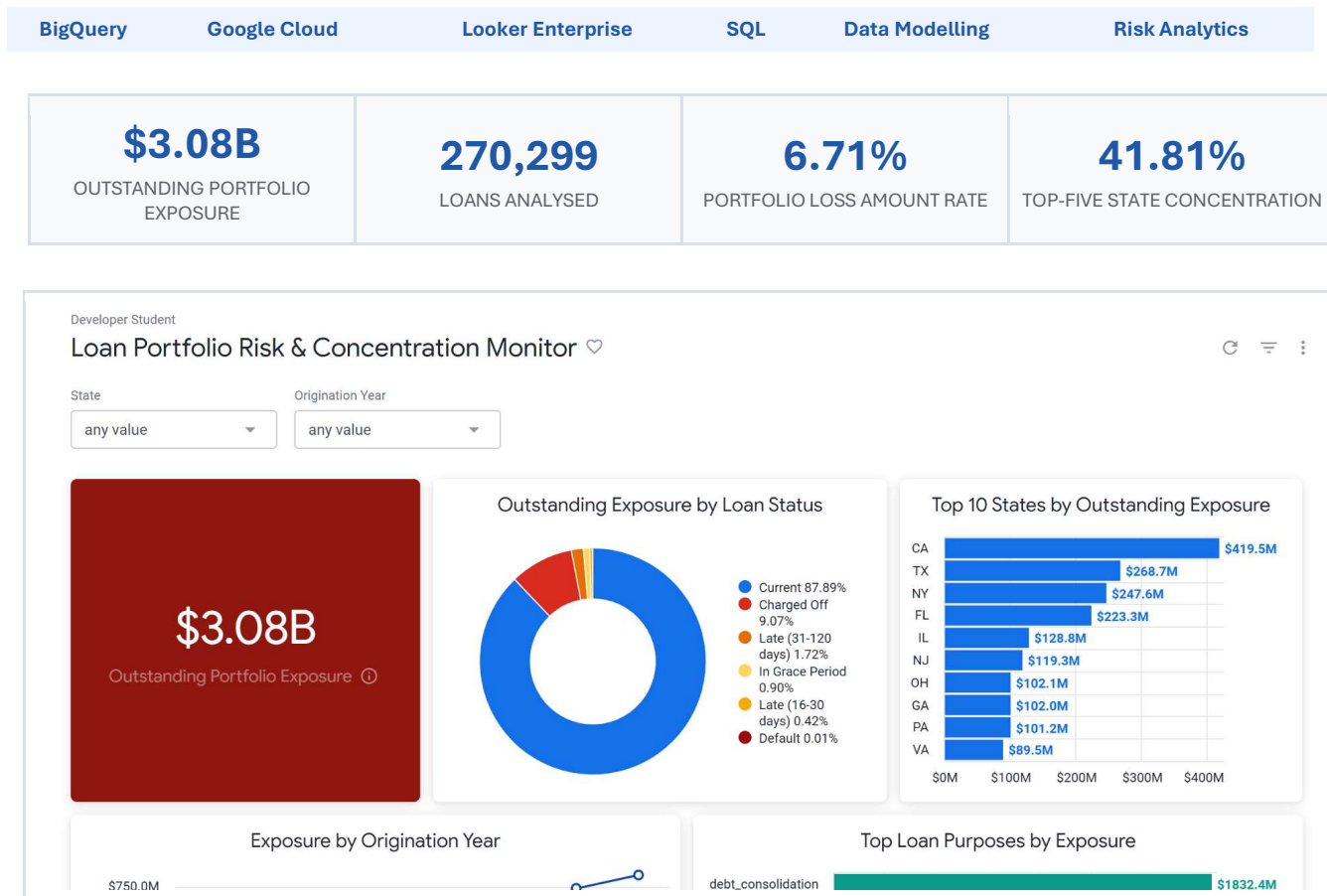


Loan Portfolio Risk & Concentration Monitor

An end-to-end cloud analytics case study using BigQuery and Looker Enterprise to transform lending data into a reusable risk mart, executive metrics and an interactive monitoring dashboard.



Final Looker dashboard: portfolio exposure, status composition and state concentration.

From temporary cloud data to an executive monitoring product

The project simulates a Treasury analytics assignment for a growth-stage fintech lender. The objective was to build a reliable analytical layer, validate the source data and translate the results into a dashboard that supports exposure, concentration and descriptive credit-risk monitoring.

01 Monitor exposure Quantify unpaid balances, compare them with an internal threshold and track how exposure changes across origination years.	02 Detect concentration Measure geographic dependence by state and region so large concentrations can be reviewed before they become systemic vulnerabilities.	03 Prioritise review Compare loan status and purpose to separate high-volume concentration from segments with higher observed loss rates.
--	--	---

DATA MODEL

Source	Relevant fields	Analytical role
fintech.loan	loan_id, loan_status, loan_amount, state, int_rate, issue_year, nested application.purpose	Primary loan-level fact table
fintech.customers	customer_id, annual income, home ownership and borrower attributes	Customer-level Looker segmentation
state_region CSV	state, subregion, region	Geographic enrichment and concentration analysis
Derived marts	risk flags, rates, ranks, quartiles and executive summaries	Reusable reporting layer

1. Inspect Schema, data types, nested records and source coverage.	2. Validate Rows, distinct keys, missing values and mapping quality.	3. Transform Status flags, nested extraction and geographic joins.	4. Analyse Year, status, purpose, state and regional metrics.	5. Activate Interactive Looker dashboard, filters and cross-filtering.
--	--	--	---	--

Validation result: 270,299 source rows matched 270,299 distinct loan IDs. The derived risk mart preserved every row and returned no missing region mapping or unknown purpose values.

Scale accelerated while most outstanding exposure remained current

193,938

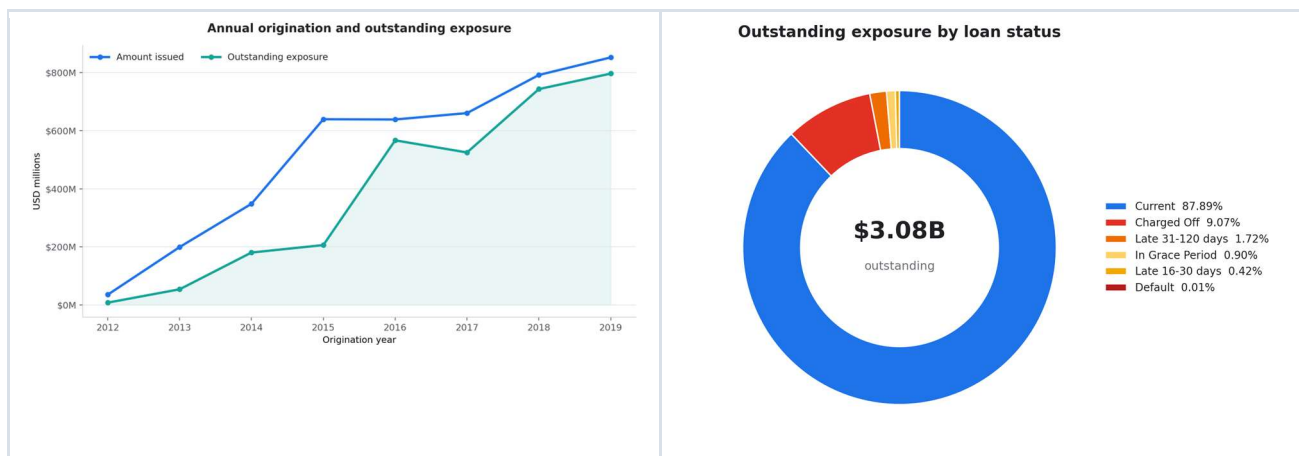
OUTSTANDING LOANS

\$279.67M

LOSS-CLASSIFIED AMOUNT

13.08%

AVERAGE INTEREST RATE



Growth pattern: Outstanding exposure increased from \$206.15M for 2015 originations to \$566.84M for 2016, then reached \$796.72M in 2019.

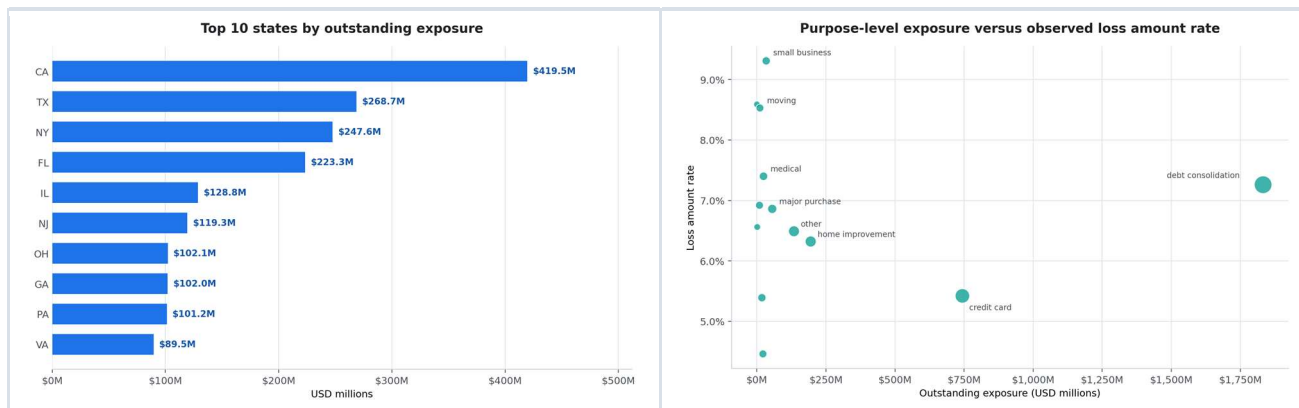
Threshold signal

The dashboard uses a \$3.00B internal monitoring threshold. The observed \$3.08B exposure is above that level, creating a clear executive alert without implying that all outstanding loans are impaired.

Status interpretation

Current loans account for 87.89% of outstanding-status exposure. Charged-off exposure remains material at 9.07%, while grace and late-payment categories form a smaller but operationally relevant delinquency pipeline.

High-volume exposure and high-rate loss segments require different controls



Geographic concentration

California contributes \$419.53M of outstanding exposure, or 13.62% of the portfolio. The five largest states together account for 41.81%.

Purpose-level divergence

Debt consolidation dominates outstanding exposure at approximately \$1.83B. Small-business loans have a much smaller balance but the highest observed loss amount rate among purposes with at least 100 loans: 9.31%.

Risk lens	Leading segment	Finding	Management implication
Portfolio concentration	California	\$419.53M outstanding	Review state-level concentration limits and stress scenarios.
Purpose concentration	Debt consolidation	\$1.83B outstanding	Monitor volume concentration and status migration.
Purpose loss rate	Small business	9.31% observed loss amount rate	Review underwriting and pricing separately from volume controls.
Loss category	Charged Off / Default	\$279.67M loss-classified amount	Track recoveries and vintage performance when event data is available.

RECOMMENDATIONS

Use a tiered monitoring framework rather than one portfolio-wide limit

A single exposure threshold is useful for executive alerting, but it is not sufficient for managing the different sources of risk identified in the analysis. Treasury should monitor aggregate exposure together with geographic concentration, origination vintage, status migration and purpose-level loss rates.

1

Investigate sustained threshold exceedance

Review the drivers of the \$3.08B outstanding balance and define escalation rules based on duration and rate of change, not only a one-time snapshot.

2

Monitor California and the top-five states

Introduce state-level concentration triggers and stress-test the effect of correlated deterioration in the largest geographic markets.

3

Separate volume controls from segment-risk controls

Treat debt-consolidation exposure as a concentration issue and small-business lending as a higher loss-rate segment requiring underwriting review.

4

Add vintage and migration monitoring

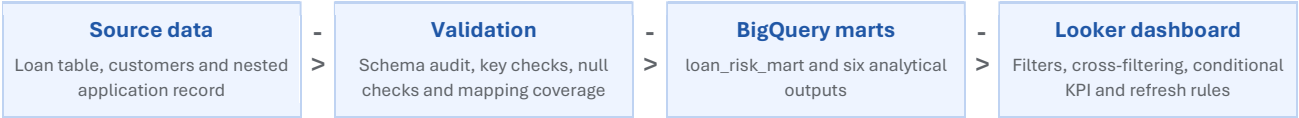
Extend the model with monthly status transitions, delinquency roll rates, recoveries and cohort repayment curves when event-level history becomes available.

DECISION FRAMEWORK

Portfolio level	Concentration level	Segment level
Total outstanding exposure, growth rate and threshold duration.	State shares, top-five concentration and regional stress scenarios.	Status migration, purpose loss rates and borrower composition.

Strategic conclusion: The Looker dashboard provides the executive monitoring surface, while the BigQuery marts preserve reusable and auditable analytical logic underneath it.

Reusable analytical architecture and demonstrated skills



SQL and modelling techniques	Delivered analytical outputs
Nested data: dot notation for application.purpose	loan_risk_mart Loan-level enriched model and risk flags
Reusable tables: CTAS and CREATE OR REPLACE TABLE	portfolio_executive_summary Portfolio KPIs
Window functions: LAG and cumulative SUM	loan_growth_by_year Annual growth and exposure
Quality assurance: row preservation and distinct-key checks	loan_risk_by_purpose Purpose volume and loss rates
Data integration: LEFT JOIN to state-region mapping	state_risk_watchlist Concentration and priority tiers
Safe metrics: SAFE_DIVIDE and explicit rate definitions	region_status_mix Regional status composition
Prioritisation: DENSE_RANK, ROW_NUMBER and NTILE	portfolio_key_findings Presentation-ready findings
Activation: dashboard filters and cross-filtering	

SCOPE LIMITATIONS

The dataset supports static exposure, concentration and status analysis, but it does not include full repayment-event history, recoveries, collateral values, macroeconomic variables or validated probability-of-default labels. This is a descriptive monitoring project, not a predictive credit-risk model.

Download package The portfolio page provides the complete BigQuery SQL scripts, derived CSV outputs and this executive summary.	Author Karla Ruiz de Matos Data analytics, SEO and multilingual content strategy.
---	--