

# **LENDING CLUB CASE STUDY**

Pilakkil, Nisar

# INTRODUCTION

A consumer finance company named 'Lending Club' specializes in lending various types of loans to urban customers. When the company receives a loan application, it has to make a decision for loan approval based on the applicant's profile.

There are two types of risks associated with the bank's decision:

- (i) If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company.
- (ii) If the applicant is not likely to repay the loan, i.e., he/she is likely to default, then approving the loan may lead to a financial loss for the company

Based on the dataset supplied company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default.

# LOADING THE LOAD DOCUMENT

1. Load Loan document to a data frame
2. Confirming the encoding of the document.
3. Listdown all statistical result by python command describe()

	id	member_id	loan_amnt	funded_amnt	funded_amnt_inv	installment	annual_inc	dti	delinq_2yrs	inq_last_6mths	mt
count	3.971700e+04	3.971700e+04	39717.000000	39717.000000	39717.000000	39717.000000	3.971700e+04	39717.000000	39717.000000	39717.000000	
mean	6.831319e+05	8.504636e+05	11219.443815	10947.713196	10397.448868	324.561922	6.896893e+04	13.315130	0.146512	0.869200	
std	2.106941e+05	2.656783e+05	7456.670694	7187.238670	7128.450439	208.874874	6.379377e+04	6.678594	0.491812	1.070219	
min	5.473400e+04	7.069900e+04	500.000000	500.000000	0.000000	15.690000	4.000000e+03	0.000000	0.000000	0.000000	
25%	5.162210e+05	6.667800e+05	5500.000000	5400.000000	5000.000000	167.020000	4.040400e+04	8.170000	0.000000	0.000000	
50%	6.656650e+05	8.508120e+05	10000.000000	9600.000000	8975.000000	280.220000	5.900000e+04	13.400000	0.000000	1.000000	
75%	8.377550e+05	1.047339e+06	15000.000000	15000.000000	14400.000000	430.780000	8.230000e+04	18.600000	0.000000	1.000000	
max	1.077501e+06	1.314167e+06	35000.000000	35000.000000	35000.000000	1305.190000	6.000000e+06	29.990000	11.000000	8.000000	

# DATA CLEANING

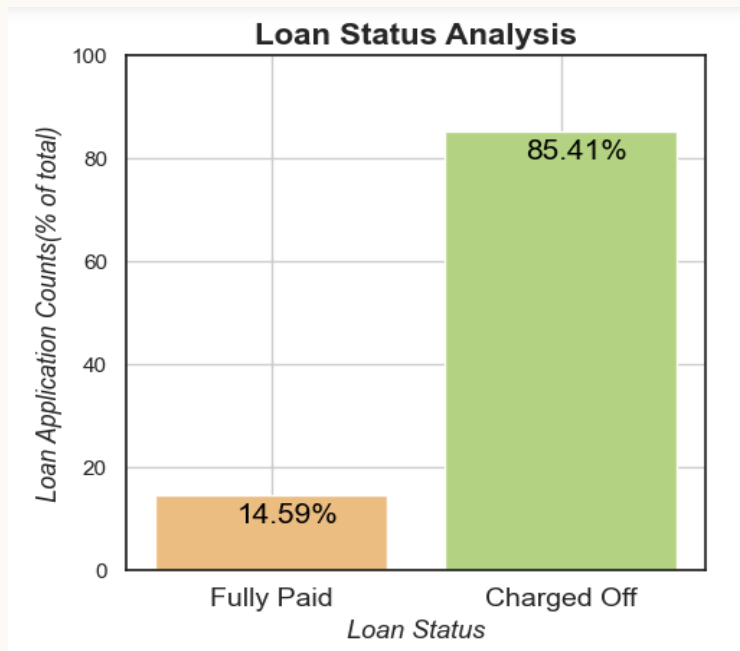
1. eliminate columns having more than 50% missing values
2. Remove single value unique columns from the dataframe, by checking `dataFrame.nunique() == 1`
3. Removing duplicate rows (if any), using `DataFrame.drop_duplicates()` function.
4. Dropping some columns as they are not of much use for analysis.
5. Checking % of NaNs in columns ( $>0$ )
6. Creating new columns like, month and year based on 'issue\_d' column
7. Converting some columns to lowercase strings and stripping off blank spaces around the content

# DATA ANALYSIS

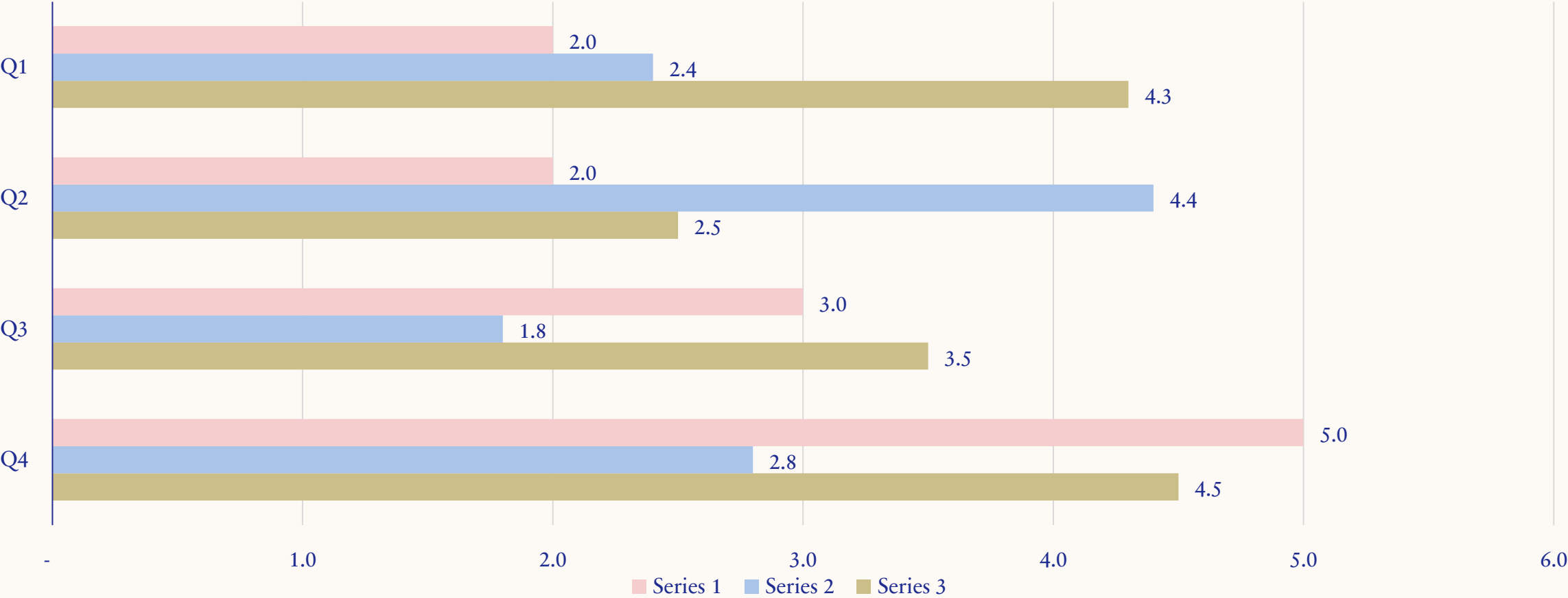
1. Check count of columns fully paid, charged off, Current
2. Calculate the percentage of counted columns
3. Considering only the candidates with 'fully paid' and 'charged off' loan\_status. Removing records for 'current state', loan status.
4. Check for assigned '0' to 'fully paid' and set 'charged off' to '1' (as our target is to determine 'defaulter' factors)
5. List down of continuous variables
6. List of categorical variables

# UNIVARIATE ANALYSIS AND SEGMENTED UNIVARIATE ANALYSIS

1. Lets check how 'loan\_status' (%) is distributed

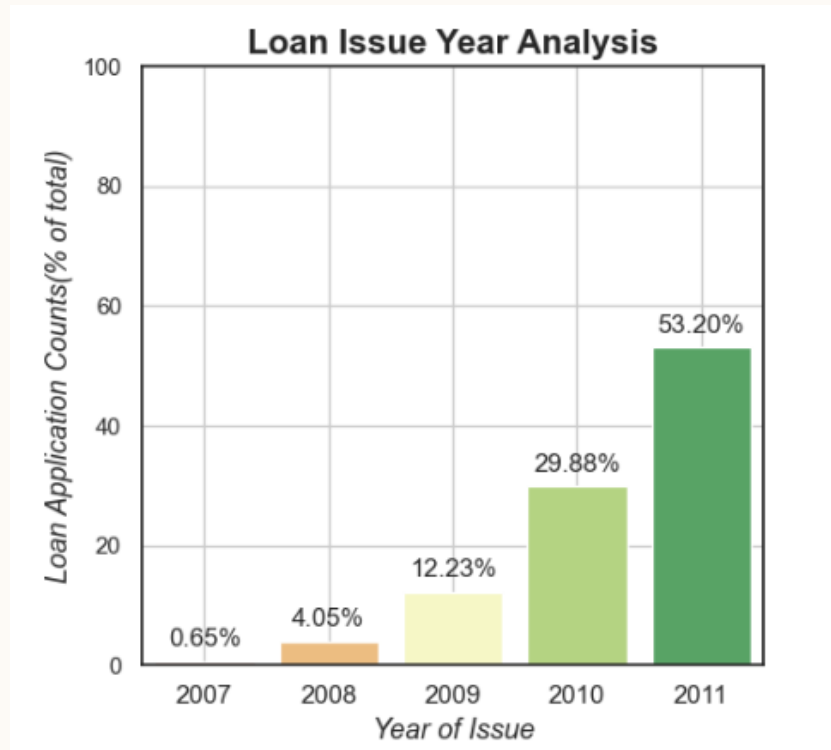


# QUARTERLY PERFORMANCE



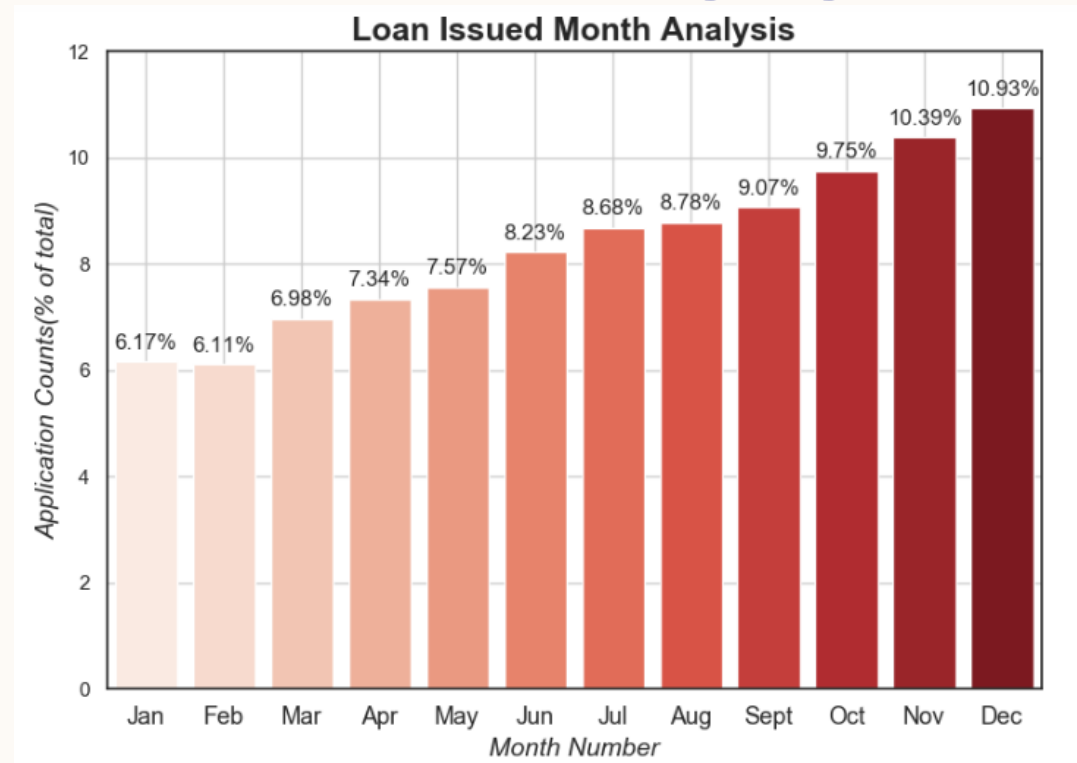
# LOAN ISSUE YEAR ANALYSIS

1. Lets check the total number of approved applications by year

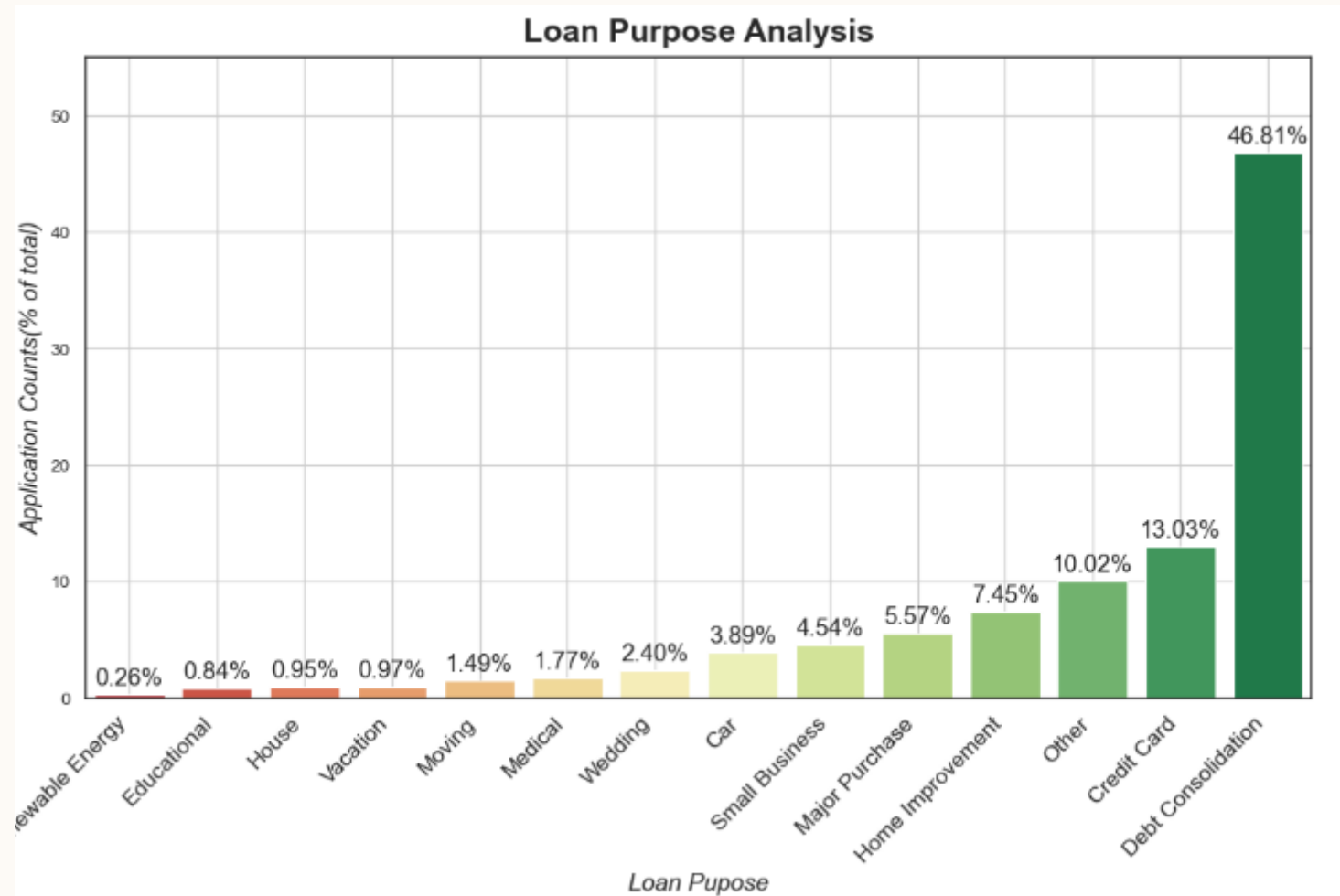




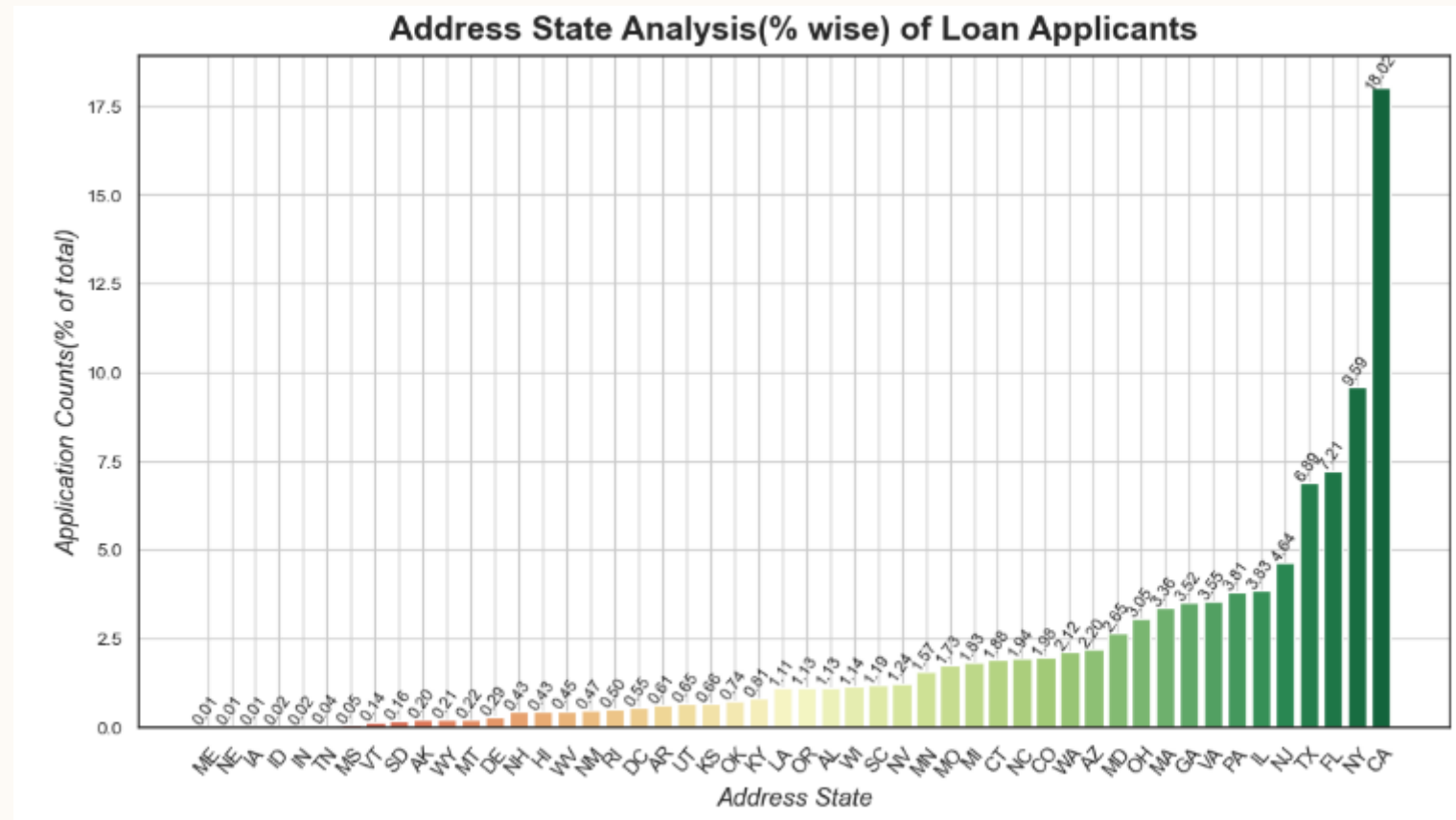
# TOTAL% OF APPROVED LOAN APPLICATIONS BY MONTH ANALYSIS:



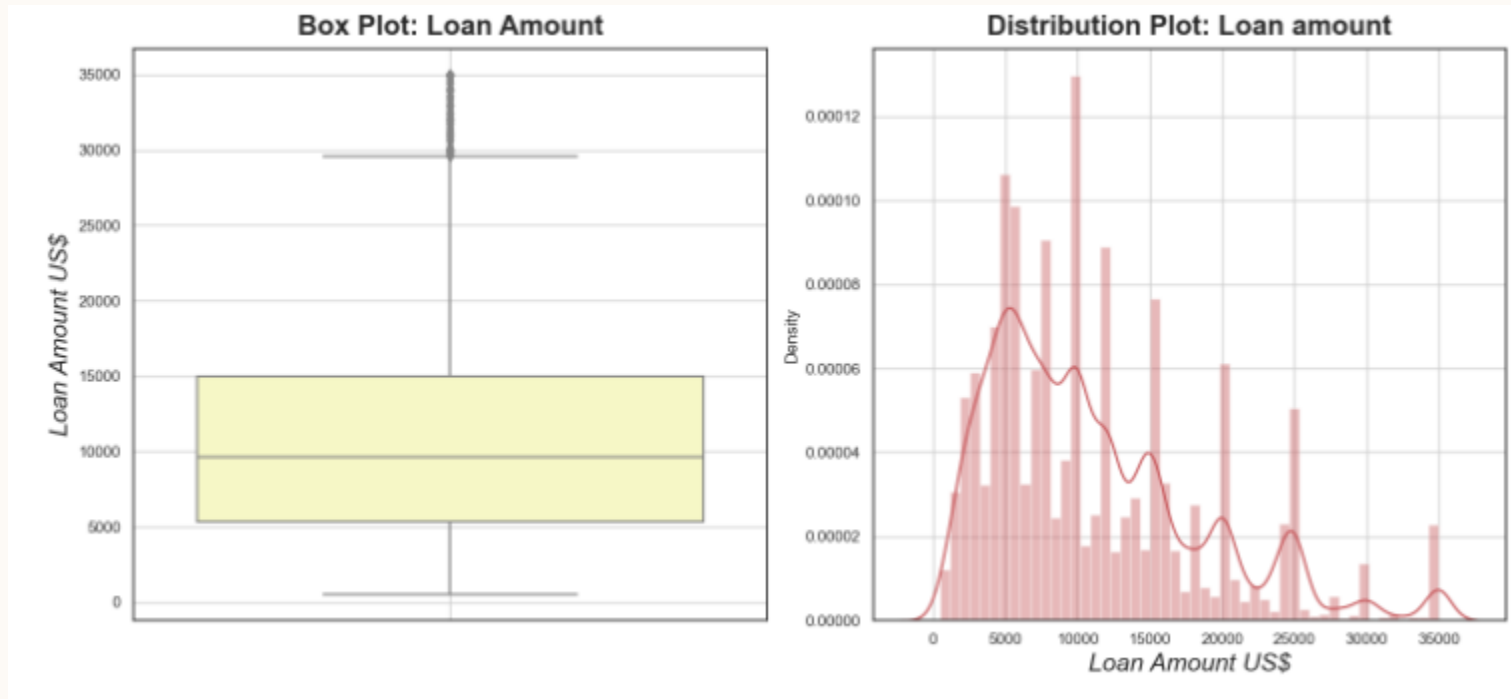
# LOAN PURPOSE ANALYSIS



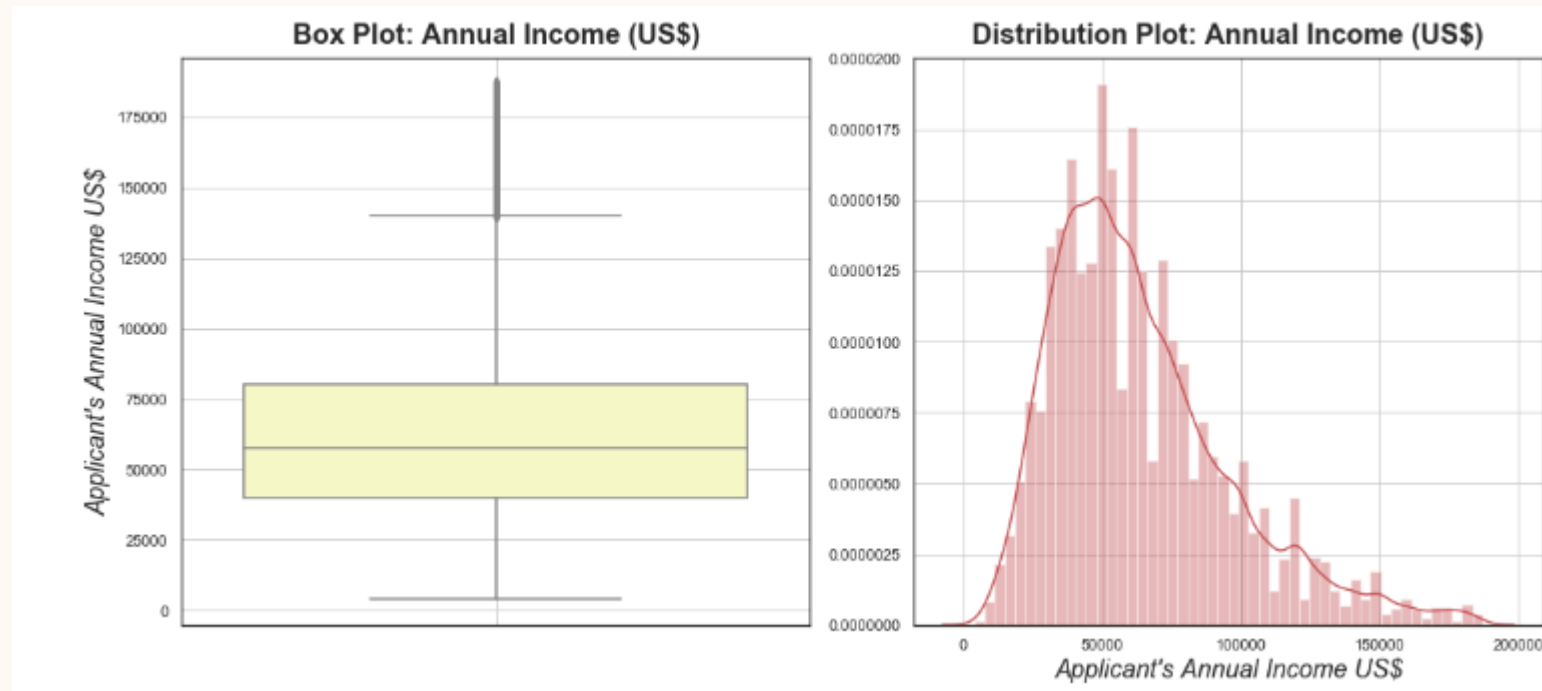
# ADDRESS STATE ANALYSIS(% WISE) OF LOAN APPLICANTS



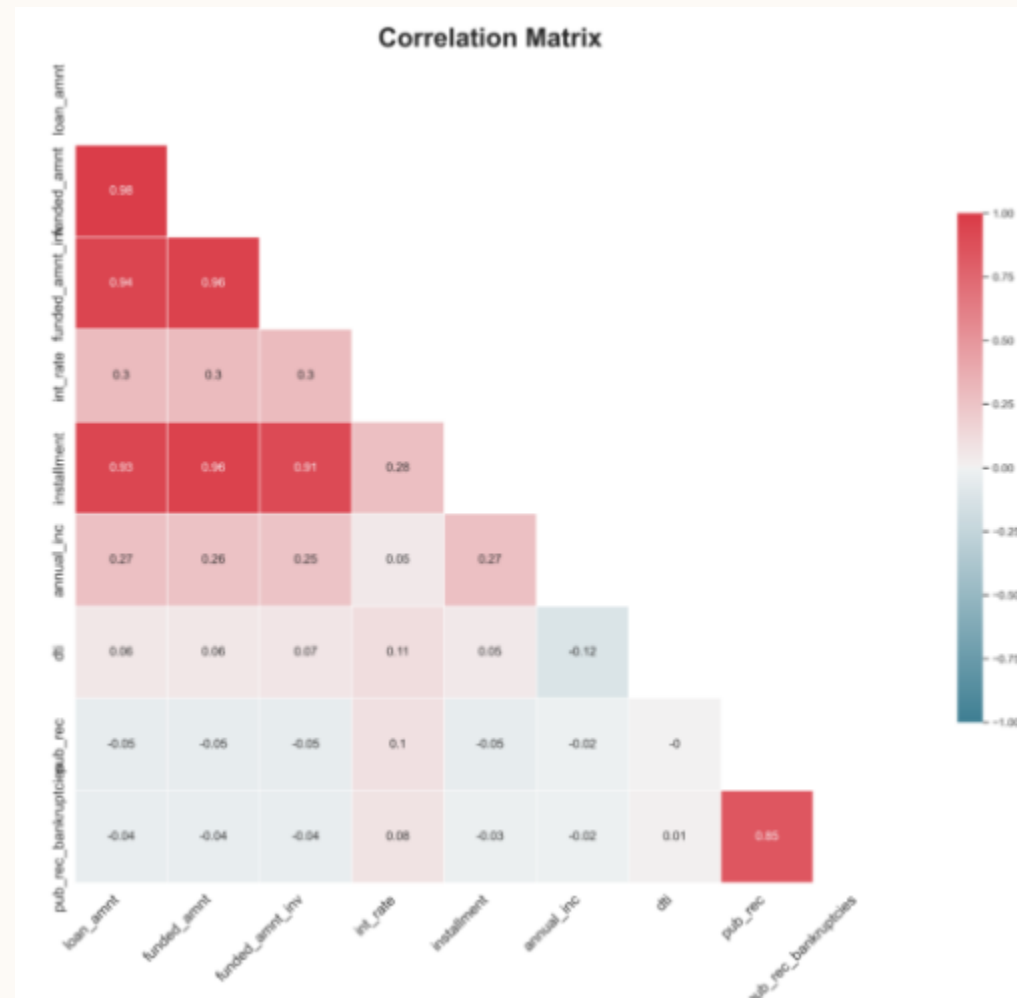
# APPLICANT'S LOAN AMOUNT' ANALYSIS



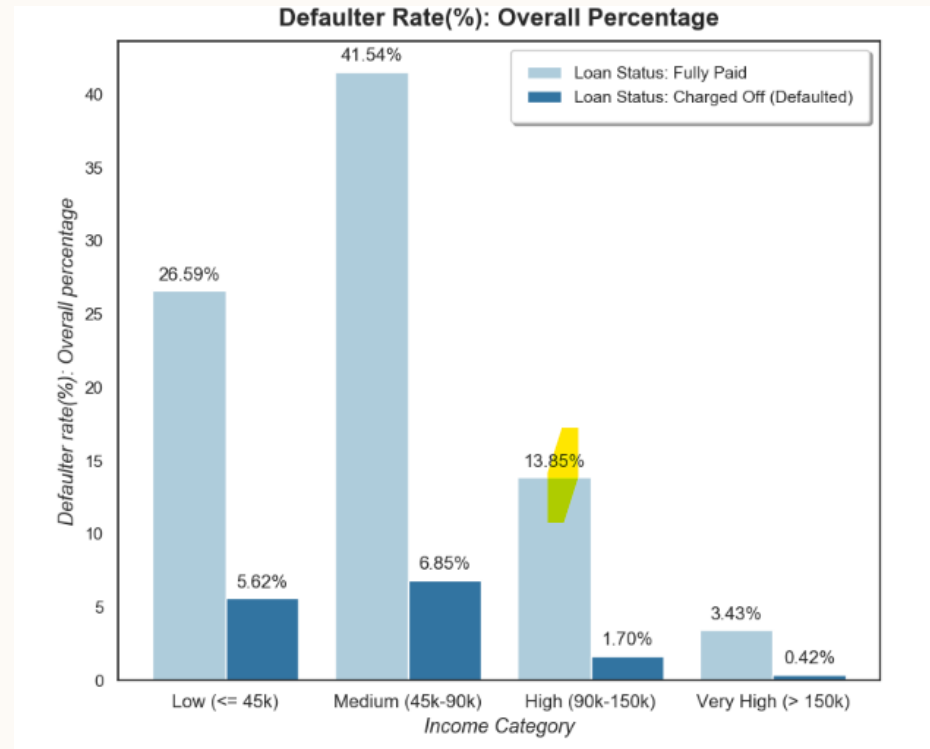
# APPLICANT'S ANNUAL INCOME ANALYSIS



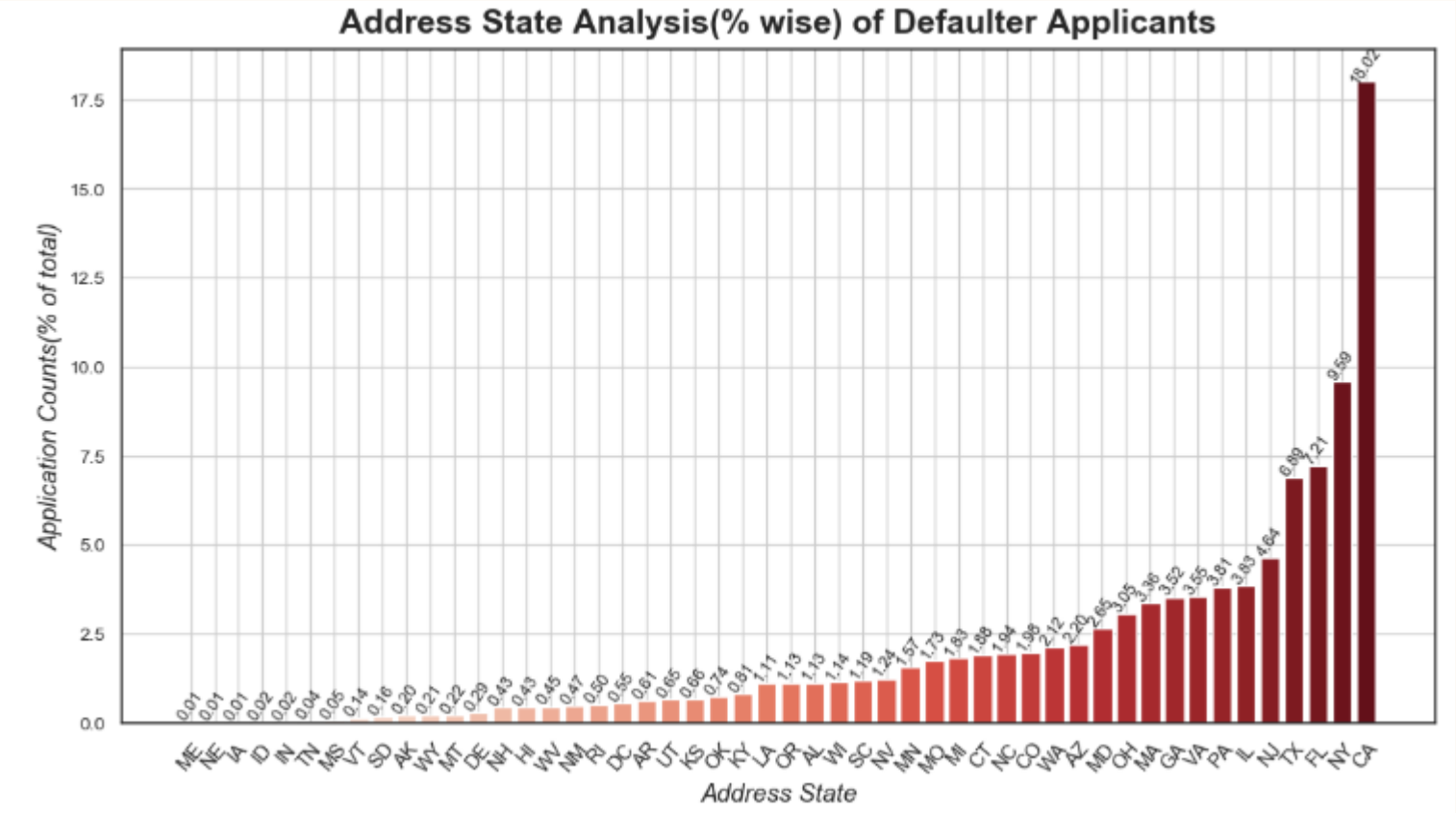
# BIVARIATE ANALYSIS



## DEFAULTERS RATE(%) ANALYSIS BASED ON ANNUAL INCOME OF APPROVED LOAN APPLICANTS



# ADDRESS STATE ANALYSIS(% WISE) OF DEFAULTERS





# SUMMARY

At Contoso, we believe in giving 110%. By using our next-generation data architecture, we help organizations virtually manage agile workflows. We thrive because of our market knowledge and great team behind our product. As our CEO says, "Efficiencies will come from proactively transforming how we do business."



# **THANK YOU**

Pilakkil, Nisar

[nisarop@gmail.com](mailto:nisarop@gmail.com)