CREDIT RISK ANALYSIS

Presented By Team A

Team Members

- NAMRATA TODKAR (TEAM LEAD)
- DIVYA CHARITHA AMBALLA
- (CO-TEAM LEAD)
- D LOKESH REDDY
- SARNALINA DATTA
- HINA SHAH
- MAYANK
- RONAK

Introduction To Credit Risk Analytics

- Credit risk is the potential for borrower loan default.
- Critical for banks, influencing profitability and stability.
- Effective credit risk management ensures a healthy loan portfolio.
- Data analytics accurately assesses borrower risk profiles.
- Informed decisions lower default risks.
- Supports strategic planning and compliance.

Business Problem

- Identifying high-risk borrowers before issuing loans.
- Predicting default probability using historical data.
- Implementing preventive measures to mitigate risks.
- Minimizing losses while maximizing profitable lending.
- Ensuring sustainable growth by managing credit exposure.
- Achieving a competitive edge through effective risk management.

Dataset Overview

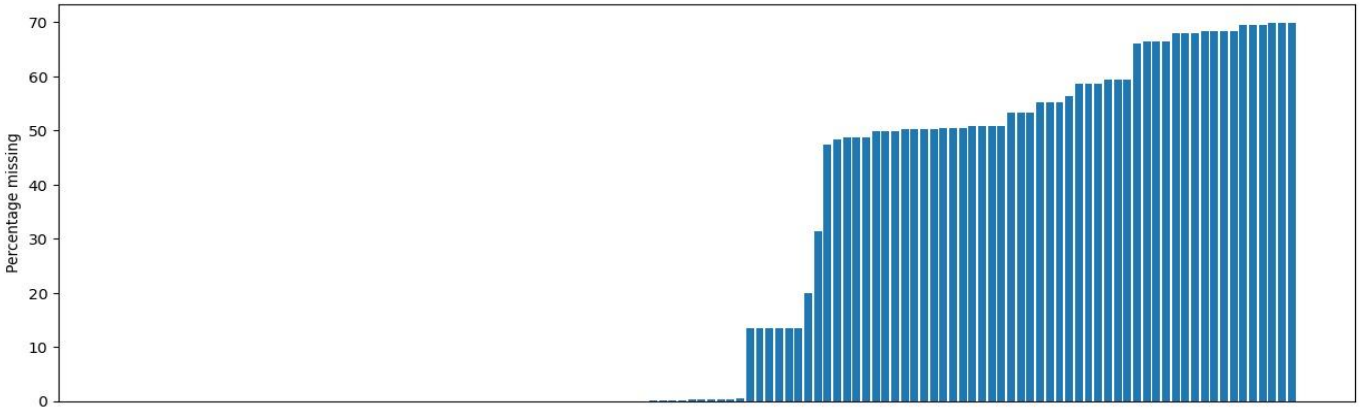
The dataset consists of three files:

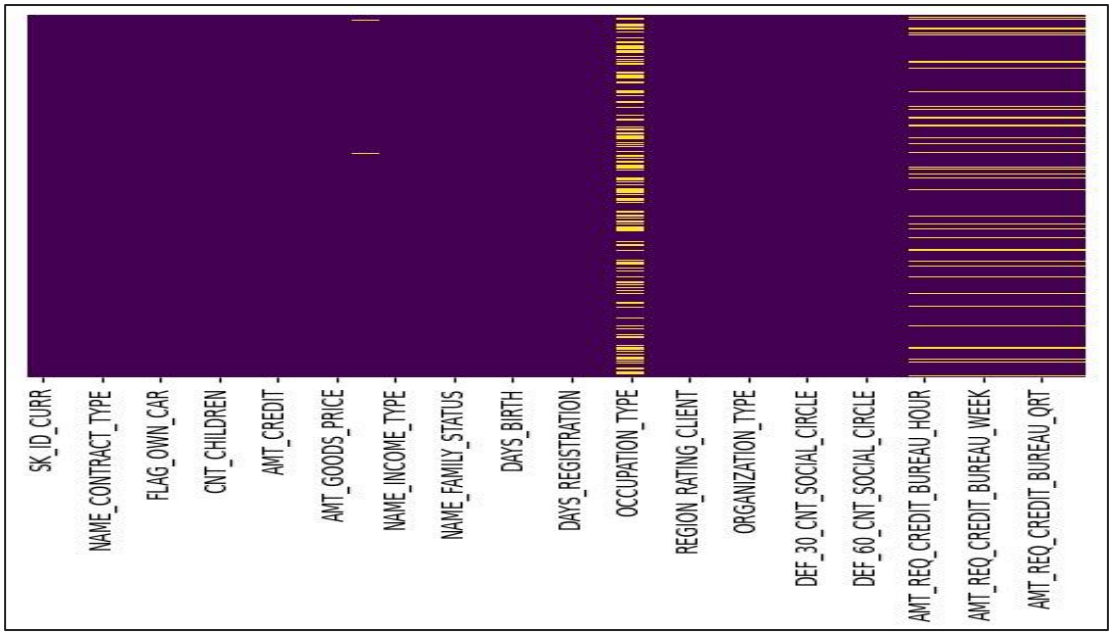
- **application_data.csv:** Contains client information at the time of application, including payment difficulties.
- **previous_application.csv:** Details previous loan applications, including outcomes like approval, cancellation, refusal, or unused offers.
- **columns_description.csv:** Serves as a data dictionary, explaining the variables.

Data Pre-Processing for EDA

- **Missing Values:** Removed columns with >50% missing data; imputed remaining missing values with modes while finding outliers.
- **Column Filtering:** Eliminated irrelevant columns and retaining 36 columns , 307511 rows for analysis.
- **Identifying Data Imbalance:** Analysing data imbalance using displot and pie-chart.
- **Correlation Analysis:** Identified key columns influencing the target variable.
- **Final Dataframe:** Prepared for Exploratory Data Analysis (EDA).

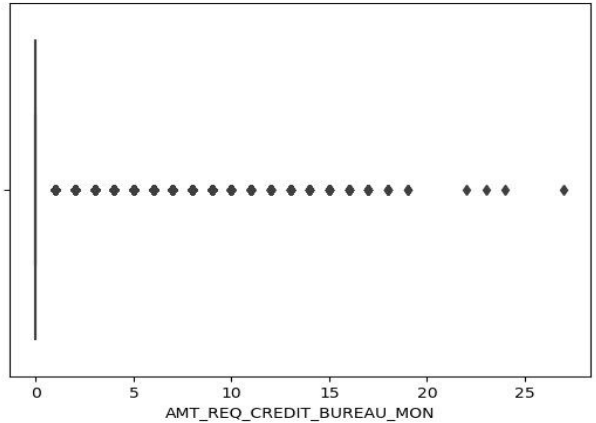
- Calculating Percentage of missing Values

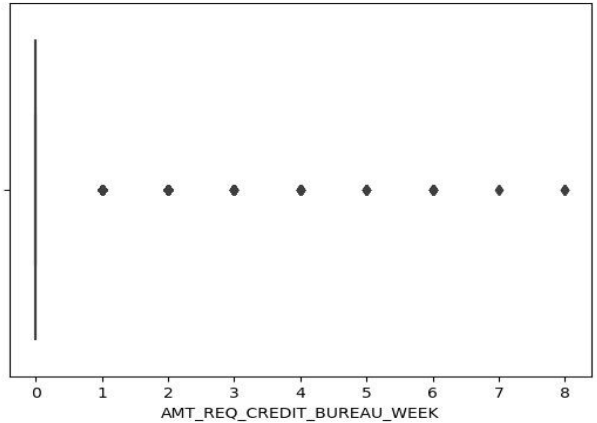


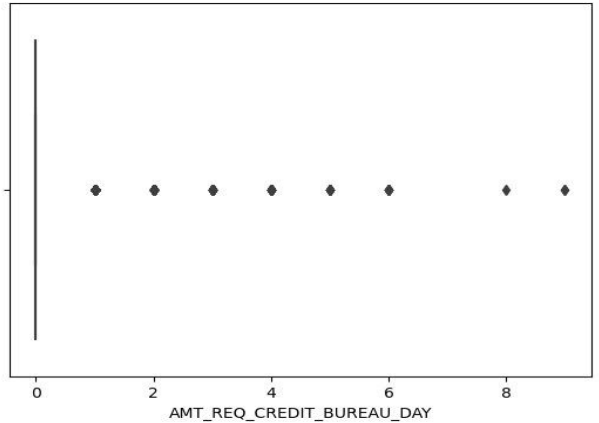


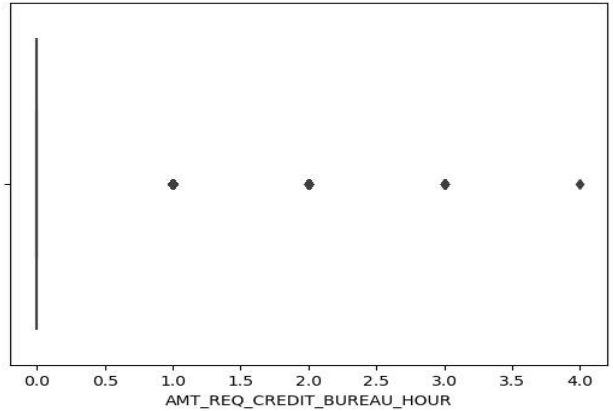
- Dropping columns which has more than 50% of missing values.
- Removing irrelevant columns.

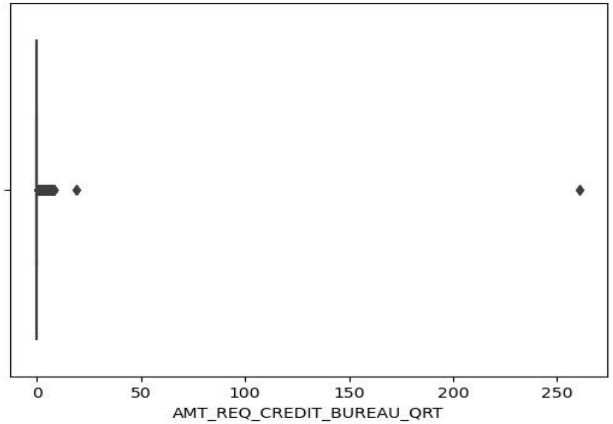
Outlier of Missing Data – Column wise

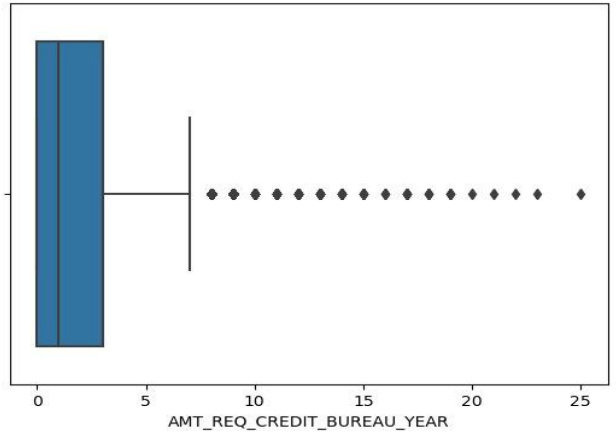




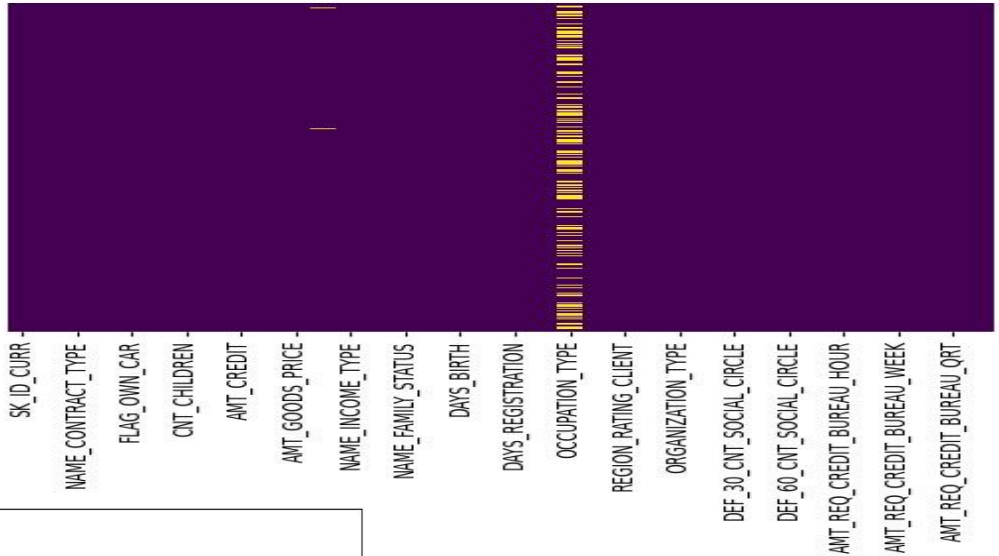


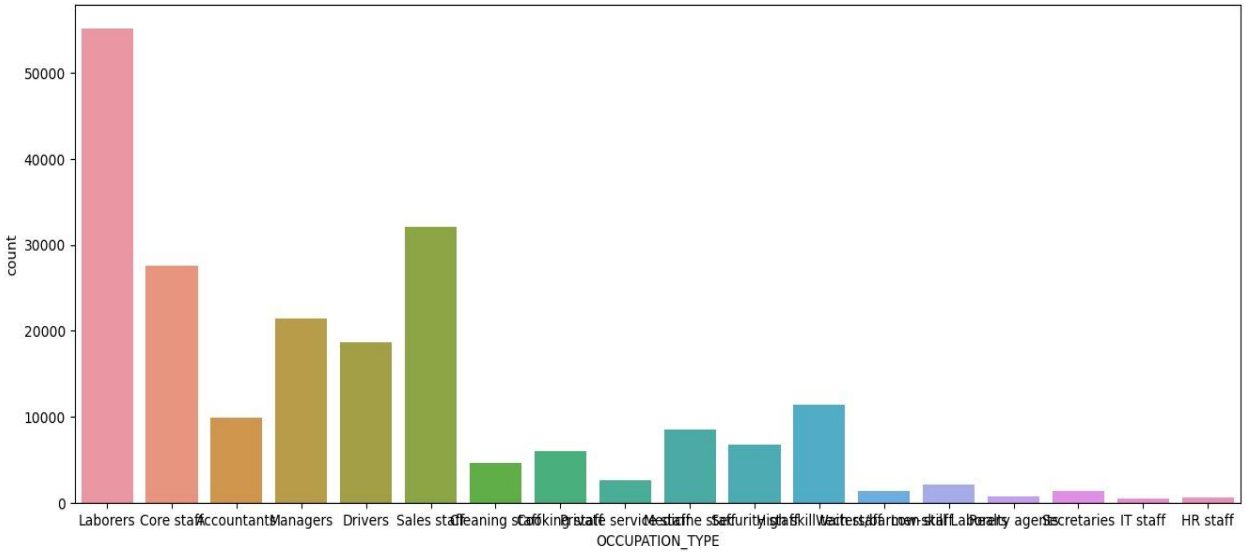




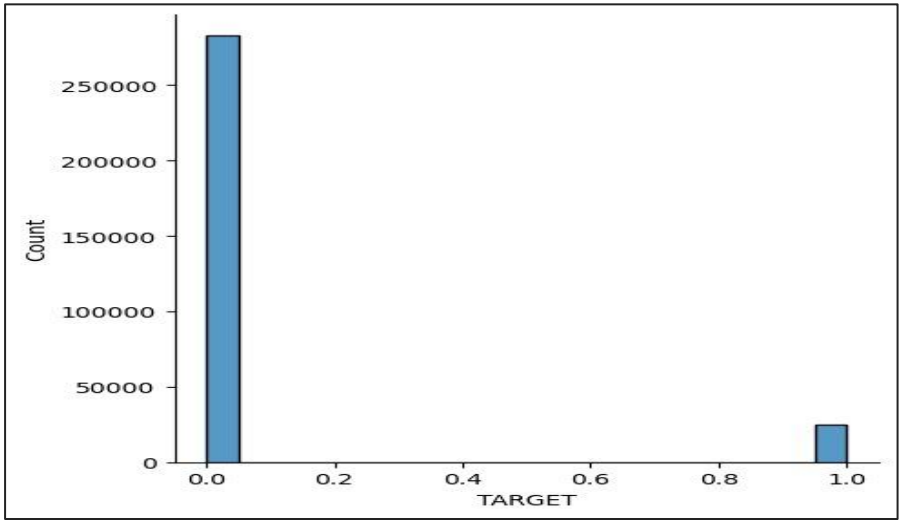


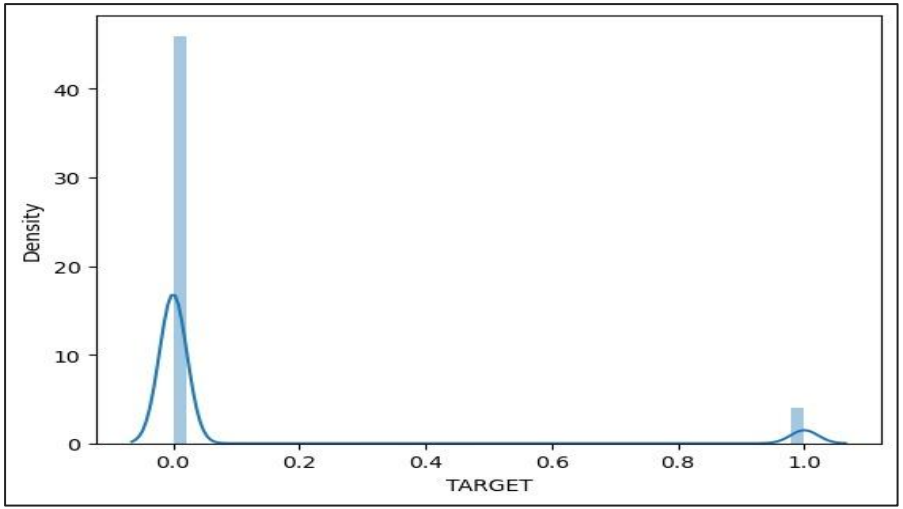
After performing Imputation





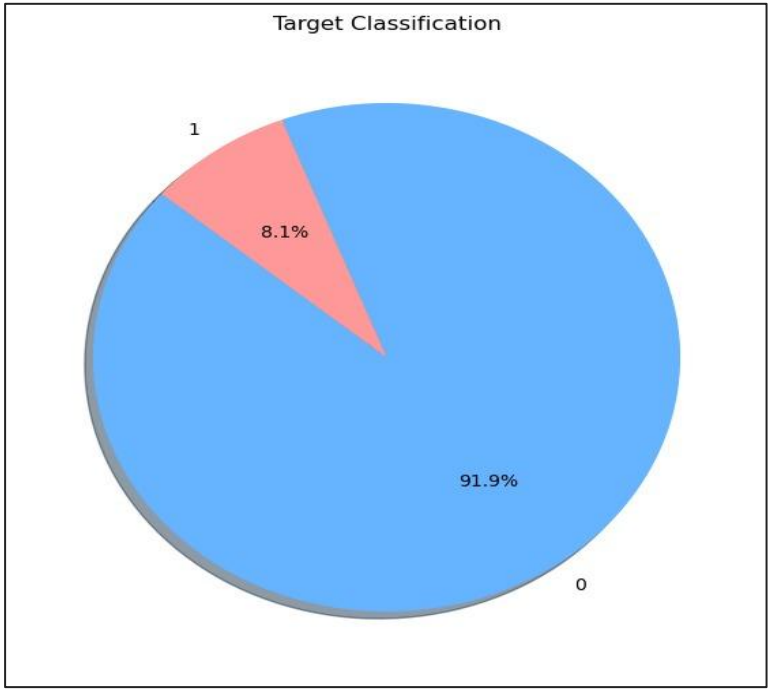
Bar Chart of Occupation Data





Checking for Data - Imbalance

Target Value

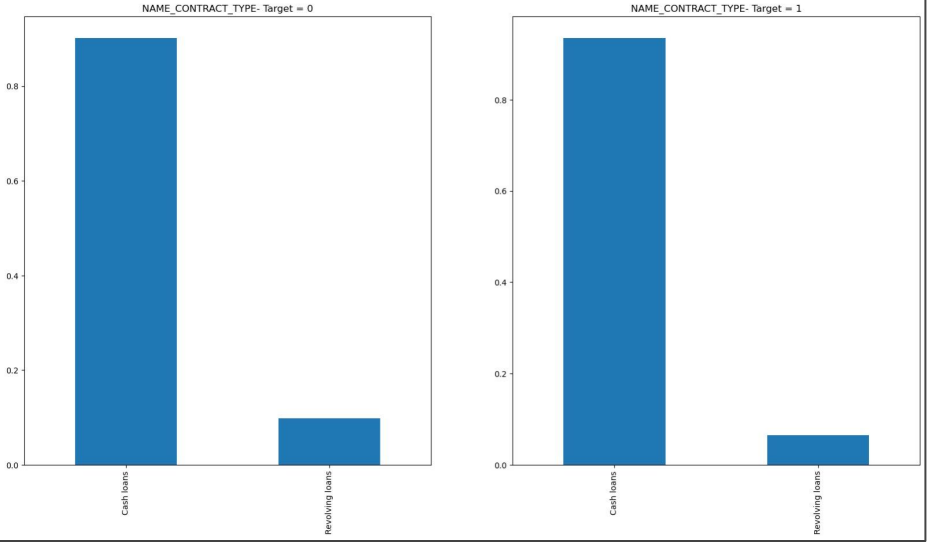


Exploratory Data Analysis(EDA)

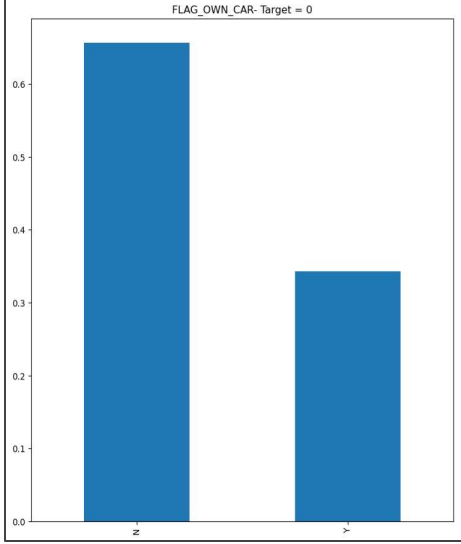
- Crucial step in data science projects.
- Analyzes and visualizes data to understand key characteristics.
- Identifies patterns, outliers, and relationships between variables.
- Conducted before formal statistical analyses or modeling.

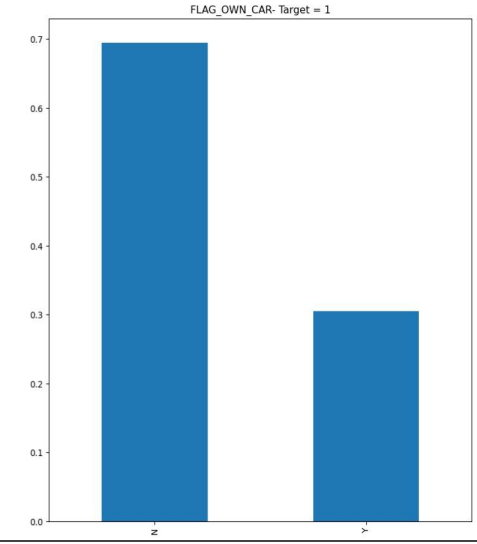
We carried out a thorough exploratory data analysis (EDA) to gain deeper insights into the patterns and relationships within the loan application data.

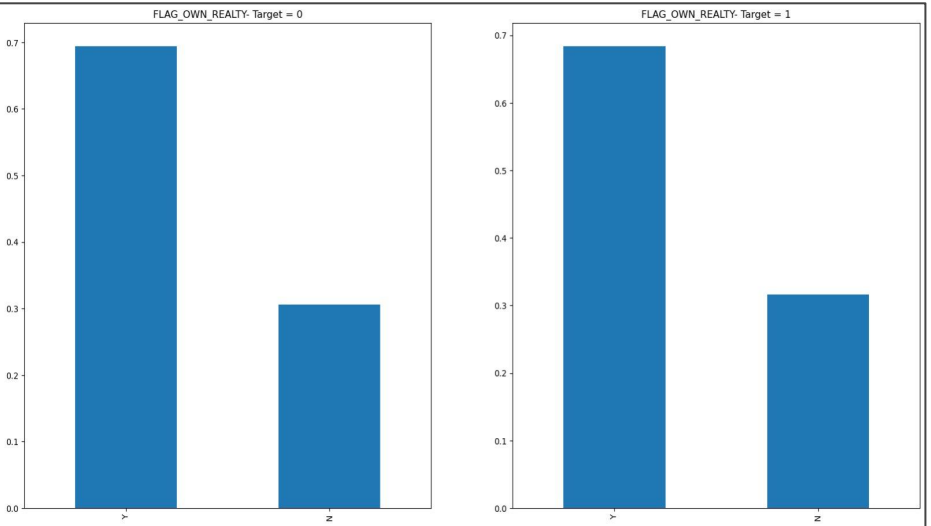
Now we will show Categorical column Univariate Analysis for Defaulters and Non-Defaulters for each Column


NAME_CONTRACT_TYPE- Target = 0
0.0
0.2
0.4
0.6
0.8
Cash loans
Revolving loans

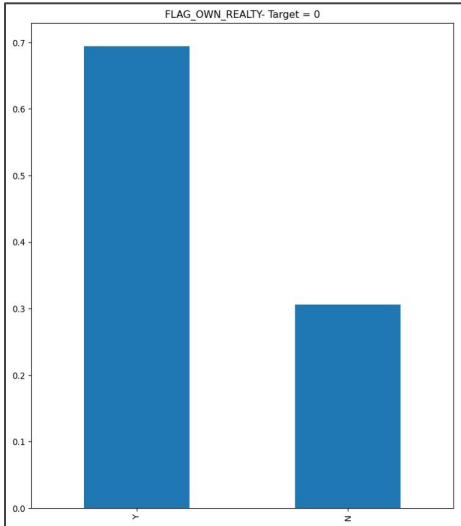
NAME_CONTRACT_TYPE- Target = 1
0.0
0.2
0.4
0.6
0.8
Cash loans
Revolving loans

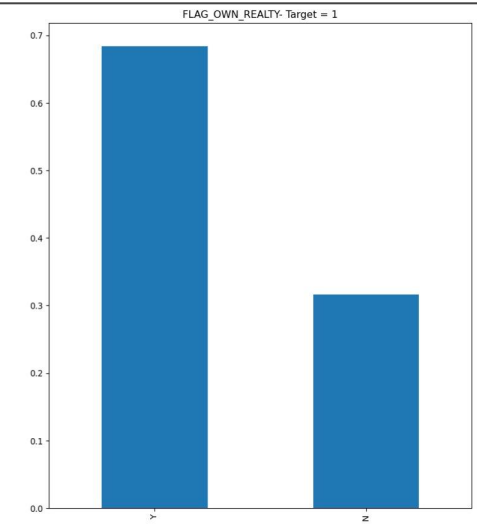

FLAG_OWN_CAR- Target = 0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
N
Y

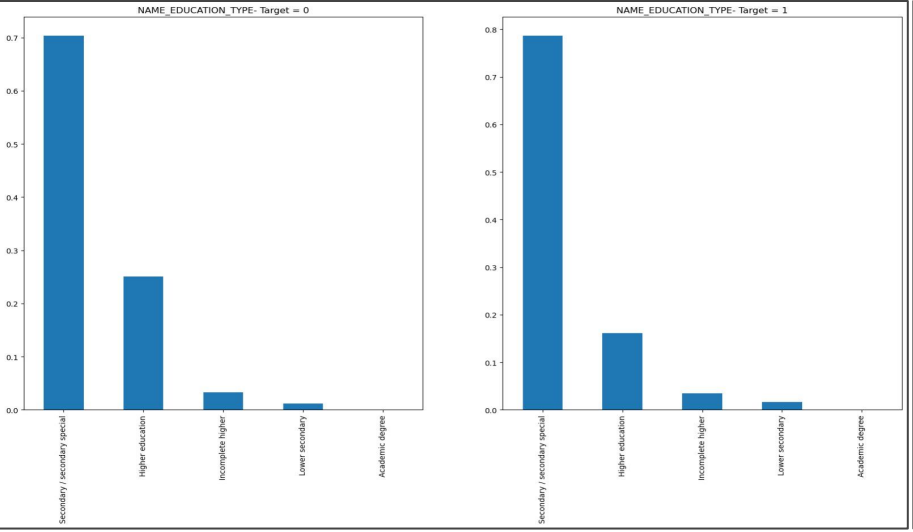

FLAG_OWN_CAR- Target = 1
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
N
Y


FLAG_OWN_REALTY- Target = 0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Y
N

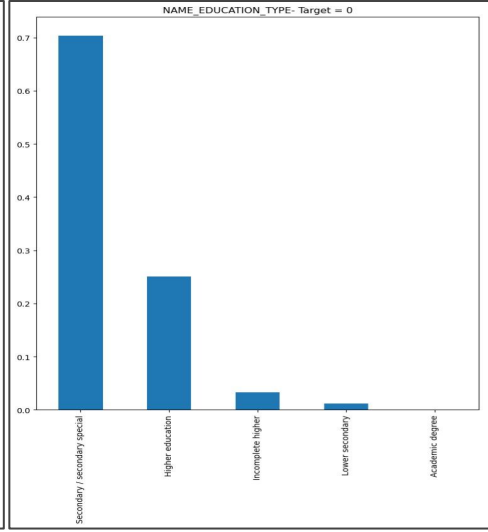
FLAG_OWN_REALTY- Target = 1
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Y
N

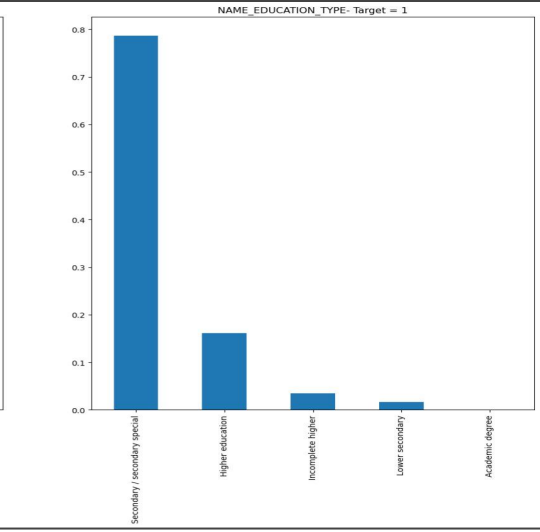

FLAG_OWN_REALTY- Target = 0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Y
N

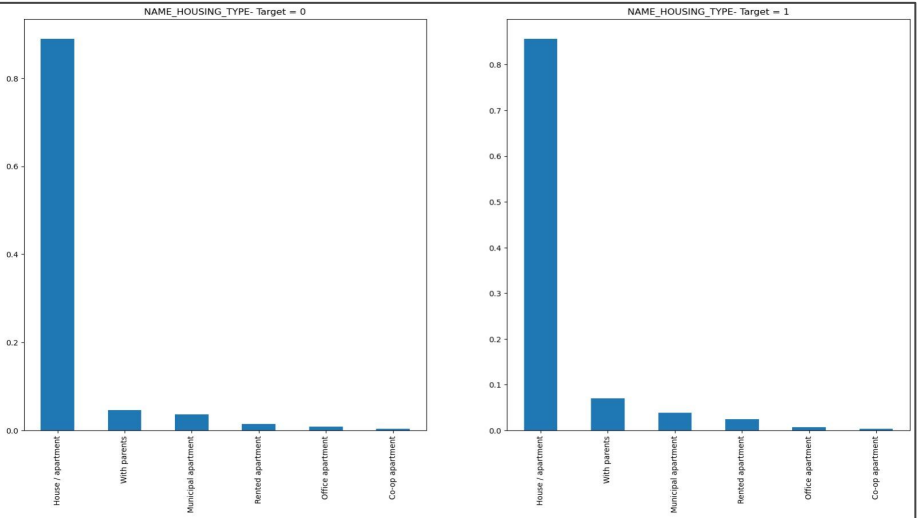

FLAG_OWN_REALTY- Target = 1
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Y
N


NAME_EDUCATION_TYPE- Target = 0
Secondary / secondary special
Higher education
Incomplete higher
Lower secondary
Academic degree

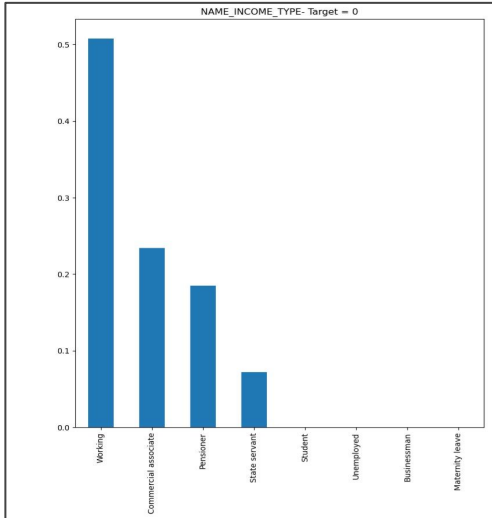
NAME_EDUCATION_TYPE- Target = 1
Secondary / secondary special
Higher education
Incomplete higher
Lower secondary
Academic degree

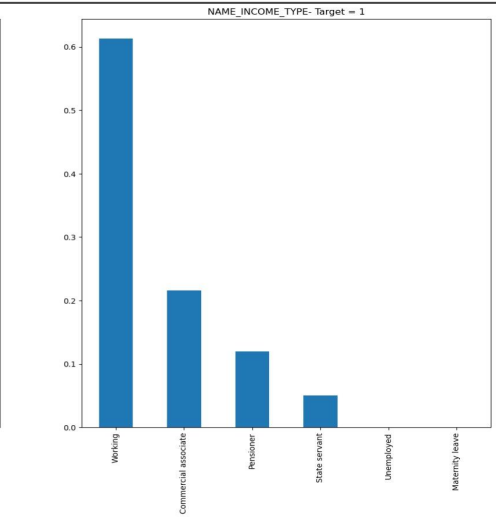

NAME_EDUCATION_TYPE- Target = 0
Secondary / secondary special
Higher education
Incomplete higher
Lower secondary
Academic degree

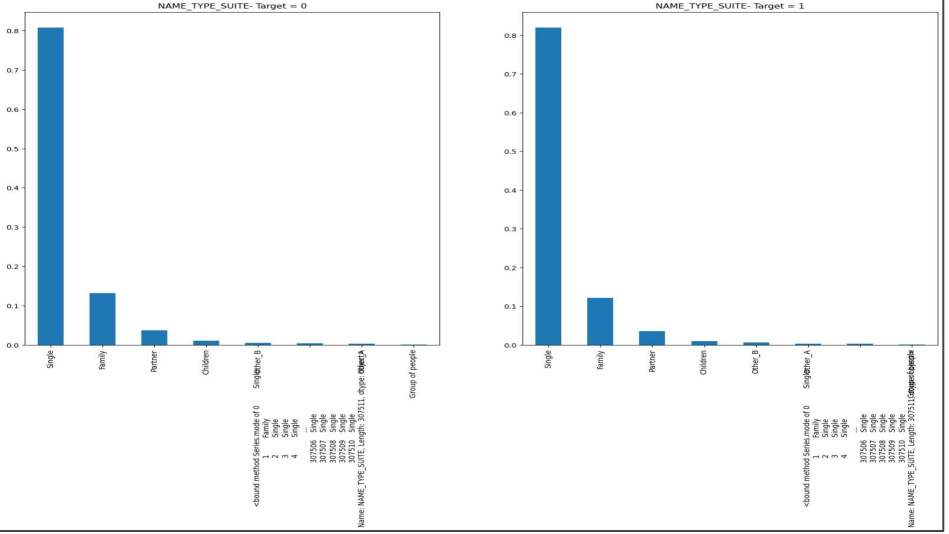

NAME_EDUCATION_TYPE- Target = 1
Secondary / secondary special
Higher education
Incomplete higher
Lower secondary
Academic degree


NAME_HOUSING_TYPE- Target = 0
House / apartment
With parents
Municipal apartment
Rented apartment
Office apartment
Co-op apartment

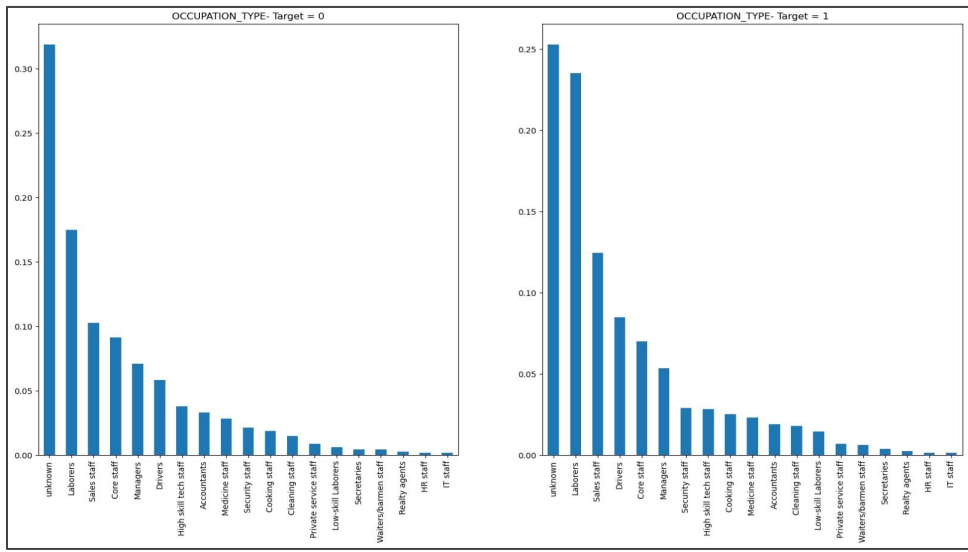
NAME_HOUSING_TYPE- Target = 1
House / apartment
With parents
Municipal apartment
Rented apartment
Office apartment
Co-op apartment


NAME_INCOME_TYPE- Target = 0
Working
Commercial associate
Pensioner
State servant
Student
Unemployed
Businessman
Maternity leave

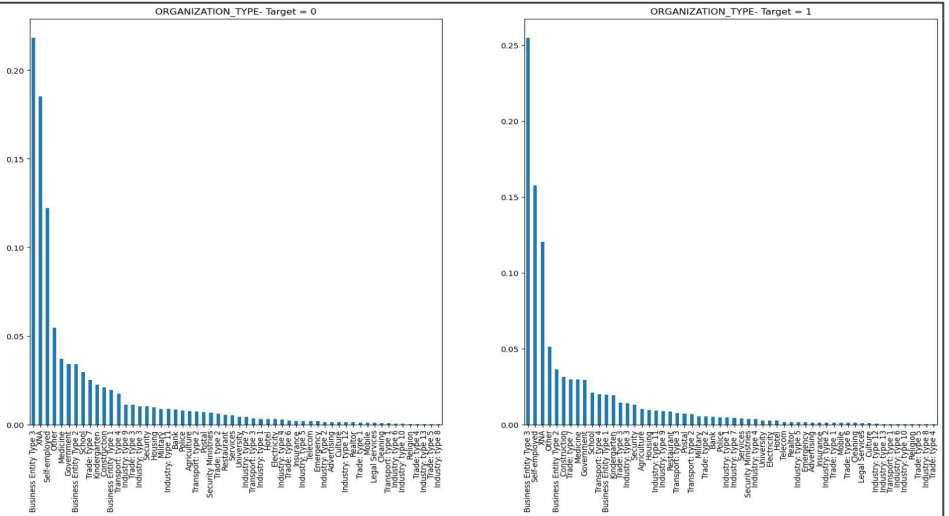

NAME_INCOME_TYPE- Target = 1
Working
Commercial associate
Pensioner
State servant
Unemployed
Maternity leave


NAME_TYPE_SUITE- Target = 0
Single
Family
Partner
Children
Other_B
Other_A
Group of people

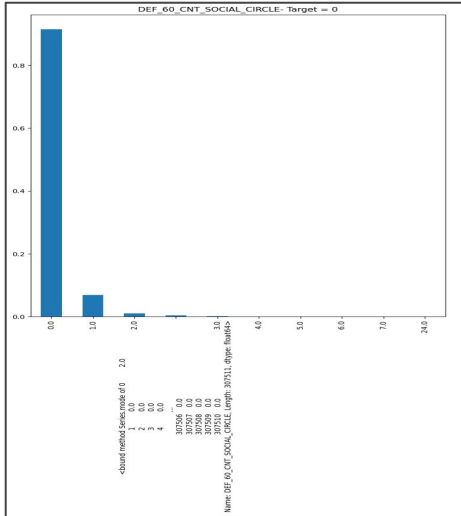
NAME_TYPE_SUITE- Target = 1
Single
Family
Partner
Children
Other_B
Other_A
Group of people

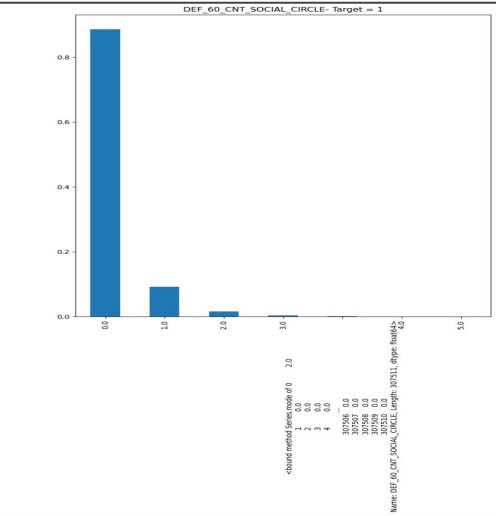

OCCUPATION_TYPE- Target = 0
unknown
Laborers
Sales staff
Core staff
Managers
Drivers
High skill tech staff
Accountants
Medicine staff
Security staff
Cooking staff
Cleaning staff
Private service staff
Low-skill Laborers
Secretaries
Waiters/barmen staff
Realty agents
HR staff
IT staff

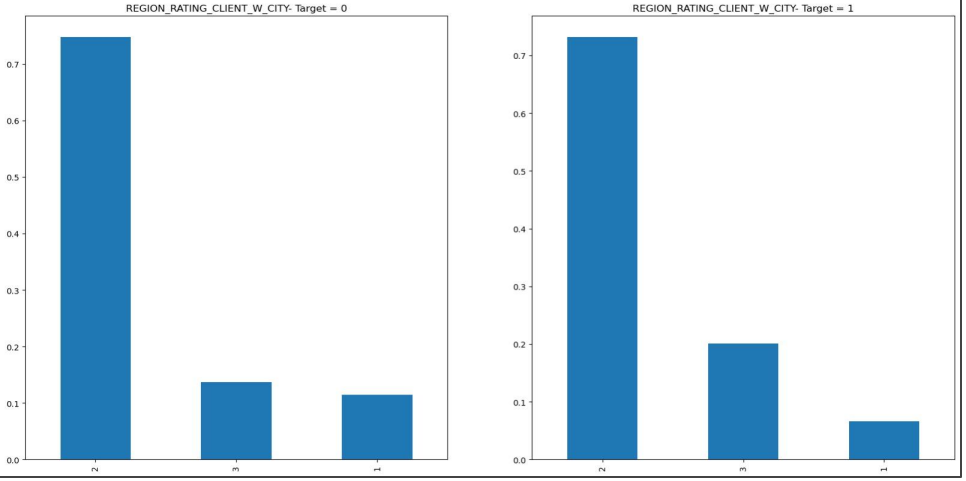
OCCUPATION_TYPE- Target = 1
unknown
Laborers
Sales staff
Drivers
Core staff
Managers
Security staff
High skill tech staff
Cooking staff
Medicine staff
Accountants
Cleaning staff
Low-skill Laborers
Private service staff
Waiters/barmen staff
Secretaries
Realty agents
HR staff
IT staff


ORGANIZATION_TYPE- Target = 0
Business Entity Type 3
Self-employed
Other
Medicine
Government
Business Entity Type 2
Trade: type 7
Kindergarten
Construction
Transport: type 4
Trade: type 1
Industry: type 3
Security
Housing
Industry: type 11
Military
Bank
Agriculture
Police
Transport: type 2
Postal
Security Ministries
Trade: type 2
Mobile
Legal Services
Transport: type 1
Industry: type 9
Industry: type 5
Insurance
Industry: type 7
Industry: type 1
Electricity
Industry: type 6
Trade: type 6
Industry: type 12
Hotel
Telecom
Emergency
Advertising
Realtor
Culture
Industry: type 10
Religion
Industry: type 13
Trade: type 5
Industry: type 8

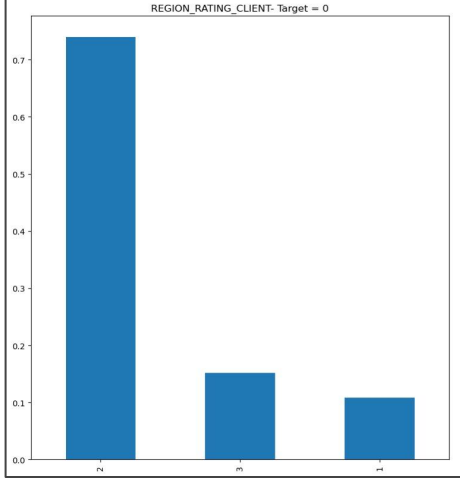
ORGANIZATION_TYPE- Target = 1
Business Entity Type 3
Self-employed
Other
Business Entity Type 2
Construction
Government
Medicine
Kindergarten
School
Transport: type 4
Business Entity Type 1
Trade: type 7
Industry: type 3
Security
Housing
Agriculture
Military
Police
Bank
Postal
Security Ministries
Industry: type 9
Industry: type 11
Transport: type 3
Industry: type 4
Trade: type 2
Mobile
Industry: type 7
Industry: type 1
Transport: type 1
Hotel
Electricity
Industry: type 6
Telecom
Trade: type 6
Industry: type 12
Legal Services
Emergency
Advertising
Realtor
Culture
Industry: type 10
Religion
Industry: type 13
Trade: type 3
Trade: type 4
Trade: type 5
Industry: type 8

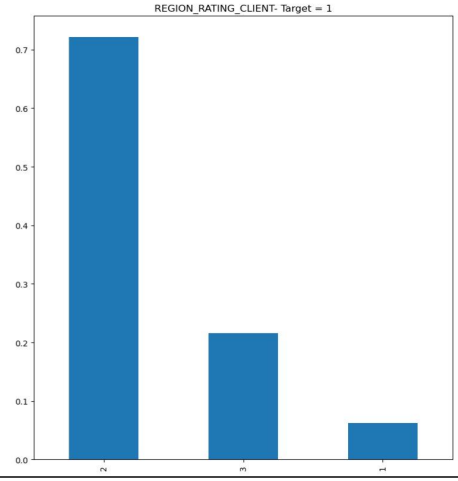

DEF_60_CNT_SOCIAL_CIRCLE- Target = 0
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
24.0

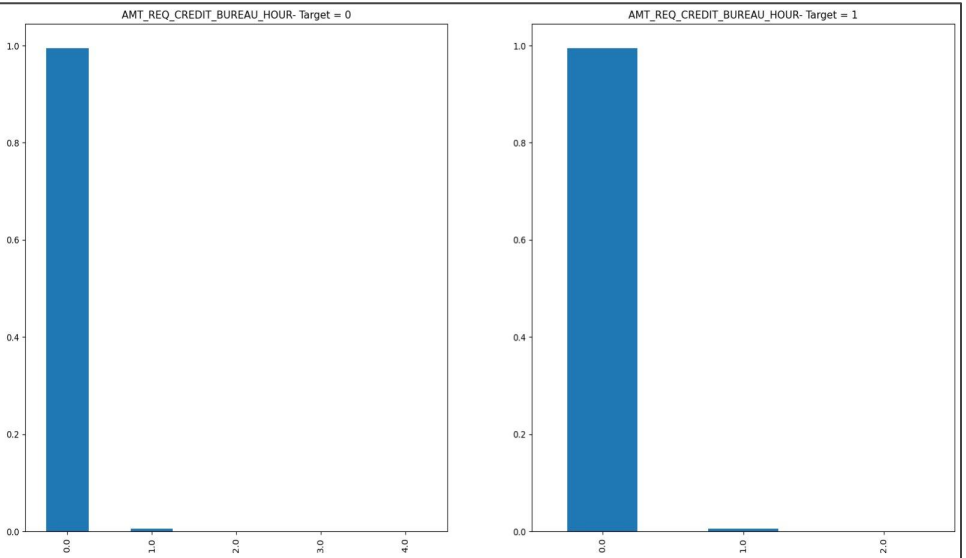

DEF_60_CNT_SOCIAL_CIRCLE- Target = 1
0.0
1.0
2.0
3.0
4.0
5.0


REGION_RATING_CLIENT_W_CITY- Target = 0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
2
3
1

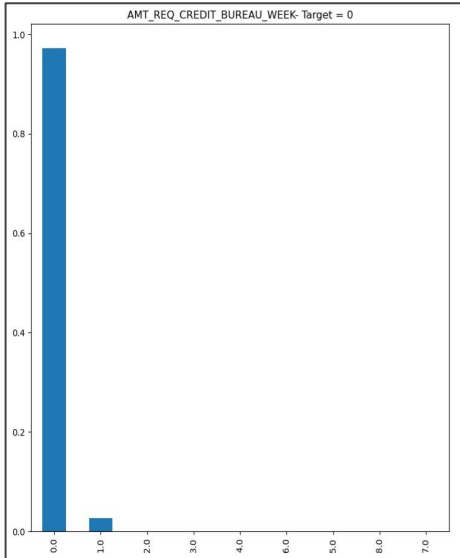
REGION_RATING_CLIENT_W_CITY- Target = 1
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
2
3
1

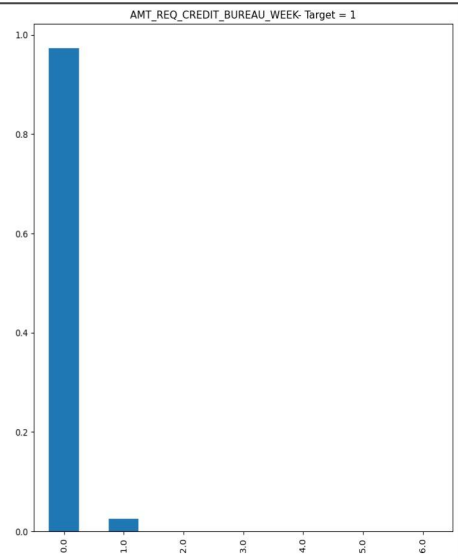

REGION_RATING_CLIENT- Target = 0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
2
3
1

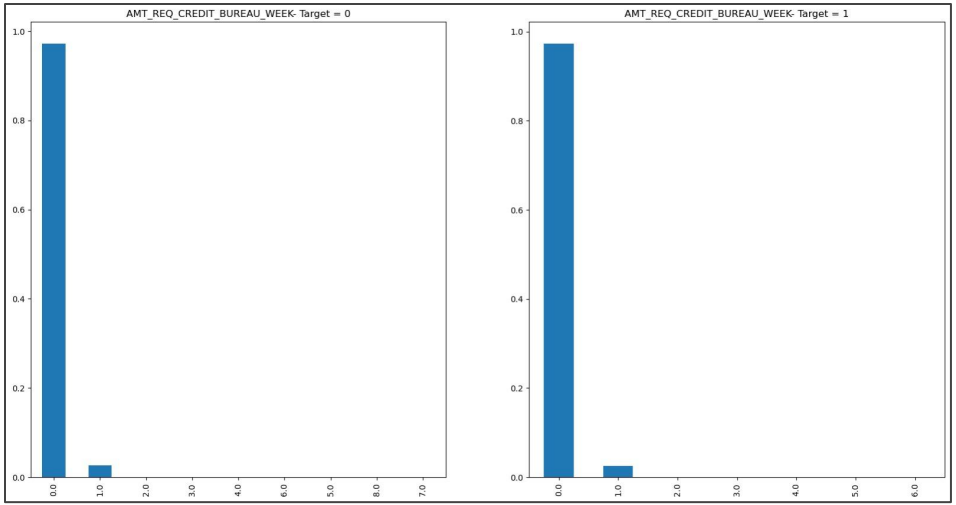

REGION_RATING_CLIENT- Target = 1
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
2
3
1

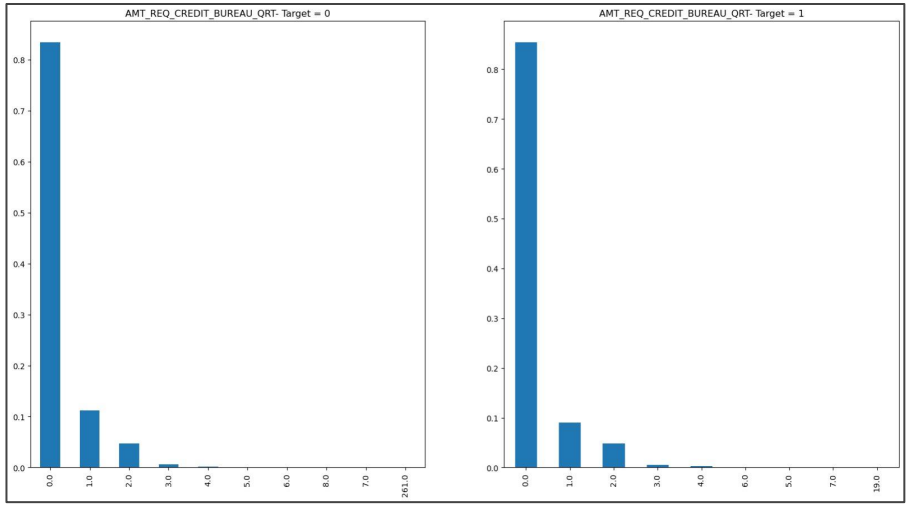

AMT_REQ_CREDIT_BUREAU_HOUR- Target = 0
0.0
0.2
0.4
0.6
0.8
1.0
0.0
1.0
2.0
3.0
4.0

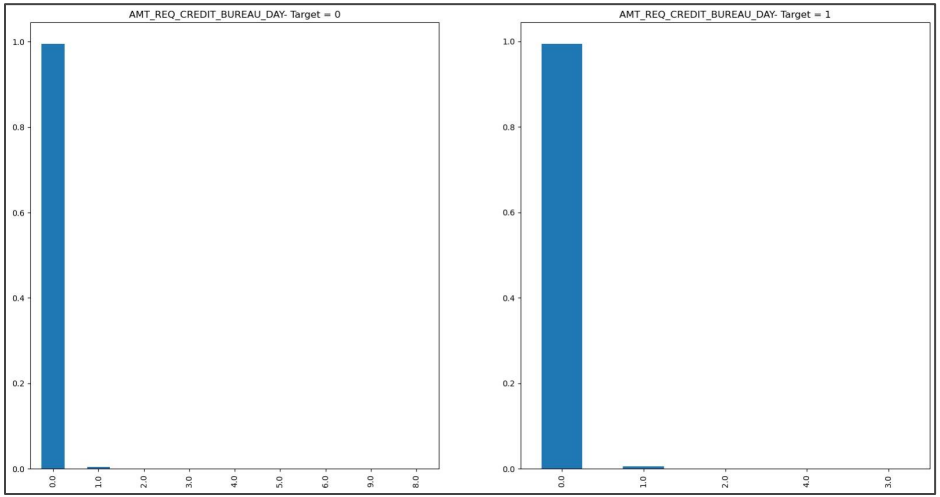
AMT_REQ_CREDIT_BUREAU_HOUR- Target = 1
0.0
0.2
0.4
0.6
0.8
1.0
0.0
1.0
2.0

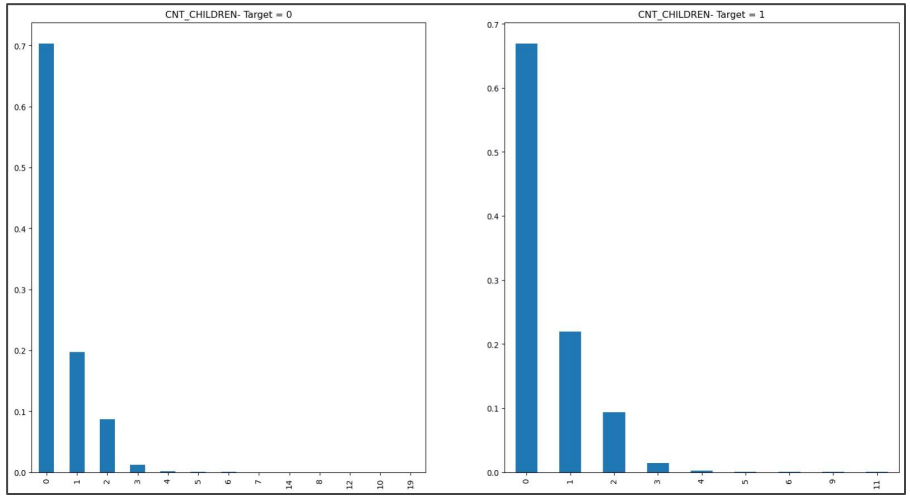

AMT_REQ_CREDIT_BUREAU_WEEK- Target = 0
0.0
0.2
0.4
0.6
0.8
1.0
0.0
1.0
2.0
3.0
4.0
6.0
5.0
8.0
7.0

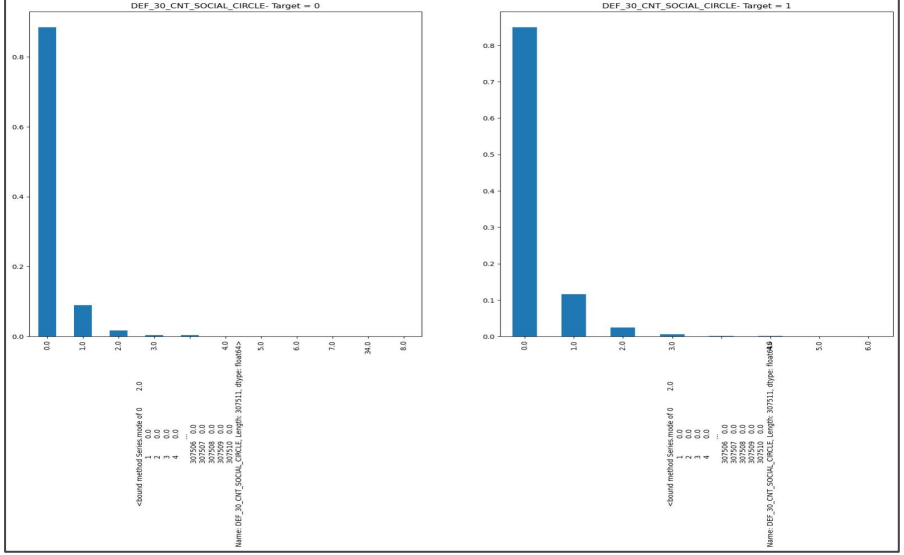

AMT_REQ_CREDIT_BUREAU_WEEK- Target = 1
0.0
0.2
0.4
0.6
0.8
1.0
0.0
1.0
2.0
3.0
4.0
5.0
6.0


AMT_REQ_CREDIT_BUREAU_WEEK- Target = 0
AMT_REQ_CREDIT_BUREAU_WEEK- Target = 1

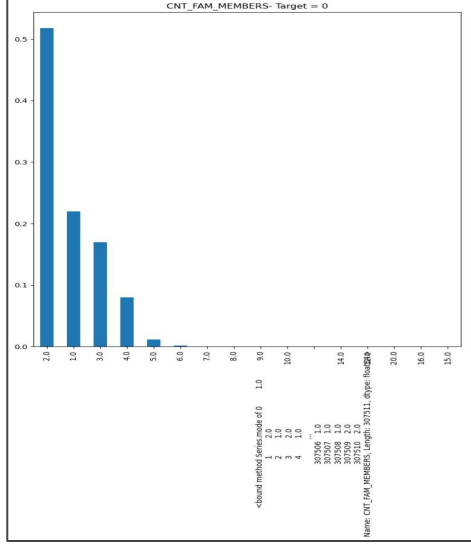

AMT_REQ_CREDIT_BUREAU_QRT- Target = 0
AMT_REQ_CREDIT_BUREAU_QRT- Target = 1

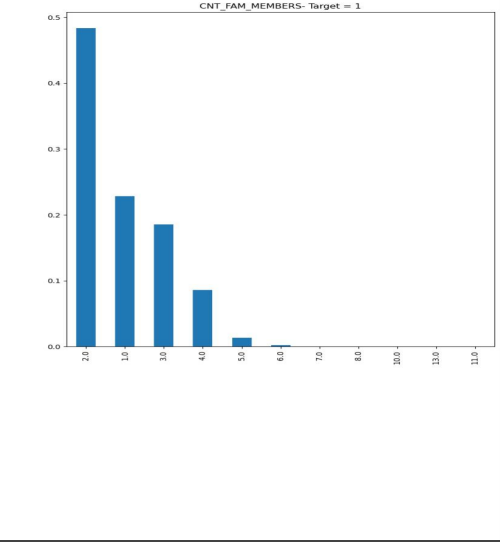

AMT_REQ_CREDIT_BUREAU_DAY- Target = 0
AMT_REQ_CREDIT_BUREAU_DAY- Target = 1

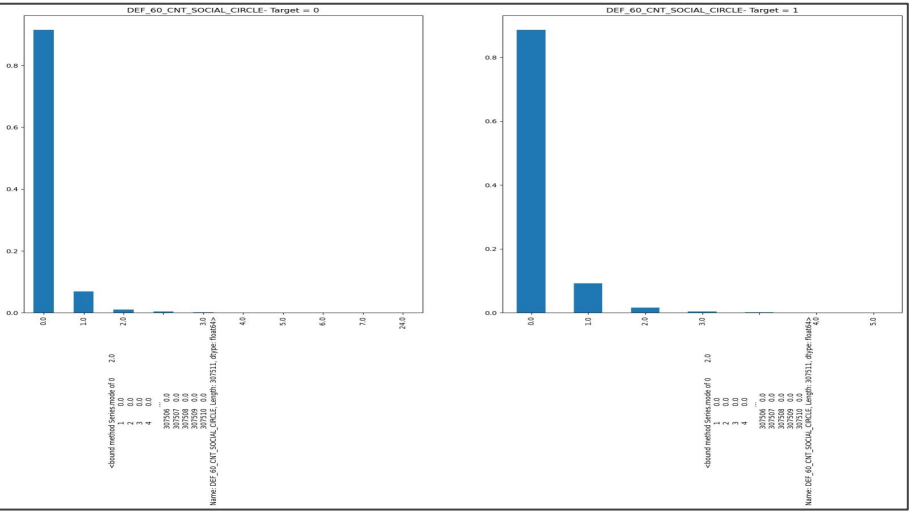

CNT_CHILDREN- Target = 0
CNT_CHILDREN- Target = 1


DEF_30_CNT_SOCIAL_CIRCLE- Target = 0

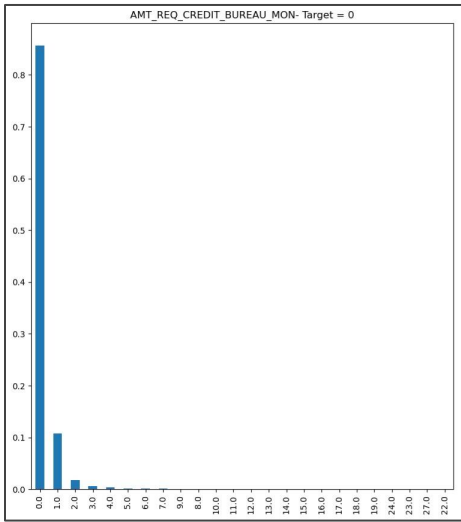
DEF_30_CNT_SOCIAL_CIRCLE- Target = 1

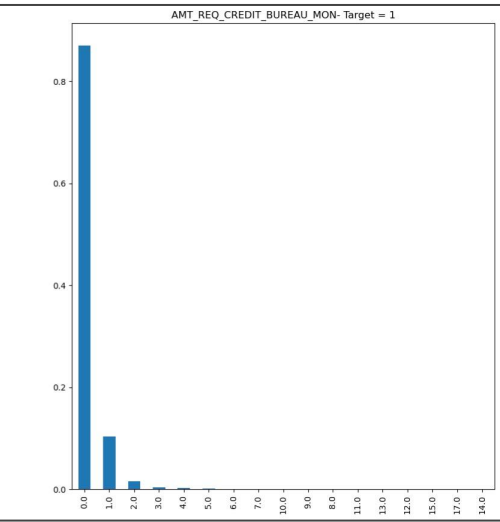

CNT_FAM_MEMBERS- Target = 0

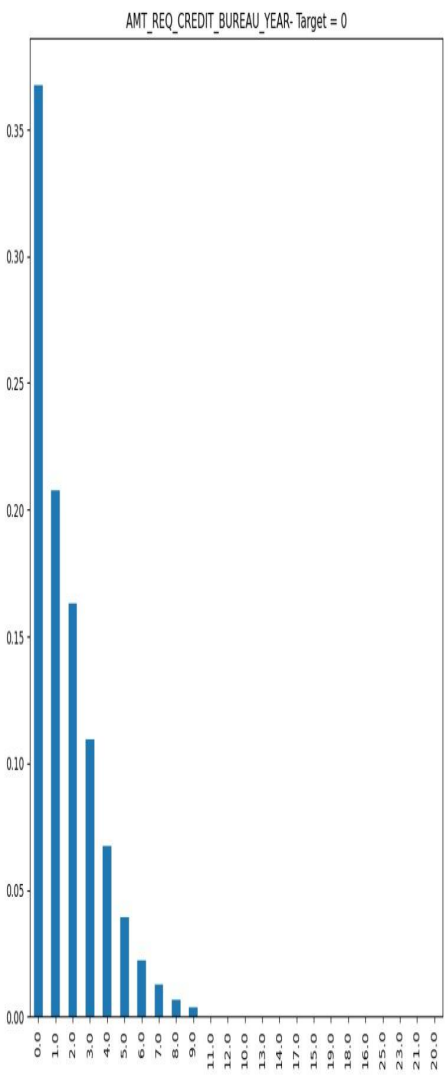

CNT_FAM_MEMBERS- Target = 1

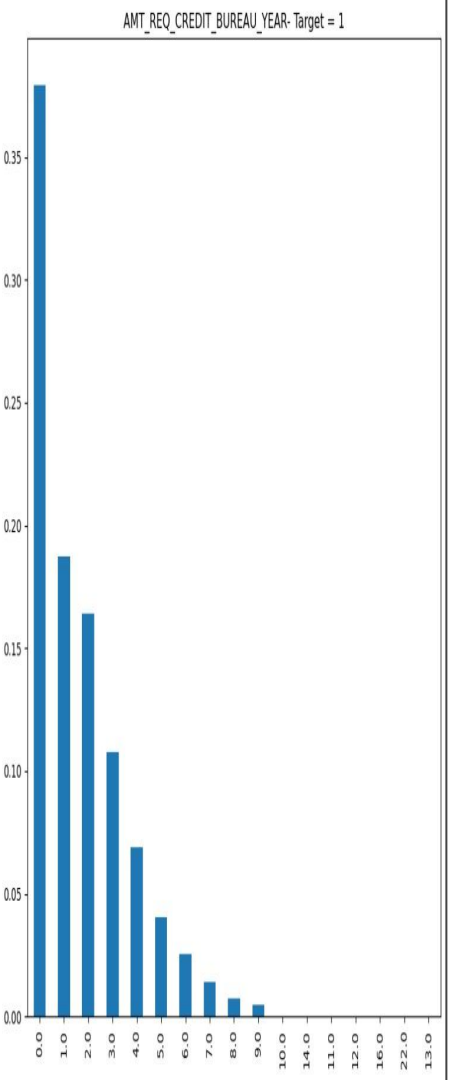

DEF_60_CNT_SOCIAL_CIRCLE- Target = 0

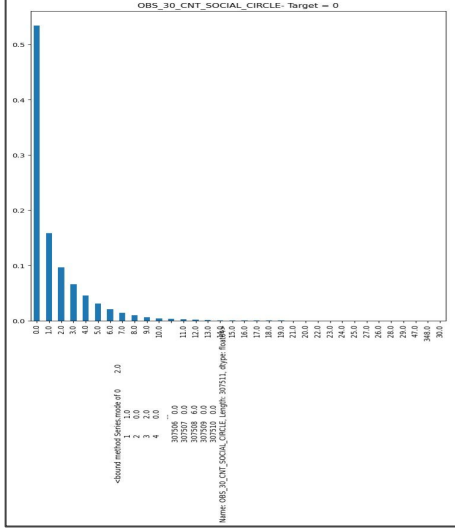
DEF_60_CNT_SOCIAL_CIRCLE- Target = 1

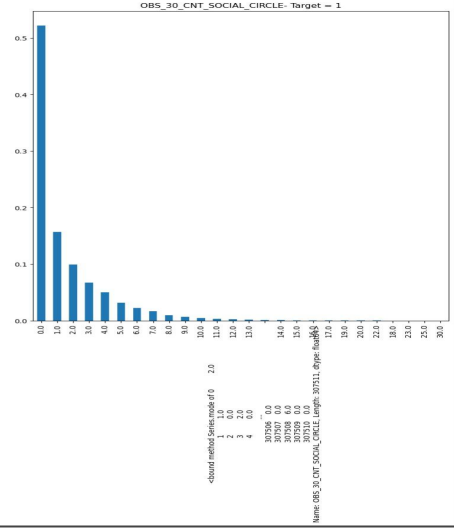

AMT_REQ_CREDIT_BUREAU_MON- Target = 0

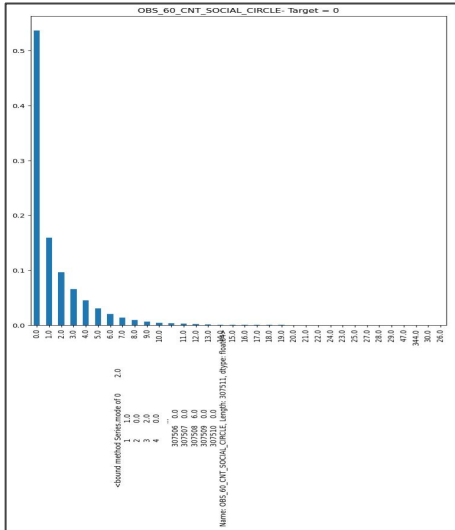

AMT_REQ_CREDIT_BUREAU_MON- Target = 1

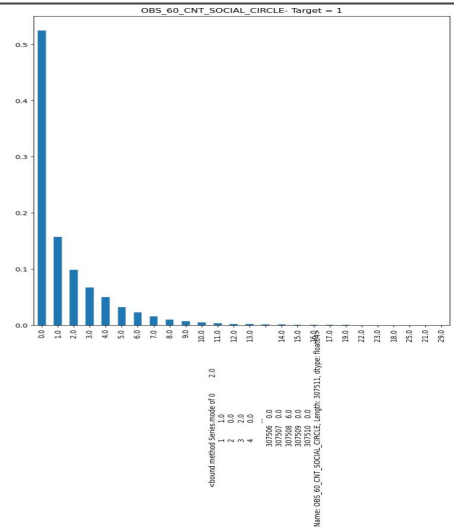

AMT_REQ_CREDIT_BUREAU_YEAR- Target = 0


AMT_REQ_CREDIT_BUREAU_YEAR- Target = 1


OBS_30_CNT_SOCIAL_CIRCLE- Target = 0

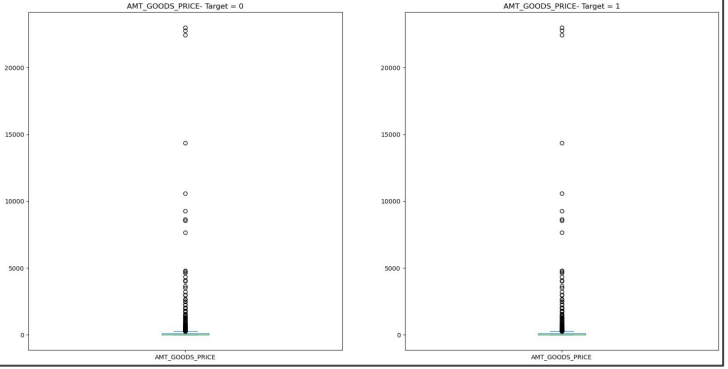

OBS_30_CNT_SOCIAL_CIRCLE- Target = 1


OBS_60_CNT_SOCIAL_CIRCLE- Target = 0

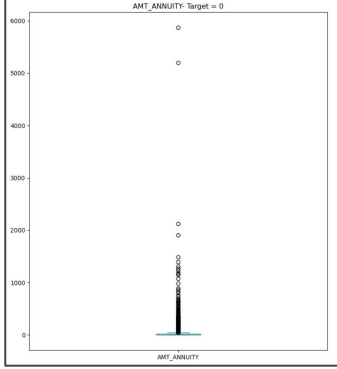

OBS_60_CNT_SOCIAL_CIRCLE- Target = 1

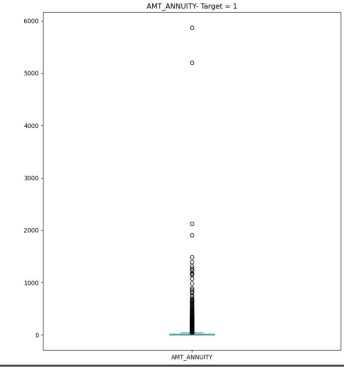
Box Plots

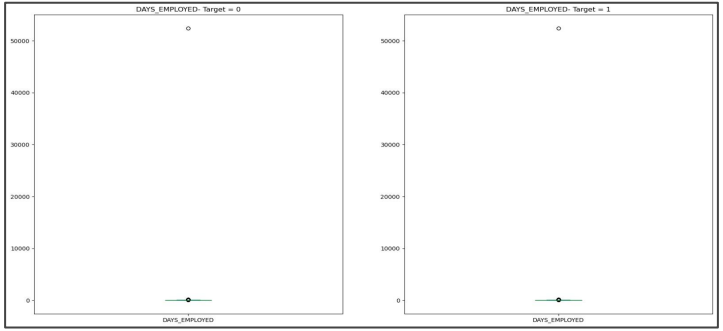
Univariate Analysis for Numerical Columns

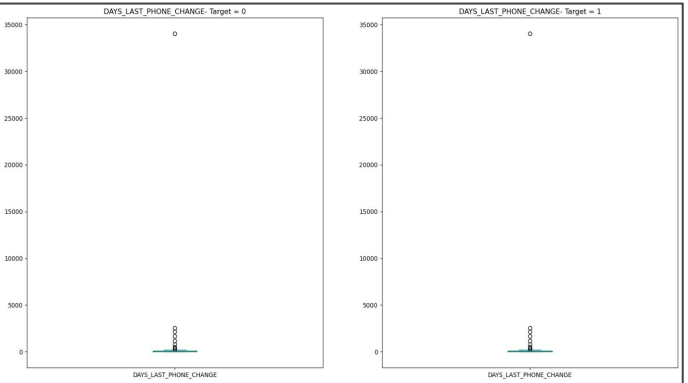

AMT_GOODS_PRICE- Target = 0
AMT_GOODS_PRICE

AMT_GOODS_PRICE- Target = 1
AMT_GOODS_PRICE

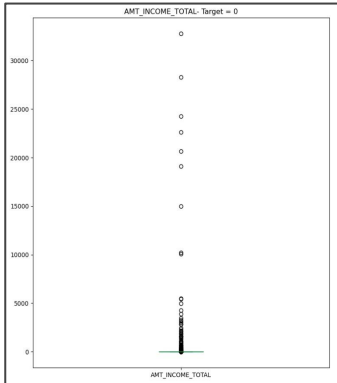

AMT_ANNUITY- Target = 0
AMT_ANNUITY

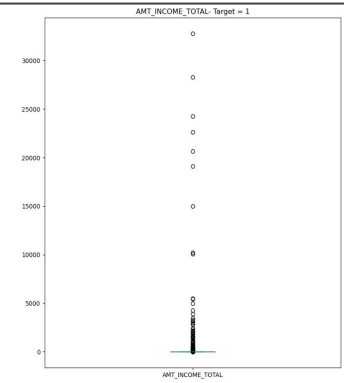

AMT_ANNUITY- Target = 1
AMT_ANNUITY

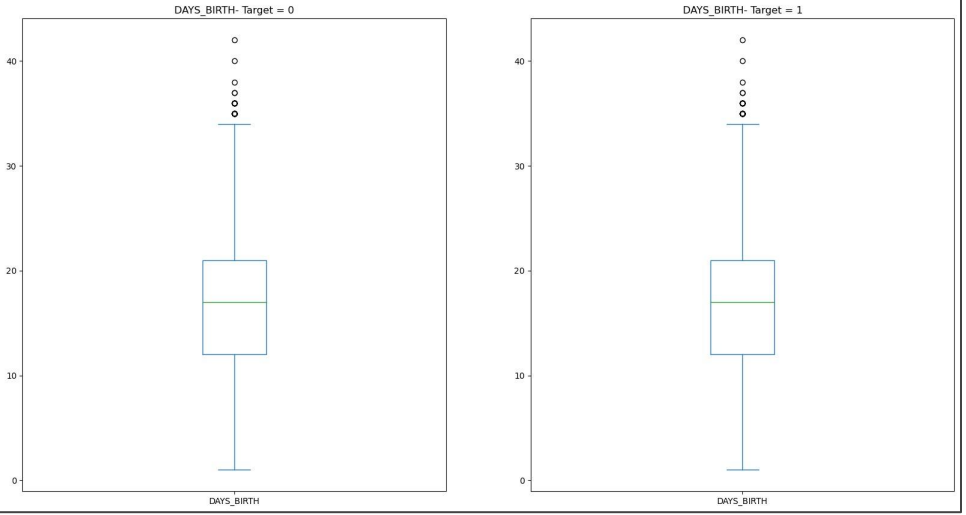

DAYS_EMPLOYED- Target = 0
DAYS_EMPLOYED
DAYS_EMPLOYED- Target = 1
DAYS_EMPLOYED


DAYS_LAST_PHONE_CHANGE- Target = 0
DAYS_LAST_PHONE_CHANGE

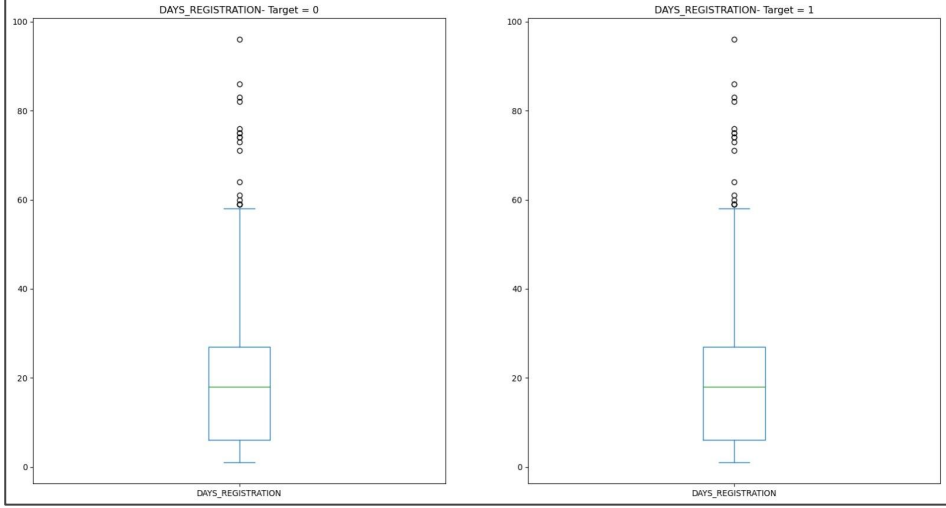
DAYS_LAST_PHONE_CHANGE- Target = 1
DAYS_LAST_PHONE_CHANGE


AMT_INCOME_TOTAL- Target = 0
AMT_INCOME_TOTAL

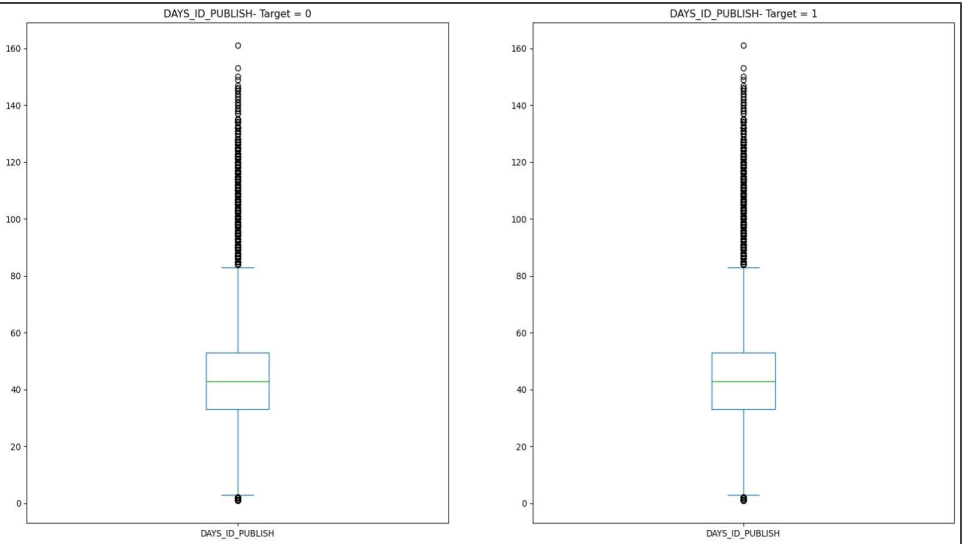

AMT_INCOME_TOTAL- Target = 1
AMT_INCOME_TOTAL


DAYS_BIRTH- Target = 0
DAYS_BIRTH

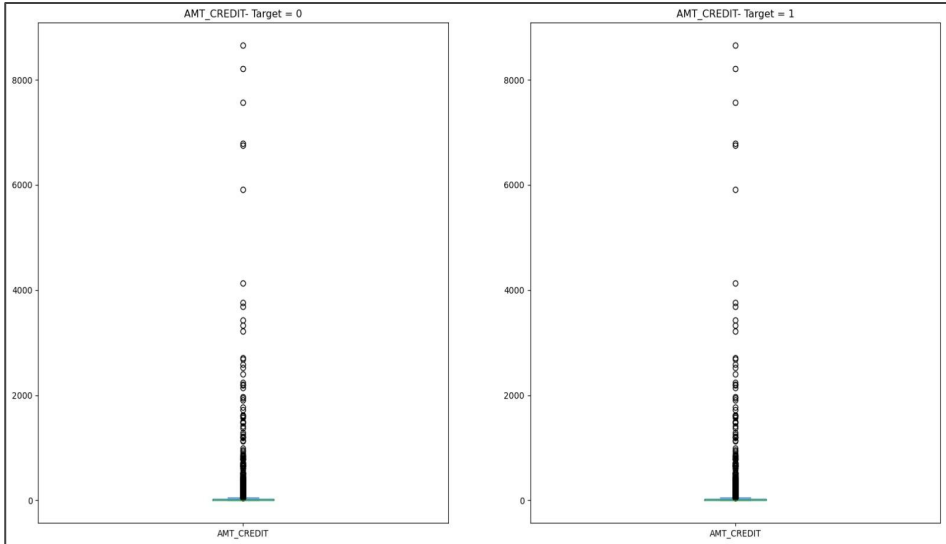
DAYS_BIRTH- Target = 1
DAYS_BIRTH


DAYS_REGISTRATION- Target = 0
DAYS_REGISTRATION

DAYS_REGISTRATION- Target = 1
DAYS_REGISTRATION


DAYS_ID_PUBLISH- Target = 0
DAYS_ID_PUBLISH

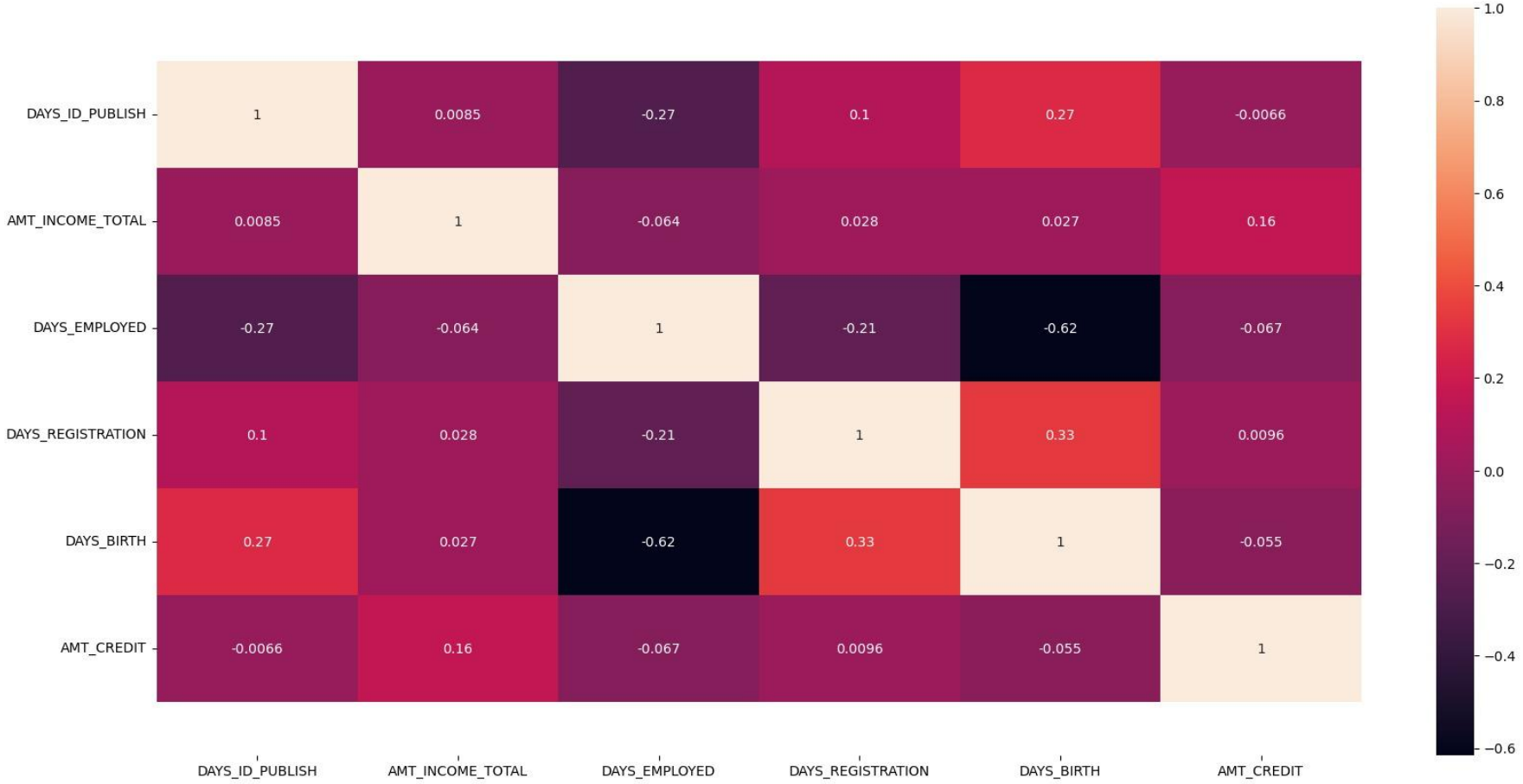
DAYS_ID_PUBLISH- Target = 1
DAYS_ID_PUBLISH


AMT_CREDIT- Target = 0
AMT_CREDIT

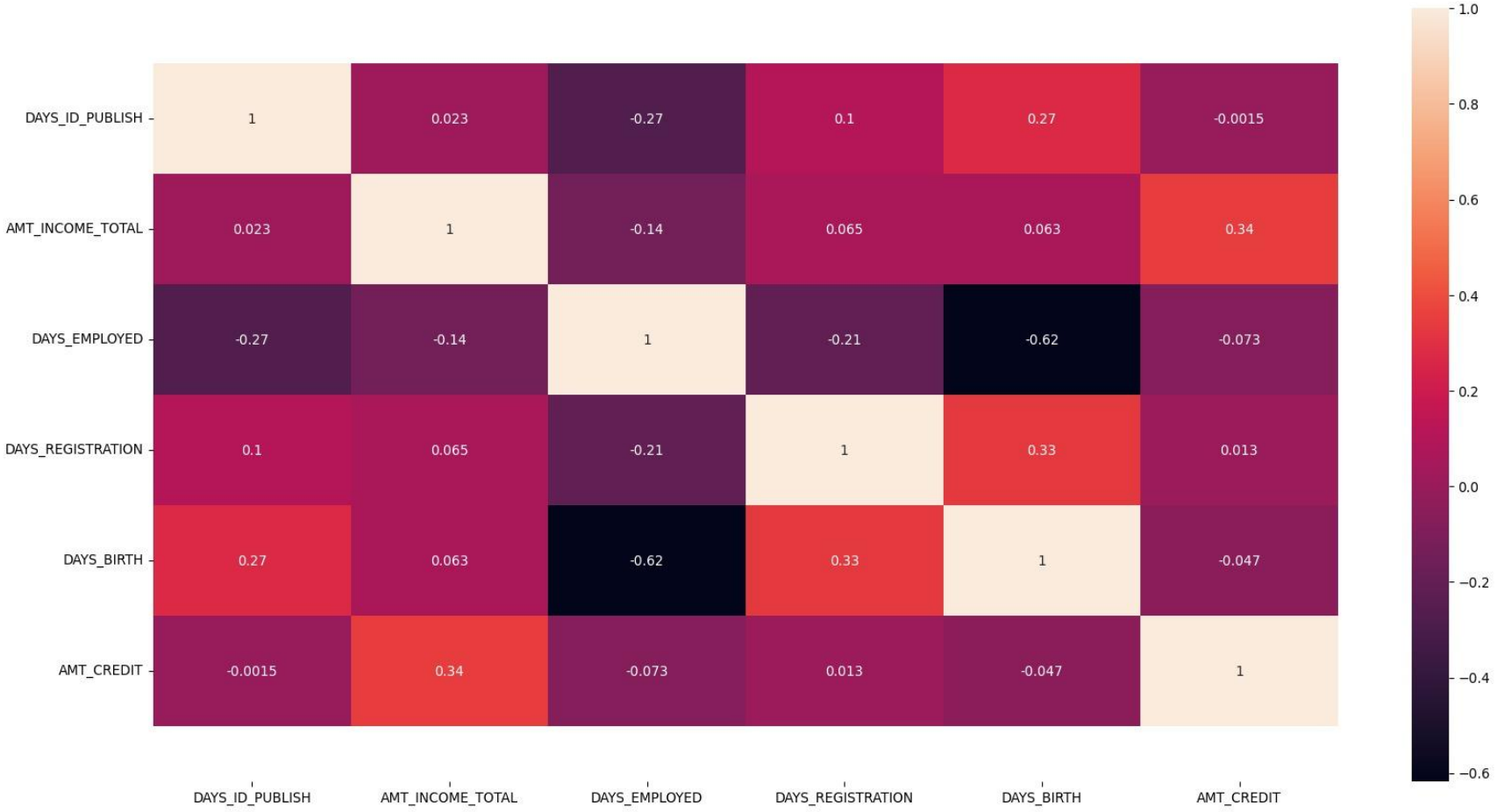
AMT_CREDIT- Target = 1
AMT_CREDIT

Correlation Analysis

Correlation Analysis for Numerical Data with Target Column



Correlation Analysis of Non-Defaulters with Numerical Column



Correlation Analysis of Clients with Payment Difficulties

	Dimension1	Dimension2	Correlation	Correlation_abs
179	REGION_RATING_CLIENT_W_CITY	REGION_RATING_CLIENT	0.95	0.95
107	DAYS_EMPLOYED	DAYS_BIRTH	-0.62	0.62
71	AMT_CREDIT	AMT_INCOME_TOTAL	0.34	0.34
87	DAYS_BIRTH	CNT_CHILDREN	0.34	0.34
124	DAYS_REGISTRATION	DAYS_BIRTH	0.33	0.33
141	DAYS_ID_PUBLISH	DAYS_BIRTH	0.27	0.27
142	DAYS_ID_PUBLISH	DAYS_EMPLOYED	-0.27	0.27
104	DAYS_EMPLOYED	CNT_CHILDREN	-0.24	0.24
215	AMT_REQ_CREDIT_BUREAU_DAY	AMT_REQ_CREDIT_BUREAU_HOUR	0.23	0.23
233	AMT_REQ_CREDIT_BUREAU_WEEK	AMT_REQ_CREDIT_BUREAU_DAY	0.22	0.22

Top 10 highly correlated dimensions for non defaulter

	Dimension1	Dimension2	Correlation	Correlation_abs
179	REGION_RATING_CLIENT_W_CITY	REGION_RATING_CLIENT	0.96	0.96
107	DAYS_EMPLOYED	DAYS_BIRTH	-0.58	0.58
124	DAYS_REGISTRATION	DAYS_BIRTH	0.29	0.29
87	DAYS_BIRTH	CNT_CHILDREN	0.26	0.26
215	AMT_REQ_CREDIT_BUREAU_DAY	AMT_REQ_CREDIT_BUREAU_HOUR	0.25	0.25
141	DAYS_ID_PUBLISH	DAYS_BIRTH	0.25	0.25
142	DAYS_ID_PUBLISH	DAYS_EMPLOYED	-0.23	0.23
104	DAYS_EMPLOYED	CNT_CHILDREN	-0.19	0.19
233	AMT_REQ_CREDIT_BUREAU_WEEK	AMT_REQ_CREDIT_BUREAU_DAY	0.19	0.19
125	DAYS_REGISTRATION	DAYS_EMPLOYED	-0.19	0.19

Top 10 highly correlated dimensions for defaulter, clients with payment difficulties

Risk Mitigation Strategies

Loan Approval Guidelines

Establish stricter criteria by prioritizing stable income and strong credit history for high-risk borrowers.

Dynamic Interest Rates

Implement flexible rates, charging higher interest for clients with elevated default risk.

Risk-Based Loan Amount

Adjust loan sizes based on risk, offering smaller amounts to higher-risk clients.

Summary And Recommendations

Summary of Findings

Identified key default risk drivers and segmented borrowers into risk categories.

Action Plan

Tighten loan criteria, apply risk-based pricing, and regularly update risk models.

Future Research

Explore machine learning, analyze client behavior, and assess market trends impacting default risk.