## **Project Overview**

The project aims to analyze credit utilization ratios among a sample of customers to understand their impact on credit scores and assess overall financial health. By employing statistical methods and visualizations, we investigate the distribution of credit utilization, segment customers based on their credit scores, and identify patterns in credit behaviors. This analysis provides valuable insights for financial institutions seeking to improve their lending practices and tailor their services to different customer segments.

## **Method**

The analysis employs a combination of exploratory data analysis (EDA) and statistical visualizations. Key steps include:

1. **Data Exploration:** Initial examination of the dataset to understand its structure, types of variables, and missing values.
2. **Descriptive Statistics:** Calculation of measures such as mean, median, and standard deviation for the credit utilization ratio and other relevant features.
3. **Visualizations:** Creation of various visual representations, including box plots and scatter plots, to illustrate the distribution of credit utilization ratios and customer segmentation based on credit scores.
4. **Correlation Analysis:** Investigation of relationships between numerical features to identify significant predictors of credit score.

## **Data Processing**

Data processing involves several key steps to prepare the dataset for analysis:

1. **Data Cleaning:** Handling missing values, outliers, and incorrect data types. The 'Segment' column was converted to categorical data types to facilitate segmentation analysis.
2. **Feature Engineering:** Mapping numerical segment identifiers to meaningful categorical labels (e.g., "Very Low," "Low," "Good," "Excellent") for better interpretability.
3. **Normalization/Standardization:** Scaling numerical features, if necessary, to ensure comparability across different ranges.
4. **Correlation Matrix:** Calculation of a correlation matrix to identify relationships among numerical features, aiding in feature selection for modeling.

## **Results**

The analysis yielded the following key findings:

1. **Credit Utilization Ratio Distribution:** The box plot revealed the central tendency, dispersion, and outliers in credit utilization ratios among the customer base. Notably, most customers exhibited a healthy credit utilization ratio, while a small proportion displayed excessive utilization, indicating potential financial strain.
2. **Customer Segmentation:** Scatter plots demonstrated distinct segments within the customer base, correlating with credit scores. Customers were effectively categorized into "Very Low," "Low," "Good," and "Excellent" segments, with varying credit behaviors observed across these groups.
3. **Correlation Insights:** The correlation heatmap highlighted significant relationships between credit utilization and credit score, indicating that higher credit utilization often correlated with lower credit scores.

## **Conclusion**

The project successfully demonstrated the importance of credit utilization ratios in assessing customer creditworthiness. The findings reveal that while many customers maintain healthy credit utilization levels, a subset exhibits risky behaviors that could lead to financial instability. The segmentation of customers based on credit scores enables financial institutions to target interventions more effectively, promoting responsible credit use and potentially improving customer financial outcomes.

## **Summary**

In summary, this project analyzed the impact of credit utilization ratios on customer credit scores through EDA and statistical visualizations. By processing the data to ensure clarity and accuracy, the project provided valuable insights into customer segments and their credit behaviors. These findings can inform financial institutions' strategies for improving lending practices and enhancing customer service. Future work could explore predictive modeling to anticipate credit score changes based on utilization trends.

References:

Credit Scoring and Segmentation using Python https://thecleverprogrammer.com/2023/07/03/credit-scoring-and-segmentation-using-python/