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

**1. Introduction**

This document provides an exploratory review of Geldium’s dataset, with the goal of checking data quality, finding useful patterns, and spotting factors that raise the risk of credit default. The main aim is to clean and prepare the data for building accurate prediction models and assessing financial risk.

**2. Dataset Summary**

The dataset has details from 500 customers of Geldium, covering key information linked to credit issues. It includes both number-based and category-based data such as income, credit use, missed payments, and the debt-to-income ratio.  
Key points:  
• Total records: 500  
• Main features: Age, Income, Credit Score, Credit Utilization, Missed Payments, Debt-to-Income Ratio  
• Types of data:  
o Categorical: Job Status, Credit Card Type  
o Numerical: Income, Loan Amount

**3. Missing Data Evaluation**

Some important fields, especially income and loan balance, have missing values. Ignoring them could lower the accuracy of any model built from the data.  
Findings:  
• Columns with missing values:  
o Income: 50 missing  
o Loan Balance: 30 missing  
• Planned fixes:  
o Fill missing numbers using the median value  
o Use AI-generated values to fill some loan balance gaps

**4. Key Insights and Risk Factors**

The review shows that customers who use a large part of their credit or miss payments often are more likely to default.  
Key findings:  
• Those using over 50% of their credit limit are more at risk.  
• People with 3 or more missed payments in 6 months are more likely to default.  
• Some unusual cases were found, like customers with high income but poor credit scores. These need closer review.

**5. Role of AI & GenAI**

AI tools helped identify patterns, spot missing data, and analyze risks. The AI findings were checked against standard financial risk guidelines to ensure correctness.  
Example AI prompts used:  
• "Find trends and show where data is missing."  
• "Analyze default risk based on how credit is used and payment history."

**6. Conclusion & Future Actions**

This analysis gave useful information about Geldium’s data, such as missing values, behavior patterns related to credit risk, and unusual cases that should be explored further.  
Key points:  
• Missing data: Gaps in income and loan fields may affect model results.  
• Default signs: High credit usage and frequent missed payments are strong risk factors.  
• Odd data: Some high earners have low credit scores—this needs to be double-checked.

Suggestions:  
• Use the best methods to fill missing income and loan details to avoid errors.  
• See if the same risk factors apply across different groups of customers.  
• Look deeper into strange data entries to confirm they’re correct and not signs of financial trouble.

These steps will help Geldium improve its credit risk evaluations and ensure data is ready for future analysis.