Report

Knime Analysis

In our analysis of customer data, which was initially performed using KNIME, we identified several key issues and took deliberate actions to improve segmentation quality and effectiveness.

Data Issues & Solutions

The first issue involved customers who had returned all purchased items, leading to `Monetary` values less than or equal to zero. These customers did not contribute positively to revenue. For this reason, we removed them to prevent misleading results.

The second issue related to the effectiveness of our clustering methods. Automated clustering or binning led to inconsistent groupings, failing to capture the significance of differences in customer behavior (`Recency`, `Frequency`, `Monetary`). To address this, we manually analyzed the distribution of each RFM metric to identify meaningful trends and defined custom bin intervals for each metric using the `Numeric Binner` node. This approach ensured that segmentation accurately represented distinct customer behaviors.

Power BI Dashboard Summary

Our Power BI dashboard provided insights into customer segmentation, spending trends, and category performance:

- **RFM Analysis**: We chose to segment customers into five distinct groups—Inactive, Occasional, Regular, High-Value, and VIP—based on their purchasing behavior. The majority of our customers are from Europe, particularly the UK, which represents approximately 91% of the total customer base. This geographic insight helped us better understand the distribution of our customer base and informed our decisions. By defining these segments, we intended to focus marketing efforts on the most promising customer groups while identifying those requiring re-engagement strategies.
- **Spending Trends**: We analyzed spending patterns over time to understand the impact of specific events or promotional activities on customer behavior. For instance, identifying peaks in spending during November 2018 helped us recognize the influence of events like Black Friday.
- **Category Analysis**: We analyzed sales across various product categories to identify top-performing and underperforming segments. The goal was to understand customer preferences and allocate marketing resources effectively.