Mention the problem statement and the analysis approach briefly

Problem Statement

The primary objective of this case study is to utilize Exploratory Data Analysis (EDA)
to identify key patterns and driving factors that influence the likelihood of loan default among applicants.
As a part of a consumer finance company specializing in lending various types of loans to urban customers, it is crucial to mitigate the risks associated with loan approvals. The company faces two significant risks: rejecting loans to applicants who are likely to repay, which results in a loss of potential business, and approving loans to applicants who are likely to default, leading to financial losses.

The dataset provided contains historical information about past loan applicants and whether they defaulted. By analyzing this data, the goal is to uncover insights that can help predict the likelihood of default, enabling the company to make more informed decisions when approving or rejecting loan applications.
This analysis aims to reduce the incidence of credit loss by identifying high-risk applicants and implementing strategies such as denying the loan, reducing the loan amount, or lending at higher interest rates to mitigate potential losses.
Through this case study, the company seeks to understand the key consumer and loan attributes that are strong indicators of default. The findings from this analysis will be instrumental in refining the company's portfolio and risk assessment processes, ultimately minimizing financial losses and improving overall lending decisions.

Explain the results of univariate, bivariate analysis etc. in business terms

- **1. Univariate Analysis:**
- **Definition:** Univariate analysis involves examining each variable individually to understand its distribution, central tendency (mean, median, mode), and dispersion (variance, standard deviation).
- **Business Explanation:**
- **Loan Amount Distribution:** If the analysis reveals that most loans are for amounts between $5,000 and $10,000, the company can focus its risk assessment strategies on loans within this range, ensuring that the terms are appropriately set to minimize risk.
- **Credit Score Distribution:** If the analysis shows that a significant portion of applicants have a credit score below 600, the company may need to reevaluate its approval criteria to avoid lending to high-risk applicants, thereby reducing the likelihood of defaults.
- **2. Bivariate Analysis:**
- **Definition:** Bivariate analysis examines the relationship between two variables to determine how they influence each other.
- **Business Explanation:**
- **Income vs. Loan Default Rate:** Suppose the bivariate analysis reveals that applicants with lower incomes are more likely to default. This insight suggests that income is a key factor in loan repayment capability. The company might consider imposing stricter lending criteria or higher interest rates for lower-income applicants to offset the higher risk.
- **Loan Amount vs. Default Rate:** If the analysis shows that larger loan amounts are associated with a higher default rate, the company might decide to impose more stringent approval processes or limit the maximum loan amount offered to certain applicants.
- **Credit Score vs. Interest Rates:** If a strong relationship is observed between credit scores and interest rates, where lower credit scores are correlated with higher interest rates, the company might use this information to optimize pricing strategies, ensuring that higher-risk borrowers are charged appropriately to mitigate potential losses.
- **3. Multivariate Analysis:**
- **Definition:** Multivariate analysis looks at multiple variables simultaneously to understand complex relationships and interactions among them.
- **Business Explanation:**
- **Interest Rate, Loan Tenure, and Default Rate:** If a significant interaction is found between high-interest rates and longer loan tenures leading to higher default rates, the company might consider revising its loan products, perhaps by offering shorter tenure options for high-interest loans to reduce the risk of default.

- Include visualisations and summarise the most important results in the presentation
- **Visualisation:**
- **Flowchart:** Simplified flowchart showing the loan approval process, highlighting key risks at each stage.
- **Summary:**
- **Key Risks:**
    - Rejecting creditworthy applicants results in lost business opportunities.
    - Approving loans for high-risk applicants leads to financial losses due to defaults.
- **Objective:** Understand the factors that drive loan defaults to enhance the company's risk assessment and lending strategies.
- **Univariate Analysis**
- **Bivariate Analysis**
- **HeatMap**
- **MultiVariate Analysis**