

# LENDING CLUB ANALYSIS



## Presenters:



Larry Chuon



Venkata Prasath

<https://vprasath1902.github.io/>

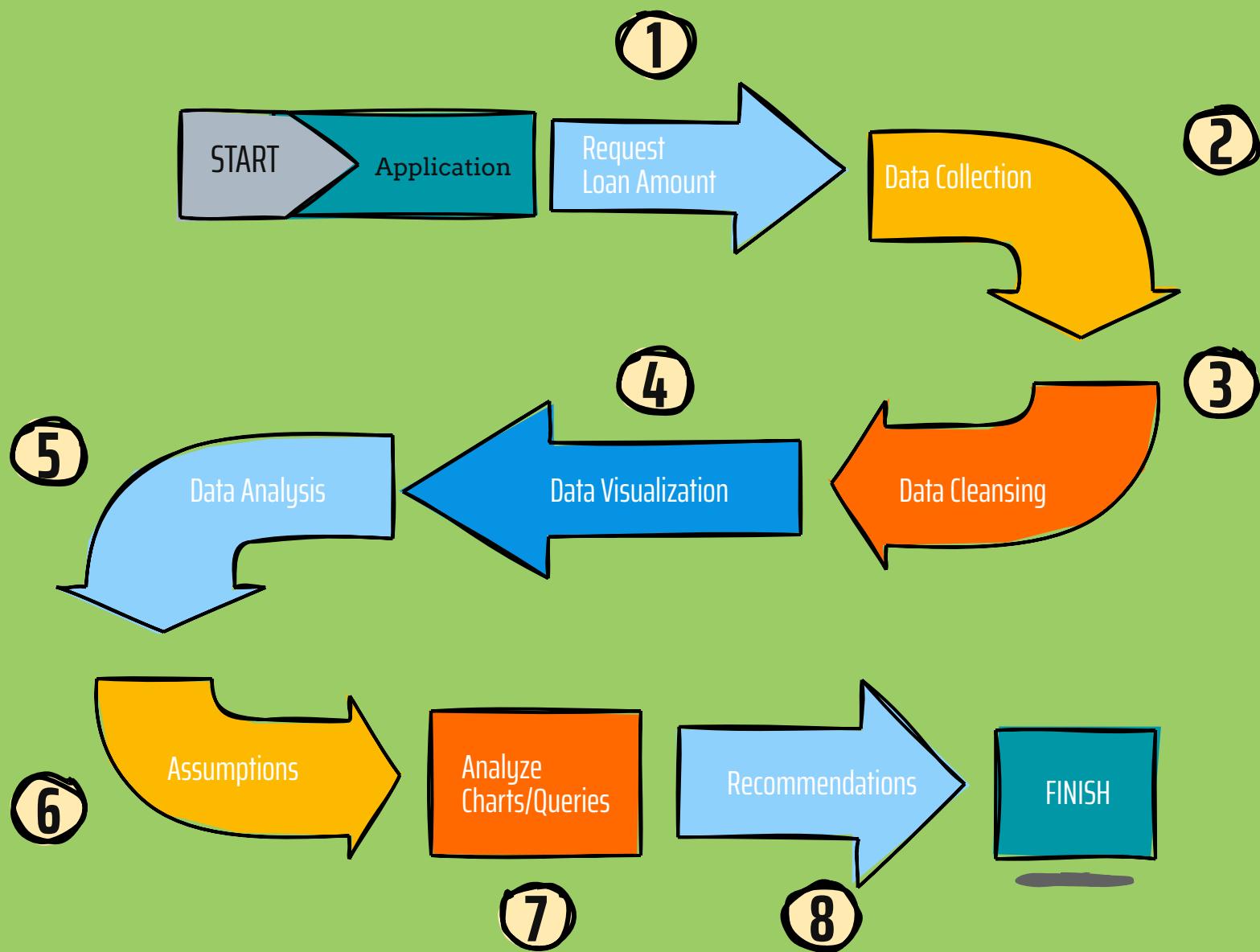
# Problem Statement:

Which variables provide a strong indication to predict loan defaults?

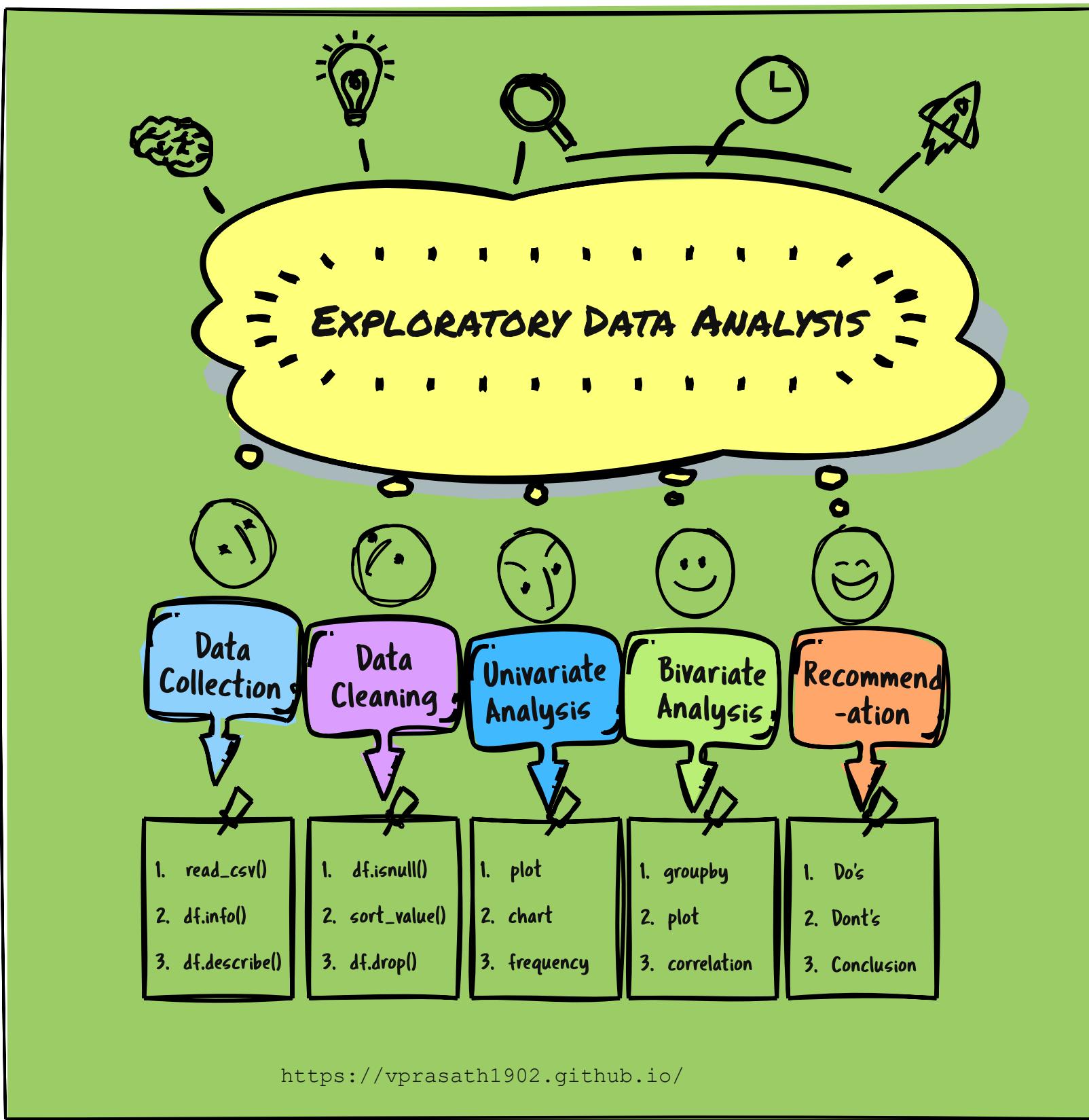


<https://vprasath1902.github.io/>

# Decision Making Process



<https://vprasath1902.github.io/>



# Data Collection

- ✓ 39,717 rows & 110 columns
- ✓ 56 columns have NO data
- ✓ 14 columns have MISSING data
- ✓ dtypes: float64(74), int64(13), object(24)

# Data Cleansing



Drop Columns with 0 data



Impute missing data

> Categorical: Mode

> Numerical w/ Outlier: Median

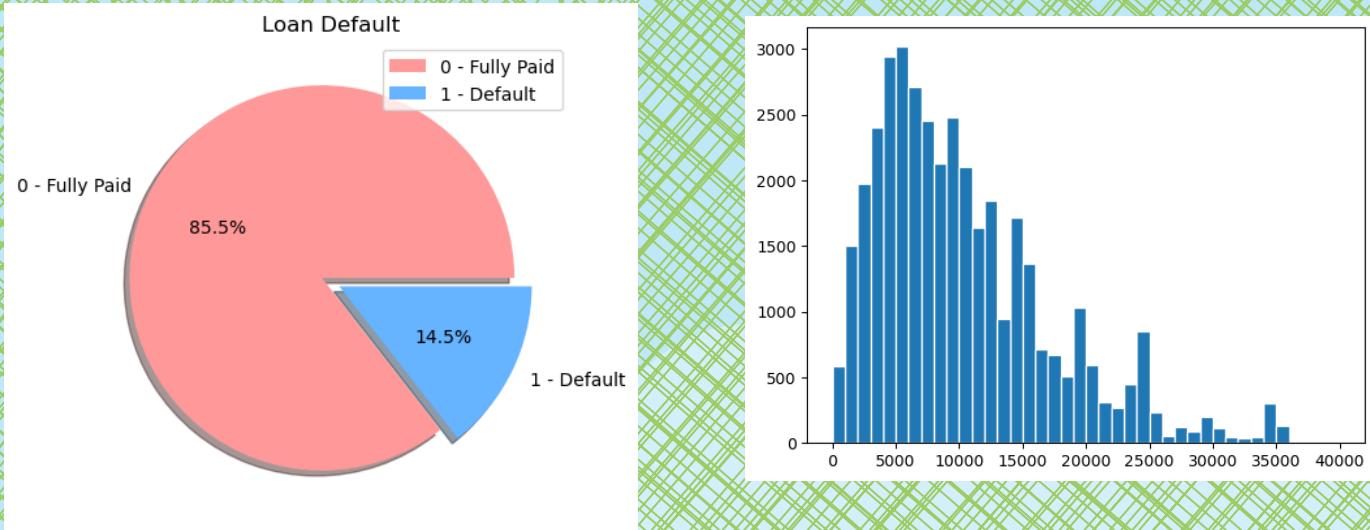
> Numerical w/o Outlier: Mean



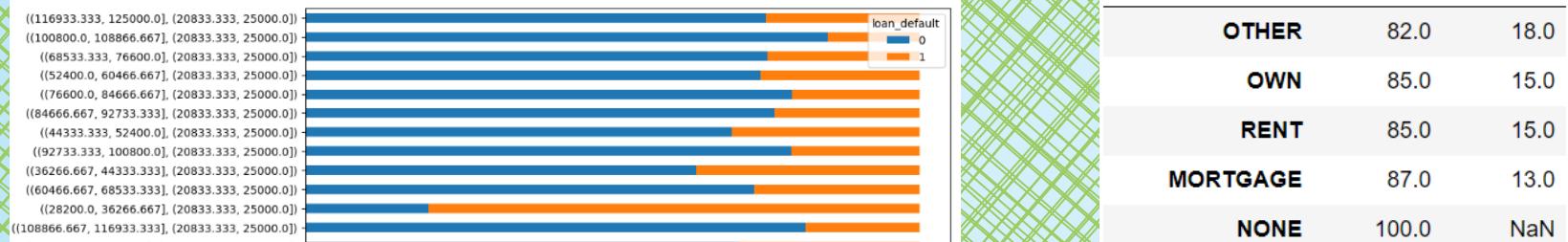
Datatype Conversion

# Data Analysis

## Univariate Analysis



## Bivariate Analysis



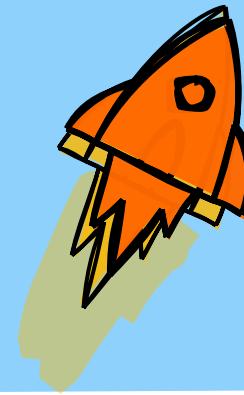
<https://vprasath1902.github.io/>



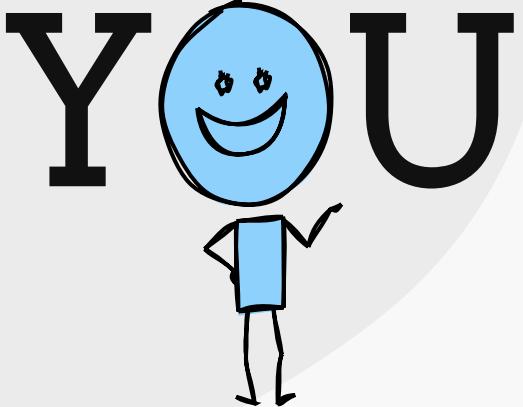
# Recommendations

-  If the Loan Amount is greater than 20833 and the loan grade is C or D or E or F or G and the applicants salary is less than 34250 then do not approve.
-  If the Loan grade is B or C or D or E with Interest rate between 5.403 and 7.797, then approve the loan.
-  When a new maximum interest rate is introduced, minimize the approval of loan applications in this interest rate so as to minimize the number of loan defaults.
-  Don't approve loan with the following criterias:

<b>States</b>	<b>Loan Amnt</b>	<b>Income</b>
WA, VA, SC, NJ, MD, AZ	16666	34250
NC	20833	34250
DC	12500	34250



# THANK YOU



## Presenters:



Larry Chuon



Venkata Prasath

<https://vprasath1902.github.io/>