Credit Risk Analysis Dashboard - Power BI & Python Integration

Developed by **Jose Antonio Ayon Wu**

Master of Data Analytics - Niagara Region, Ontario, Canada June

2025

This report presents a professional Power BI dashboard for Credit Risk Analysis, integrating Python-driven PCA and Clustering to identify customer risk segments. The dataset used is the publicly available UCI Credit Card Default dataset.

Executive Summary

This project delivers an interactive Power BI dashboard designed to assess credit risk and customer behavior using the UCI Credit Card Default dataset. It combines Power BI's data modeling, DAX-based KPIs, and AI-driven insights through Python (PCA and clustering) to support informed business decision-making in financial risk management. The dashboard identifies risk patterns, customer segments, and behavioral insights that can guide strategies for credit policy and customer engagement.

Data Architecture & Model Design

The dataset includes over 30,000 credit card clients with demographic, billing, and payment information. Data was normalized into multiple tables for performance optimization and scalability. A Date Table was created for time intelligence analysis, while auxiliary tables were implemented to support segmentation and clustering visualizations. Python-generated clustering and PCA results were re-integrated into Power BI for unified visualization and analysis, ensuring seamless integration between analytical and reporting layers.

Key Metrics & Insights

The dashboard provides a clear overview of key financial risk indicators: (1) Default Rate: 22.1% overall, with the highest default among customers aged 60+. (2) Healthy-Use Ratio (0.5–1.0): Lowest default rate at 15.7%. (3) High-Risk Ratio (>5): Default rate of 26.0%. (4) Education & Age Impact: Higher education correlates with lower default probability. These findings highlight behavioral differences that can inform targeted risk mitigation strategies.

Python Integration for AI Insights

Python was used to enhance Power BI's analytical capabilities by performing Principal Component Analysis (PCA) and K-Means Clustering on customer behavior data. The models identified four customer segments differing in default probability, credit utilization, and payment consistency. (a) Tools: Python (pandas, scikit-learn, matplotlib) (b) Integration: Model outputs exported as CSV and merged back into Power BI. (c) Result: Cluster 3 showed a 0.56 predicted default probability, identifying a high-risk segment requiring strategic monitoring.

Business Recommendations

Based on the analytical findings, the following recommendations are proposed: (1) Develop behavior-based segmentation to tailor credit monitoring and communication strategies. (2) Use predictive clustering outputs for proactive default prevention. (3) Leverage Power BI for ongoing portfolio tracking and visualization of high-risk clusters. (4) Integrate advanced analytics for periodic recalibration of risk models and policy adjustments.

Conclusion

This project demonstrates an end-to-end data analytics workflow combining Power BI for visualization and Python for AI-driven modeling. It reveals that customer age, education, and credit usage behavior are key determinants of default probability. The integration of Python-based insights into Power BI dashboards represents a scalable approach to data-driven financial risk management and business intelligence.