Revolutionizing Loan Approvals: A

Data-Driven ApproachCredit EDA

# Assignmen@

In today'sfas&paced consumerfinance sector, making accurate lending decisions is crucial. Traditional methods often struggle to assess creditworthiness effectively, leading to missed opportunities. However, with the rise of data science and advanced analytics, a new approach is emerging.

**Created by: Amit Kumar Singh**

Importance of Risk Handling in the Banking

Sector

Financial Stability

Managing risk is crucial for financial stability and long-term success.

## Trust Compliance and

Regulatory Requirements

Reputation

Follow risk management guidelines Effective risk management builds to avoid penalties, legal issues, and trust and credibility, and leads to disruptions to business operations. continued growth.

Overview of Loan Approval Process

1 Application Submission

Customers submit loan applications through various channels such as online platforms, mobile apps, or iFperson at branch locations. The information provided includes personal details, income, employment history, and the requested loan amount.

## 2 Evaluation and Verification

Loan officers and automated systems review the application data to assess the applicant's creditworthiness, income stability, and debt-to-income ratio. Verification of documents and financial information is conducted to ensure accuracy.

3 Approval and Funding

If the applicant meets the eligibility criteria and passes the risk assessment, the loan is approved. The approved funds are then disbursed to the borrower, and the terms and conditions of the loan are communicated.

Common Challenges Faced by Loan Providing

Companies

## Economic Downturns

Loan providing companies often face challenges during economic downturns when the default rates tend to increase. Fluctuating interest rates, inflation, and recession can lead to higher defaults, impacting the company's profitability and stability.

## Regulatory Compliance

Loan providing companies must comply with a myriad of regulations. Staying abreast of the changing regulatory landscape and ensuring compliance can be challenging and resource-intensive, impacting the operational efficiency of the company.

## Risk Management

Effective risk management is essential, but many companies struggle with accurately assessing the risk associated with each loan application. Without robust risk assessment models, the company may approve risky loans, leading to increased defaults.

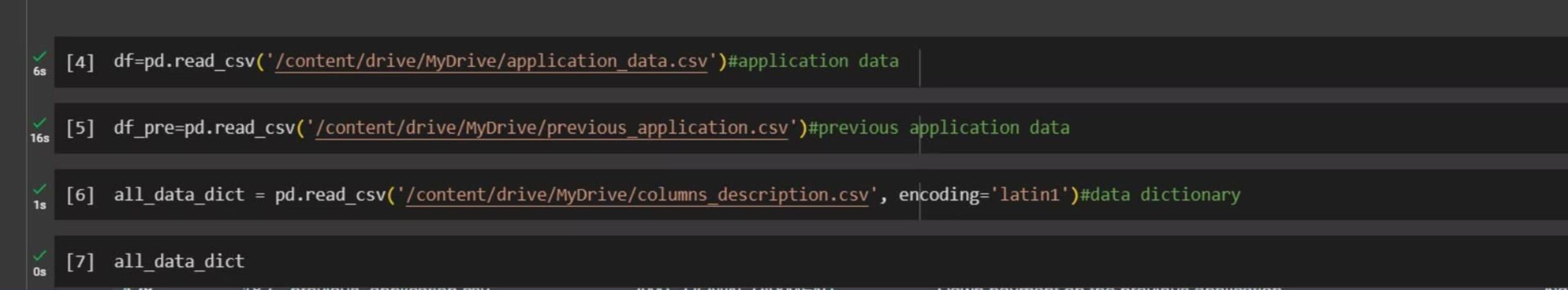
Customer Default Behaviour

Understanding and predicting customer default behavior is a complex challenge. It requires thorough analysis of historical data, customer profiles, and market trends. Failing to accurately predict default behavior can result in significant financial losses for the company.

Essence of EDA for analying the data

Exploratory Data Analysis (EDA) is a crucial first step in analyzing data. It helps us understand the dataset's structure, clean the data, and identify patterns using visualizations. EDA also involves creating new features and generating hypotheses for further analysis. Overall, EDA plays a vital role in uncovering insights and informing decisiorrmaking.

# Data sources for risk assessment

* External Data:This includes social media data, news articles, economic indicators, weather data, and geopolitical events. External data provides valuable insights into the broader environment that may impact loan repayment.
* Internal Data:lnternal data sources like customer financial records, credit scores, loan application details, and transaction history aid in assessing the creditworthiness of applicants and the overall health of the loan portfolio.
* Behavioral Data:Tracking customer behavior and payment patterns helps in identifying early warning signals of potential default, allowing proactive risk management and mitigation strategies.univariate analysis

univariate analysis and risk handeling through it

Univariate analysis involves examining a single variable at a time to understand its distribution and characteristics. Irdüext of risk handling, univariate analysis can be used to assess the risk associated with individual variables.

univariate categorical variables analysis

NAME INCOME TYPE :



The safest segments are typically those comprising individuals who are employed, working as commercial associates, or are pensioners.

## NAME EDUCATION TYPE -

The safest segment for loan approval is individuals with higher education, as they have a default rate of less than 5%.

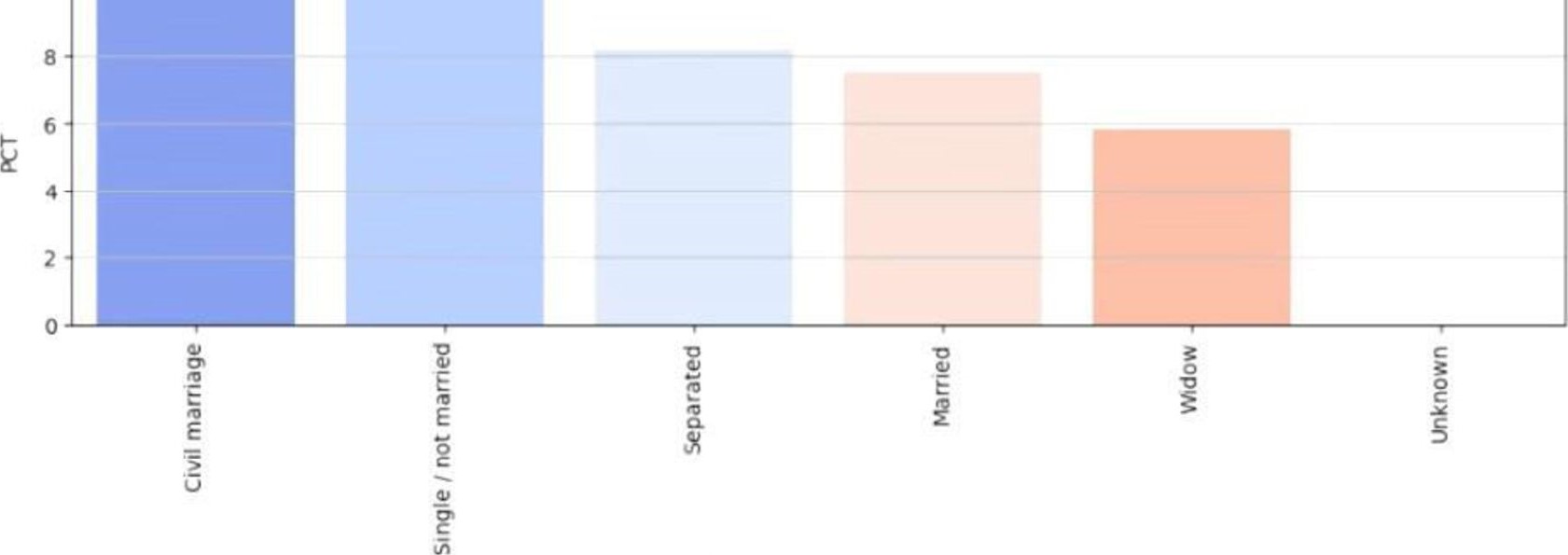
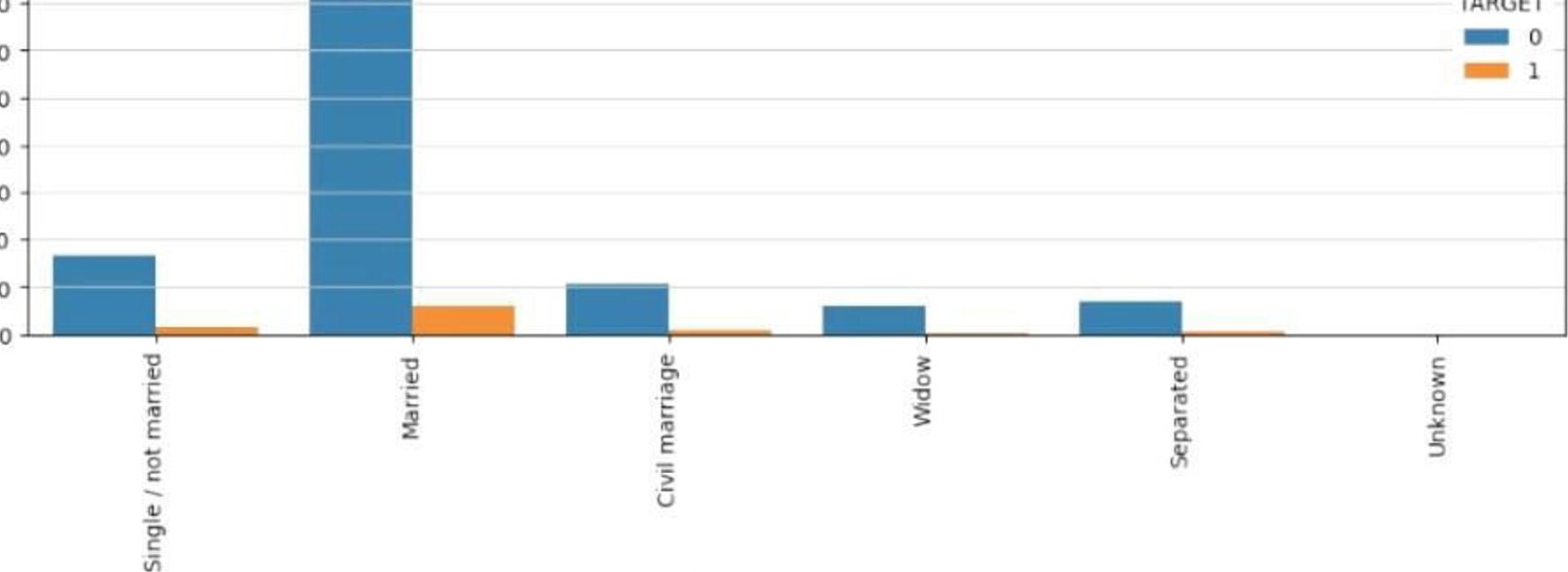
125000

100000

75000

50000

25000

 NAME FAMILY STATUS NAME FAMILY STATUS

## NAME FAMILY STATUS -



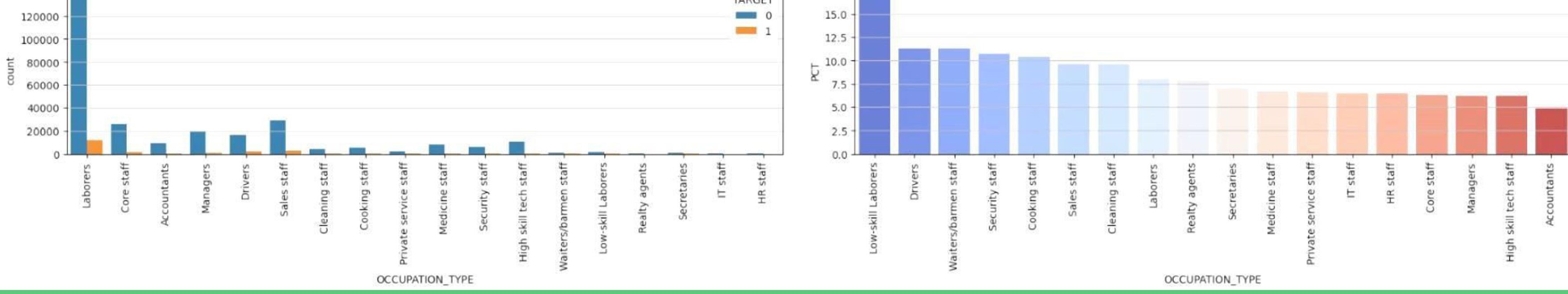
' Married individuals present a favorable prospect for lending, characterized bya default rate of approximately 8%.

### NAME HOUSING TYPE-

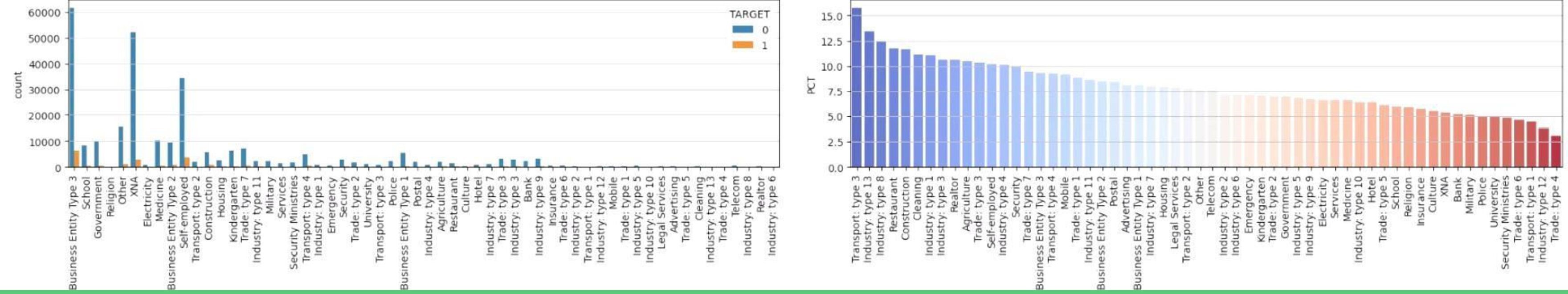
' Individuals who own houses or apartments exhibit a comparatively lower default rate of approximately8%, making them favorable candidates for loan approval.

## OCCUPATION TYPE-



* Low-skill laborers and drivers exhibit the highest default rates.
* Accountants demonstrate a lower tendency to default on loans.
* Core staff, managers, and laborers are relatively safer targets, with default rates ranging from 7.5% to 10%.

### ORGANIZATION TYPE-

 Transport type 3 individuals demonstrate the highest default rates, warranting cautious considerationfor loan approvals.  Meanwhile, individuals associated with Business Entity Type 3 and s#ployed individuals exhibit a relatively lower default rate of around 10%, making them viable candidatesfor loan approval.

# univariate numeric variables



le-6

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | TARGET | | | | | | |
|  |
|  | |  |  |  |  |  |  |  |

## AMT INCOME TOTAL

' Mostly the customers have income between O to 1 million came out to be non defaulter

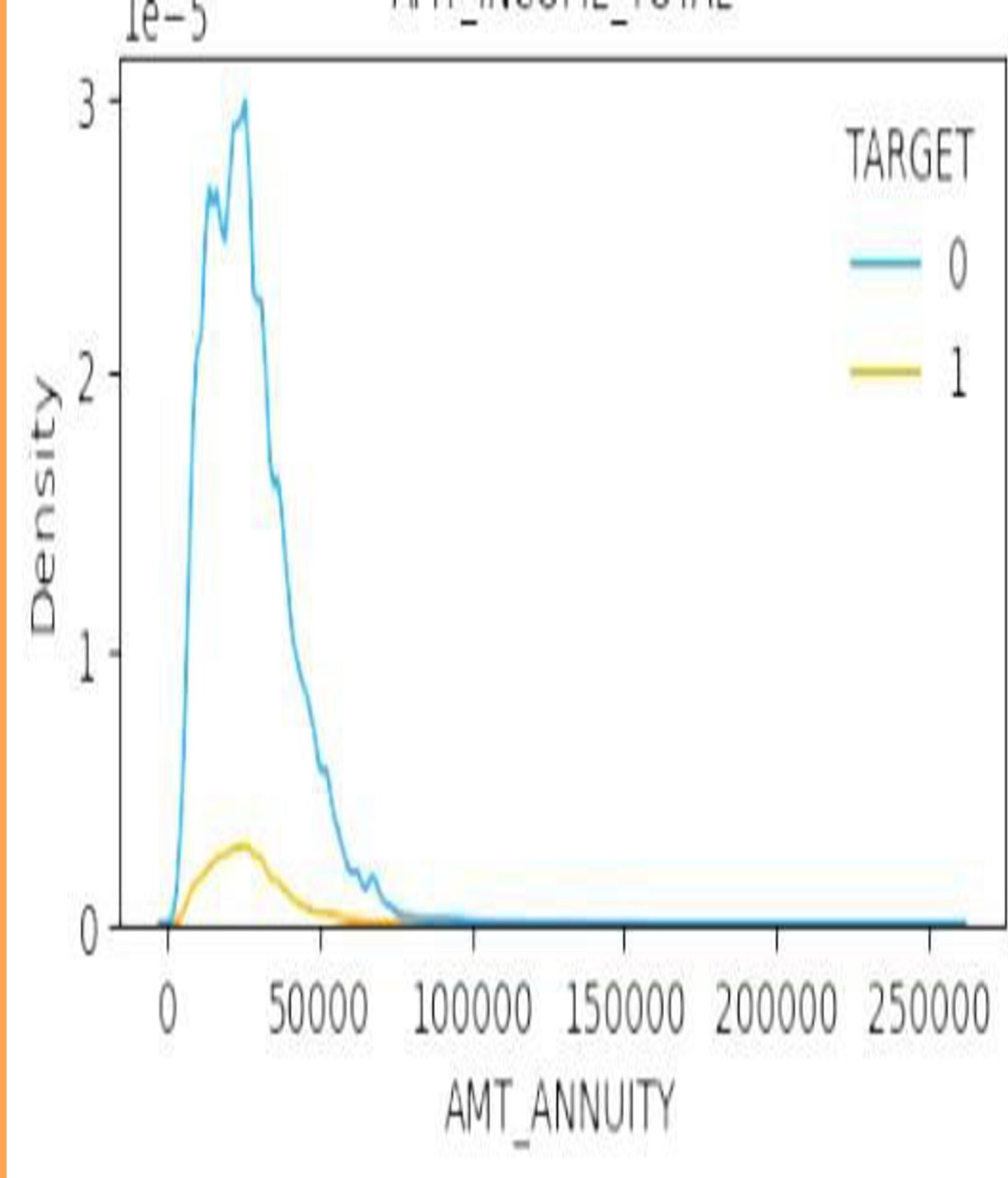
INCOM TOTA

le-6

## AMT CREDIT

 most of the loans were given for the credit amount of O to 1 ml are turn out to be non\_defaulter

CREDIT

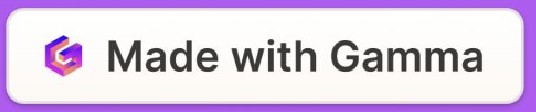
AMT ANNUITY

' most of the customers are paying annuity ofO to

50 K are turn out to be non defaulter

### AMT GOODS PRICE

 most of the loans were given for the goods price ranging between O to 1 mollion turn out to be non defaulter

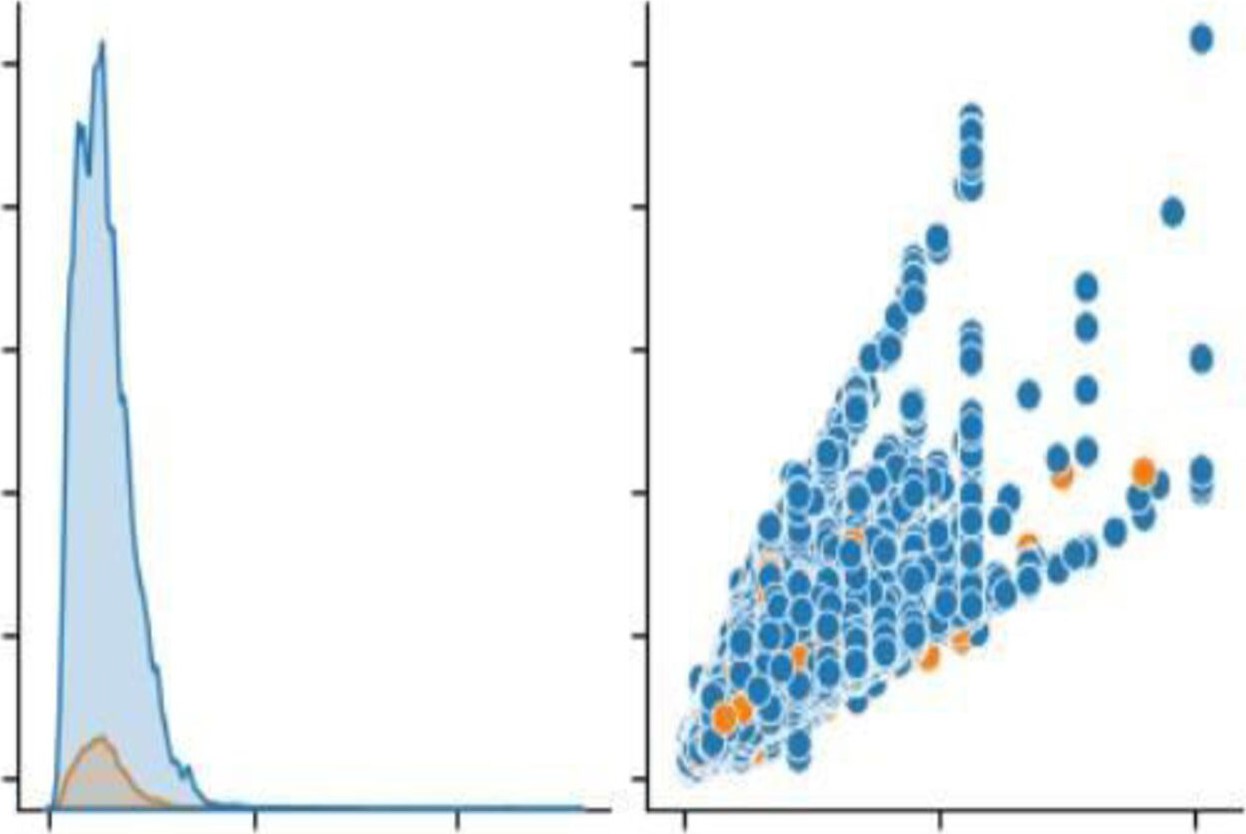
Bivariate analysis

AMT CREDIT vs

#### AMT GOODS PRICE

 The variables AMT CREDIT and

AMT GOODS PRICE exhibit a negative linear correlation, As the amount of credit requested (AMT CREDIT) increases, there is a corresponding decrease in the likelihood of default.

250000

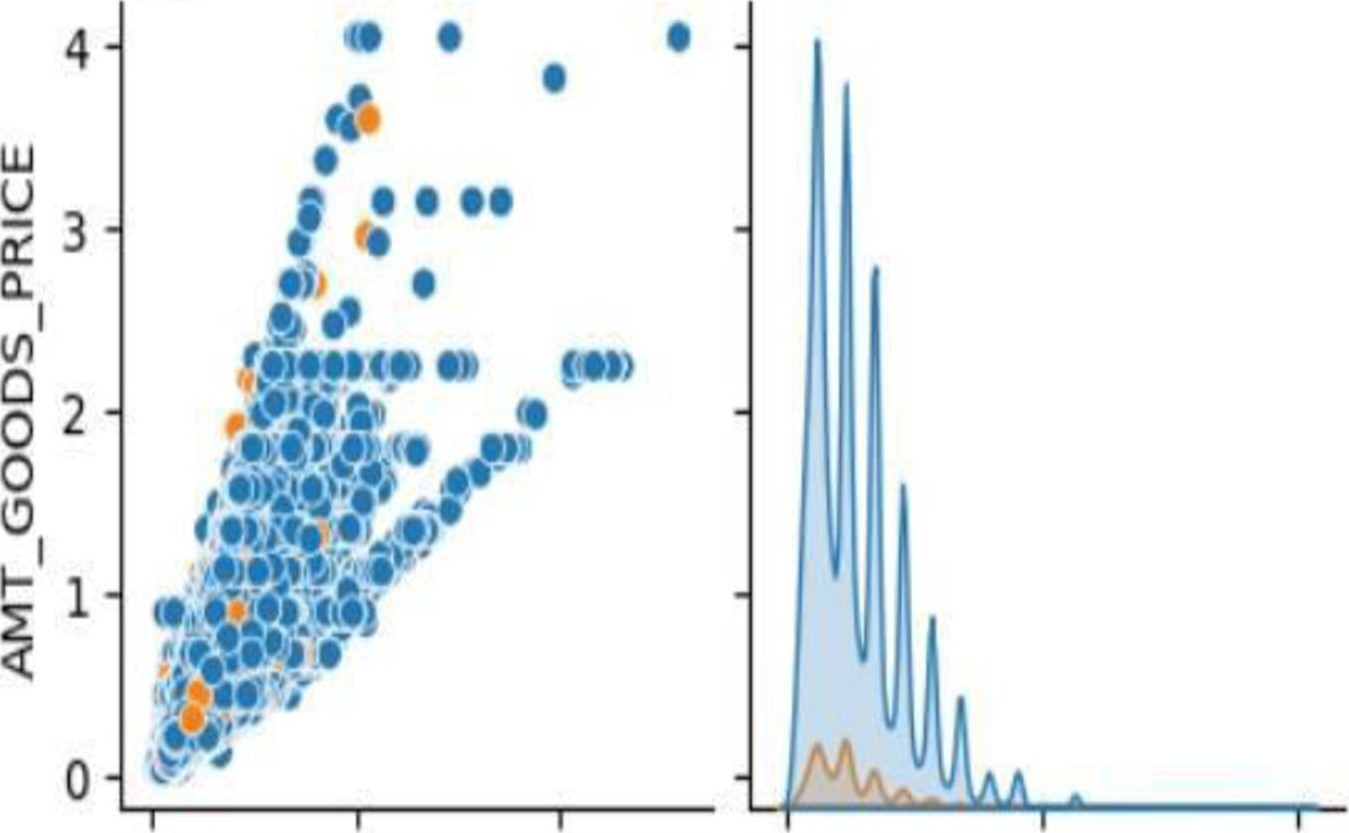
200000 z 150000

100000

50000

0

le6 TARGET



0 100000 200000 0 2 4

AMT ANNUITY AMT GOODS PRICE le6

#### AMT CREDIT vs CNT CHILDREN

 Applicants with one to less than five children exhibit a more favorable risk profile for loan approval.

|  |  |  |
| --- | --- | --- |
|  | 00 0 |  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |

17.5

15.0

12.5 z 9 10.0

u

7.5

u

5.0

2.5

0.0

0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0

AMT CREDIT le6

##### AMT ANNUIT vs AMT CREDIT

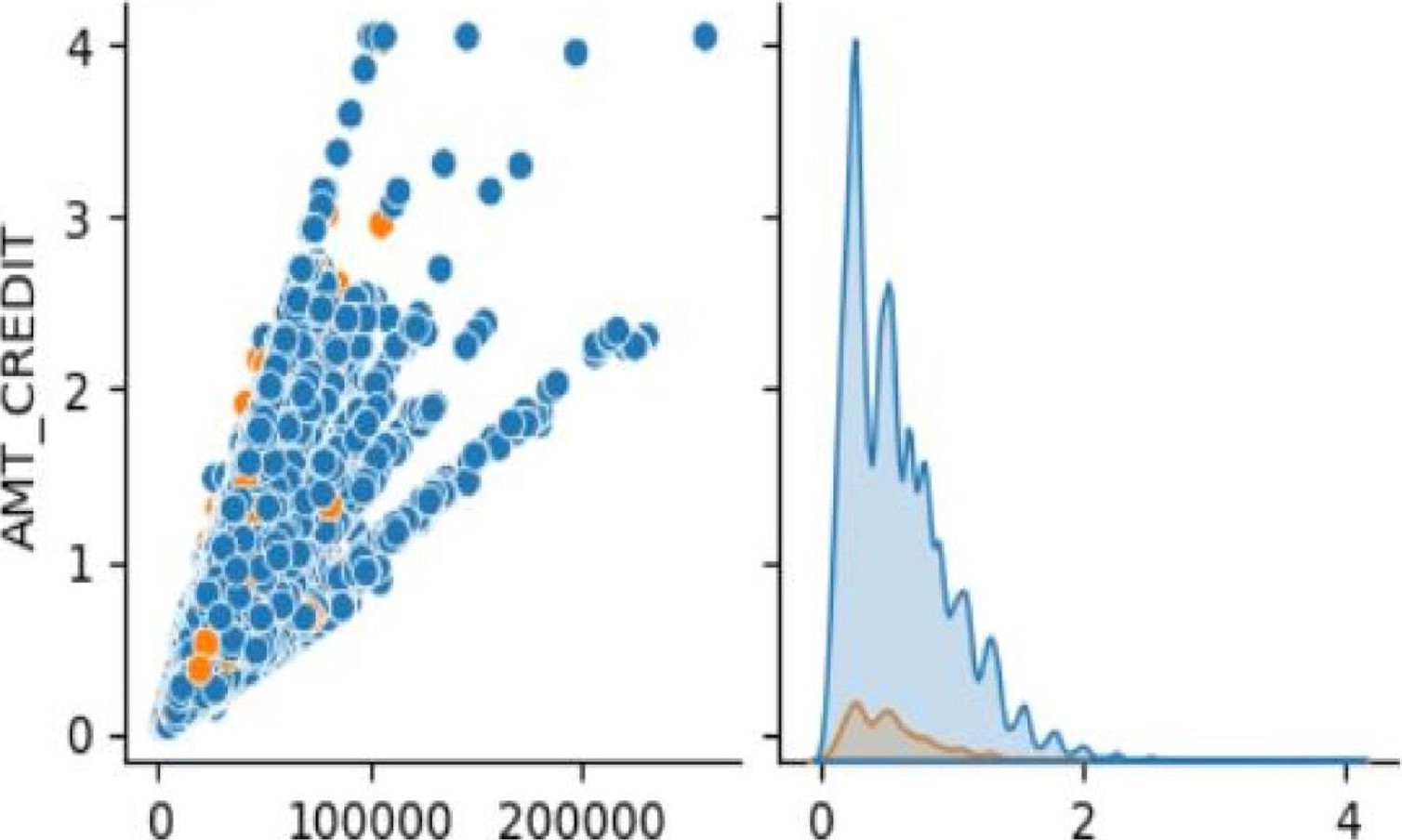
' Individuals demonstrating the ability tole6 TARGET

consistently meet an annuity payment of

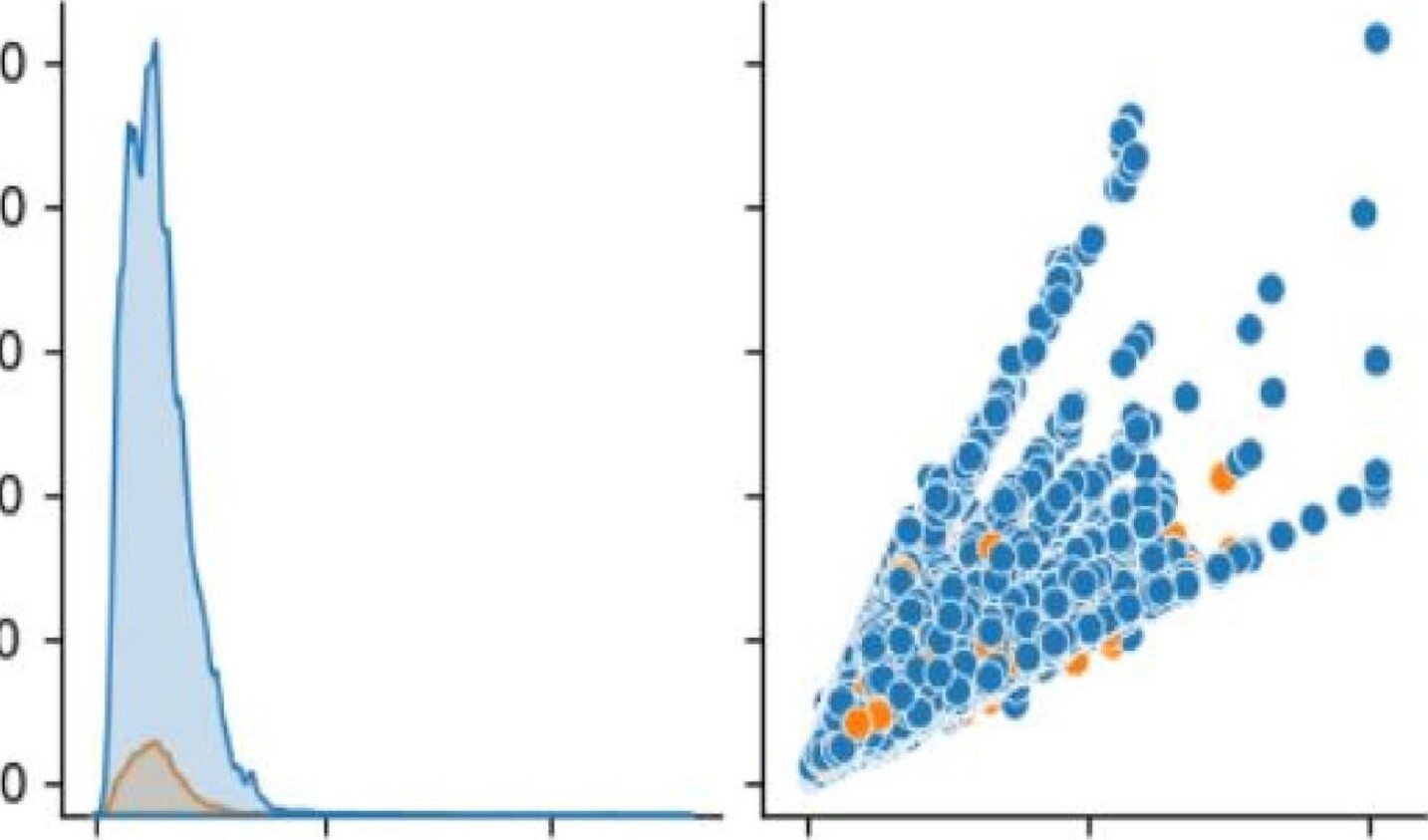
100,000 are more likely to be approved for loans up to amounts less than 2 million, representing segment characterized by lower risk profiles.



z

100000 200000

AMT ANNUITY AMT CREDIT le6

250000

200000

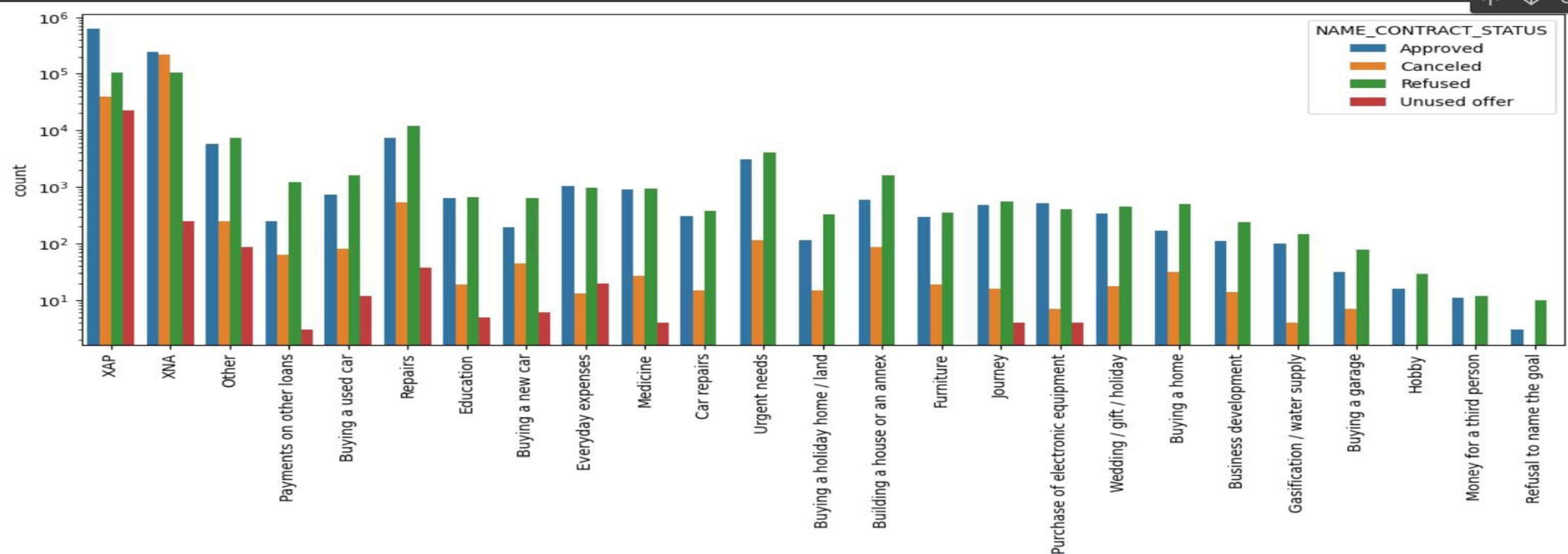
150000

100000

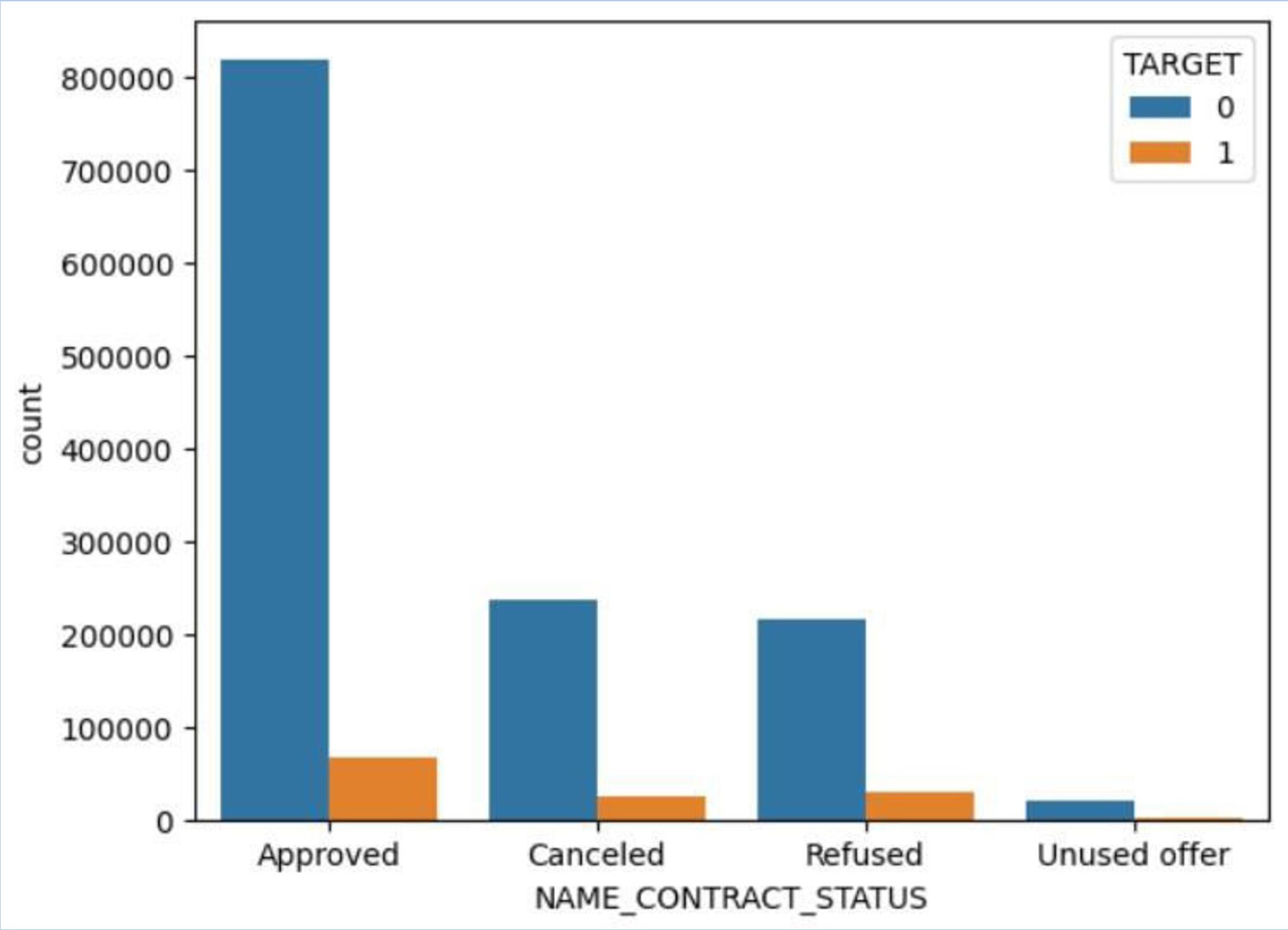
50000



###### NAME CASH LOAN PURPOSE vs NAME CONTRACT STATUS

' The repair categoryhas the highest numberofloan applications, indicating a significant demandfor re—ntedfinancing. However, it also experiences the highest number of cancellations, suggesting a need for further investigation into the underlying reasons and potential mitigation strategies.

Name contact status

 Approximately 8090% of loan applications that were previously canceled or refused are from customers who have a history of repaying their loans in the current dataset,

|  |
| --- |
| 230000  220000  F- 210000  200000 z  190000  180000  170000  Approved Canceled Refused Unused offer  NAME CONTRACT STATUS |

###### NAME CONTRACT STATUS vs AMT INCOME TOTAL

 Loan offers that were previously unused are now associated with the highest number of defaulters, even among customers with high income levels.

Final Conclusion/ Insights

Bank should target the customers

—SApplicants with an annual income below 1 million are more likely to default.

—Individuals workiq in industries classified as "Others " "Business Entit Type 3 " orsefiployed are athigher risk of defa

—pccupations such as accountants, core staff, managers, and laborers are associated with higher default rates pplicants who are married, have children numberingfive orfewer, and own a house or apartment are more likel to de ault i her levels of education do not significantly impact default rates.

' emale applicants are more likely to default.

naccom anied individuals have a lower default rate, approximately 8.5%.

## Amount segment recommended

* The credit amount should not exceed 1 million.
* An annuity of up to 50,000 is feasible, subject to eligibility.
* The income bracket should be below 1 million,
* Analysis reveals that 800% ofcustomers previously canceled or refused are reliable repayers, suggesting potential loan approvalfor this segment.

precautions

* Organizations classified as Transport type 3 should be excluded.

' Individuals in lowskill labor and driver roles should be excluded.

* Unused offers targeting highincome customers should be avoided.

