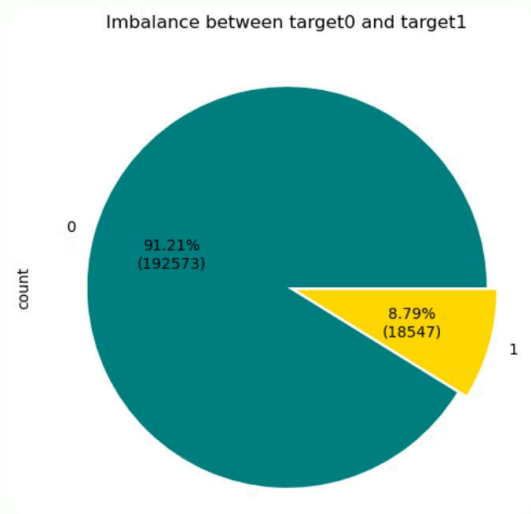# Credit Risk Case Duty

Presented by: Nitesh Pradhan

Introduction

- Objective: Understand risk analytics in banking.

- Focus: Minimize the risk of losing money.

- Scope: Identify patterns of loan repayment difficulties.

# Business Objectives

Imbalance between target0 and target1

- Identify patterns indicating repayment difficulties.

- Enable actions: Denying, reducing, or higher rates.

- Goal: Avoid rejecting capable borrowers.

Data Overview

- Datasets: application_data.csv, previous_application.csv.

- Previous Applications: 1,670,214 rows, 37 columns.

- Metadata: columns_description.csv.

- Key Libraries: Pandas, NumPy, Matplotlib, Seaborn.

Data Structure - Previous Applications
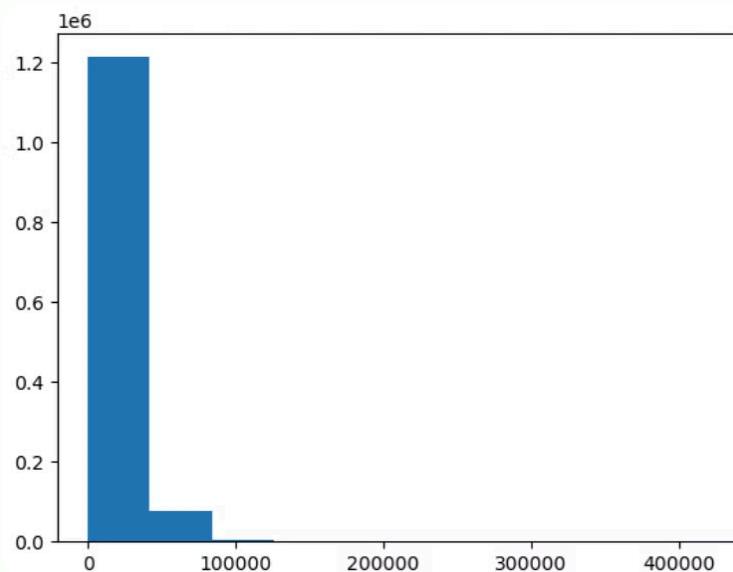
- Shape: 1,670,214 rows, 37 columns.

- Identifiers: SK_ID_PREV, SK_ID_CURR.

- Financial metrics: AMT_ANNUITY, AMT_APPLICATION, AMT_CREDIT.

- Missing Data: High null values in RATE_INTEREST_PRIMARY.

Correlation Analysis

Focus: Top 10 variables correlated with loan default. Heatmap generated using Seaborn to visualize correlations.

Key Inferences from Correlation Analysis (Part 1)
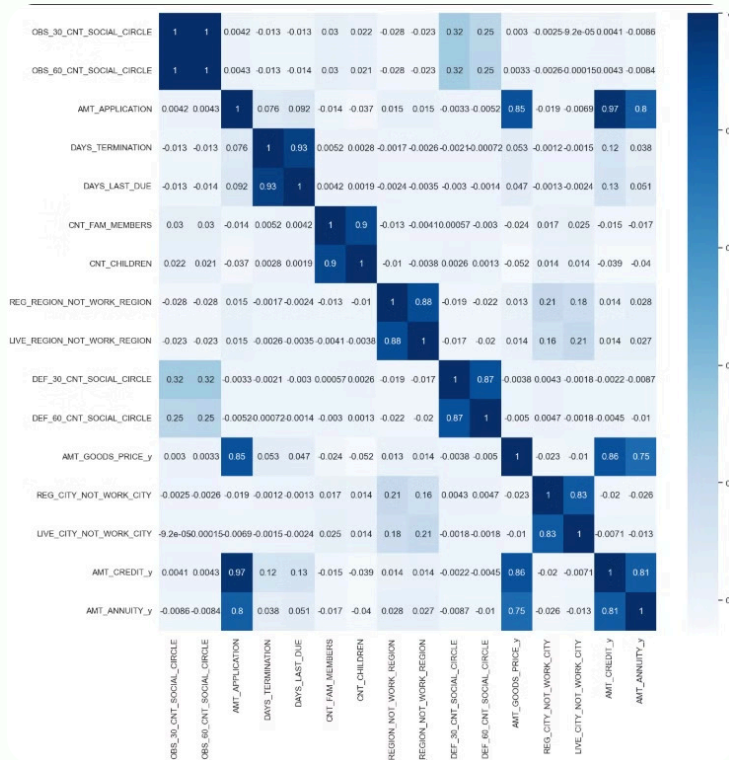
- " AMT_GOODS_PRICE & AMT_APPLICATION: High correlation - Credit requested is proportional to goods price."

- " AMT_ANNUITY & AMT_APPLICATION: Higher annuity correlates with higher goods price."

- " Address Mismatch: Contact address not matching work address strongly correlates with permanent address mismatch."

Key Inferences (Part 2)

- Goods Price & Credit: Higher goods price leads to higher credit.

- Timing Metrics: DAYS_FIRST_DUE correlated with DAYS_TERMINATION.

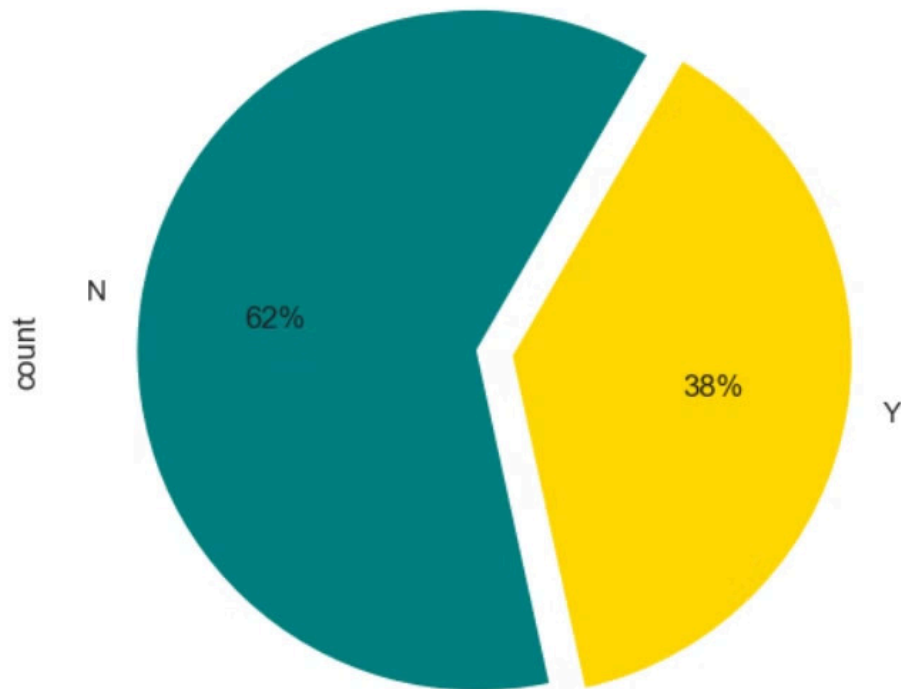- Family Metrics: CNT_CHILDREN and CNT_FAM_MEMBERS correlated.
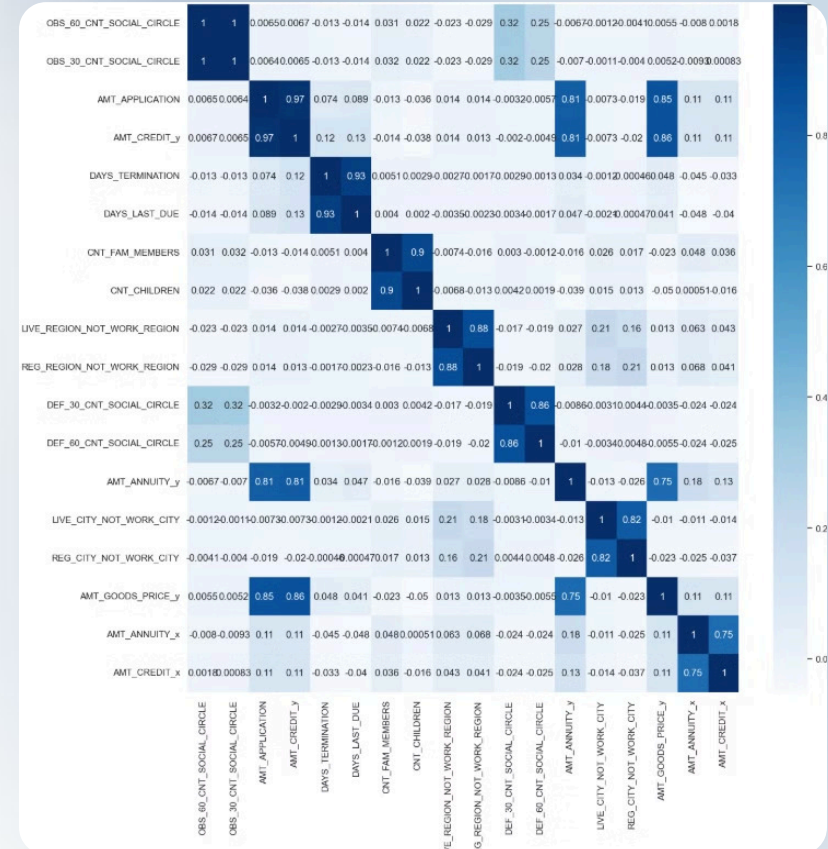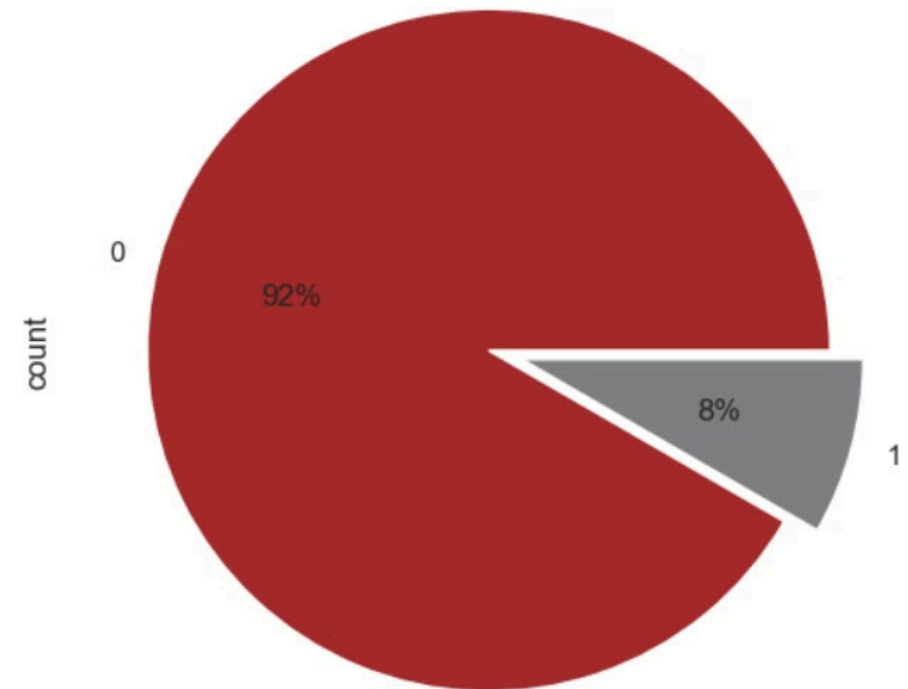
Conclusion - Targeting Strategies (Part 1)
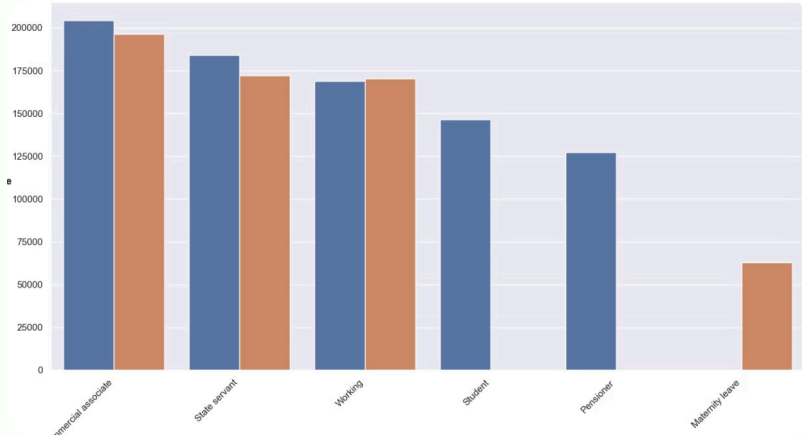
- Target: Students, Pensioners, Commercial Associates.

- Avoid: Females on maternity leave.

- Low-Risk: Clients living with parents.

- Reduce Focus: Working clients.

Conclusion - Targeting Strategies (Part 2)

**Distribution of Client by car ownership**

Distribution of Client by car ownership based on repayment status

- Inconclusive: Education type alone insufficient.
- Target: Car owners.
- Avoid: Revolving loans.
- Target: Clients with no children.

Conclusion - Targeting Strategies (Part 3)

- " Selective Targeting: Loans for 'Repairs' purpose (high defaulters and repayers; target low-risk clients)."

- " Target: Female clients (nearly double repayment rate compared to males)."

Distribution of Gender by Target Variable
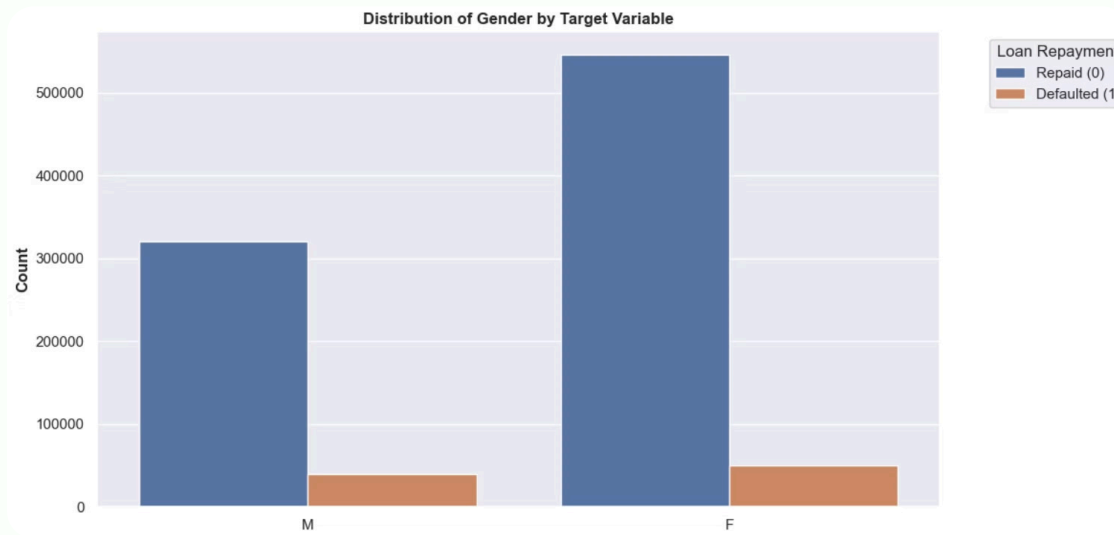
Summary

1. Clients who are Students, Pensioners and Commercial Associates with a housing type such as office/co-op/municipal apartments NEED TO BE TARGETED by the bank for successful repayments. These clients have the highest amount of repayment history.

2. Female clients on maternity leave should NOT be targeted as they have no record of repayments (therefore they are highly likely to default and targeting them would lead to a loss)

3. While clients living with parents have the least amount of repayers, they also have the least amount of defaulters. So, in cases where the risk is less, such clients can be TARGETED.

4. Clients who are working need to be targeted LESS by the bank as they have the highest amount of defaulters.

5. Clients should NOT be targeted based on their education type alone as the data is very inconclusive.

6. Banks SHOULD target clients who own a car.

7. There are NO repayers/negligible repayers when the contract type is of revolving loan.

8. Banks SHOULD target more people with no children.

9. 'Repairs' purpose of loan is the one with the most defaulters and repayers. Therefore, clients with very low risk SHOULD be given loans for such purpose to yield high profits.

10. Banks SHOULD also target female clients as they are the highest repayers (almost as double as males) amongst both the genders.