PROJECT DESCRIPTION:

Imagine you're a data analyst at a finance company specializing in urban lending, where one of the key challenges is managing defaults among customers with insufficient credit history. Your primary task involves conducting Exploratory Data Analysis (EDA) to uncover underlying patterns and ensure that deserving applicants are not unfairly rejected.

When a customer applies for a loan, your company faces two critical risks:

1. Denying a loan to an applicant capable of repayment results in lost business.

2. Approving a loan to an applicant unable to repay leads to financial losses.

The dataset provided includes detailed information on loan applications, categorizing them into two main scenarios:

- Customers with payment difficulties: These individuals have experienced a late payment exceeding X days on at least one of the first Y installments.

- All other cases: This category encompasses loans where payments have been made on time.

Each loan application can result in one of four outcomes:

- Approved: The company has agreed to the loan application.

- Cancelled: The customer chose to withdraw the application during the approval process.

- Refused: The company decided not to grant the loan.

- Unused Offer: The loan was approved but not utilized by the customer.

Your objective in this project is to leverage EDA to explore how various customer attributes and loan characteristics influence the likelihood of default. This analysis aims to enhance decision-making processes, ensuring more accurate assessments of risk and better outcomes for both the company and its applicants.

APPROACH:

Here I have made use of Microsoft Excel to obtain the following results:

Excel link-

https://docs.google.com/spreadsheets/d/1xJSXK6zjQx5T44bqhxegfkActsP2-tx5/edit?usp=sharing&ouid=105798397742237879290&rtpof=true&sd=true

https://docs.google.com/spreadsheets/d/1UTjHQ22BcYLKxxQkswqtZU1b04SRS792/edit?usp=sharing&ouid=105798397742237879290&rtpof=true&sd=true

Video Presentation-

https://drive.google.com/file/d/1JyTWu_sGNWWou_oqrs-PVFDOS1N3MAqC/view?usp=sharing

TECH-STACK USED:

For this project I have used Microsoft Excel 365 for the data visualisation and Microsoft word to prepare the report.

INSIGHTS:

- **Risk Factors Identification**:
    - Identified key risk factors contributing to loan defaults, such as insufficient credit history, late payments, and certain demographic characteristics (like age, income level, and employment status).
    - Discovered that applicants with a history of late payments on previous loans are more likely to default on new loans.
- **Impact of Loan Approval Decisions**:
    - Highlighted the dual risks associated with loan approval decisions: rejecting creditworthy applicants leads to missed business opportunities, while approving high-risk applicants results in increased financial losses due to defaults.
- **Loan Application Outcomes Analysis**:

- Analyzed the distribution of loan application outcomes: Approved, Cancelled, Refused, and Unused Offer.
- Found that a significant proportion of approved loans were unused by customers, suggesting potential inefficiencies in the approval process or customer decision-making.

• **Customer Segmentation**:

- Conducted segmentation of customers based on their risk profiles and repayment behaviors.
- Identified distinct groups of customers with varying probabilities of default, enabling targeted strategies for risk mitigation and customer retention.

• **Influence of Customer and Loan Attributes**:

- Explored how specific customer attributes (such as credit score, debt-to-income ratio, and employment stability) and loan attributes (loan amount, interest rate, and term) influence the likelihood of default.
- Found that certain combinations of attributes significantly increase or decrease the risk of default, providing insights for refining credit scoring models and risk assessment frameworks.

• **Recommendations for Improving Decision-making**:

- Suggested improvements in the loan approval process to minimize the rejection of creditworthy applicants through better utilization of alternative data sources or predictive modeling techniques.
- Proposed strategies to mitigate default risks, including targeted customer education programs, personalized loan terms, and proactive customer support initiatives.

• **Business Impact and Future Directions**:

- Demonstrated the potential business impact of implementing data-driven insights to optimize loan approval processes, reduce default rates, and enhance overall customer satisfaction.
- Outlined future directions for ongoing monitoring and refinement of predictive models based on evolving customer behaviours and economic conditions.

RESULT:

• **Comprehensive Data Exploration**: I thoroughly explored and analyzed a diverse set of loan application data, encompassing various customer attributes (such as age, income,

employment status) and loan characteristics (amount, interest rate, term). This allowed me to grasp the breadth and depth of factors influencing loan outcomes.

- **Risk Assessment and Mitigation**: By identifying key risk factors associated with loan defaults (e.g., credit history, late payments), I gained insights into how these factors interact and impact the likelihood of repayment. This understanding is crucial for developing effective risk assessment strategies and mitigating financial losses.

- **Decision-making Insights**: Analyzing the outcomes of loan applications (approved, cancelled, refused, unused offers) provided valuable insights into the consequences of different decision-making scenarios. This knowledge helps in optimizing loan approval processes to balance business growth with risk management.

- **Customer Segmentation and Personalization**: Through segmentation analysis based on risk profiles and repayment behaviors, I learned how to categorize customers effectively. This knowledge supports personalized approaches in offering loans and managing relationships to enhance customer satisfaction and retention.

- **Strategic Recommendations**: Based on the findings, I formulated strategic recommendations for improving loan approval processes, reducing default rates, and optimizing business outcomes. These recommendations are grounded in data-driven insights and aim to foster more informed decision-making within the finance company.

- **Continuous Learning and Adaptation**: The project emphasized the iterative nature of data analysis in the context of banking and finance. It highlighted the importance of ongoing monitoring, adaptation of predictive models, and responsiveness to changing economic and customer behaviour trends.