**Exploratory Data Analysis (EDA) Report**

**Dataset**: Delinquency Prediction Dataset

**1. Key Patterns and Anomalies**

* **Age**: Customer ages range from 18 to 74 years, with an average of around 46.
* **Income**: Ranges widely from ₹15,404 to ₹199,943, with a mean of ₹1.08 lakh. Some records had missing values.
* **Credit Score**: Values span from 301 to 847. A lower score is typically associated with higher risk.
* **Credit Utilization**: The average is approximately 49%. A few customers show values above 100%, which is a strong anomaly.
* **Missed Payments**: Average is 3, with some customers reaching the maximum of 6 — a strong indicator of risk.
* **Loan Balance**: The loan amount ranges up to ₹99,620. Balances above ₹75,000 are considered financially burdensome for many.
* **Debt-to-Income Ratio**: Most values cluster between 0.23 and 0.36. Higher ratios (>0.4) indicate reduced repayment capacity.
* **Account Tenure**: Ranges from 0 to 19 months, with newer accounts (<6 months) lacking sufficient history for risk prediction.
* **Delinquency Rate**: 16% of customers have a delinquent account, highlighting a significant at-risk segment.

**2. Missing Values and Handling Strategy**

The following columns had missing data:

* **Income**: 39 missing values
* **Credit Score**: 2 missing values
* **Loan Balance**: 29 missing values

**Imputation Methods Used**:

* **Income** and **Loan Balance** were imputed using the **median** to reduce the influence of outliers.
* **Credit Score** was imputed using the **mean**, assuming a roughly normal distribution.
* All missing data was successfully handled, ensuring a complete dataset for further analysis or modeling.

**3. High-Risk Indicators and Their Impact**

1. **High Missed Payments**  
   Frequent missed payments directly signal financial instability and delinquency risk.
2. **High Credit Utilization (>60%)**  
   Heavy usage of credit limits suggests over-leveraging and financial pressure.
3. **Low Credit Score (<500)**  
   A poor credit score indicates past repayment issues and raises future risk.
4. **High Debt-to-Income Ratio (>0.4)**  
   High financial obligations compared to income reduce the customer’s ability to pay future installments.
5. **Loan Balance Above ₹75,000**  
   Large outstanding amounts increase the burden on a customer’s financial planning.
6. **Short Account Tenure (<6 months)**  
   New customers have less historical data, making them harder to assess and riskier to trust.
7. **Employment Status = "Unemployed"**  
   Lack of a stable income source increases the probability of delinquency.