Financial Risk Analysis

This project focuses on Bank Risk Analysis, aiming to evaluate the probability of a client representing a financial risk for the bank.
It addresses key questions such as:
- Will the customer default on a loan?
- Does the client have enough assets to cover liabilities?
- Is the financial behaviour stable or risky?
- What is the exposure across financial products?

By answering these questions, financial institutions can make better decisions on:
- Loan approvals
- Credit card limits
- Customer loyalty strategies
- Interest rates & loan conditions

▪ Project Objective

The goal is to enable banking institutions to minimize the risk of financial losses while develop an end-to-end risk analytics solution using:

- Python → Data cleaning, feature engineering, exploratory analysis
- Power BI → KPI tracking, interactive dashboards, creditworthiness evaluation.

▪ Project Structure

1. Data loading & exploratory data analysis (EDA)

- About dataset

This comprehensive banking dataset integrates information from multiple source tables including banking relationship, client-banking, gender, investment advisor, and period data. The consolidated dataset provides a holistic view of client profiles, financial behaviors, and banking relationships.

- Data quality assessment

- Null & error validation: systematic identification and handling of missing or erroneous data points.
- Date conversion: standardization of joined bank dates, total loans, total assets, Income band, Risk level for temporal analysis.
- Variable typing: proper classification of numerical and categorical variables for appropriate analysis

2. descriptive analysis & customer segmentation

- Demographic distribution analysis

- Age distribution and generational cohort analysis.
- Income level segmentation and stratification.
- Account type prevalence across client base.
- Property ownership patterns and correlations.

- Customer segmentation framework

- Age-based segmentation: generational behavior patterns.
- Income-tier analysis: financial behavior across income brackets
- Risk profiling: client classification by risk appetite and financial health
- Loyalty segmentation: behavior analysis by banking relationship duration

3. risk metric development & scoring system

- Risk assessment framework

Development of a comprehensive risk scoring system based on:

- Debt-to-income ratio analysis
- Loan-to-income ratio analysis
- Loan-to-deposit ratio analysis
- Asset-to-liability ratios
- Banking product engagement levels

- Target variable construction

creation of a composite risk score incorporating:

- Financial leverage indicators
- Payment behavior metrics
- Product concentration risk
- Demographic risk factors

- ## <u>Key Performance Indicators (KPIs)</u>

- Customer Metrics

- Total Client Count: Overall customer base size
- Average Estimated Income: Mean income across all clients
- Average Tenure: Mean duration of banking relationship (years)
- Age group segmentation
- Nationality composition
- Loyalty Tier Distribution: Client classification by loyalty status

- Product Engagement Metrics

- Active Loan Penetration: Percentage of clients with outstanding loans
- Average Credit Card Balance: Mean credit card debt per client
- Product Portfolio Density: Average number of banking products per client (accounts, loans, cards)
- Product Mix Analysis: Distribution across account types (checking, savings, foreign currency)

- Risk Exposure KPIs

- High-Risk Client Percentage: Proportion of clients classified as high-risk
- Regional Risk Concentration: Risk score distribution by geographic region
- Debt-to-Deposit Ratio: Average loan-to-deposit ratio across client base
- Credit-to-Income Ratio: Total credit exposure relative to income
- Property Ownership Gap: Percentage of clients without real estate assets

---

▪ Advanced Risk Indicators

- Loan-to-Income Ratio (LTI): The Loan-to-Income Ratio is a critical financial metric used to assess an individual's debt burden relative to their earnings. It measures the proportion of income that is used to service debt obligations.

LTI = [bank_loans] / [estimated_income]

- Financial Reference – Typical LTI Thresholds

Note: Thresholds vary slightly by country and institution. These are broad guidelines used in mortgage and consumer-credit risk assessments in markets such as the U.S., U.K., and Canada.

LTI Range	General Interpretation	Credit Risk
< 3	Very healthy. Debt is less than three times annual income.	Low
3 – 4.5	Acceptable and generally sustainable for mortgages/loans.	Moderate
4.5 – 6	Caution zone: leverage is rising and requires closer monitoring.	Medium–High
> 6	High leverage; often classified as "high LTI" by regulators.	High
> 10	Extreme leverage; significant default probability.	Very High

- Loan to deposit ratio (LDR): the loan to deposit ratio shows how much the client owes in loans relative to what they have deposited in the bank.
LDR = ['bank_loans'] / ['bank_deposits']

- Financial Reference – Typical LDR Thresholds

Note: These guidelines are based on banking and credit-risk practices used by financial institutions and regulators (values expressed as a ratio, not percentage).

LDR Range	General Interpretation	Credit Risk
< 0.8	Strong liquidity: loans are well covered by deposits.	Low
0.8 – 1.0	Optimal operating zone for most banks/clients.	Low–Moderate
1.0 – 1.2	Acceptable but watch for rising leverage.	Moderate
1.2 – 2.0	High lending relative to deposits; monitor funding stability.	Medium–High
> 2.0	Very high leverage; potential liquidity or repayment concerns.	High

- **Debt-to-Income Ratio (DTI):** The Debt-to-Income Ratio measures how many years (or annual income multiples) it would take someone to pay off all their debt if they used their entire annual income for debt repayment. It's a measure of debt burden relative to annual earning capacity.

DTI = ['Total_debts'] / ['estimated_income']

DTI Range	DTI Annual	Interpretation	Risk Level
< 0.36 (36 %)	**<4.3**	Generally healthy; typical mortgage approval threshold.	Low
0.36 – 0.49	**4.3 – 5.9**	Manageable but needs monitoring.	Moderate
0.50 – 0.75	**6.0 – 9.0**	High debt burden: lenders become cautious.	High
> 0.75	**> 9.0**	Extremely high risk; very limited capacity for new credit.	Very High

- Debt-to-Asset Ratio (DTA): The Debt-to-Asset Ratio is a financial metric that measures the proportion of a client's total assets that are financed by debt. It indicates the degree of leverage and financial risk a client is taking.

DAR = ['Total_debts'] / ['total_liquid_assets']

- **Financial Reference – Debt-to-Asset Ratio Guidelines**

Used in corporate finance and personal credit assessments

DAR Range	General Interpretation	Risk Level
< 0.3	Very conservative capital structure; strong ability to cover liabilities.	Low
0.3 – 0.6	Healthy balance; manageable debt load.	Low–Moderate
0.6 – 1.0	Leveraged but typically acceptable for many industries or individuals.	Moderate
1.0 – 2.0	Assets are less than or equal to debt; liquidity pressure likely.	High

DAR Range	General Interpretation	Risk Level
> 2.0	Debt far exceeds asset value; potential insolvency or severe stress.	Very High

---

▪ Possibles Insights from the dataset

- Loan-to-Income Ratio (LTI)

- **Mean: 4.72** – on average, clients hold loans equal to about 4.7 times their annual income.
- **Median: 3.34** – half of the clients have an LTI below 3.3, showing that most maintain moderate debt levels.
- **Range:** from **0** (no debt) up to **41.2** (extreme outliers with debt 41 times income).
- **75 %** of clients have an LTI ≤ **6.1**, indicating that credit risk is concentrated in a small, highly leveraged group.
- This pattern suggests **credit risk is driven by a limited set of outliers** rather than the overall portfolio.

Recommendations

- Segment and monitor clients with LTI > 6 for proactive risk-mitigation strategies.
- Tighten lending criteria or require stronger collateral for new loans in this segment.

- Loan-to-Deposit Ratio (LDR)

- **Mean: 1.87** – on average, total loans are about **187 %** of deposits.
- **Median: 0.99** – half of the clients maintain loans roughly equal to their deposits, indicating a balanced core group.
- **Range:** from **0** (no loans) up to **27.78** (extreme outliers where loans are 27× deposits).
- **75 %** of clients have an LDR **≤ 1.95**, suggesting most stay within a sustainable range, while a small minority drives the very high mean.
- This shows **healthy deposit backing for the majority**, but a **tail of highly leveraged customers** that may require tighter credit controls.

Recommendations

- **Focus monitoring** on clients with **LDR > 2.0** as they represent the highest funding and credit risk.
- For strategic lending, **encourage deposit growth** or require additional collateral for clients above the 1.2–2.0 zone.

- **Debt-to-Income Ratio (DTI)** was calculated on an **annual** basis. Standard monthly benchmarks were multiplied by 12 to define risk bands.

- **Mean: 11.62** – on average, clients carry **over 1,160 %** of their monthly income in debt payments, which is extremely high.
- **Median: 8.66** – half of the clients spend almost **nine times** their income servicing debt, showing a highly leveraged population.
- **Interquartile Range (IQR): 5.24 – 14.65** – even the middle 50 % of clients exceed common affordability standards.
- **Max: 93.52** – some clients' debt obligations are **ninety-plus times** their income, indicating critical risk.
- Compared with standard lending benchmarks, nearly the entire client base demonstrates **severe over-extension**, signaling a high probability of default if additional credit is granted.

Recommendations

- **Flag** all clients with **DTI > 0.75** (75 %) as critical; in your data, that's essentially **everyone**.
- Focus on combined metrics (e.g., **Risk Level** and **Debt-to-Asset Ratio**) to identify the rare exceptions who might still qualify for credit.

- Debt-to-Asset Ratio (DAR)

- **Mean: 3.15** – on average, clients hold **315 % more debt than assets**, but this mean is heavily influenced by outliers.
- **Median: 1.23** – half of the clients owe slightly more than their total assets, indicating a moderate leverage profile for the core population.
- **Interquartile Range: 0.77 – 2.30** – most clients stay within this band, suggesting a reasonable debt position for the majority.
- **Max: 544.44** – extreme outliers signal clients whose debt is hundreds of times their asset value; these are significant risk flags.
- The typical client is moderately leveraged, but the **long right tail** of the distribution reveals a small group of highly indebted customers who could pose severe credit risk.

Recommendations

- **Flag** any client with **DAR > 2.0** for immediate review.
- Investigate the few cases with ratios above **10**, as these represent extraordinary leverage and may require special credit terms or denial of new credit.

▪ Python Data Preparation & Exploratory

Before building the Power BI dashboards, all raw banking data was **cleaned, transformed, and merged using Python** (Pandas, NumPy, and Matplotlib/Seaborn). Key steps included:

- **Data integration:** Multiple relational tables were combined into a single, well-structured dataset for analysis.
- **Pre-visualization:** Histograms, boxplots, and heatmaps provided an initial understanding of distributions and outliers**.**

- **Correlation analysis:** Python's statistical libraries helped identify significant relationships between financial ratios (e.g., Loan-to-Income vs. Debt-to-Income), which guided the selection of KPIs and the design of the Power BI dashboards.

These Python insights **shaped the business rules** and thresholds later implemented in DAX for the interactive Power BI reports.

Power BI Dashboard & Insights

After data preparation in Python, **Power BI** was used to design an interactive, multi-page dashboard**.** It combines the cleaned dataset with custom DAX measures to deliver real-time risk evaluation and client segmentation.

- **Client Demographics & Portfolio**
From Page 1

- Customer base is **balanced by gender**, but income varies by region, with North America showing the highest average income (~$178 K)**.**
- Loyalty tiers reveal that **silver and gold** clients form more than **70 %** of the portfolio, useful for targeting retention campaigns.

- **Financial Products & Risk**
From Page 2

- **Loans outweigh deposits** in high-risk groups: High-risk clients hold **over $2.4 B** in loans compared to only **$1.0 B** in deposits**,** signalling liquidity concerns.
- Certain products—**business lending and credit cards**—are more concentrated among high-risk clients, guiding product-level risk strategies.

- **Creditworthiness Evaluation**
From Page 3

- Final classification shows **35 % Approved, 23 % Rejected, and 41 % Review** using combined ratios (LDR, LIR, DTI, DAR).
- **Critical clients** are flagged when **Debt-to-Asset Ratio ≥ 5**, allowing instant drill-down for underwriting teams.
- Low-risk clients with property ownership (≈1,574 customers) represent the best opportunity for cross-selling and preferential credit terms.

- Credit Metrics Interpretation: The column 'Credit_Decision' classifies each client as Approved**,** Review**,** or Rejected based on four key risk ratios derived from the dataset:

Metric	Why It Matters	Threshold Choice
Loan-to-Deposit Ratio (LDR)	Measures how much the client borrows relative to their deposits. A high value suggests low internal liquidity.	**< 2 (Approved)** keeps the client's borrowing comfortably below their own deposits (≈75th-percentile of our data = 1.95). **≥ 5**

Metric	Why It Matters	Threshold Choice
		(Rejected) flags extreme leverage (≈beyond 98th-percentile).
Loan-to-Income Ratio (LTI)	Captures how many times the client's annual income equals their total loans.	**< 6 (Approved)** aligns with our 75th-percentile (≈6.1). **≥ 10 (Rejected)** is well above the 90th-percentile and indicates very heavy debt vs. income.
Debt-to-Income Ratio (DTI – annual)	Shows total debt obligations relative to annual income. Industry monthly guidance (< 0.36) translates to annual < 4.3 for "low risk." Because our median was 8.7, we set **< 8.7 (Approved)** near our dataset's median and **≥ 20 (Rejected)** to capture extreme cases (≈top 10%).	
Debt-to-Asset Ratio (DAR)	Indicates what portion of assets is financed by debt.	**< 2.3 (Approved)** matches the 75th-percentile. **≥ 10 (Rejected)** isolates clients with very high leverage or minimal assets (well beyond normal ranges).

- The DAX formula:

Credit_Decision =

```
SWITCH(
   TRUE(),
   -- Approved: all key ratios within conservative bands
   Banking_final[Loan_to_Deposit_Ratio] < 2
      && Banking_final[Loan_to_Income_Ratio] < 6
      && Banking_final[Debt_to_Income_Ratio] < 8.7
      && Banking_final[Debt_to_Asset_Ratio] < 2.3,
   "Approved",

   -- Rejected: any single ratio exceeds high-risk thresholds
   Banking_final[Loan_to_Deposit_Ratio] >= 5
      || Banking_final[Loan_to_Income_Ratio] >= 10
      || Banking_final[Debt_to_Income_Ratio] >= 20
      || Banking_final[Debt_to_Asset_Ratio] >= 10,
   "Rejected",

   -- Otherwise, send for manual review
   "Review"
)
```

- **Conclusion**

Empowered by the latest data visualization techniques, Power BI dashboards are among the most effective resources for using in banking sector. As outlined in this write-up, a banking operations dashboard in Power BI can be developed with key banking related metrics and KPIs.

- How to Use the Dashboard for Future Projects

- **Scenario Testing:** Adjust DAX thresholds (e.g., Debt-to-Income or Loan-to-Deposit limits) to simulate new lending policies.
- **Portfolio Monitoring:** Apply date filters or incremental data refresh to track risk migration over time.
- **Cross-department Collaboration:** Export visuals or embed the report in Power BI Service to share insights with credit, marketing, and compliance teams.
- **Reusable Template:** The same model—Python for data prep + Power BI for visualization—can be adapted to other domains such as insurance risk, retail credit scoring, or investment analysis.

- Personal Takeaway

Completing this project has been a significant personal achievement. Designing meaningful KPIs for bank risk analysis, integrating Python for data preparation, and building an interactive Power BI dashboard gave me hands-on experience with real financial data challenges.

This work has sharpened my analytical thinking, strengthened my technical skills in both Python and Power BI, and deepened my understanding of risk evaluation in the banking industry. It represents a clear step forward in my professional growth as a data analyst and demonstrates the value I can bring to future financial analytics projects.