

Based on EDA + Data Analysis :

→ Columns to KEEP (for Feature Engineering):

```
[  
    'Age',  
    'Salary_Net_Income',  
    'Social_security_net_income',  
    'Self_employed_net_income',  
    'Salary_frequency',  
    'Months_current_job',  
    'Total_experience_months',  
    'Ever_Bankrupt',  
    'Loan_Amount',  
    'Interest_rate',  
    'Tenure',  
    'Fee',  
    'Recurring_payment_amount',  
    'Penalty',  
    'Current_delinquent',  
    'Current_default',  
    'Past_delinquencies',  
    'Credit_Score',  
    'Past_to_be_collected',  
    'Past_Collection_amount',  
    'Refinance',  
    'Extended_payment_plan',  
    'Settlement_plans_with_waive_off'  
]
```

→ Columns to Drop (for Feature Engineering):

```
[  
  'Customer_ID',  
  'Loan_ID',  
  'Current_Bankruptcy_status',  
  'Tenure_type',  
  'Accepted_option',  
  'Promised_to_pay',  
  'Promised_to_pay_status',  
  'Past_Collection_Success',  
  'Total_Due',  
  'No_of_days_since_last_payment',  
  'Missed_payments',  
  'On_time_payments'  
]
```

Why to keep ?

selected features that have a direct or logical relationship with a customer's ability to repay or settle their dues — things like income, job stability, loan details, credit score, and past payment behavior.

Why to Drop ?

Dropping columns like like (**Customer_ID**, **Loan_ID**) are redundant (like **Current_Bankruptcy_status** overlapping with Ever_Bankrupt), textual/categorical fields with too many missing values, or columns that could potentially leak future information (**Total_Due**, **Missed_payments**, etc.).