

Exploratory Data Analysis (EDA) Summary Report Template

1. Introduction

[Insert purpose and goal(s) of the report.]

2. Dataset Overview

This section summarizes the dataset, including the number of records, key variables, and data types. It also highlights any anomalies, duplicates, or inconsistencies observed during the initial review.

Key dataset attributes:

- Number of records: [Insert count]
- Key variables: [List key columns and descriptions]
- Data types: [Categorical, Numerical, etc.]

3. Missing Data Analysis

Identifying and addressing missing data is critical to ensuring model accuracy. This section outlines missing values in the dataset, the approach taken to handle them, and justifications for the chosen method.

Key missing data findings:

- Variables with missing values: [List affected columns]
- Missing data treatment: [Deletion, Imputation, Synthetic Data, etc.]

4. Key Findings and Risk Indicators

This section identifies trends and patterns that may indicate risk factors for delinquency. Feature relationships and statistical correlations are explored to uncover insights relevant to predictive modeling.

Key findings:

- Correlations observed between key variables: [Summarize findings]
- Unexpected anomalies: [Highlight data points requiring further investigation]

5. AI & GenAI Usage

Generative AI tools were used to summarize the dataset, impute missing data, and detect patterns. This section documents AI-generated insights and the prompts used to obtain results.

Example AI prompts used:

- 'Summarize key patterns in the dataset and identify anomalies.'
- 'Suggest an imputation strategy for missing income values based on industry best practices.'

6. Conclusion & Next Steps

[Summarize key findings and outline the recommended next steps.]