BANKING CLIENT ANALYSIS – PROJECT REPORT

1. Project Overview

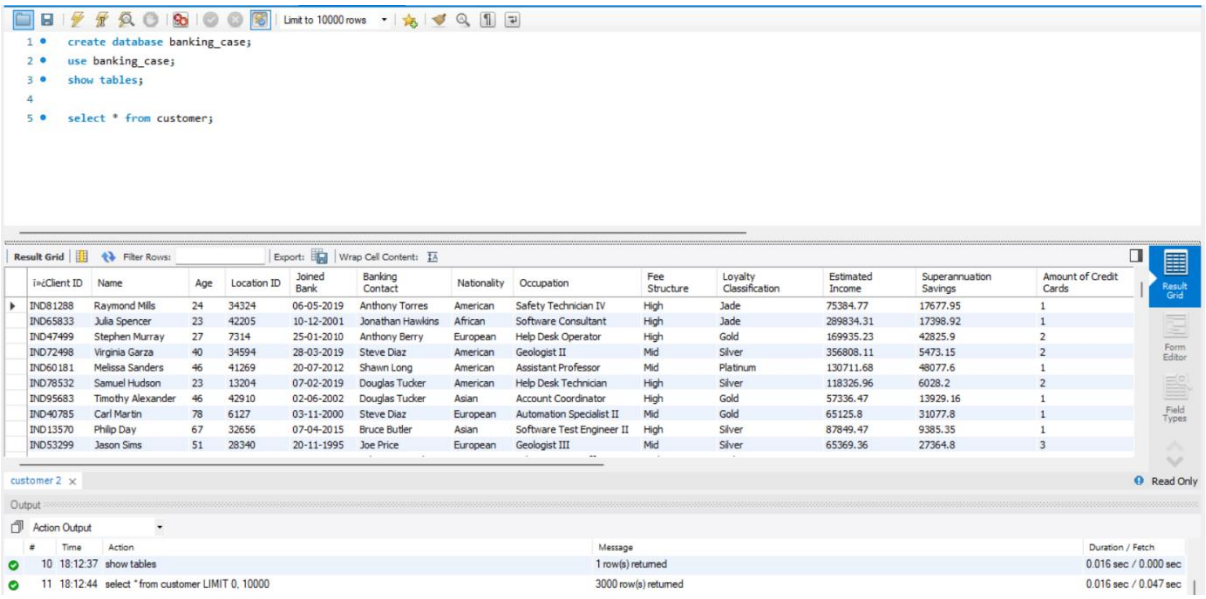
The goal of this project was to analyze client banking data to understand customer profiles, product usage, and financial behavior patterns. This analysis helps identify key customer segments, advisor performance, and opportunities for improving client engagement and profitability.

The project involved three main steps:

1. Data Preparation: Merging multiple sheets (Clients, Gender, Banking Relationship, and Investment Advisor).
2. Exploratory Data Analysis (EDA): Performed in Jupyter Notebook using Python.
3. Interactive Visualization: Built in Power BI for easy interpretation and business insights.

2. Data Preparation

- Raw data: Contained 4 sheets — Clients, Gender, Banking Relationship, and Investment Advisor.
- These were merged into a single client-banking file and saved as CSV with 3000 rows and 25 columns.
- The data was imported into MySQL for integration with both Python (EDA) using sqlalchemy and Power BI.



After merging and cleaning:

- Missing values were handled.

```
# Rename the columns as required
df = df.rename(columns = {"ï»¿Client ID":"Client ID","BRId":"Banking Relationship","GenderId":"Gender","IAId":"Investment Advisor Id"})
```

```
df.head()
```

	Client ID	Name	Age	Location ID	Joined Bank	Banking Contact	Nationality	Occupation	Fee Structure	Loyalty Classification	...	Checking Accounts	Saving Accounts	Foreign Currency Account	Business Lending	Properties Owned
0	IND81288	Raymond Mills	24	34324	06-05-2019	Anthony Torres	American	Safety Technician IV	High	Jade	...	603617.88	607332.46	12249.96	1134475.30	1
1	IND65833	Julia Spencer	23	42205	10-12-2001	Jonathan Hawkins	African	Software Consultant	High	Jade	...	229521.37	344635.16	61162.31	2000526.10	1
2	IND47499	Stephen Murray	27	7314	25-01-2010	Anthony Berry	European	Help Desk Operator	High	Gold	...	652674.69	203054.35	79071.78	548137.58	1
3	IND72498	Virginia Garza	40	34594	28-03-2019	Steve Diaz	American	Geologist II	Mid	Silver	...	1048157.49	234685.02	57513.65	1148402.29	0
4	IND60181	Melissa Sanders	46	41269	20-07-2012	Shawn Long	American	Assistant Professor	Mid	Platinum	...	446644.25	128351.45	30012.14	1674412.12	0

5 rows × 26 columns

```
# Generating rows and columns
df.shape
```

```
(3000, 25)
```

```
# Generating descriptive statistics for the dataframe
df.describe()
```

	Age	Location ID	Estimated Income	Superannuation Savings	Amount of Credit Cards	Credit Card Balance	Bank Loans	Bank Deposits	Checking Accounts	Saving Accounts	Foreign Currency Account
count	3000.000000	3000.000000	3000.000000	3000.000000	3000.000000	3000.000000	3.000000e+03	3.000000e+03	3.000000e+03	3.000000e+03	3000.000000
mean	51.039667	21563.323000	171305.034263	25531.599673	1.463667	3176.206943	5.913862e+05	6.715602e+05	3.210929e+05	2.329084e+05	29883.529993
std	19.854760	12462.273017	111935.808209	16259.950770	0.676387	2497.094709	4.575570e+05	6.457169e+05	2.820796e+05	2.300078e+05	23109.924010
min	17.000000	12.000000	15919.480000	1482.030000	1.000000	1.170000	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	45.000000
25%	34.000000	10803.500000	82906.595000	12513.775000	1.000000	1236.630000	2.396281e+05	2.044004e+05	1.199475e+05	7.479440e+04	11916.542500
50%	51.000000	21129.500000	142313.480000	22357.355000	1.000000	2560.805000	4.797934e+05	4.633165e+05	2.428157e+05	1.640866e+05	24341.190000
75%	69.000000	32054.500000	242290.305000	35464.740000	2.000000	4522.632500	8.258130e+05	9.427546e+05	4.348749e+05	3.155750e+05	41966.392500
max	85.000000	43369.000000	522330.260000	75963.900000	3.000000	13991.990000	2.667557e+06	3.890598e+06	1.969923e+06	1.724118e+06	124704.870000

```
Client ID                  0
Name                       0
Age                        0
Location ID                0
Joined Bank                0
Banking Contact            0
Nationality                0
Occupation                 0
Fee Structure              0
Loyalty Classification     0
Estimated Income           0
Superannuation Savings     0
Amount of Credit Cards     0
Credit Card Balance        0
Bank Loans                 0
Bank Deposits              0
Checking Accounts          0
Saving Accounts            0
Foreign Currency Account   0
Business Lending           0
Properties Owned           0
Risk Weighting             0
BRId                       0
GenderId                   0
IAId                       0
Income Band                0
dtype: int64
```

- Gender, Branch, and IA (Investment Advisor) IDs were replaced with meaningful labels.

```
gender_map = {1:"Male",2:"Female"}
branch_map = {1:"Retail",2:"Institutional",3:"Private Bank",4:"Commercial"}

df['Gender'] = df['Gender'].map(gender_map)
df['Banking Relationship'] = df['Banking Relationship'].map(branch_map)
```

- The final dataset represented 3,000 clients with demographic and financial details.

3. Exploratory Data Analysis (EDA)

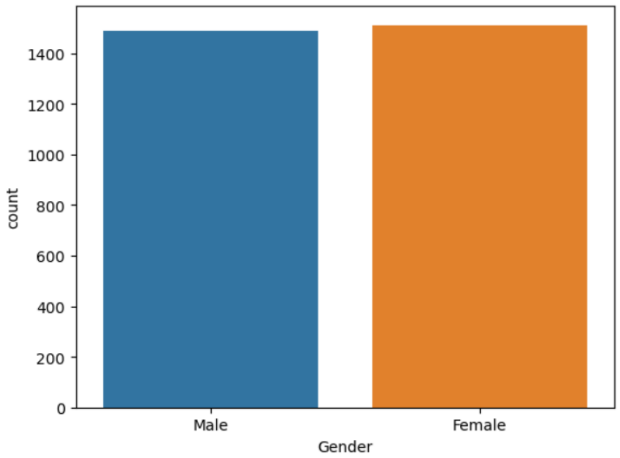
Performed in Jupyter Notebook using Pandas, Matplotlib, and Seaborn.

Data Cleaning

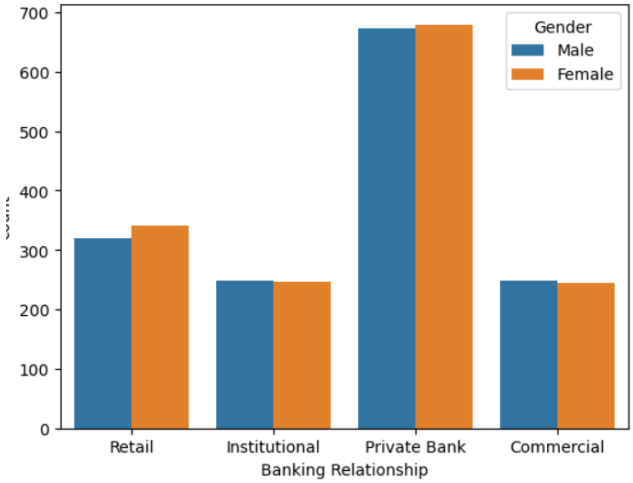
- Standardized categorical values (Gender, Branch names, etc.).
- Removed or filled missing values.
- Verified data consistency and numerical accuracy.

Univariate & Bivariate Analysis

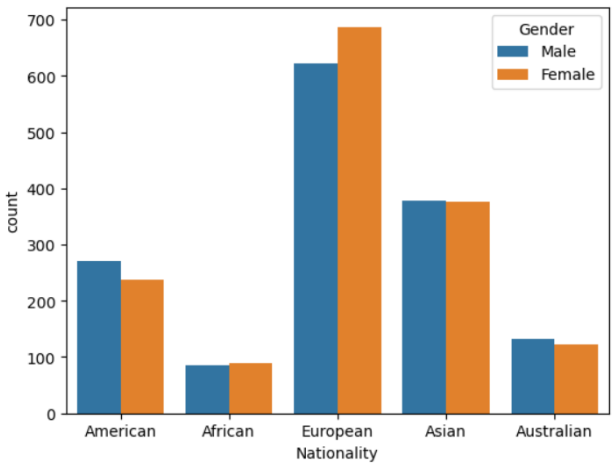
- Gender distribution: Males slightly outnumber females.



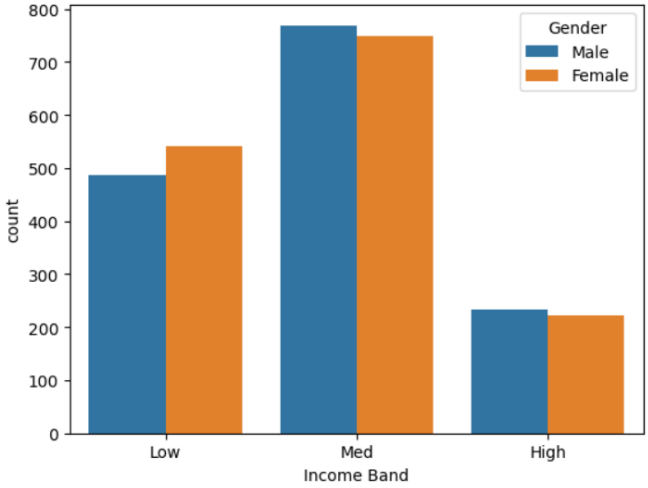
- Banking Relationship: Private banking clients form the majority.



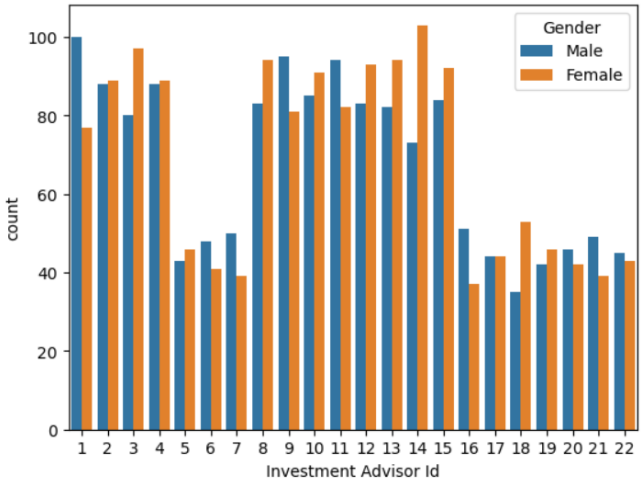
- Nationality: European and Asian clients dominate.



- Income Bands: Med-income clients form the largest segment.

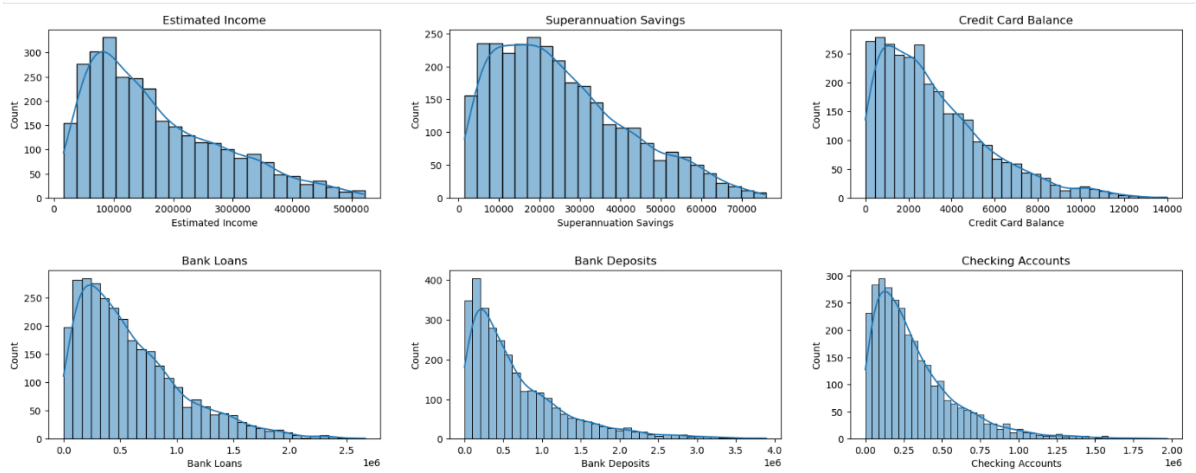


- Advisor Mapping: Around 22 Investment Advisors were identified.

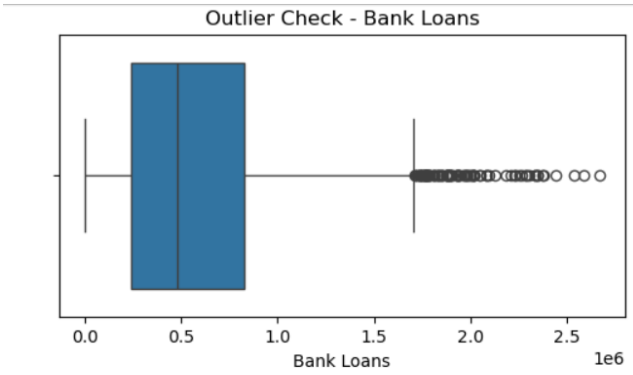


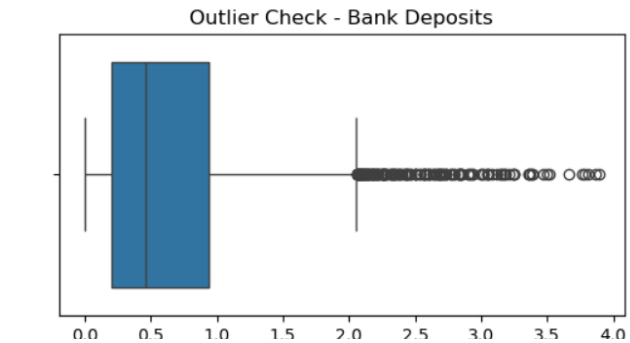
Numerical Analysis

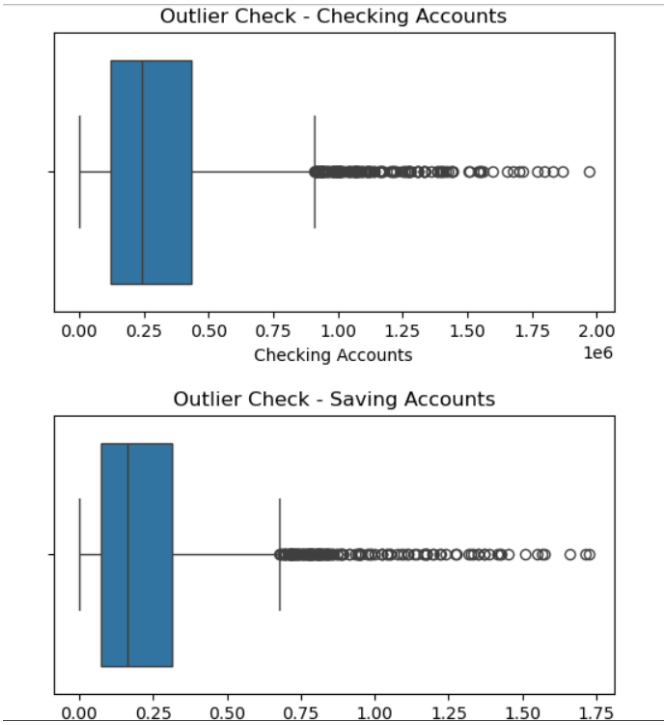
- Estimated Income and Superannuation Savings show a positive correlation with Deposits.
- Credit Card Balance and Bank Loans are moderately correlated.



- Outliers were identified in checking, saving accounts and loan values, typical for high-net-worth clients.

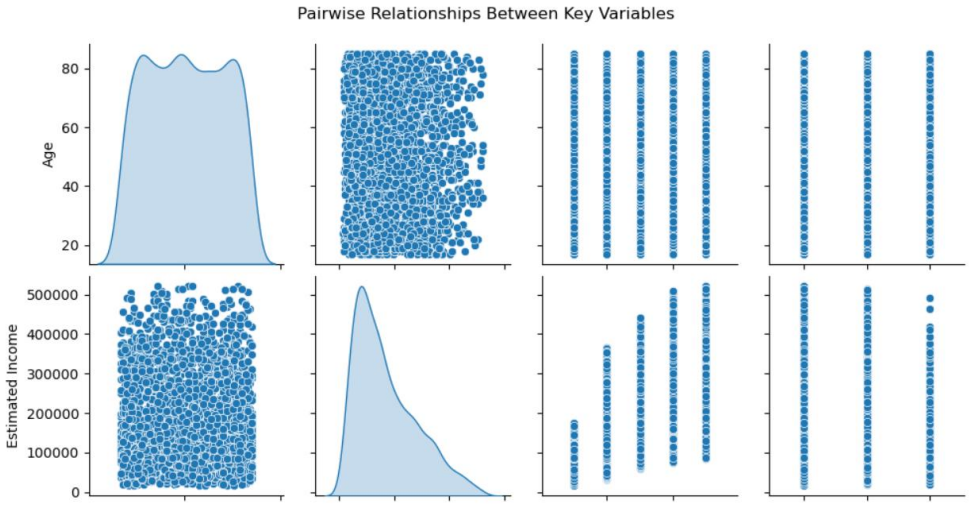


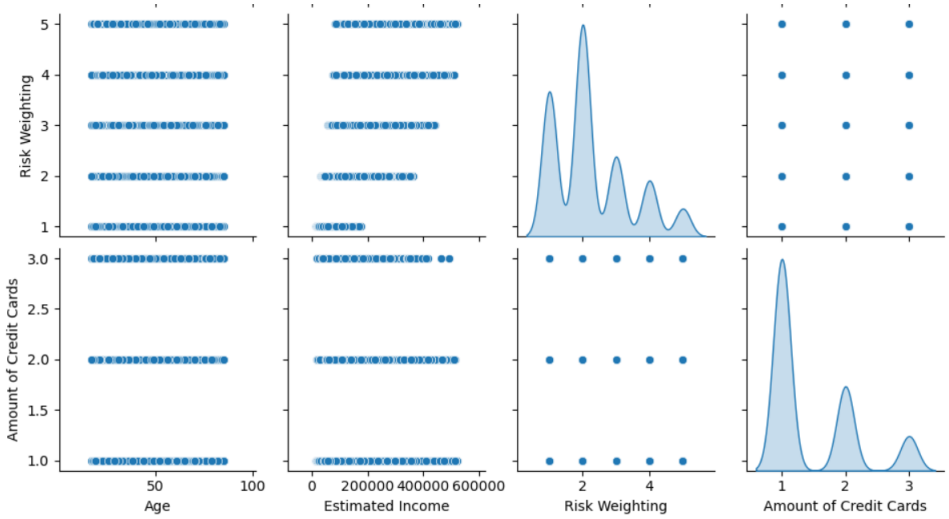


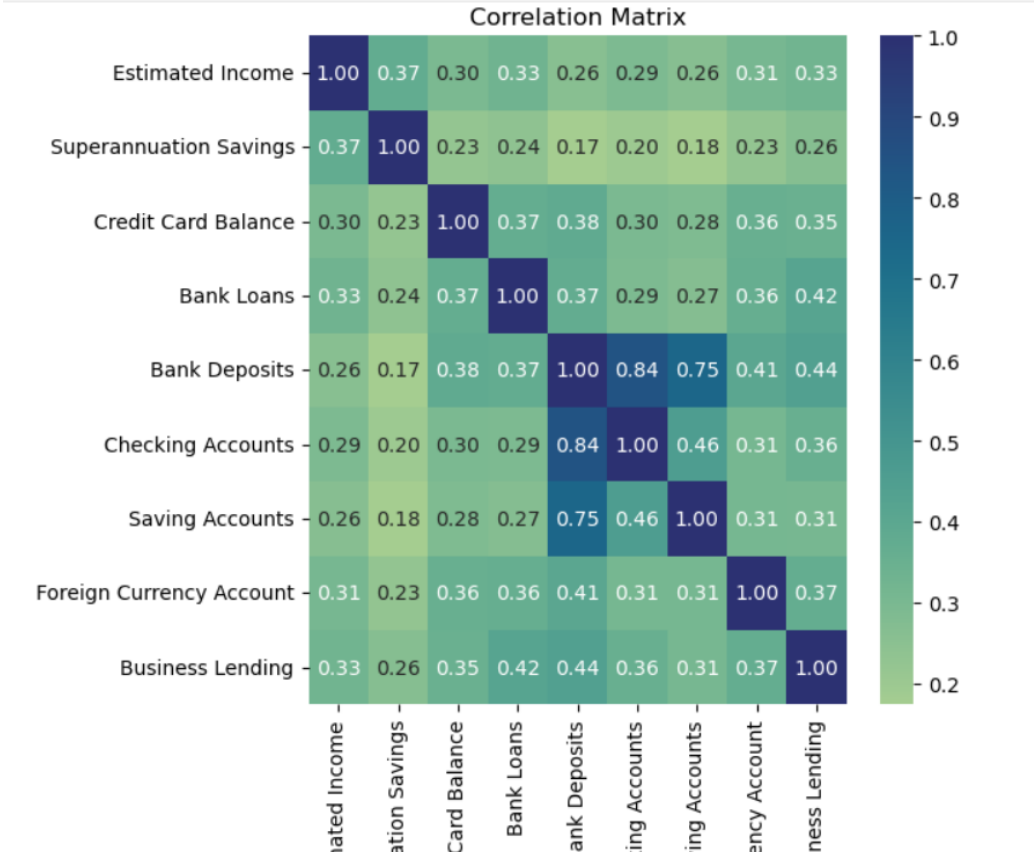


Correlation Heatmap

- Strong relationship observed between Deposits, Loans, and Income.
- Weak correlation between Savings Accounts and Business Lending, showing separate client bases.







EDA Summary:

- **Data Quality**: No missing values; columns cleaned and renamed for clarity.
- **Categorical Trends**: Majority male clients; Retail and Institutional banking dominate. Private Banking likely serves high-income individuals.
- **Numerical Patterns**: Income, deposits, and savings are right-skewed. Income and deposits show strong correlation; credit card balance moderately relates to loans.

- **Outliers & Correlation**: High-value outliers in income and deposits. Financial metrics are interrelated, while demographics show weak correlation — behavior drives financial patterns.

Key Takeaways:

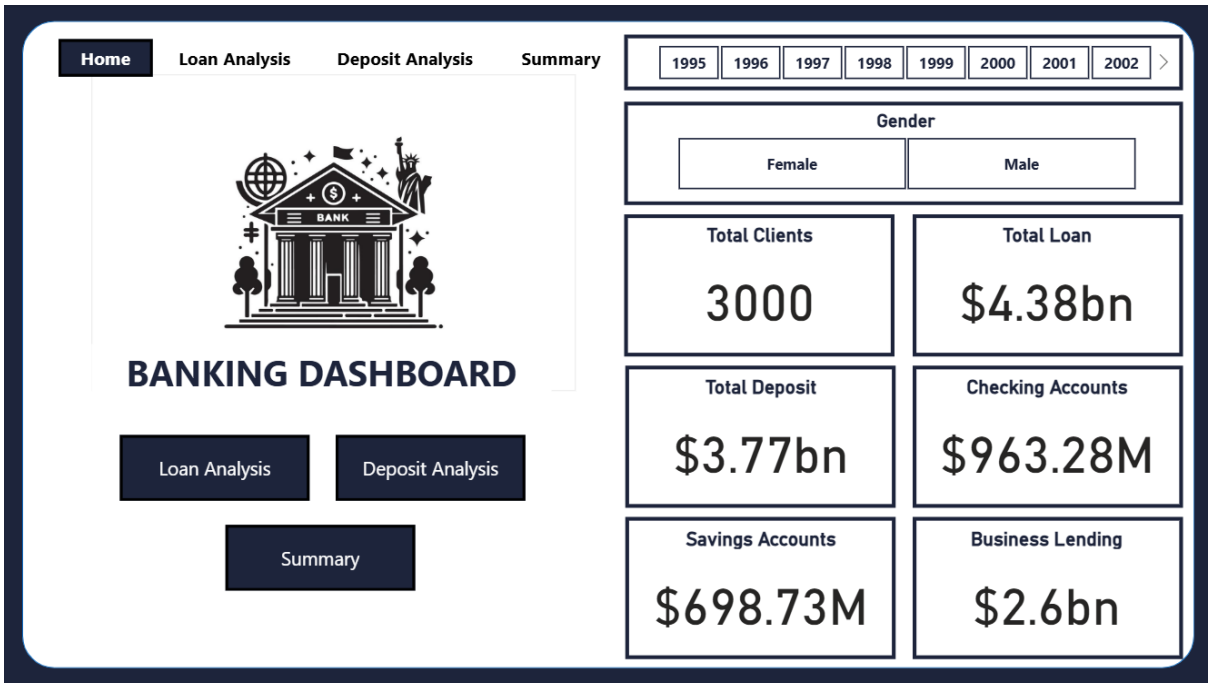
1. Retail and Institutional segments form the core customer base.
2. High-income clients drive most deposits and savings.
3. Financial behavior is shaped more by income than gender.
4. Clean, reliable dataset — ready for modeling and segmentation.

4. Power BI Dashboard Insights

Home Page Overview

- Total Clients: 3,000
- Total Loan: $4.38B
- Total Deposit: $3.77B
- Business Lending: $2.6B
- Savings Accounts: $698.7M
- Checking Accounts: $963.2M

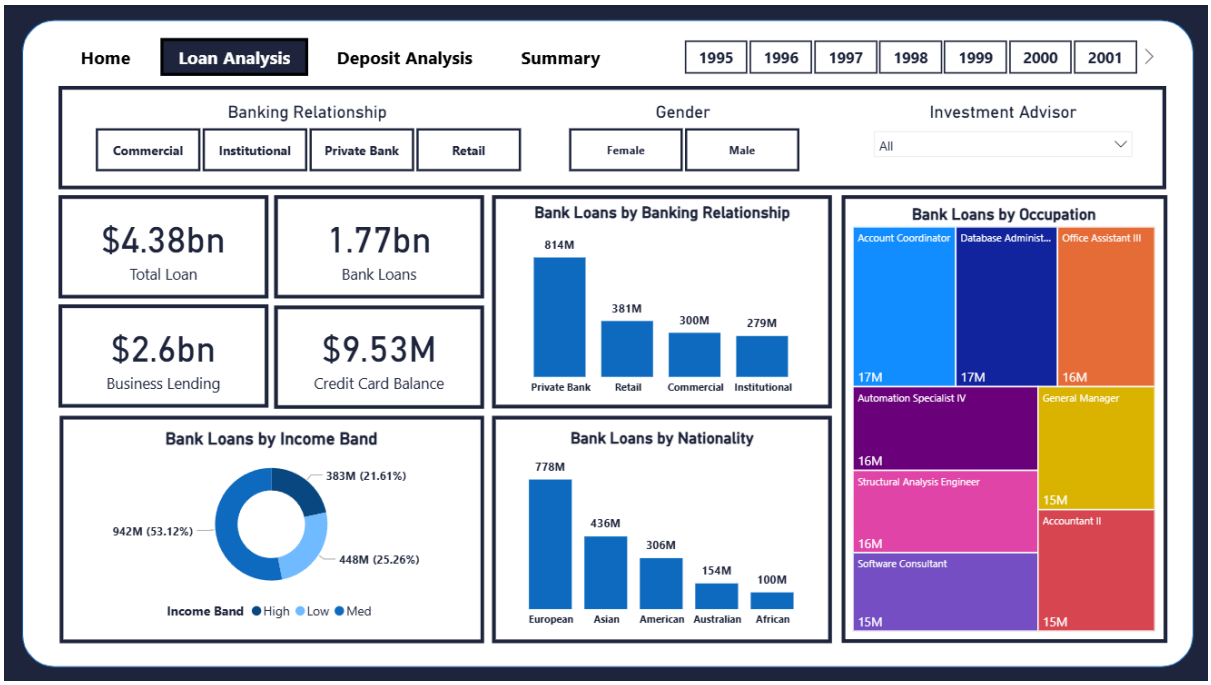
Filters allow insights by Year, Gender, Banking Relationship, and Investment Advisor.



5. Loan Analysis Dashboard

Key Insights:

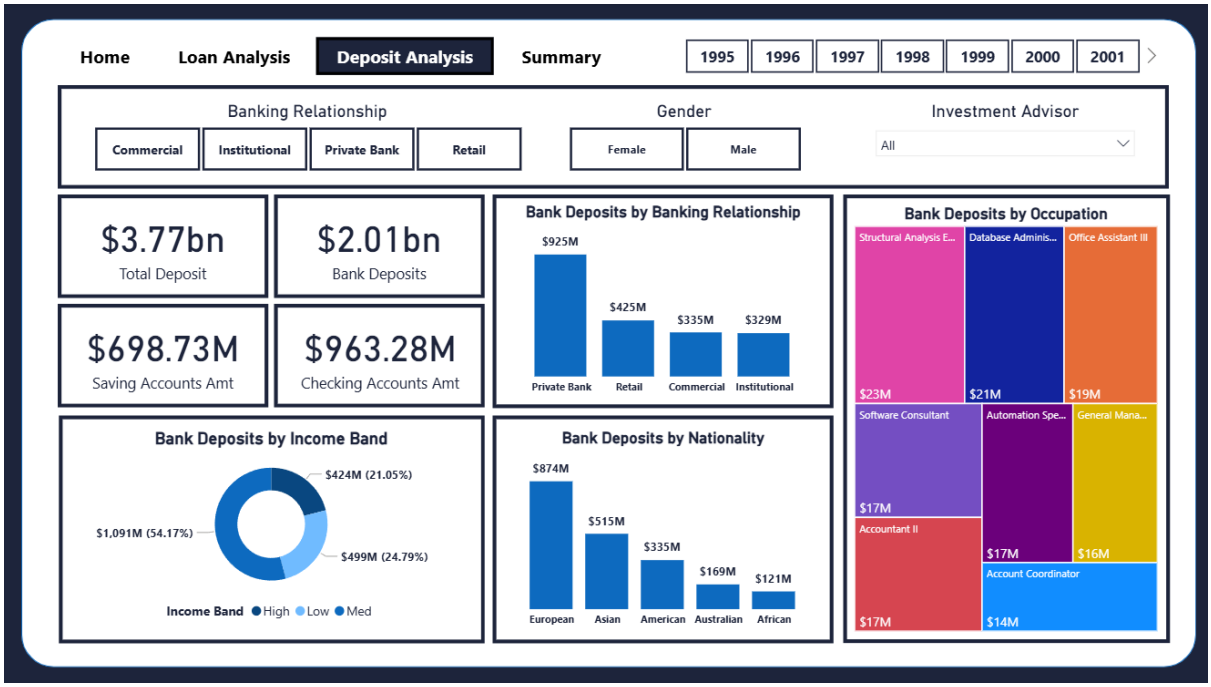
- Private Bank clients hold the largest share of loans ($814M).
- Medium-income group dominates lending (53% of total).
- Europeans contribute the most to loan value ($778M).
- Top occupations include Account Coordinators, Database Admins, and Office Assistants.



6. Deposit Analysis Dashboard

Key Insights:

- Private Bank leads in deposits ($925M).
- Medium-income clients contribute 54% of total deposits.
- European clients again dominate with $874M in deposits.
- Consistency in top contributing occupations between loans and deposits suggests stable, loyal clients.

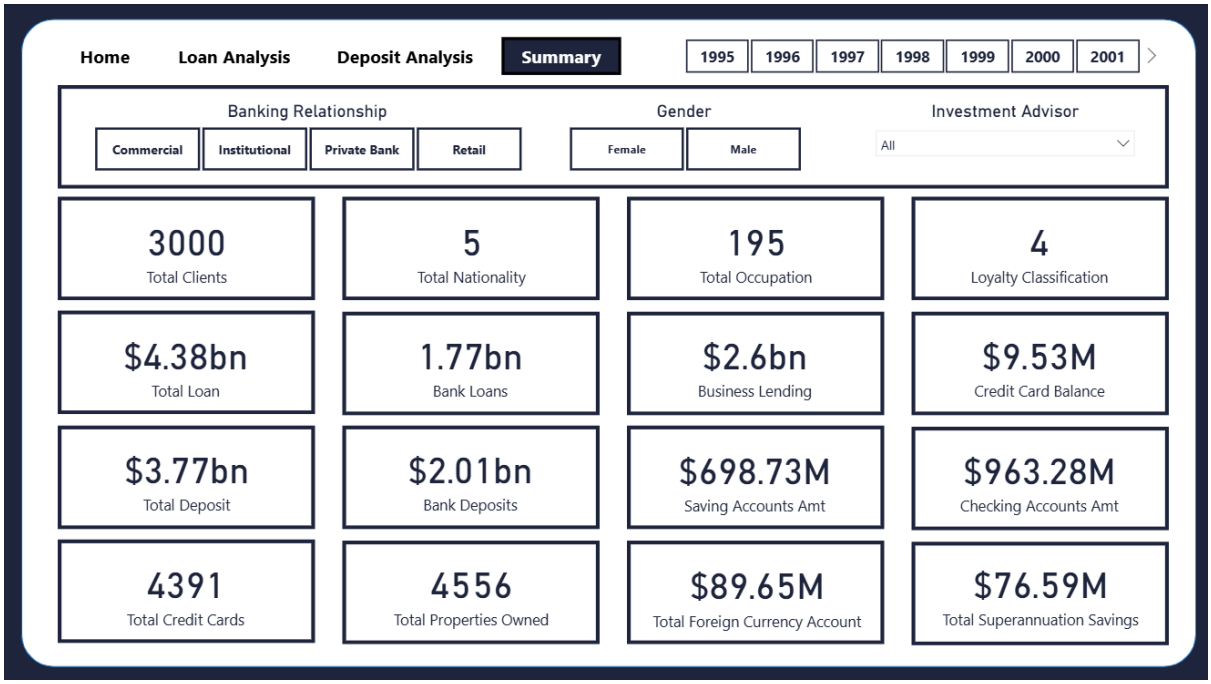


7. Summary Dashboard

Comprehensive Overview:

- 5 Nationalities and 195 Occupations represented.
- 4 Loyalty Classifications tracked.
- 4,391 Credit Cards and 4,556 Properties Owned.
- Foreign Currency Accounts: $89.65M
- Superannuation Savings: $76.6M

This unified summary helps monitor total assets and liabilities by category and segment.



8. Business Insights

- **Private Banking** clients are the most valuable segment in both loans and deposits.
- **High-income European clients** form the bank's core customer base.
- **Gender balance** is fairly even, showing no major bias in financial behavior.
- **Investment Advisors** play a key role — some manage significantly higher-value clients.
- **Credit card and business lending growth** present cross-selling opportunities.

9. Tools & Technologies

Step	Tool Used	Purpose
Data Merging	Excel	Combined sheets
Data Storage	MySQL	Centralized database
Data Analysis	Python (Pandas, Seaborn, Matplotlib)	EDA & Cleaning
Visualization	Power BI	Dashboard creation

10. Conclusion

This project successfully demonstrates how integrated data analysis using Python, MySQL, and Power BI can deliver actionable insights into customer banking behavior. Through a complete end-to-end workflow — from data preparation to interactive visual dashboards — the analysis transformed raw banking data into clear, decision-ready insights. The visualizations and dashboards effectively highlight key trends in client segments, income patterns, and advisor performance, supporting data-driven strategies for growth and customer retention.

The findings enable data-driven decision-making for:

- Targeted marketing
- Advisor performance tracking
- Product optimization
- Customer retention strategies