

Project Title: Personal Loan Campaign Analysis – Thera Bank

Data Description:

The dataset Bank.xls contains information on **5,000 customers**, including demographic features (such as age and income), banking relationships (like mortgage status and securities account), and customer responses to a past **personal loan marketing campaign**. Out of 5,000 customers, **only 480 (9.6%) accepted the personal loan offer**, indicating an imbalanced classification problem.

Project Objective:

This project focuses on helping **Thera Bank** improve the effectiveness of future personal loan campaigns by identifying patterns in customer behavior and selecting better target groups, with the goal of increasing conversion rates while minimizing marketing costs.

Project Steps & Analysis Workflow:

1. Data Import & Basic Inspection:

- Loaded the dataset and examined its shape, size, and basic summary statistics.
- Identified missing values, skewness, and data types to understand initial data quality.

2. Age-wise Loan Acceptance Analysis:

- Explored how age impacts loan acceptance.
 - Found that **customers over 55** rarely accept loans—likely due to retirement.
 - Customers **under 23** also show low interest—possibly students.
- Detected and handled **negative or invalid age values**.

3. Income-Based Analysis:

- Analyzed how income influences loan eligibility and approval.
- Segmented customers by income range to determine their likelihood of accepting loans.

4. Family Size & Education Level Analysis:

- Studied how **family size (1–5 members)** and **education levels** impact loan decisions.
- Compared patterns across education vs. income.

5. Experience Column Fix & Distribution Plot:

- Noted that the **‘Experience’** column had negative values, which were not logical.
- Visualized and handled these outliers appropriately.

6. Correlation & Feature Selection:

- Checked correlation among variables.
- Found **high correlation between Age and Experience**—decided to **drop Experience**, as age alone was more reliable for modeling loan decisions.

7. Additional Feature Insights:

- Performed segment-wise analysis on:
 - **Education & Income**
 - **Credit card ownership**
 - **Securities account status**
 - **Online banking usage**
 - **Account type/category**

8. Handling Skewness in Data:

- Applied two transformation methods to reduce skewness and normalize distributions:
 - **Log Normal Transformation**
 - **Power Transformer**

Outcome:

This analysis provided actionable insights into customer segments that are more likely to respond to loan campaigns. The refined understanding helps Thera Bank focus its marketing efforts on the right audience while improving campaign efficiency.