Financial Risk: A Deep Dive into Fraud Detection & Credit Assessment

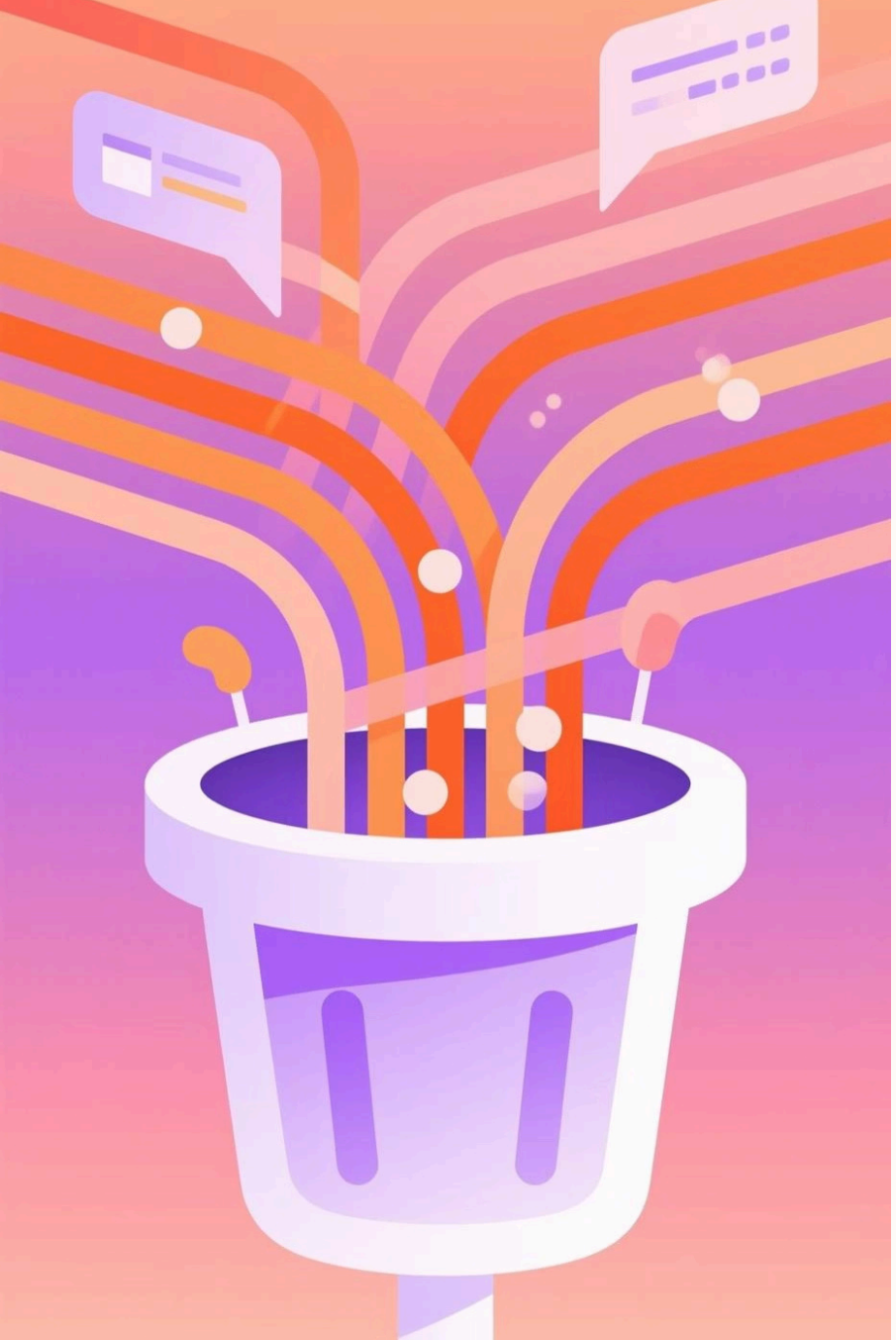
The financial industry faces escalating fraudulent activities and credit risks. This project leverages a large-scale dataset from Paisabazaar to analyse customer demographics, financial details, credit utilization, and payment behaviours. Through extensive Exploratory Data Analysis (EDA) and visualizations, our aim is to uncover crucial insights for detecting fraudulent behaviour, assessing financial risk, and strengthening customer credit profiling.

Dataset Overview: A Wealth of Financial Insights

The dataset comprises 81,782 records with 28 attributes, offering a comprehensive view of customer identity, financial history, and credit scores. Key variables include:

- **Annual Income & Monthly Inhand Salary:** Core indicators of financial capacity.
- **Number of Bank Accounts & Credit Cards:** Reflecting financial complexity and access to credit.
- **Interest Rate & Outstanding Debt:** Direct measures of financial burden.
- **Credit Utilization Ratio & Credit History Age:** Insights into credit management and experience.
- **Payment Behavior & Credit Score:** Crucial for risk assessment and fraud detection.

Demographic details such as Age, Occupation, and SSN provide additional layers for identifying potential fraud, such as multiple accounts linked to a single identity or inconsistent financial behaviours.



The Foundation: Data Cleaning and Preprocessing

01

Initial Data Integrity Check

The dataset was initially checked for missing values, confirming its completeness. This crucial first step ensures the reliability of subsequent analyses.

02

Duplicate Entry Examination

A thorough examination for duplicate entries or repeated customer records across multiple months was conducted. This prevents skewed results and maintains data accuracy.

03

Ensuring Data Consistency

By addressing potential inconsistencies early, we establish a robust foundation for meaningful exploratory data analysis and accurate risk modelling.

Uncovering Patterns: Exploratory Data Analysis (EDA)

The next stage focuses on EDA, investigating how different features correlate with fraudulent or risky financial behaviour. This involves a deep dive into various financial indicators.

Credit Utilization vs. Credit Score

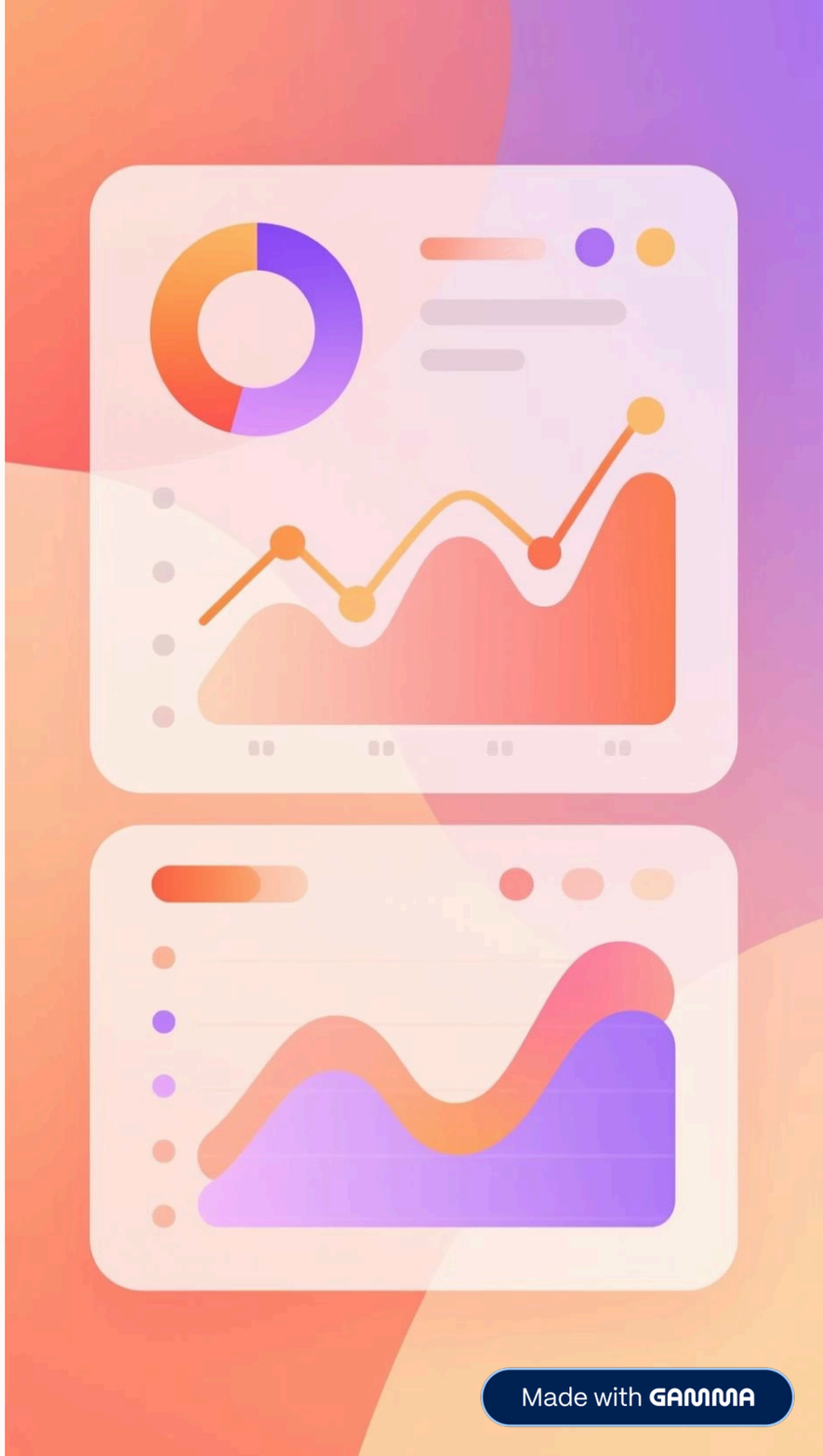
Analysing the relationship between Credit Utilization Ratio and Credit Score helps reveal if high credit usage correlates with poorer scores, indicating potential financial strain.

Delayed Payments Analysis

Examining the Number of Delayed Payments and Delay from Due Date highlights patterns of financial irresponsibility, often precursors to fraud or default.

Visualizing Trends

Visualizations such as histograms, boxplots, heatmaps, and correlation matrices are employed to identify trends, anomalies, and interdependencies between features, providing a clearer picture of financial health.

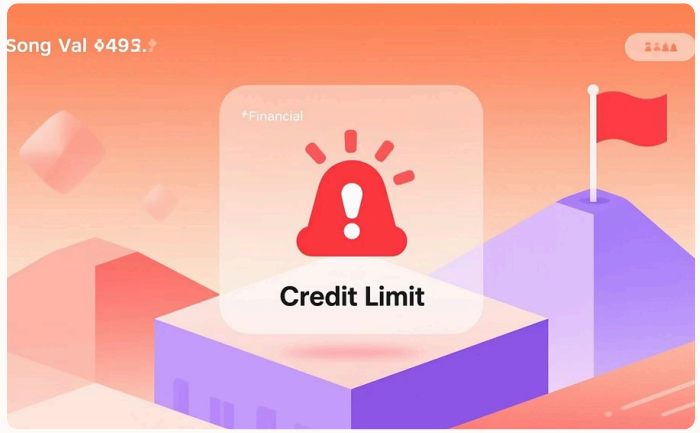


Fraud Analysis: Identifying Red Flags

Fraudulent behaviour often manifests as unrealistic or contradictory patterns within financial data. Our analysis aims to highlight these irregularities through a combination of statistical and visual exploration.

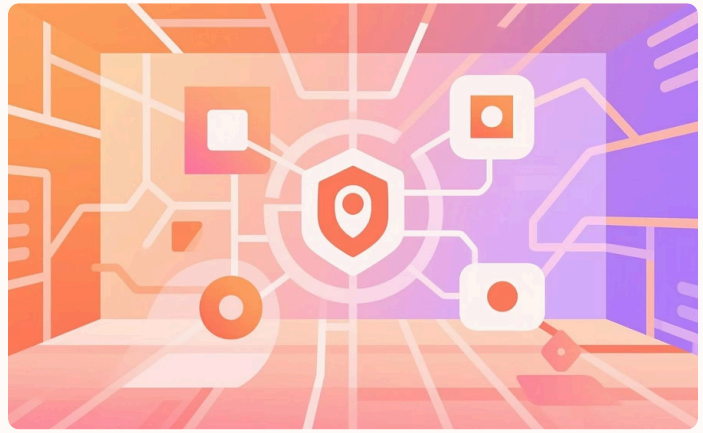
Income-Credit Discrepancies

A customer with a low annual income but multiple credit cards, high loan counts, and excessive monthly spending could indicate suspicious activity.



Sudden Financial Shifts

Sudden changes in Credit Limit, unusually high Credit Inquiries, or inconsistent Payment Behavior serve as critical red flags for potential fraud.



Pattern Recognition

By combining statistical analysis with visual exploration, we aim to pinpoint these irregularities, enabling early detection and prevention of fraudulent activities.

Credit Risk Profiling: Understanding Credit Scores

A key outcome of this project is credit risk profiling, utilizing the dataset's labeled Credit Score feature (Good, Standard, or Poor). By analysing feature distribution across these groups, we uncover factors contributing to poor credit performance.

High EMI Obligations

Customers with high Equated Monthly Instalment (EMI) obligations relative to their income are more likely to have lower credit scores, indicating financial strain.

Frequent Payment Delays

Frequent delays in payment are strongly correlated with poor credit performance, highlighting a lack of financial discipline.

Poor Credit Mix

An unbalanced or poor credit mix (e.g., too many unsecured loans) is expected to correlate strongly with lower credit scores, impacting overall creditworthiness.

These insights are invaluable for banks in developing robust scoring models and improving lending decisions, leading to more informed and responsible financial practices.

Visual Storytelling: Illustrating Financial Behaviour

Visual storytelling is paramount in this analysis, transforming complex data into understandable insights. Various chart types are employed to illustrate how financial indicators differ across customer segments.

- **Bar Charts & Scatter Plots:** Show variations in Outstanding Debt across income groups or compare Payment Behavior with Credit Utilization Ratios.
- **Trend Graphs:** Highlight evolving financial patterns over time.
- **Heatmaps & Pair Plots:** Identify multi-feature interactions and complex relationships within the data.
- **Boxplots:** Effectively highlight outliers that may correspond to fraudulent activities or unusual spending habits.

These visualizations help detect unusual spending habits and provide a clear, intuitive understanding of customer financial behaviour.

Project Conclusion: Data-Driven Insights for a Safer Financial System

This project offers a comprehensive understanding of customer financial behaviour using real-world banking data. By leveraging EDA and visualization techniques, we identify patterns of fraud, risk factors affecting credit scores, and relationships between financial attributes.

1 Fraud Detection

The insights gained aid significantly in detecting fraudulent activities by identifying anomalous patterns and red flags within customer data.

2 Optimized Risk Strategies

Financial institutions can optimize their credit risk strategies, leading to more accurate assessments and reduced exposure to bad debt.

3 Enhanced Customer Profiling

Improved customer profiling systems enable banks to offer tailored financial products and services, fostering stronger customer relationships.

Ultimately, this project demonstrates how data-driven analysis contributes to safer, more transparent, and more efficient banking systems, benefiting both institutions and customers.

Solution to Business Objective: Key Findings and Impact

This project successfully explored the Paisabazaar dataset, uncovering critical patterns in customer behaviour, financial health, and creditworthiness. After thorough data cleaning and preparation, meaningful EDA and visualizations were performed.

1

Credit Score Distribution

Most customers fall into the 'Standard' credit category, with 'Good' and 'Poor' segments exhibiting distinct financial patterns.

2

KPIs for Creditworthiness

Key Performance Indicators such as Outstanding Debt, EMI, and Utilization Ratio clearly differentiate 'Good' scorers (disciplined debt levels) from 'Poor' scorers (high debt burdens and delayed repayments).

3

Income vs. Repayment Discipline

While income is important, repayment discipline and credit utilization proved more decisive in determining credit scores, highlighting behavioural factors over sheer earning capacity.

4

Impact of Minimum Payments

Charts confirmed that customers paying only minimum dues are more likely to have 'Poor' credit scores, reinforcing the importance of full and timely repayments.

5

Late Payments as Risk Indicators

Analysis of 'Delay from Due Date' further reinforced that late payments are strong indicators of financial stress and increased risk.

Conclusion: Strengthening Financial Stability

Correlation analysis indicated strong links between income, debt, and utilization ratios, making them critical features for credit risk modelling. Radar charts further emphasized that 'Poor' credit scorers consistently performed worse across all financial indicators, while 'Good' scorers maintained balance and control.

These findings align with the business objective by highlighting how repayment behaviour, utilization ratio, and EMI obligations serve as stronger risk predictors than income alone. The project demonstrates that fraud risk and poor credit performance can be mitigated by focusing on early detection of high-risk behaviours such as:

- High credit utilization
- Minimum payments only
- Frequent payment delays

With these insights, Paisabazaar can implement targeted interventions, improve risk assessment models, and promote responsible borrowing, ultimately strengthening both customer trust and institutional stability.