**Customer Financial Behaviour Analysis – Banking Dataset**

**Objective**

The goal of this project was to explore and analyse financial customer data from a bank to understand customer segments, product usage, and account behaviours using Python and data visualization techniques**.**

**Dataset Summary**

Total Records: 3,000 customers

Total Features: 26 columns

Data Includes:

* Demographic Info (Age, Gender, Location)
* Income & Savings
* Account Holdings (Loans, Deposits, Credit Cards)
* Loyalty Classification

**Data Cleaning & Preparation**

Removed unnecessary columns and characters

Handled missing values

Converted income bands to numeric format

Created new metrics like total\_liability and total\_balance for better understanding

**Exploratory Data Analysis (EDA)**

**Bank Balance**

Total Balance : 3842929878.6000004

Bank Deposits : 2014680581.7700002

Saving Accounts : 698725060.4499999

Superannuation Savings : 76594799.02

Checking Accounts : 963278847.38

Foreign Currency Account : 89650589.98

**Total Lending’s**

Credit card balance : 9528620.83

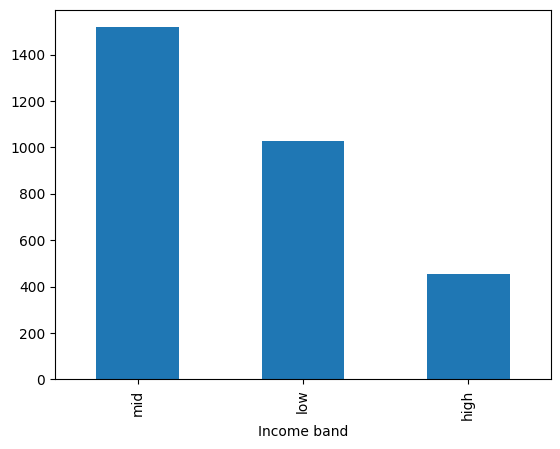
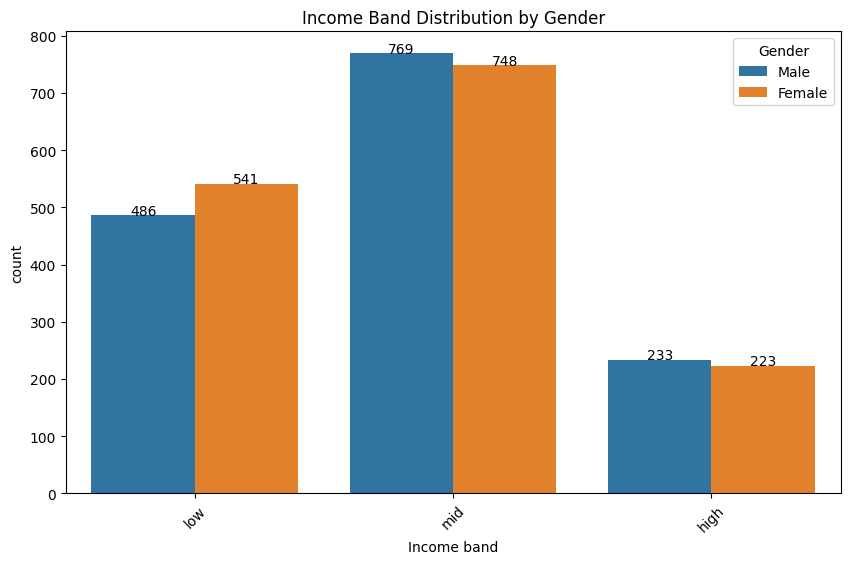
Business lending : 2600279425.2200003

Bank Loan : 1774158466.46

Total loan : 4383966512.51

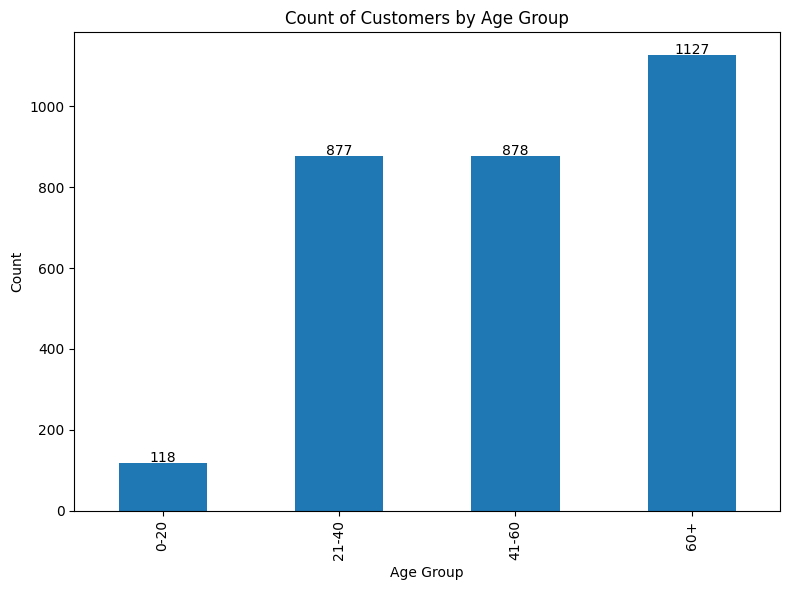
**Income band**

Low 0< x <100000 Mid 100001 < x< 300000 High 300001 < x

**Gender wise customers:**

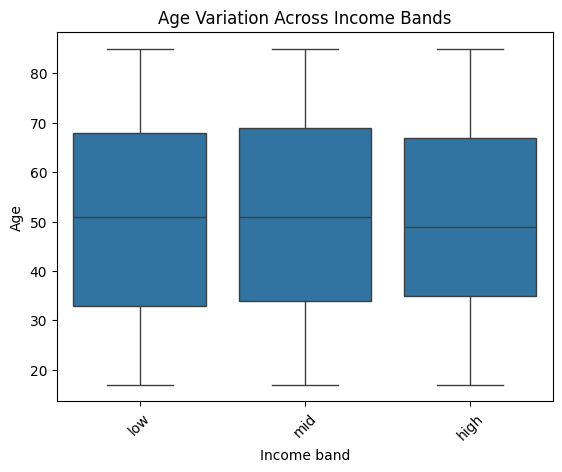
Male = 1488 , Female = 1512

**Age**

The largest customer segment belongs to the 60+ age group, with over 1,100 individuals. This may indicate decades of accumulated savings, higher product engagement, or a preference for stable, long-term financial products such as retirement accounts and fixed deposits.

The 21–60 age range forms a strong middle segment, representing working professionals with active financial behaviours.

The **0–20 age group** is minimally represented, likely due to limited financial independence or product eligibility.



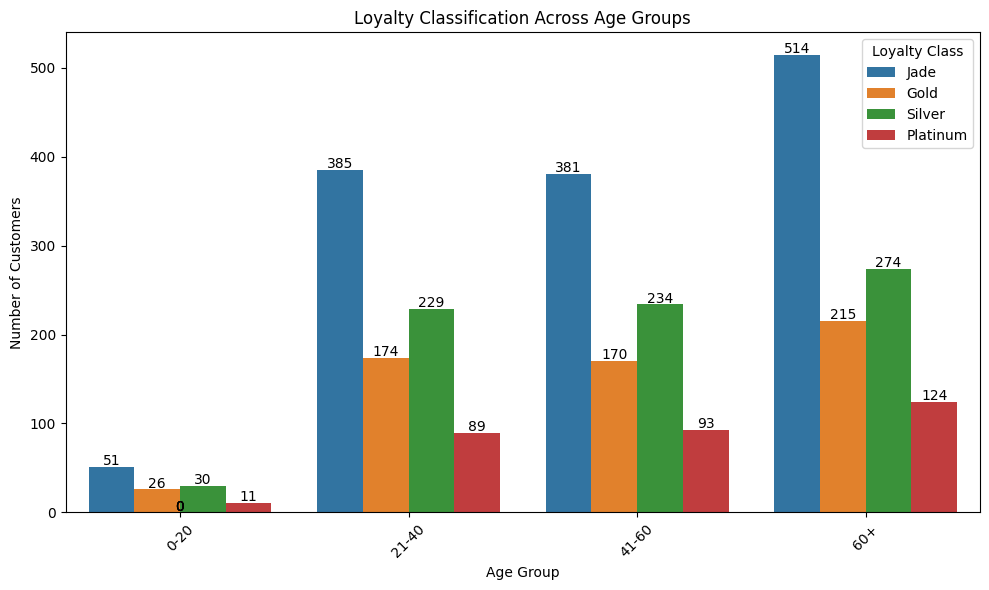
Low Income Band: Median age ~51

Mid Income Band: Median age ~52

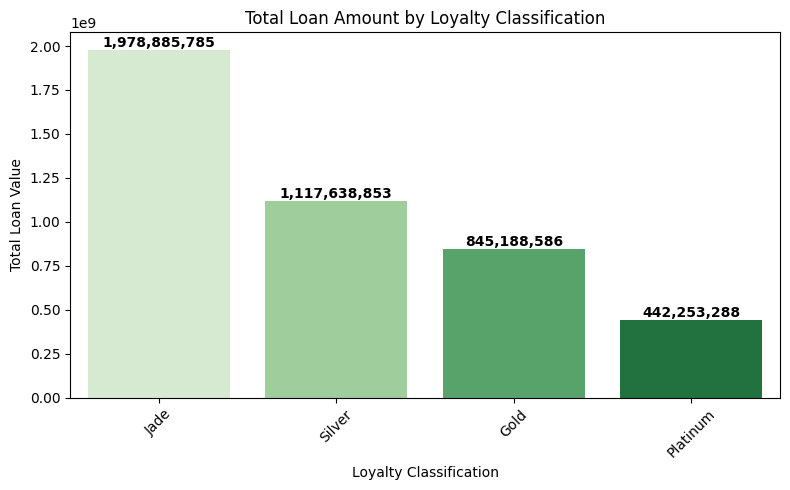
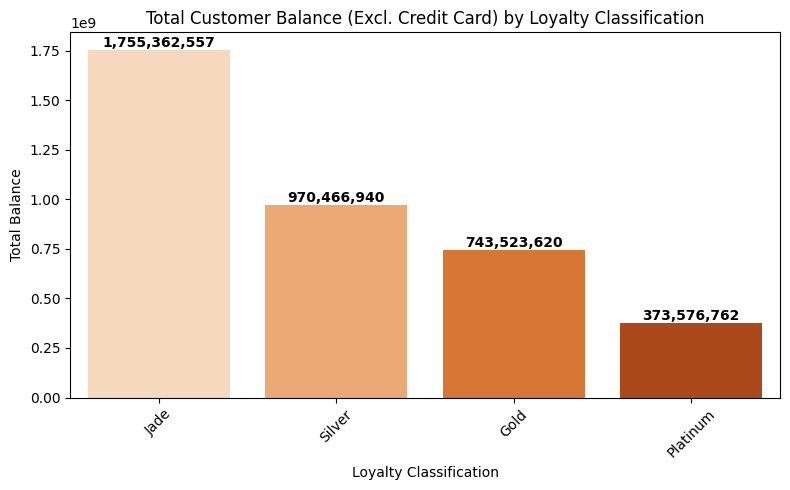
High Income Band: Median age ~49

Across all income bands, the median age is around 50 — not a huge difference, but high-income customers may be slightly younger on average. Bank customers show wide range of age. All Income Band have same range of customers

**Loyalty Classification**

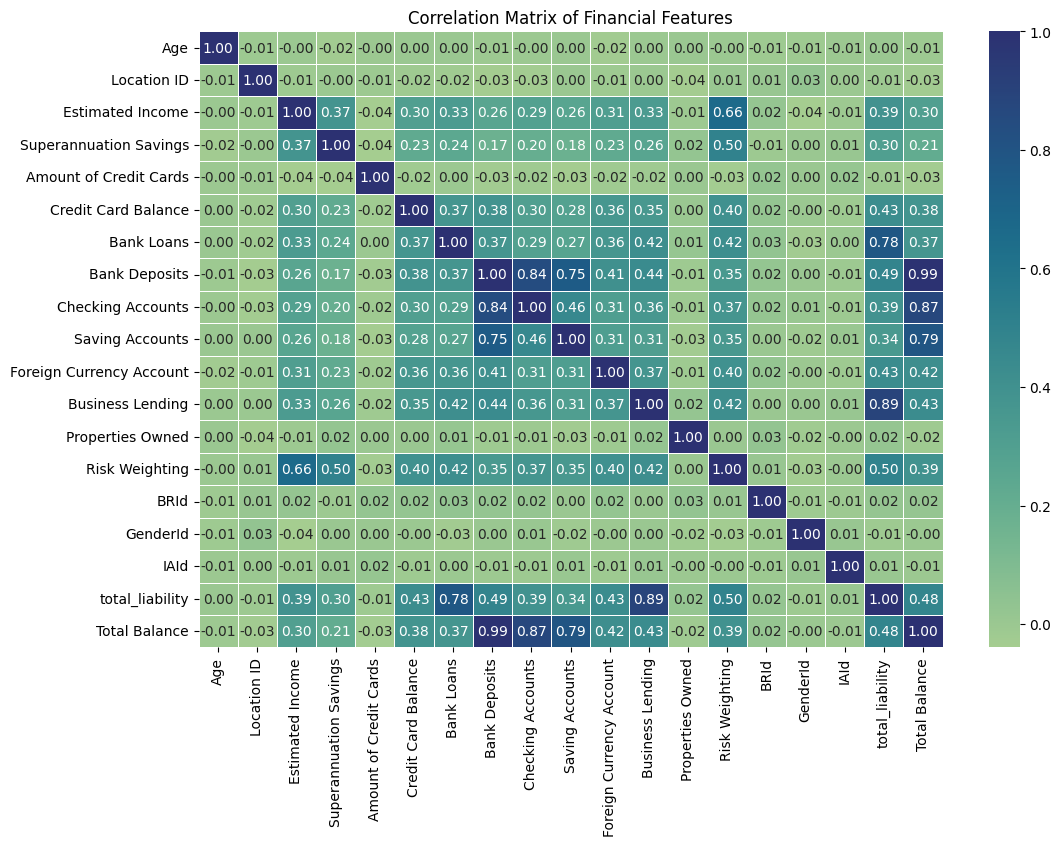


Loyalty tier increases with age. The 60+ group have the most loyal (Platinum/Silver) customers, while younger clients (0–20) are mostly in lower tiers. Mid-age groups (21–60) show strong engagement across all tiers

**Jade customers** hold the highest share of both **loans** and **total balance**, indicating they are the bank’s most financially active segment. In contrast, **Platinum customers** have the lowest totals in both, suggesting either lower engagement or conservative financial behaviour.

**Correlation between data**



* Bank Deposits and Total Balance (r = 0.99)  
  → Bank deposits are the primary contributor to the total balance held by customers.
* Checking Accounts and Total Balance (r = 0.87)  
  → Customers with high checking account balances also have significantly higher total balances, indicating active financial engagement.
* Business Lending and Total Liability (r = 0.89)  
  → Business loans are a major component of customer liabilities.
* Bank Loans and Total Liability (r = 0.78)  
  → Traditional bank loans contribute heavily to the total debt profile of customers.
* Estimated Income and Risk Weighting (r = 0.66)  
  → Higher-income customers tend to have higher financial risk, possibly due to larger exposure to high-value products.
* Credit Card Balance and Total Liability (r = 0.43)  
  → Credit card usage has a moderate impact on customer liabilities.

**Tools Used**

Python (Pandas, Matplotlib, Seaborn)

Jupyter Notebook

Microsoft Word / PowerPoint (for documentation)

**Author**

**Name:** Aniket Navele  
**Role:** Data Analyst   
GitHub:[[Link](https://github.com/Aniketnavele05)] | LinkedIn:[[Link](http://www.linkedin.com/in/aniketnavele)]