**Market Segmentation in Insurance Unsupervised**

**Abstract**  
The insurance industry generates vast amounts of customer data, which holds untapped potential for understanding customer behavior, identifying patterns, and tailoring services. This project focuses on leveraging **unsupervised learning** techniques for **market segmentation**, aiming to categorize customers into distinct groups based on their financial and transactional behavior.

The dataset includes 8,950 records and 18 attributes, such as **BALANCE**, **PURCHASES**, **CREDIT\_LIMIT**, **CASH\_ADVANCE**, and behavioral metrics like **PURCHASES\_FREQUENCY** and **PRC\_FULL\_PAYMENT**. These attributes are preprocessed through techniques such as handling missing values, scaling, and normalization to ensure compatibility with machine learning models.

The primary algorithm employed is **K-Means Clustering**, chosen for its efficiency and interpretability in segmenting large datasets. The **Elbow Method** is used to determine the optimal number of clusters. To further enhance the analysis, **Principal Component Analysis (PCA)** is applied, reducing dimensionality while retaining significant variance for visualization and interpretability.

The insights gained from this clustering process reveal distinct customer profiles, such as high-value clients with consistent full payments, frequent small-purchase customers, and cash-reliant individuals with higher credit usage. These segments offer actionable insights for:

* **Targeted Marketing**: Designing campaigns tailored to specific customer needs.
* **Risk Assessment**: Identifying financially at-risk customers based on spending and repayment patterns.
* **Resource Allocation**: Strategically deploying resources to enhance customer retention and profitability.

By implementing unsupervised learning, this project underscores the importance of data-driven decision-making in the insurance sector, demonstrating how clustering techniques can transform raw data into strategic insights.