



# Lending Club Case Study

---

SHATRUGHNA SINHA

# Exploratory Data Analysis

---

- 1) Univariate Analysis
- 2) Segmented Univariate Analysis
- 3) Bivariate Analysis

# Univariate Analysis

---

Column used for Univariate Analysis :

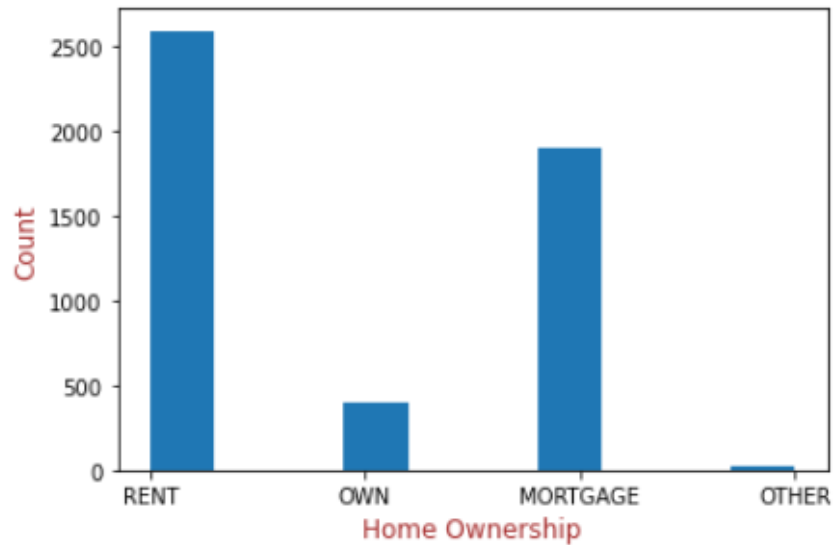
- 1) home\_ownership
- 2) int\_rate
- 3) purpose
- 4) verification\_status
- 5) emp\_length
- 6) dti
- 7) collection\_recovery\_fee
- 8) grade

# Observation:

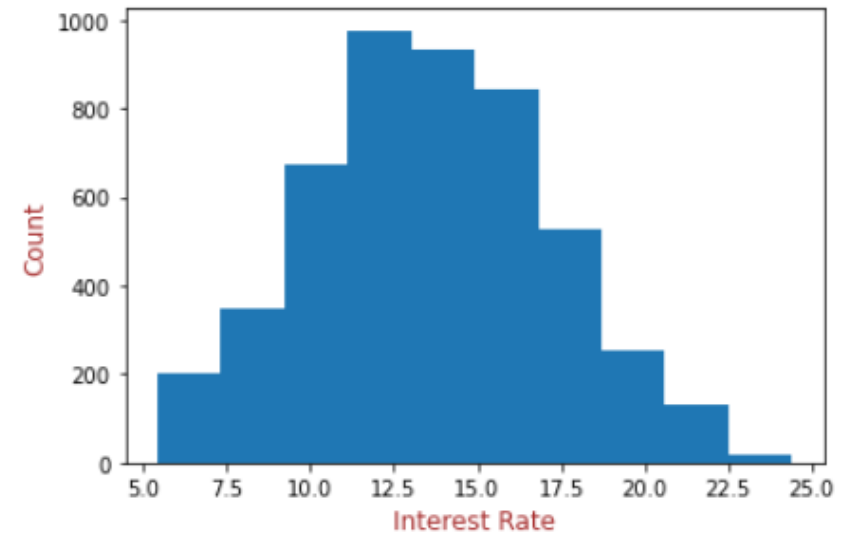
Below indicators are high volume of default.

---

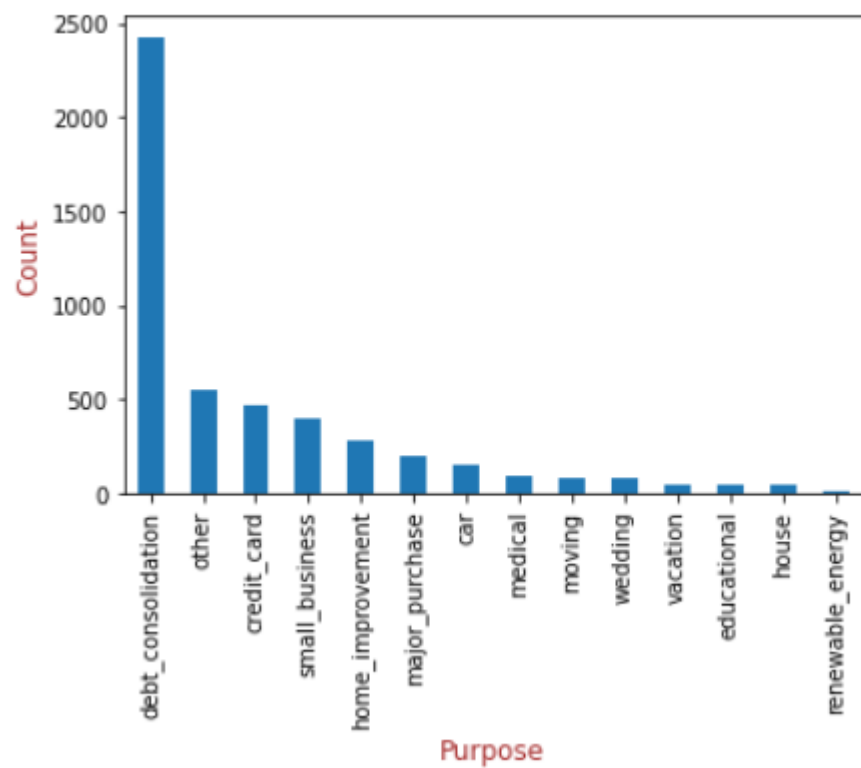
Home Ownership Column Analysis



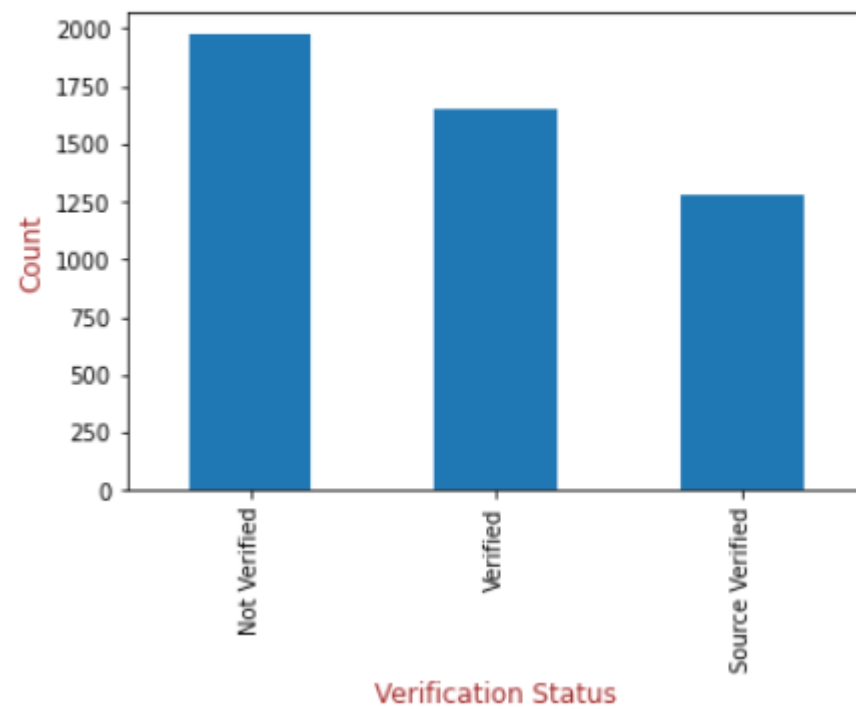
Interest Rate Column Analysis



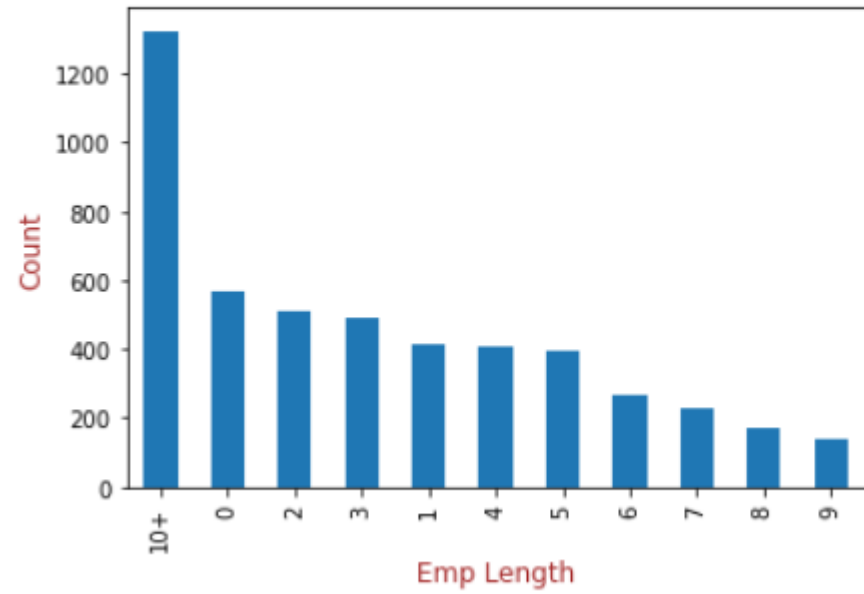
## Purpose Column Analysis



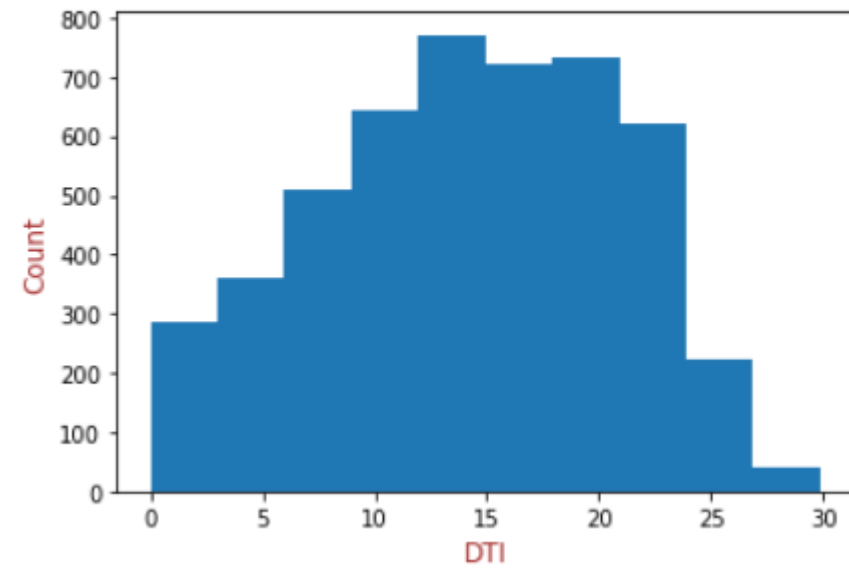
## Verification Status Column Analysis



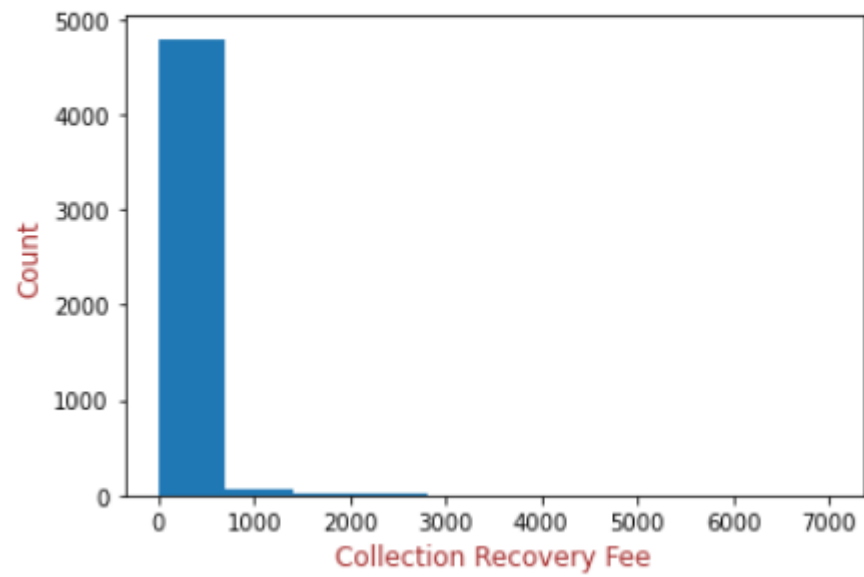
### Emp Length Column Analysis



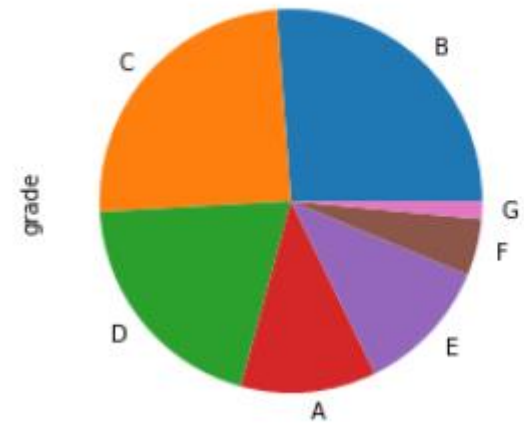
### DTI Column Analysis



## Collection Recovery Fee Column Analysis



## Grade Column Analysis



# Segmented Univariate Analysis

---

Segmented Univariate Analysis based on target column that is “loan\_status”

1. Charged Off -> status indicate default
2. Fully Paid -> status indicate not default
3. Current status -> not used in this case study,

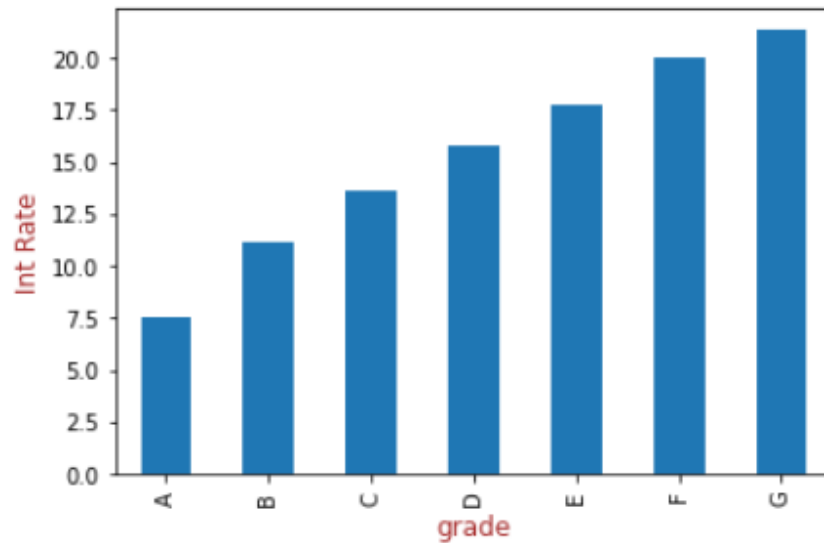


# Bivariate Analysis

---

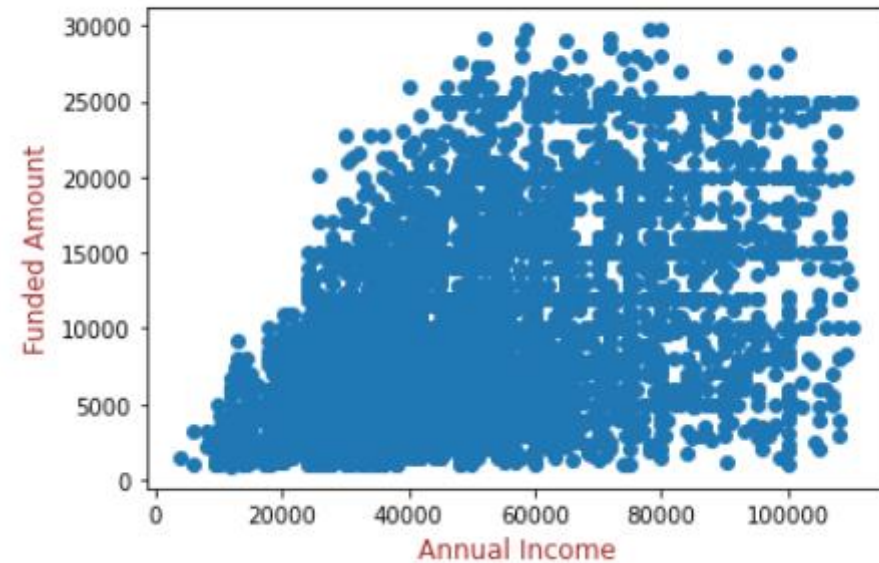
- int\_rate and grade are correlation

Bivariate Analysis for Grade and Int Rate



- annual\_inc and funded\_amnt are correlation

Annual Income versus Funded Amount

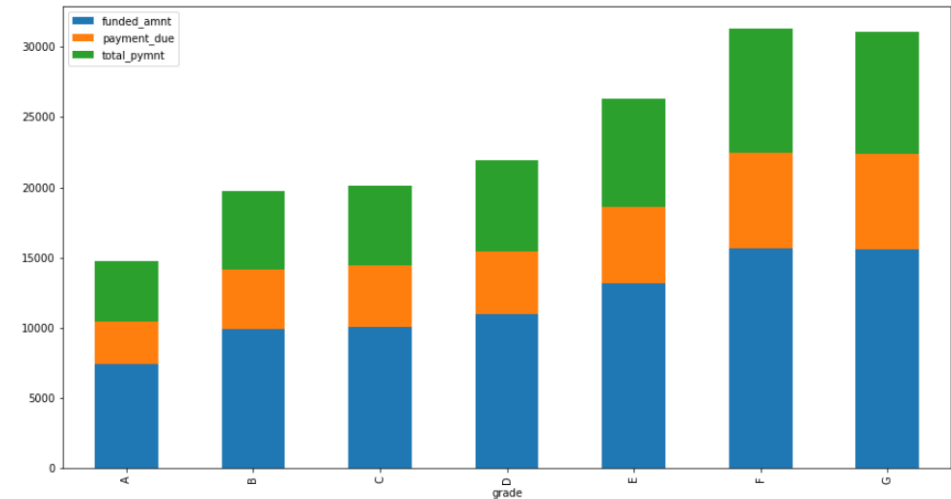
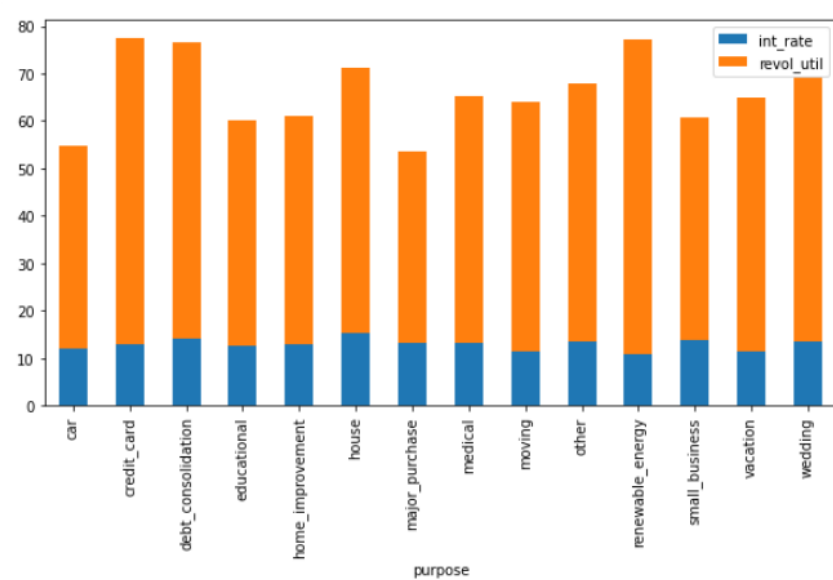


# Driver Variables

---

## Observation:

- int\_rate and revol\_util, indicate default variable.
- funded\_amnt, total\_pymnt and payment\_due, indicate default variable.



## Observation:

- installment, monthly\_income, emp\_length and home\_ownership, indicate default variable.
- annual\_inc, funded\_amnt, emp\_length and verification\_status, indicate default variable.

