**Exploratory Data Analysis (EDA) Summary Report**

**1.Introduction**

To analyze a business dataset and explore how to Gen AI can enhance the process of data analytics, insight generation and decision making. In this exploratory analysis, Geldium’s dataset is examined to ensure data integrity, uncover underlying trends and identify drivers of credit default risk. These steps are essential for creating robust predictive models and improving the accuracy of risk evaluations.

**2. Dataset Summary**

The dataset consists of 500 customer records from Geldium, each capturing key features related to key delinquency. It contains a mix of numerical and categorical variables, including earnings, credit usage, missed payments and debt-to-income ratio.

**Important Details:**

* **Total records:** 500
* **Key attributes:** Age, Income, Credit Score, Credit Utilization, Missed Payments, Debt-to-Income Ratio
* **Data types:**
  + **Categorical:** Employment Status, Credit Card Type
  + **Numerical:** Income, Loan Balance

**3.Missing Data Evaluation**

Several key variables contain missing values, particularly in the income and loan balance fields. If these gaps are not addressed, they may negatively affect the accuracy of predictive models.

**Observations:**

* **Fields with missing data:**
  + Income:50 missing entries
  + Loan Balance:30 missing entries
* **Proposed solution:**
  + Impute missing numerical values using the median to minimize bias

**4.Key insights and Risk Factors**

The analysis reveals a significant relationship between high credit utilization and credit delinquency, along with an elevated risk linked to frequent missed payments.

**Notable Findings:**

Customers utilizing over 50% of their credit limit are more likely to default.

Individuals with three or more missed payments within six months show a higher probability of credit delinquency.

Some anomalies were identified, such as high-income customers with lowcredit scores, which may require additional investigation.

**5.Role of Gen AI in Analysis**

Generative AI tools supported the identification trends, detection of missing values and examination of risk elements.

**AI tools supported the workflow by:**

Summarizing insights from raw analytics outputs into business-friendly language.

Recommending visualizations to better highlight patterns in regional and category-wise performance.

Automating report creation, saving time otherwise spent writing manual summaries.

**6.Conclusion and Future Action**

This EDA uncovered meaningful insights into Geldium’s dataset, highlighting missing entries, behavioural patterns tied to credit risk and some outlier cases worth deeper analysis.

**Takeways:**

**Data gaps:** Missing income and loan data could influence outcomes.

**Delinquency indicators:** High credit usage and repeated missed payments are strong predictors.

**Data anomalies:** Cases of high income but low credit scores need clarification

**7.Recommendations**

* Choose suitable imputation techniques for missing income and loan values to minimize bias.
* Confirm if key risk factors remain consistent across various customer groups.
* Investigate irregular data entries to ensure accuracy and detect potential financial instability.

These efforts will aid Geldium in refining its risk analysis process and enhance data reliability for further modelling.