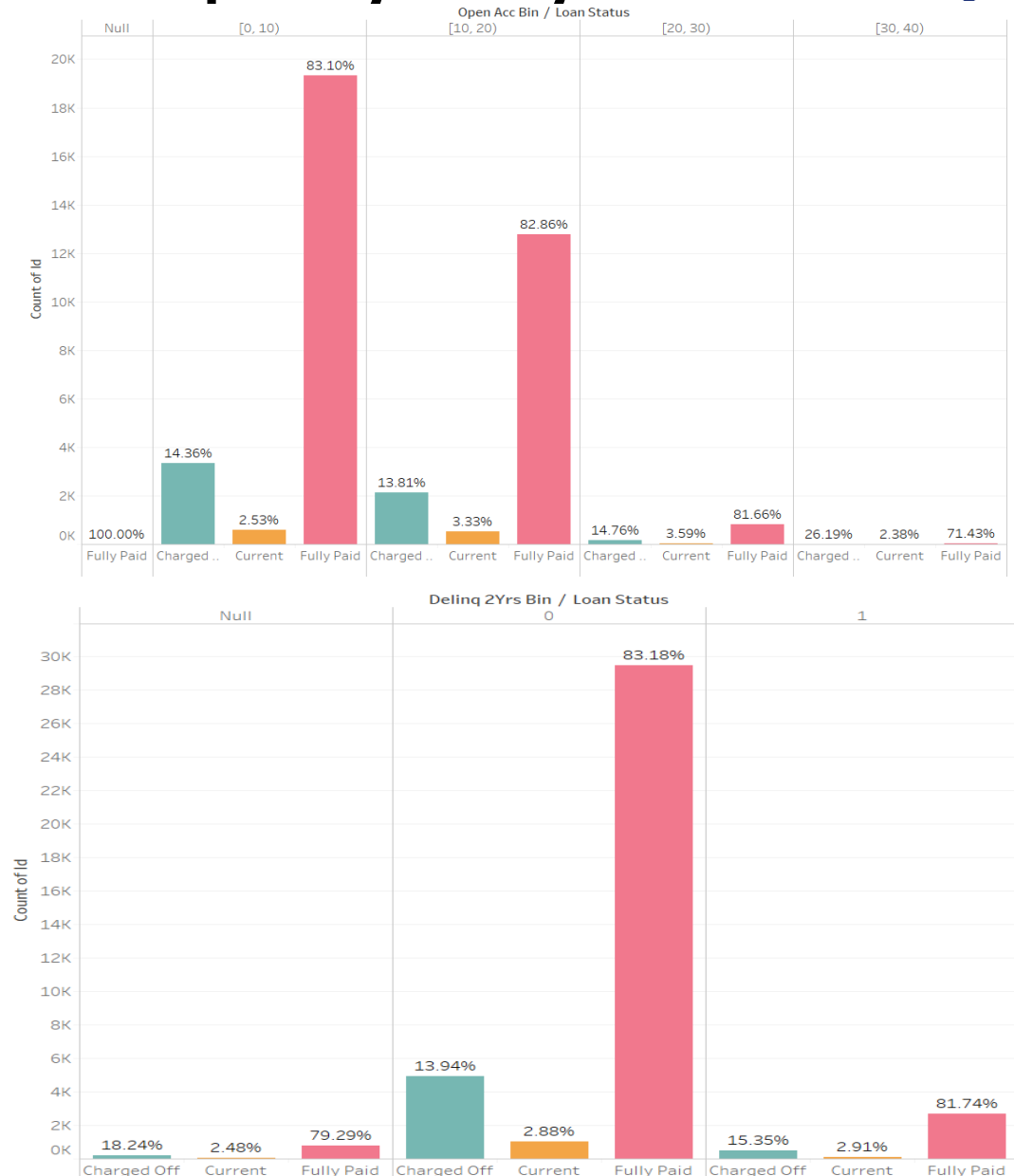
- Segmented Univariate Analysis.
- Total Percentage of Applicants with multiple Credit Lines taking loan from LC.
- Compare Charged off loan counts across various Number of Credit lines.

## How the **Delinquency Analysis** was done ?

- Segmented Univariate Analysis.
- Total Percentage of Applicants with number of Delinquency in last 2 years.
- Compare Charged off loan counts across various Number of Delinquent Incident in last 2 years.

## Findings

- The Consumer Finance company should be more **cautious** with Applicant having **Open credit Lines more than 30**.
- No significant effect of Delinquency Frequency on the Charged Off loans.



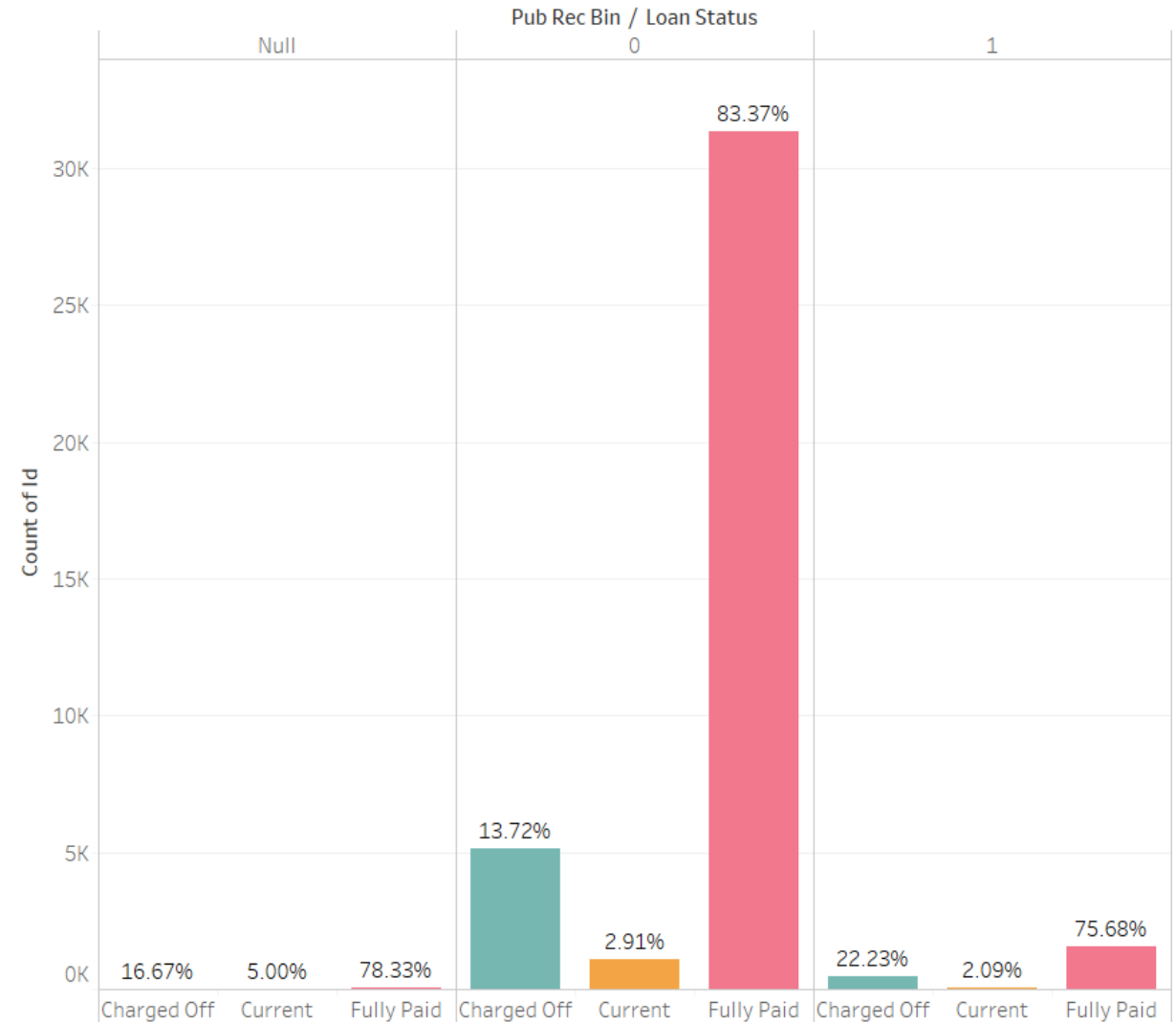
In Fig – Null represents Derogatory Public Record and Open Account = 2+

## How the Derogatory Public Record Analysis was done ?

- Segmented Univariate Analysis.
- Total Percentage of Loan Applicants with Derogatory records.
- Compare Charged off loan counts across Number of Derogatory Public records count.

## Findings

- **Charged Off loan** percentage **increases as the Number of Derogatory Public Record increases**



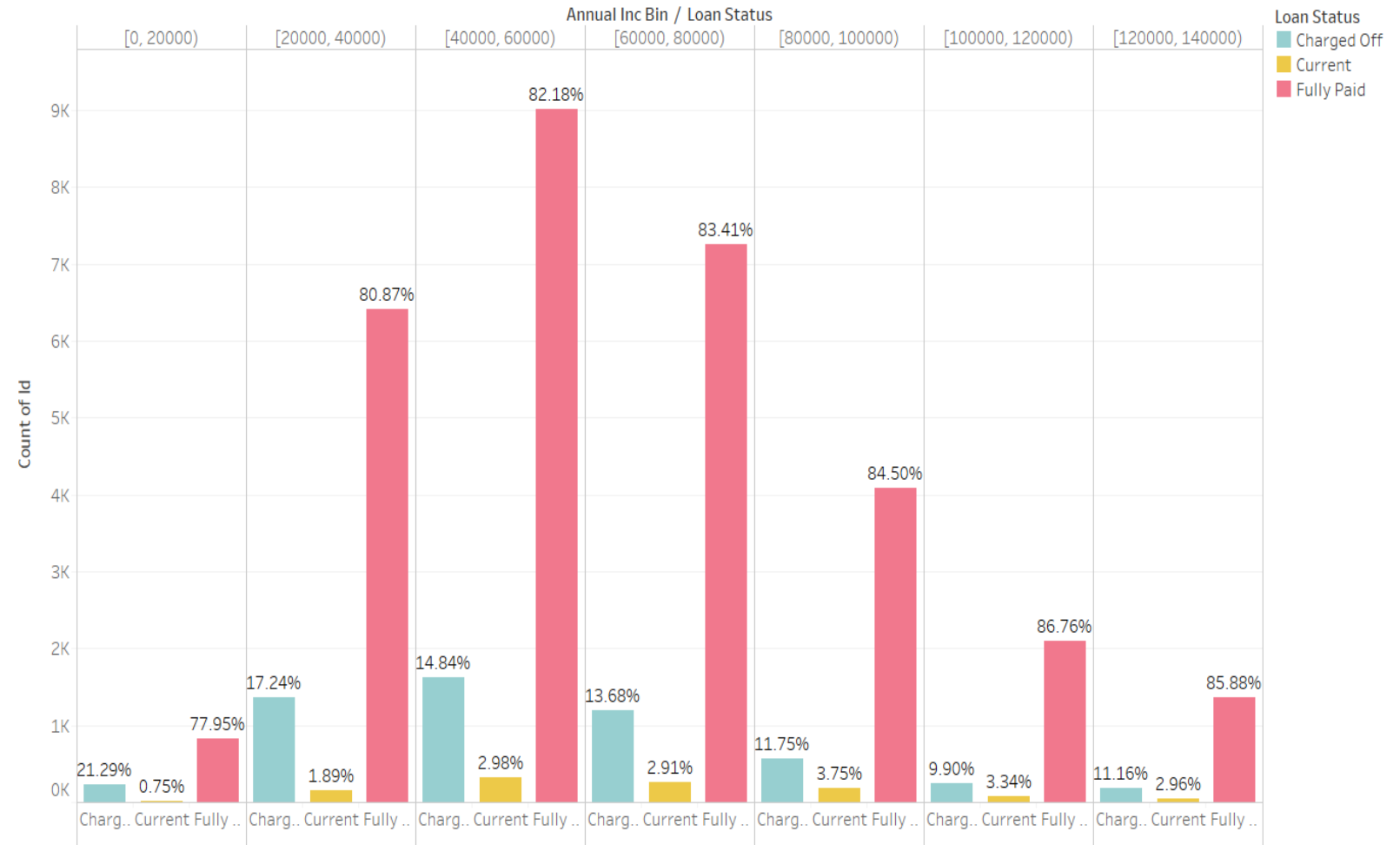
In Fig – Null represents Derogatory Public Record = 2+

## How the Annual income Analysis was done ?

- Segmented Univariate Analysis.
- Outlier treatment conducted.
- Compare Charged off loan against different income range.

## Findings

- Less the annual income of borrower more is the probability of Loan getting Charged off.

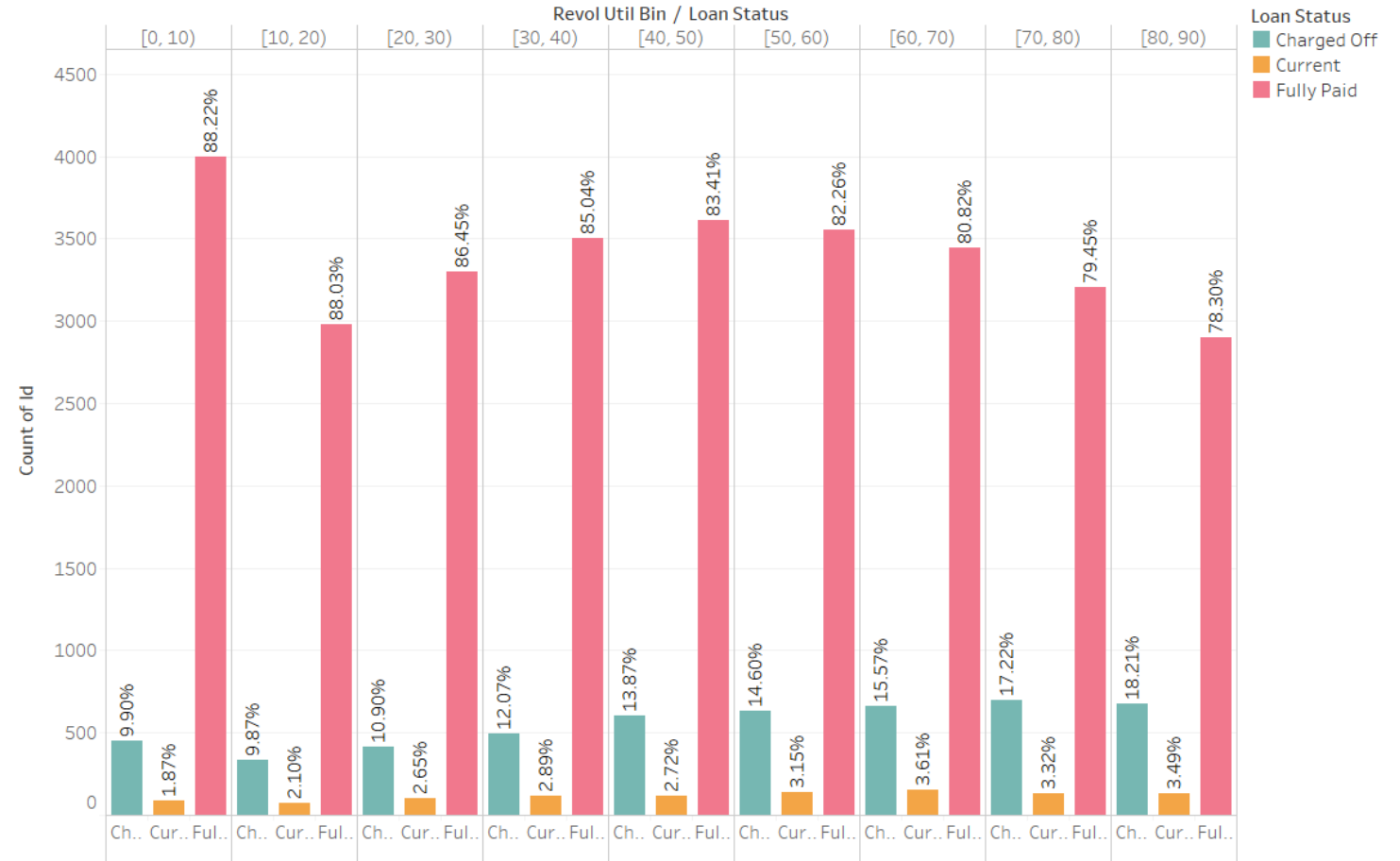


## How the **Revolving Utilization Analysis** was done ?

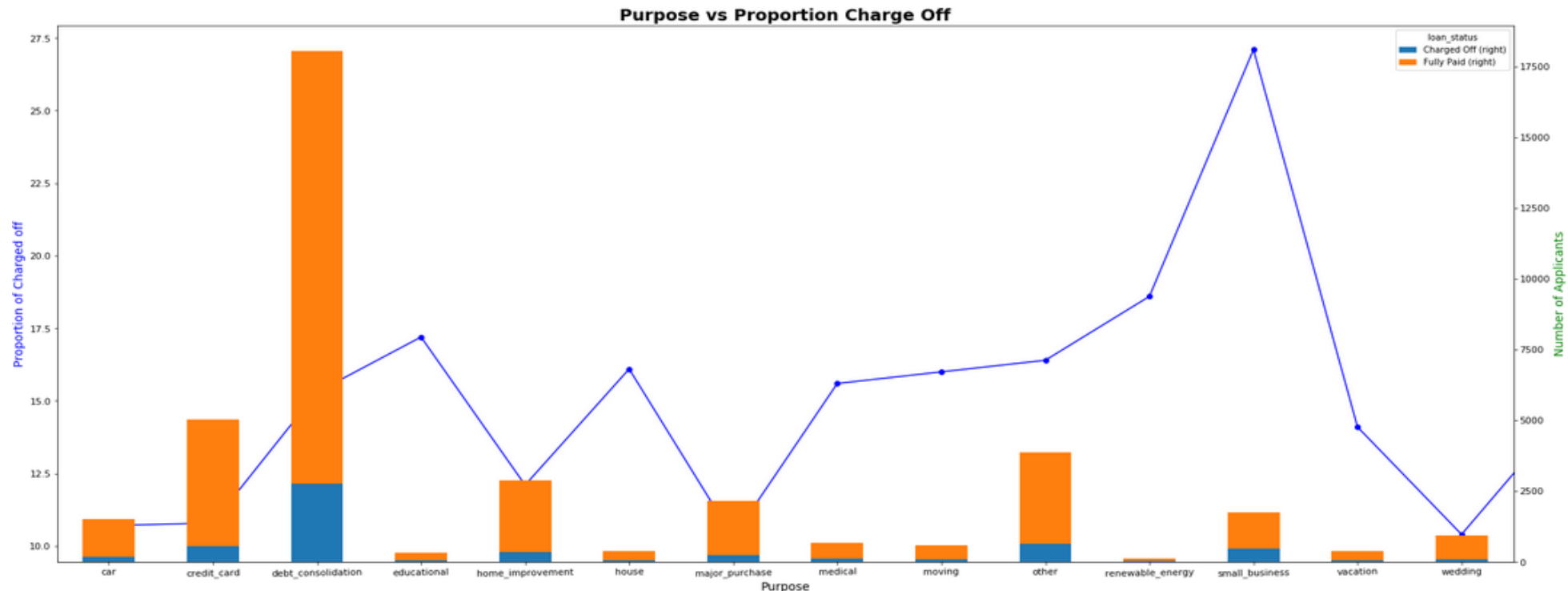
- Segmented Univariate Analysis.
- Total Percentage of Loan Applicants with Revolving Utilization Rate.
- Compare Charged off loan counts across range of revolving utilization rates.

## Findings

- Loans are more prone to **default** as the **Revolving Utilization rate increases**.



# Loan Purpose Analysis



## How the **Loan Purpose Analysis** was done ?

- Bivariate Analysis.
- Check Number of Loan Applicants going Default when Loan is requested for specific purpose.
- Get the Proportion in percentage to identify which purpose is more risky and prone to Default.

## Findings

- The **Blue Trend line** indicates the Consumer Finance company should do **more Due Diligence** for loans being requested for **Small Business Financing**.

## How the Analysis was done ?

- Multivariate Analysis on Tableau.
- Removed Annual Income Outliers above \$150K
- Binned the Annual Income of \$25K.
- Identified the % and number of Loans at Risk for each Loan Grade.

% of Loan Amount at Risk for Range of Annual Income and Loan Grade

Annual Inc (bin)	Grade						
	A	B	C	D	E	F	G
0 - 25K	15%	30%	25%	19%	7%	3%	2%
25K - 50K	10%	28%	26%	19%	12%	4%	1%
50K - 75K	7%	22%	22%	21%	17%	10%	2%
75K - 100K	5%	21%	21%	21%	18%	10%	4%
100K - 125K	3%	17%	20%	20%	19%	15%	6%
125K - 150K	5%	14%	20%	23%	20%	12%	5%
150K	13%	15%	6%	15%	24%	27%	

Number of Loans at Risk for Range of Annual Income and Loan Grade

Annual Inc (bin)	Grade						
	A	B	C	D	E	F	G
0 - 25K	68	128	112	90	35	11	6
25K - 50K	256	556	525	378	206	59	19
50K - 75K	169	419	410	353	239	121	30
75K - 100K	71	190	176	165	112	56	23
100K - 125K	19	63	74	67	50	32	14
125K - 150K	10	25	31	31	22	13	4
150K	4	5	2	4	6	5	

## Findings

- Applicants with **Income Range of \$25,000 - \$75,000** are the once more **vulnerable to default** when they are give Loans of Grade B or C.
- Perhaps, the Consumer Finance company should give Loan on Higher interest loans i.e. a lower grade loan in the range of E, F or G for such income group so the recovery of Loan is Arbitraged in case of default.



# Driving Variables

**Based on the EDA on the Loan Data Set, below are some of the Driving factors (or driver variables) identified that could be strong indicators of Loan Default**

- Loan Term (term)
- Grade (grade) and Sub-grade (sub\_grade) of Loan
- Interest rate (int\_rate)
- Debt to income (dti) ratio
- Open Account (Credit Line) (open\_acc)
- Delinquency (2 yrs) (delinq\_2yrs)
- Derogatory public record (pub\_rec)
- Revolving Utilization (revol\_util)
- Annual Income (annual\_inc)



1. Loan term shows strong relationship with loan\_status
2. Grade and Subgrade influence the loan\_status strongly.
3. Loan are getting more defaulted when income is verified.
4. Home ownership provides stable information about defaulters.
5. Loan provided to 'Small business' defaulted most ~25%.
6. Grade E , F , G shows more defaulters than any other Grades.
7. Lower the income of borrower, more the probability of loan getting defaulted.

The Consumer Finance Company can utilise above variables for their Loan portfolio and Risk assessment.

## **Recommendation :**

- In addition to the above variables the Consumer Finance Company must have proper Due Diligence on their Loan Grading mechanism so the recovery could be arbitrated by providing the loans at higher Rates if applied for Long Term.
- The Consumer Finance company should also have better Applicants income Source verification process, as we found many loan defaults in spite of Applicants income source being verified.