Case Study — Credit Risk Analysis Capstone
(Tools: SQL · Excel · Power BI)

🎯 Project Objective

The goal of this capstone project was to identify the key factors influencing loan default risk using an integrated analytics pipeline.
The workflow combined SQL for data engineering, Excel for validation and exploratory analysis, and Power BI for interactive visualization and insight storytelling.
The end objective: provide actionable recommendations to reduce portfolio default rates and strengthen lending decisions.

🧮 Data & Methodology

Dataset:
Credit risk dataset (~30 000 records) containing borrower demographics, loan details, and repayment outcomes.

Workflow Overview:

1. **SQL (ETL & Analysis):**
   - Created staging → clean tables; standardized datatypes & handled “unknown” values.
   - Generated 10 analytical queries addressing risk by grade, income, ownership, etc.
   - Exported each result as UTF-8 CSV for Excel validation.
2. **Excel (Validation & Static Dashboard):**
   - Built PivotTables for each query to confirm SQL outputs.
   - Assembled a summary dashboard of key KPIs (defaults %, avg loan amount, avg interest rate).
3. **Power BI (Visualization & Insights):**
   - Imported clean dataset
   - Created DAX measures:
     {Total Loans, Total Defaults, Default Rate %, Avg Loan Amount, Avg Interest Rate, Avg Income }
   - Built an interactive dashboard with slicers (grade, intent, ownership) and an “Insights & Recommendations” page.

🔟 Key Analytical Questions

1. Total loans & overall default rate.
2. Default rate by loan grade.
3. Default rate by home ownership.
4. Default rate by employment length.
5. Default rate by loan intent.
6. Income vs default behavior.
7. Top risky segments (ownership × grade).
8. Loan amount distribution by grade.
9. Credit history length vs default rate.
10. Average interest rate by loan status.

📊 Findings

- Overall default rate ≈ 9 %; Grades E–F exceed 15 %.
- Renters and low-income borrowers show the highest default probability.
- Shorter employment and credit history durations increase risk.
- Personal and Medical loan intents have above-average defaults.
- Higher interest rates correlate with greater default likelihood.

💡 Recommendations

1. Focus growth on Grade A–C portfolios to balance yield vs risk.
2. Tighten underwriting for high-interest and short-tenure applicants.
3. Offer financial literacy and repayment assistance to renters & low-income groups.
4. Incorporate credit history length and employment stability in internal risk scoring models.
5. Monitor portfolio mix monthly; maintain default rate < 8 %.

🧠 Impact

By prioritizing low-risk loan grades and improving borrower assessment, the analysis suggests a potential 2–4 pp reduction in default rate within the first review cycle.
The project demonstrates full-cycle analytical capability — from data extraction and cleaning in SQL to insight communication in Power BI.

🪪 Deliverables

Component	File
SQL scripts	`Credit_Risk_Quesries.sript.sql`, `credit_clean_text.sql`,
CSV outputs	`credit_clean_text.csv`, `credit_risk_raw.csv`,
Excel validation	`Credit_Risk_Analysis.xlsx`
Power BI validation	`Credit_Risk_Analysis.pbix`
Case Study (PDF)	`credit_risk_case.study.pdf`
Screenshots	`dashboard_main.png`