

Initial Plots : Credit Analysis

December 9, 2023

1. Data Aggregation for Insightful Analysis:

- **Data Aggregation for Insightful Analysis:** Instead of merely loading and cleaning the data, we strategically aggregated relevant information to gain deeper insights. Aggregations such as sum, max, and min were applied to specific columns, providing a more comprehensive understanding of the credit data. This process not only simplifies the analysis but also uncovers patterns and trends that may not be immediately apparent in raw data.
- **Handling Missing Values and Error Correction:** During the data cleaning phase, various strategies were employed to ensure a robust analysis. Null values were replaced with medians and errors were corrected to maintain data integrity.

2. Exploratory Data Analysis (EDA):

- **Descriptive Statistics:** Used simple statistics to summarize the main characteristics of the data.
- **Data Visualization:** Created easy-to-understand charts and graphs to explore patterns and trends.
- **Outlier Detection:** Identified unusual data points that may need special attention.

Credit Analysis

In this section, we explore detailed information about credit amounts and statuses, providing key insights into the bank's lending patterns:

- **Heatmap Visualization:** The heatmap represents the average credit amount by credit status and activity. Lighter shades indicate higher average credit amounts, especially for active credits and sold statuses. This visualization helps identify patterns in credit amount distribution across different statuses, potentially informing risk management and financial analysis.
- **Correlation Matrix:** The correlation matrix offers a comprehensive view of how different credit attributes relate to each other. Strong positive correlations, like `days_credit` and `days_enddate_fact`, suggest potential relationships. Negative correlations, such as between `days_credit` and `amt_credit_sum_overdue`, hint at connections between credit duration and repayment behavior.
- **Bar Plots:** Bar plots demonstrate the distribution of average credit amounts by credit type. Types like 'real estate loan' and 'loan for working capital replenishment' show significant average credit amounts, reflecting applicants' financial behavior or institutional credit policies. Another bar plot shows the count of credits by type and active status, offering insights into the portfolio's composition.

- **Interpretation and Conclusions:** The visualizations and outputs provide a detailed examination of credit amounts and statuses, revealing key insights into the bank's lending patterns. Certain credit types and statuses are more prevalent or have higher credit amounts, which could inform the bank's credit approval criteria or risk assessment models. Strong correlations identified in the matrix may prompt further analysis to understand factors influencing credit performance and repayment.

The strategic data aggregation process not only made the analysis more insightful but also facilitated a more straightforward and effective exploration of the credit data.