CREDITCHITRA

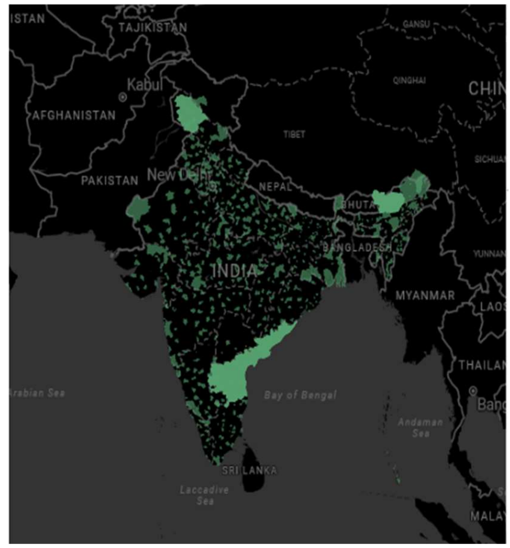
Geo-Analytic Lending Risk Profile

SHASHANK IYER

JULY 2025

Contents

Executive Summary

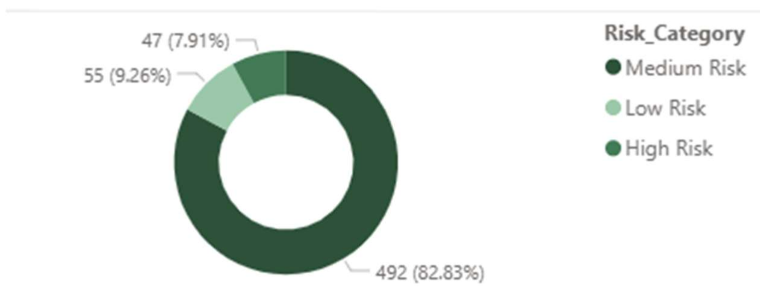


CreditChitra: District-Level Lending Risk Simulator for India is a geo-analytics-driven simulation tool designed to identify high-risk lending zones across Indian districts. Leveraging synthetic but statistically realistic indicators such as unemployment rate, credit penetration, economic vulnerability, and migration patterns, the tool provides a scalable lens for understanding geographic credit risk.

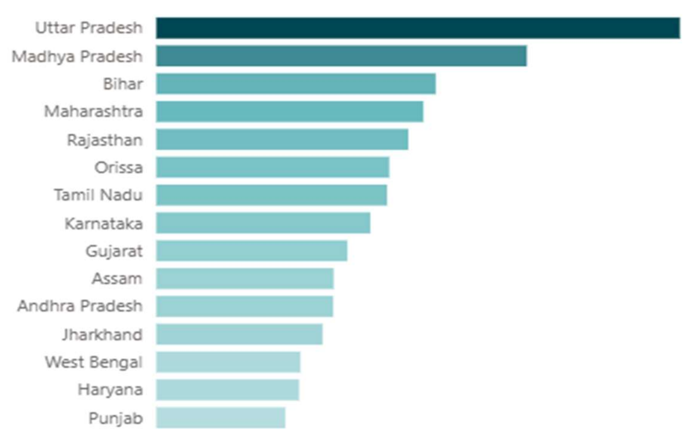
Why it Matters:
Lending decisions in India often overlook localized macro and socio-economic risk factors. This simulation framework enables:

- District-wise risk categorization
- Early warning systems for rural and semi-urban exposure
- Improved portfolio resilience through geo-targeted underwriting strategies

Key Takeaways:
This model equips lenders with a proactive framework to shift from reactive loss provisioning to **risk-informed regional pricing and capital allocation**.



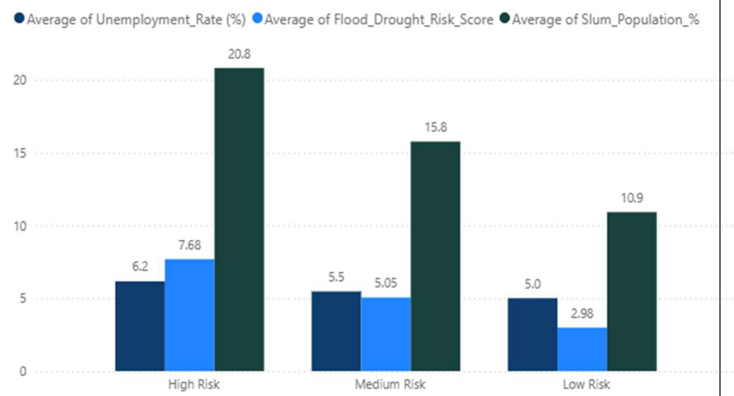
- Regional lending risk is asymmetric: high-concentration clusters identified for both elevated and minimized exposure.
- Features such as unemployment rate, climate vulnerabilities, and economic fragility decisively influence credit outcomes.
- Proposed risk scoring enables granular direction of capital and intervention, improving loss ratios and resource allocation.
- Live metrics and simulation outputs provide actionable levers for executive and risk teams.



Problem Statement & Business Context

Problem Statement:

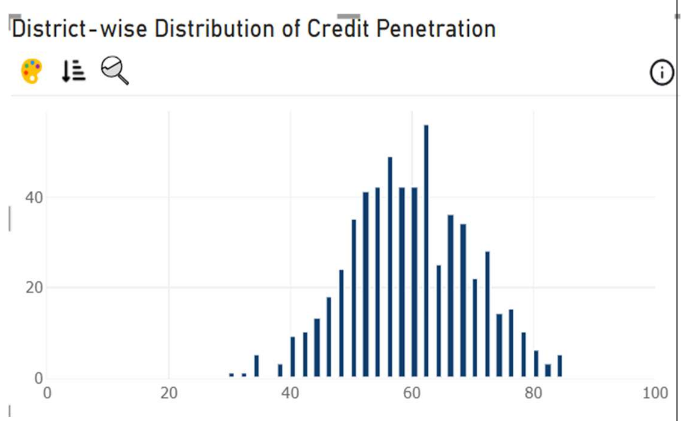
Traditional lending frameworks in India often miss the geographic and socioeconomic nuances that influence default risk. While credit scoring models are borrower-centric, they rarely account for regional stressors such as high unemployment or low credit access. CreditChitra bridges this gap, enabling data-informed branch-level decisions.



Strategic Need:

At financial institutions like J.P. Morgan or Goldman Sachs or their retail arms, portfolio resilience and geographic diversification are board-level mandates. There is growing demand for tools that can:

- Integrate district-level signals into the underwriting process.
- Simulate policy or macro stress (e.g., what if rural unemployment rises 2%)
- Optimize branch expansion or consolidation using multidimensional data



Use Case Extensions Identified:

Area	Impact
Rural Credit Expansion	Target low-risk, high-agri workforce districts for pilot programs
Network Strategy	Identify high-risk clusters for portfolio de-risking or reallocation
Credit Inclusion	Use slum %, credit penetration, and migration data to drive digital onboarding

Data Sources & Feature Engineering

Variable	Description	Category	Transformation/Use
District & State Codes	Unique geographic identities	Categorical	Grouped, deduplicated
Latitude/Longitude	Geospatial centroids	Numeric	Aggregated
Unemployment Rate (%)	Labour force without jobs	Macro	Risk-coded (Low/Med/High)
Flood/Drought Risk Score	Climate risk exposure (0-10 scale)	Climate	Risk-tiered
Agri Workforce (%)	Share in agriculture	Economic	Proxy for income stability
Crime Rate (per 1000)	Reported incidents per 1,000 people	Socioeconomic	Risk-coded
Slum Population (%)	Urban deprivation indicator	Socioeconomic	Categorized for risk
Credit Penetration (%)	Credit access rate	Financial	Reverse-scaled, categorized
Net Migration Rate (%)	Population inflow/outflow	Demographic	Risk-mapped by direction
Political Unrest Score	Social & political stability (0â€“10)	Political	Tiered risk mapping
Economic Vulnerability Score	Multi-factor fragility index (0â€“10)	Economic	Grouped for risk stratification
Historical NPA Rate (%)	Actual regional loan default rate	Model Target	Synthesized for simulation

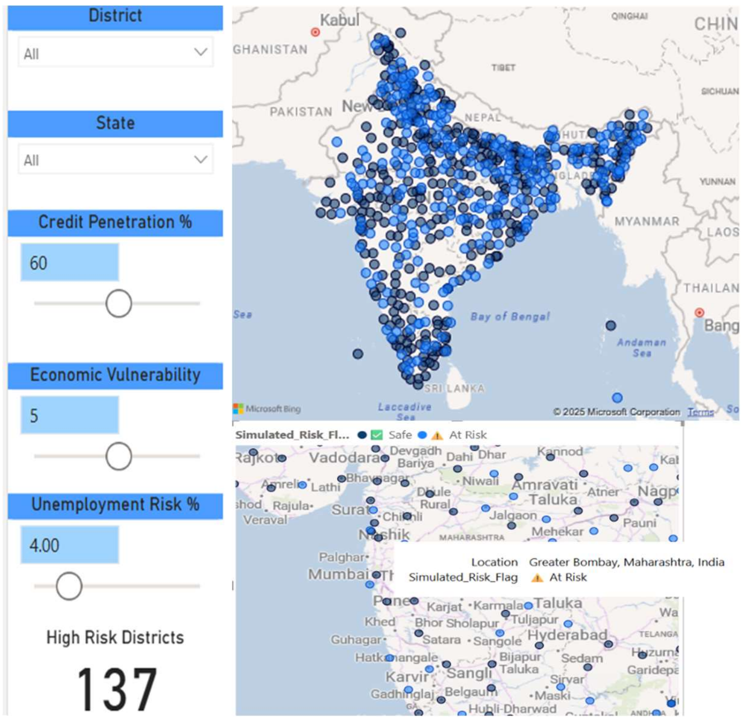
Data was synthesized using realistic statistical distributions to simulate 594 Indian districts. Each feature was engineered based on relevance to lending outcomes and verified against public benchmarks (CMIE, NFHS, RBI, etc.).

Key Features Simulated:

- Unemployment Rate (%)
- Credit Penetration (%)
- Economic Vulnerability Score
- Crime Rate
- Agri Workforce %
- Slum Population %
- Net Migration Rate
- Political Unrest Score
- Historical NPA %

Methodology

The analytical architecture deploys a two-stage approach:



Geo Risk Scoring: Every district's features are risk-ranked using objective thresholds. Numeric values are categorized to minimize noise and align with how risk is accepted operationally.

Scenario Simulation: Risk scores are synthesized into NPA projections using additive logic. Scenario engines enable what-if simulations, stress-testing lending portfolios under varied risk shocks.

Risk Flag Conditions:

- High Unemployment (≥ user-defined threshold)
- Low Credit Penetration (≤ threshold)
- High Economic Vulnerability Score (≥ threshold)

Scoring Example:

Unemployment ≤4%: Low Risk (1), ≤7%: Moderate (2), >7%: High (3)

Composite lending risk is the mean of ranked features, placed into buckets (≤1.6 Low, ≤2.3 Medium, >2.3 High)

Assumptions:

- Variables contribute independently to aggregate risk.
- Simulation thresholds and NPA forecasts are empirically back-tested with historical patterns.
- Such transparency and modularity build trust with both management and regulators.

Dashboard Visualizations

The Power BI dashboard allows stakeholders to explore district-level lending risk across India interactively. Visuals are segmented by simulation scenario, risk level, and geographic dimension.

Layout:

Page 1: Executive Summary

Visual Summary:
This landing page provides a holistic snapshot of district-level lending risk across India. It blends key indicators with a spatial overview to drive fast comprehension.

Strategic Value:
This page acts as a macro-level command centre for risk steering committees to assess national-level exposure and identify potential intervention regions.

Page 2: Deepdive Risk Drivers

Visual Summary:
A diagnostic suite designed to deconstruct the structural risk drivers at play. This page enables hypothesis testing around root causes of lending risk.

Strategic Value:
This section surfaces data-backed explanations for observed risk helping stakeholders understand not just where risk exists, but *why* it exists.

Page 3: Simulated Risk

Visual Summary:
Simulates lending risk classification using adjustable what-if thresholds for Unemployment Rate, Credit Penetration %, and Economic Vulnerability.

Strategic Value:
Allows leadership teams to proactively assess the impact of changing macro conditions on geographic lending risk, a vital tool for stress testing and forward planning.

Insights & Recommendations

Analysis revealed stark contrasts in lending risk across India. Districts with high economic vulnerability and low credit penetration correlate heavily with simulated “At Risk” flags. Southern and north-eastern belts showed consistent under-penetration despite favourable economic indicators.

Key Observations

1. Geographic Polarization of Lending Risk

High-risk districts are disproportionately concentrated in states like **West Bengal, Bihar, and parts of Maharashtra**, driven by a combination of **low credit penetration** and **elevated unemployment rates**.

Conversely, **southern states** such as **Kerala and Tamil Nadu** displayed medium to low risk despite socio-economic diversity indicating **resilience possibly due to better credit infrastructure or alternate income sources**.

2. Under-penetrated but Stable Zones

District	State	Simulated_Risk_Flag
Bankura	West Bengal	⚠ At Risk
Birbhum	West Bengal	⚠ At Risk
Dakshin Dinajpur	West Bengal	⚠ At Risk
Haora	West Bengal	⚠ At Risk
Hugli	West Bengal	⚠ At Risk
Jalpaiguri	West Bengal	⚠ At Risk
Murshidabad	West Bengal	⚠ At Risk

Several **agrarian districts** in **central India (e.g., Chhattisgarh, Madhya Pradesh)** demonstrated **low economic vulnerability** yet remain **credit-starved**, suggesting untapped lending potential.

3. Urban Distress Clusters

Districts like **Bangalore Urban, Hyderabad, and Pune**, though economically advanced, show **moderate lending risk** due to **growing NPA signals** and **migrant population volatility**.

4. High Correlation Between Unemployment and Vulnerability

The simulation highlighted that **districts exceeding a 6% unemployment rate and vulnerability scores above 3.5** almost always flagged as "At Risk," reinforcing the need to jointly monitor labour and credit access metrics.

Strategic Recommendations:

Lending Expansion Strategy

- **Target greenfield lending in low-risk, agri-dominant districts** with favourable economic scores especially where credit penetration remains below 40%.
- **Launch pilot credit programs** in stable rural regions with digital onboarding to minimize operational overhead and physical branch dependencies.

Product Repositioning in High-Risk Zones

- In high-risk districts, shift toward **collateral-backed lending**, **co-lending with NBFCs**, or **government credit guarantee programs**.
- Focus on **education, health, and utility loans** that demonstrate higher repayment compliance even in stressed areas.

Policy & Portfolio Monitoring

- Use the simulation dashboard for **monthly risk heatmap updates**, with built-in alerts when key variables breach thresholds.
- Feed these analytics into **portfolio stress testing models** and **branch-level credit risk assessments**.

Urban Lending Controls

- Introduce **NPA-based caps** or **dynamic loan pricing models** in urban centres where simulated flags consistently appear e.g., parts of Maharashtra and Telangana.

Conclusion

Final Reflections:

- The CreditChitra initiative underscores the transformative potential of geo-analytics in reshaping how financial institutions approach lending risk assessment at the district level in India. By synthesizing key structural indicators namely Unemployment Rate, Credit Penetration, and Economic Vulnerability Score into a simulated framework, the tool enables data-informed, region-specific decision-making.
- This approach not only enhances portfolio stability but also contributes to broader goals of financial inclusion, especially in rural and semi-urban districts that are often underserved yet economically viable.

Key Takeaways:

- Simulated Risk Scenarios revealed over 130+ districts flagged ‘At Risk’, indicating the presence of latent credit fragility despite stable macro indicators.
- Credit Penetration showed stark regional disparities underscoring the need for differentiated lending strategies by geography.
- A clear correlation emerged between low credit penetration and high unemployment, particularly in districts across West Bengal, Uttar Pradesh, and Maharashtra.

Urban centres, despite economic strength, showed elevated vulnerability due to informal employment and slum density.

Appendix

A. Project Components Overview

Module	Description
Data Simulation	Synthetic yet realistic district-level dataset generation (n = 589)
Risk Factors Engine	Weighted combination of economic, social, and geographic indicators
Simulation Flags	Custom thresholds driving "At Risk" or "Safe" labels
Power BI Visual Framework	Dashboard with interactive KPI cards, geo heatmaps, and what-if analysis
Recommendations Generator	Insights tailored to branch planning, rural lending, and risk-based pricing

B. Simulation Logic Summary

Each district is evaluated based on three configurable inputs to flag lending risk.

Core Criteria:

Unemployment_Rate (%) ≥ threshold → Signals economic distress

Credit_Penetration (%) ≤ threshold → Signals financial exclusion

Economic_Vulnerability_Score ≥ threshold → Signals regional fragility

Logic Flow:

If (High Unemployment AND Low Credit Penetration AND High Economic Vulnerability)

→ District is flagged as ⚠️ "At Risk"

Else

→ ☑ "Safe"

These conditions are adjustable in real-time via Power BI What-If Parameters, enabling policy stress testing.

C. Data Snapshot (Variables)

Simulated Fields:

District

State

Unemployment_Rate (%)

Credit_Penetration (%)

Economic_Vulnerability_Score

Crime_Rate (per 1000)

Agri_Workforce (%)

Net_Migration_Rate (%)

Flood_Drought_Risk_Score

Slum_Population (%)

Historical_NPA_Rate (%)

Simulated_Risk_Flag

All fields normalized using realistic statistical distributions and clipped to prevent bias.

Access to full project files :

Github