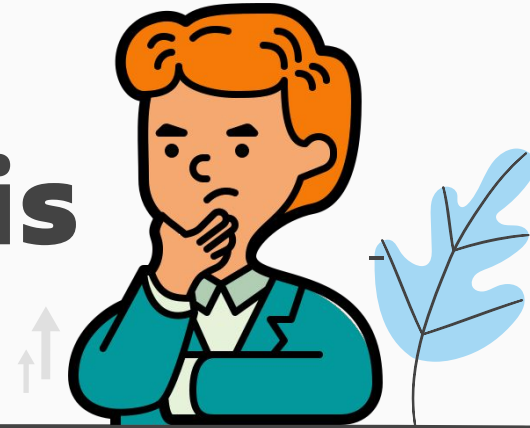


# ADCB Loans' Dataset Analysis



## OBJECTIVES

- ADCB Loans' Dataset Analysis
- Recommendations to automate Loan Process

## DATASET FEATURES



Loan Amount



Loan Term



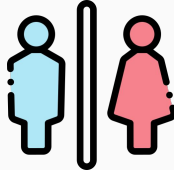
Booking Date



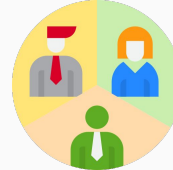
Maturity Date



Age



Gender

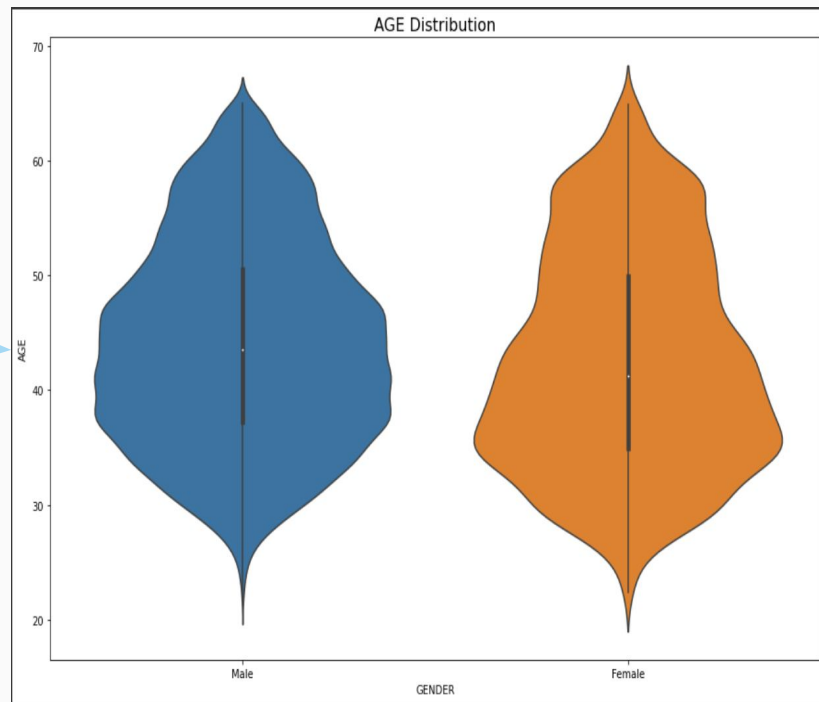
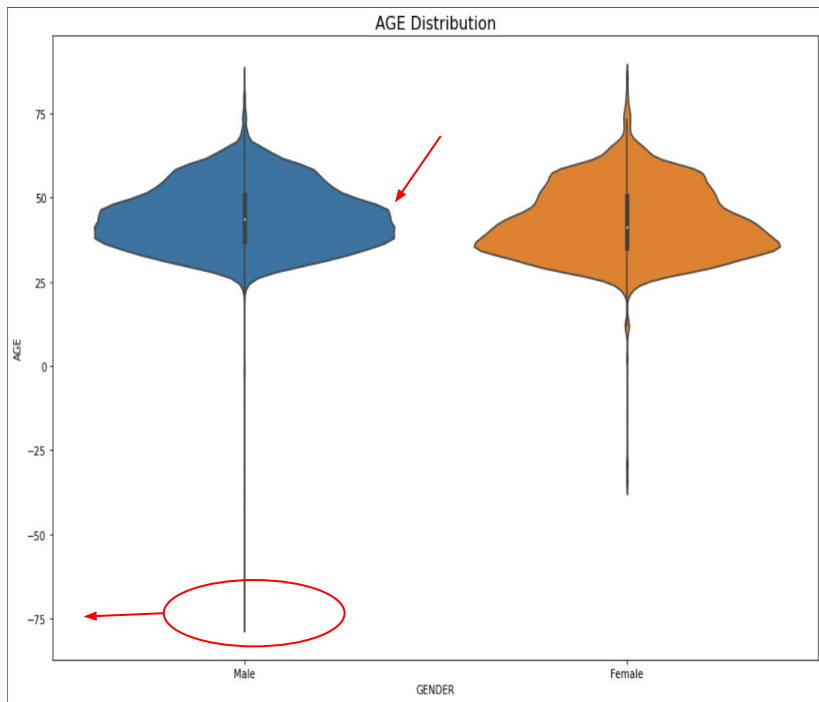


Customer  
Segment

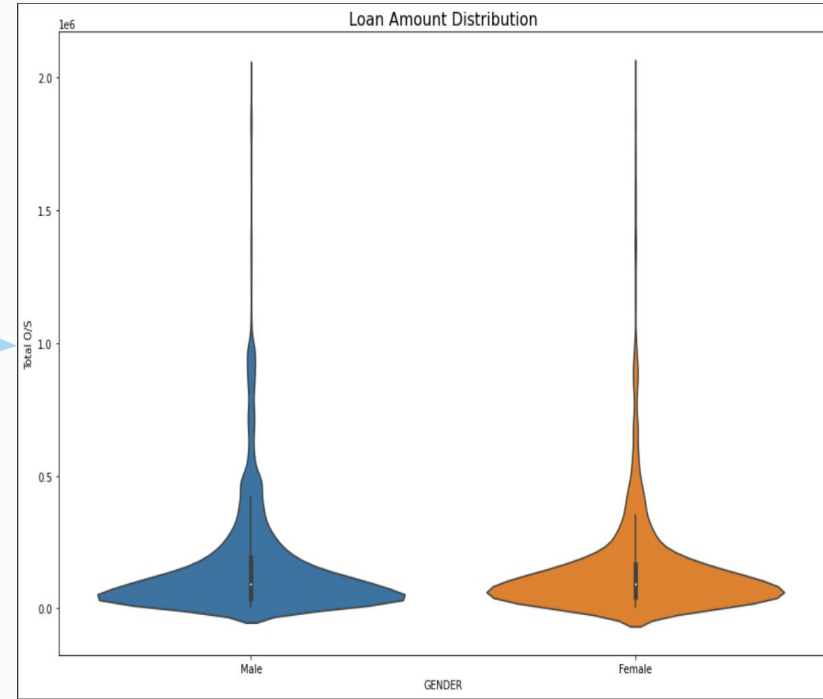
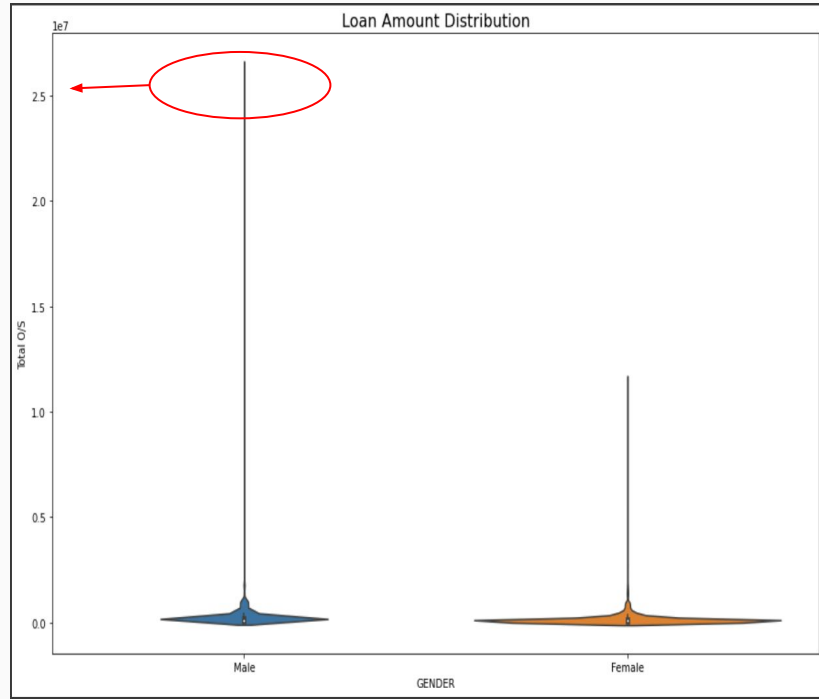


Days Past  
Dues

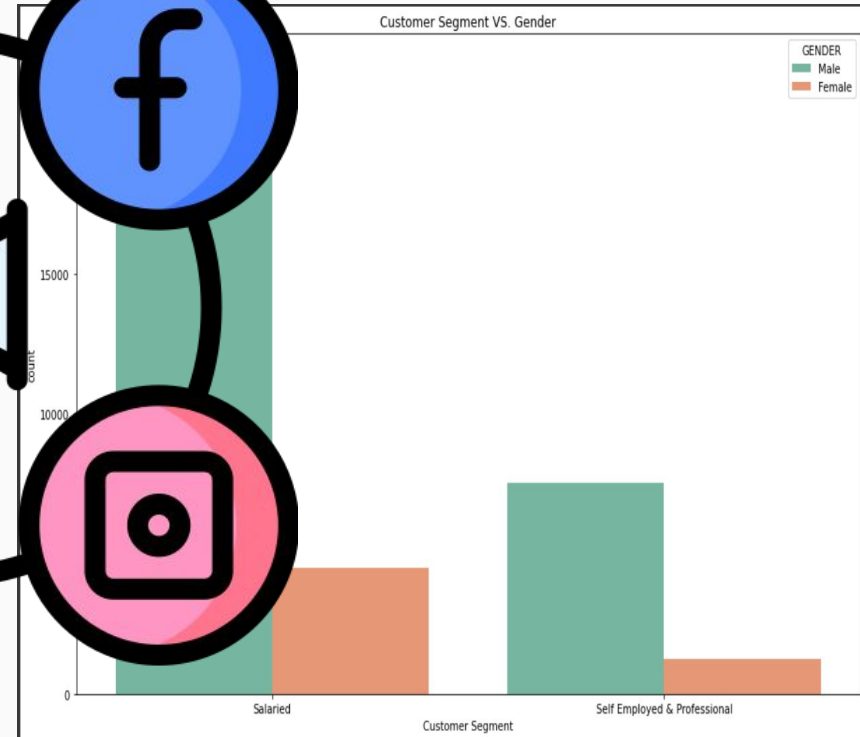
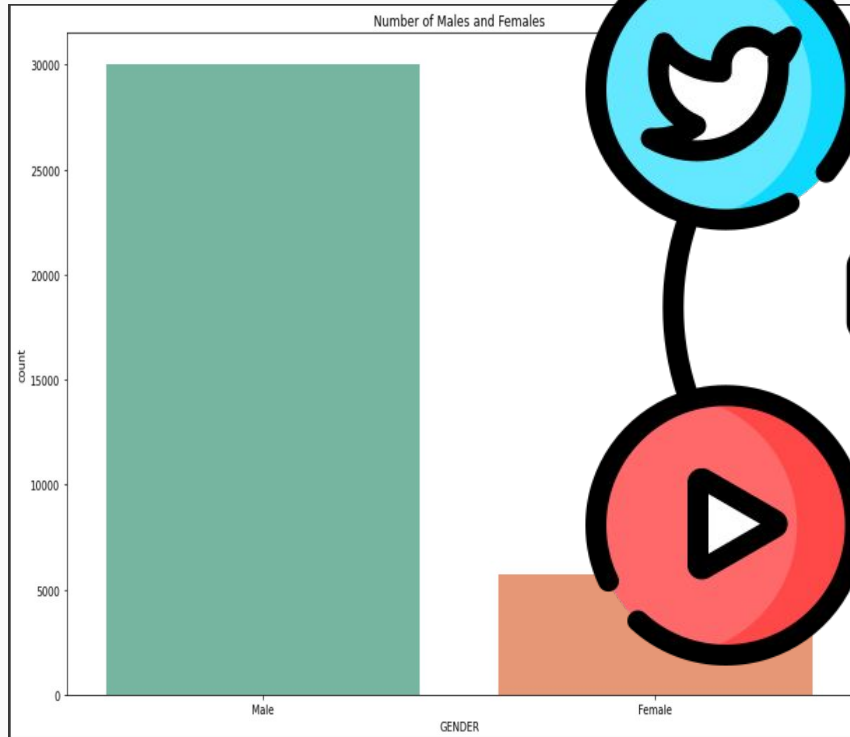
# AGE



# LOAN AMOUNT

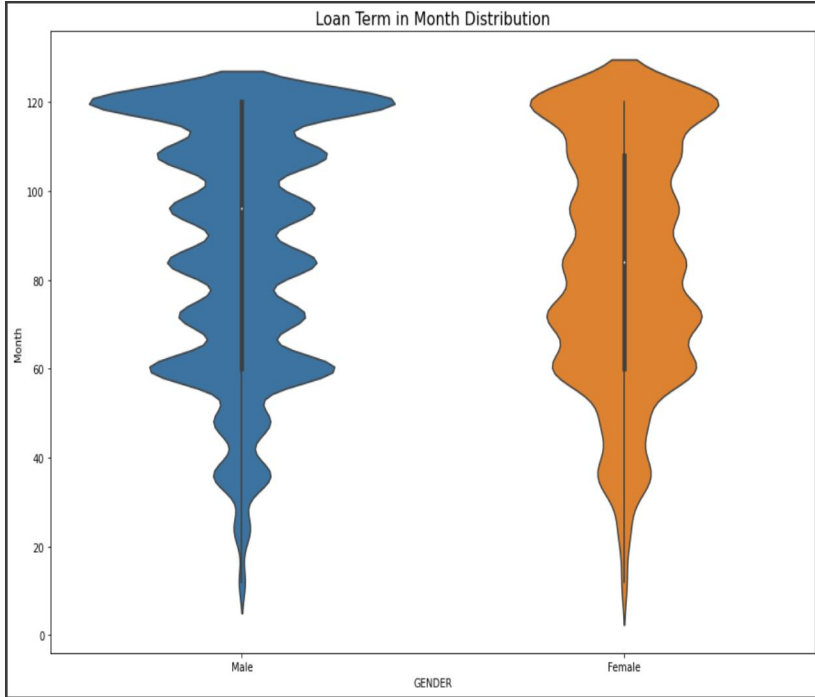


# GENDER AND CUSTOMER SEGMENT

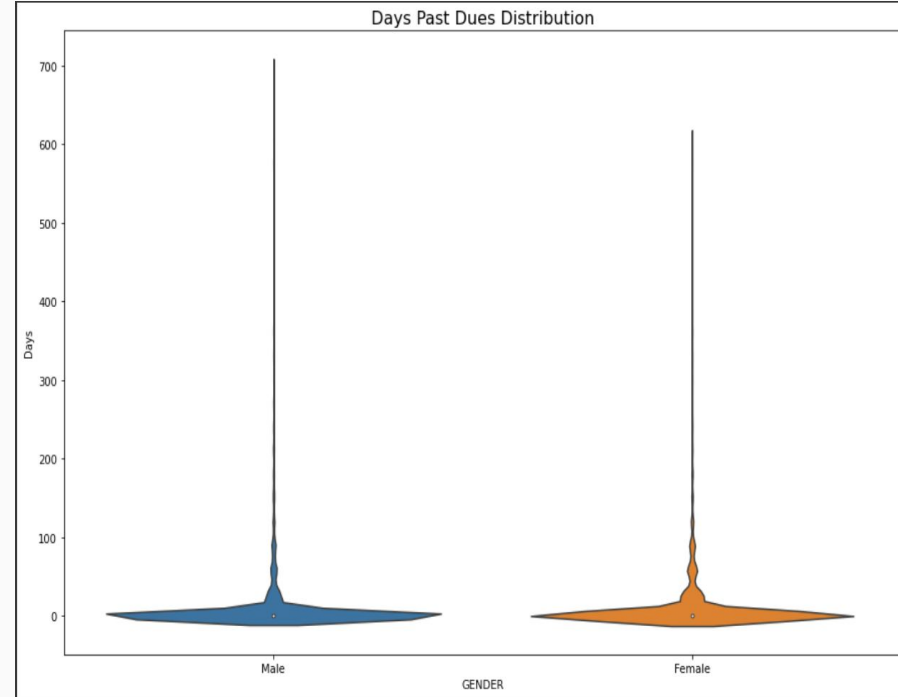


# LOAN TERM AND PAST DUES

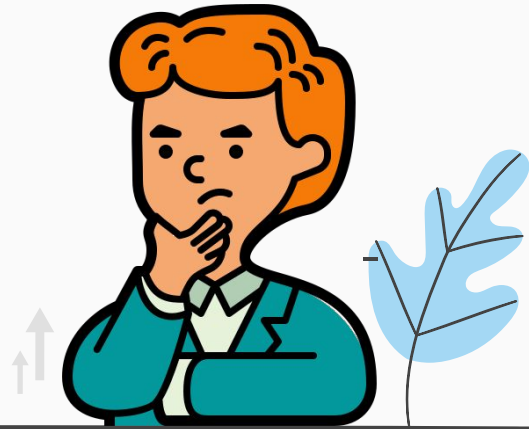
Loan Term in Month Distribution



Days Past Dues Distribution

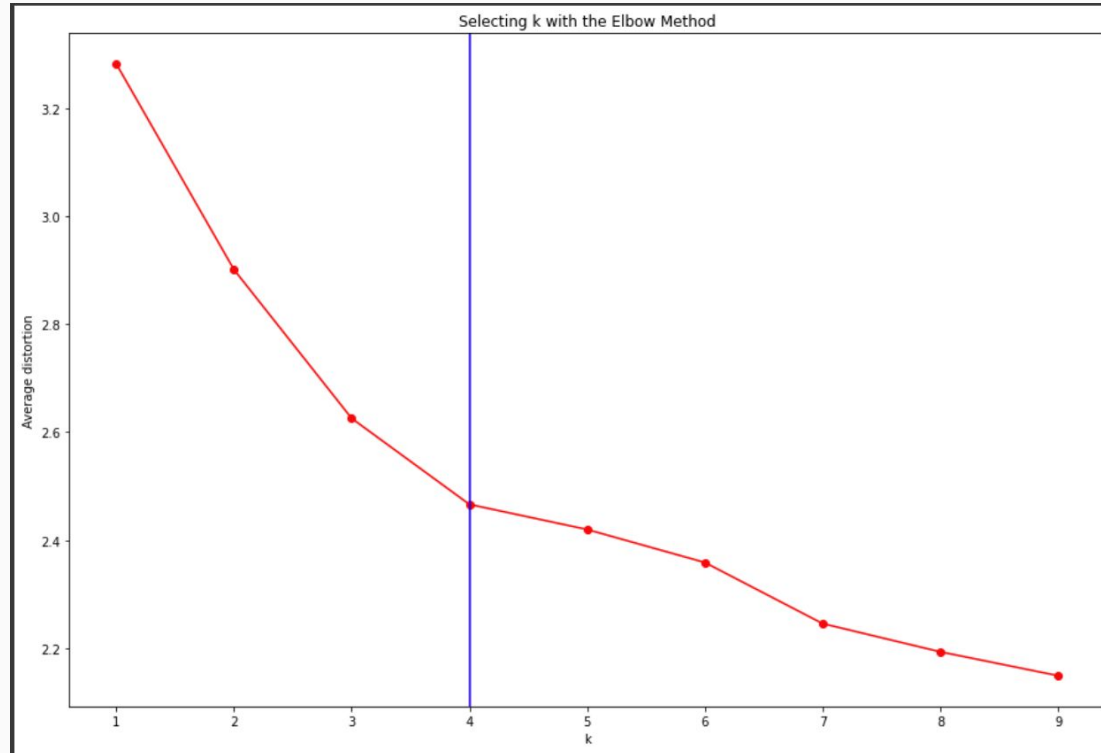


# Machine Learning Models



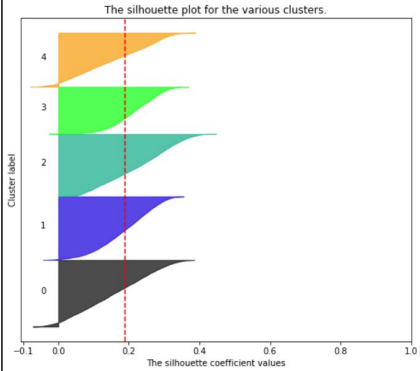


# CHOOSING NUMBER OF CLUSTERS

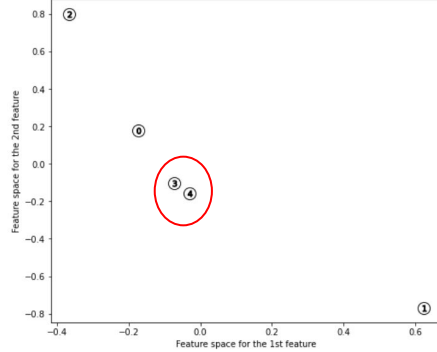


# CHOOSING NUMBER OF CLUSTERS

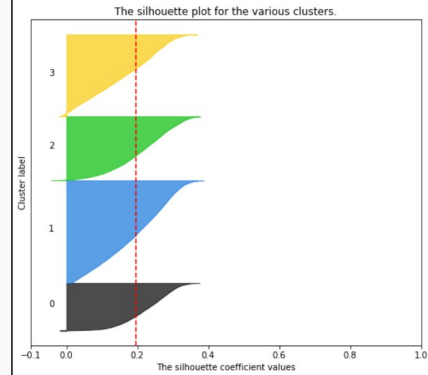
Silhouette analysis for KMeans clustering on sample data with  $n\_clusters = 5$



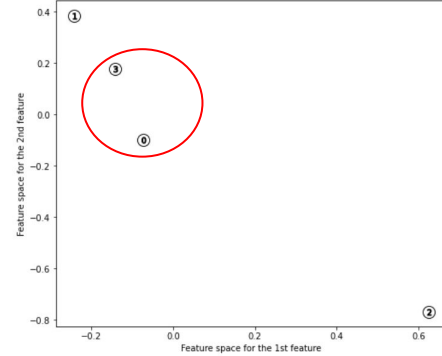
The visualization of the clustered data.



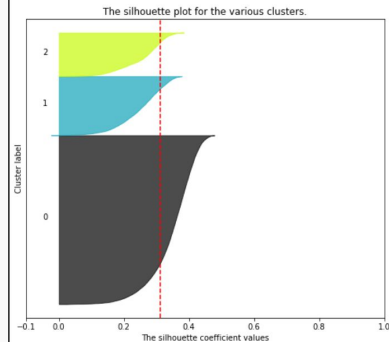
Silhouette analysis for KMeans clustering on sample data with  $n\_clusters = 4$



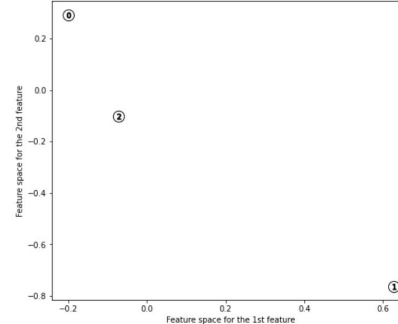
The visualization of the clustered data.



Silhouette analysis for KMeans clustering on sample data with  $n\_clusters = 3$

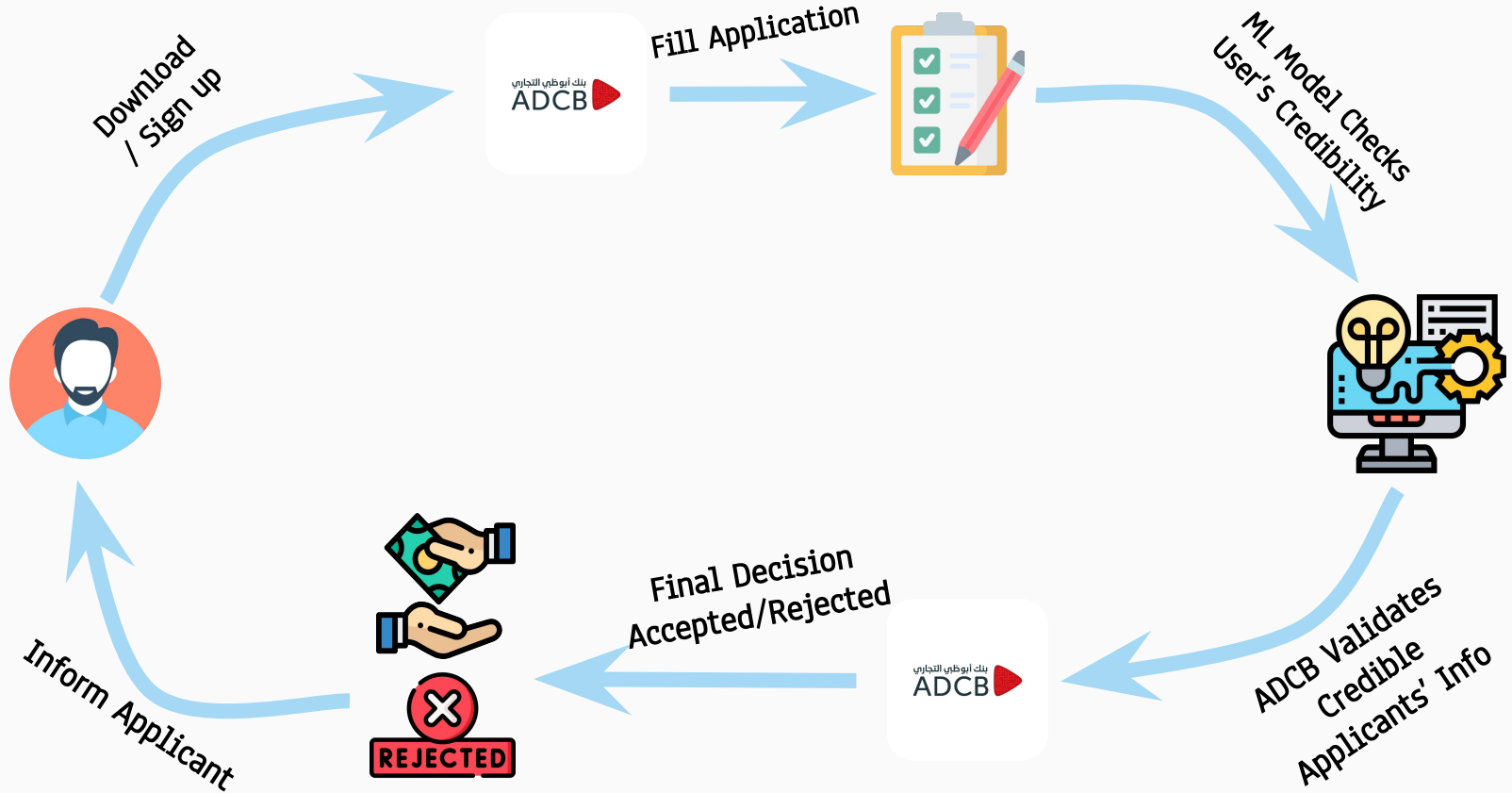


The visualization of the clustered data.



- High Correlation between **Loan Amount** and **Male Customers**
- High Correlation between **Loan Term** and **Salaried** Customers

## PROPOSED BUSINESS PROCESS



## ADDITIONAL FEATURES



Loan Amount



Loan Term



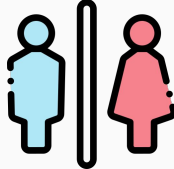
Booking Date



Maturity Date



Age



Gender

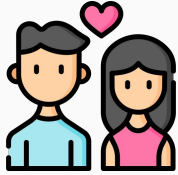


Customer  
Segment

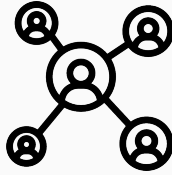


Days Past  
Dues

## ADDITIONAL FEATURES



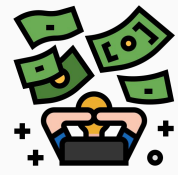
Marital Status



Dependents



Education



Applicant Income



Co-applicant  
Income



Credit  
History

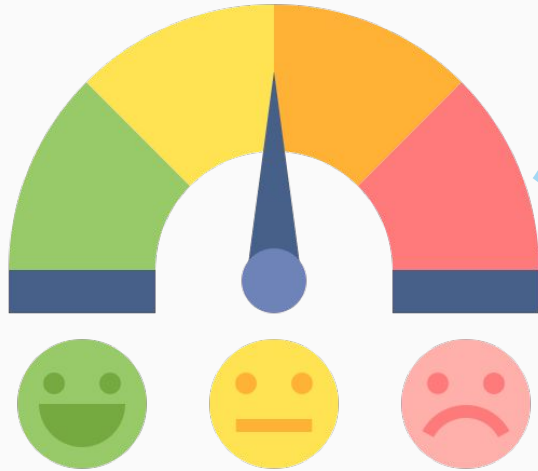


Property  
Area



Properties

## MACHINE LEARNING MODEL OUTPUT



CREDIT SCORE

Credible: 200 to 300




Risky: 100 to 200



Not-credible: 0 to 100



## DATASET SAMPLES OF IMPLEMENTED MODELS

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area	Loan_Status	
0	LP001002	Male	No	0	Graduate	No	5849	0.0	NaN	360.0	1.0	Urban	Y	
1	LP001003	Male	Yes	1	Graduate	No	4583	1508.0	128.0	360.0	1.0	Rural	N	
2	LP001005	Male	Yes	0	Graduate	Yes	3000	0.0	66.0	360.0	1.0	Urban	Y	
3	LP001006	Male	Yes	0	Not Graduate	No	2583	2358.0	120.0	360.0	1.0	Urban	Y	
4	LP001008	Male	No	0	Graduate	No	6000	0.0	141.0	360.0	1.0	Urban	Y	

1 26 2023/05/01



## RESULTS OF IMPLEMENTED MODELS

Model name	Accuracy	Cross-validation accuracy
Logistic regression	78%	89%
Random forest	79%	85%
Decision trees	75%	81%

## **Fully Automated Loan Process**

# THANKS

Does anyone have any questions?

**CREDITS:** This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik** and illustrations by **Storyset**

