



Insure Trust

Client Loyalty Project for Mutual Insurance

Team 11

Riccardo Becciolini

Marco Donatucci

Loris Catalano

Giovanni Angeli





Data-Driven Loyalty: Insights that Matter



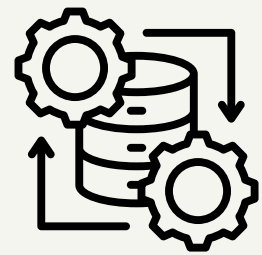
The insurance claims dataset was analyzed through a structured four-phase process:

- **Exploratory Data Analysis (EDA):** Understanding data structure, assessing quality, and detecting anomalies and missing values.
- **Preprocessing and Data Cleaning:** Building a cleaning pipeline to manage outliers and missing data, following insurance industry best practices.
- **Feature Engineering & Data Enrichment:** Creating new informative features and integrating external data sources to enhance insights.
- **Segmentation and Clustering:** Applying machine learning algorithms to identify homogeneous customer segments.

The main goal is to design personalized and differentiated marketing strategies for mutual insurance customers. Specifically, we aim to:

- Identify hidden patterns in claims data that reveal **distinct customer profiles**
- Understand factors influencing **customer risk** and value
- Maximize **Customer Lifetime Value (CLV)** while minimizing risk
- **Propose a product** recommendation system based on customer profiles

From Raw Data to Risk Intelligence



Data Preprocessing

Outliers: IQR-based detection on claim amounts and age; winsorization used to preserve data.

Missing Values: Contextual imputation using geographic (province-region) and vehicle (brand-model) relationships.

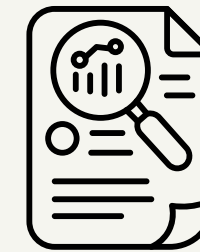


Feature Engineering

Temporal Features: Extracted year, month, day to capture seasonal trends.

Loss Ratio: Key performance metric in insurance.

Demographic Binning: Age groups and claim amount categories for clearer segmentation.

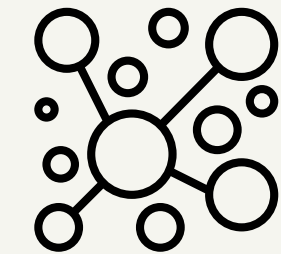


Data Enrichment

Internal (Simulated): Customer fidelity, claim history, portfolio value, acquisition channel.

External (Simulated): Income stats (ISTAT), vehicle risk (ACI), weather, market values.

Combined Indices: Risk Index & Customer Lifetime Value (CLV).



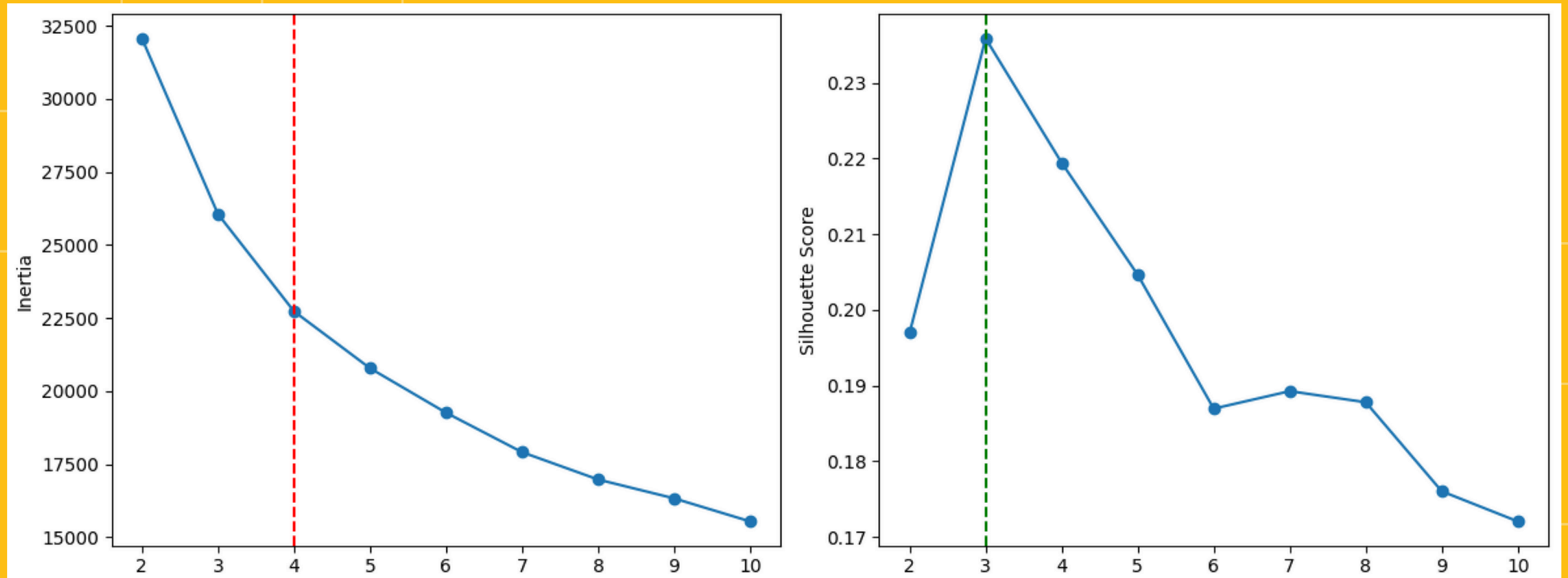
Clustering

K-Means: Efficient and interpretable for customer segmentation.

PCA: Dimensionality reduction for 2D visualization.

Silhouette & Elbow: Used to select optimal number of clusters.

Elbow Method vs Silhouette Score



Data Made Visible

Exploratory Visuals

- **Boxplots & Histograms:** To detect distribution and outliers (age, claim amounts)
- **Correlation Heatmaps:** To highlight relationships between key variables
- **Bar Charts:** For analyzing guarantee types and vehicle brands

KPI Visuals

- **Time Series Charts:** To reveal claim trends and seasonality
- **Colored Bar Charts:** To show loss ratio by guarantee type and region
- **Geographic Maps (Suggested):** For visualizing spatial distribution of claims

Cluster Visuals

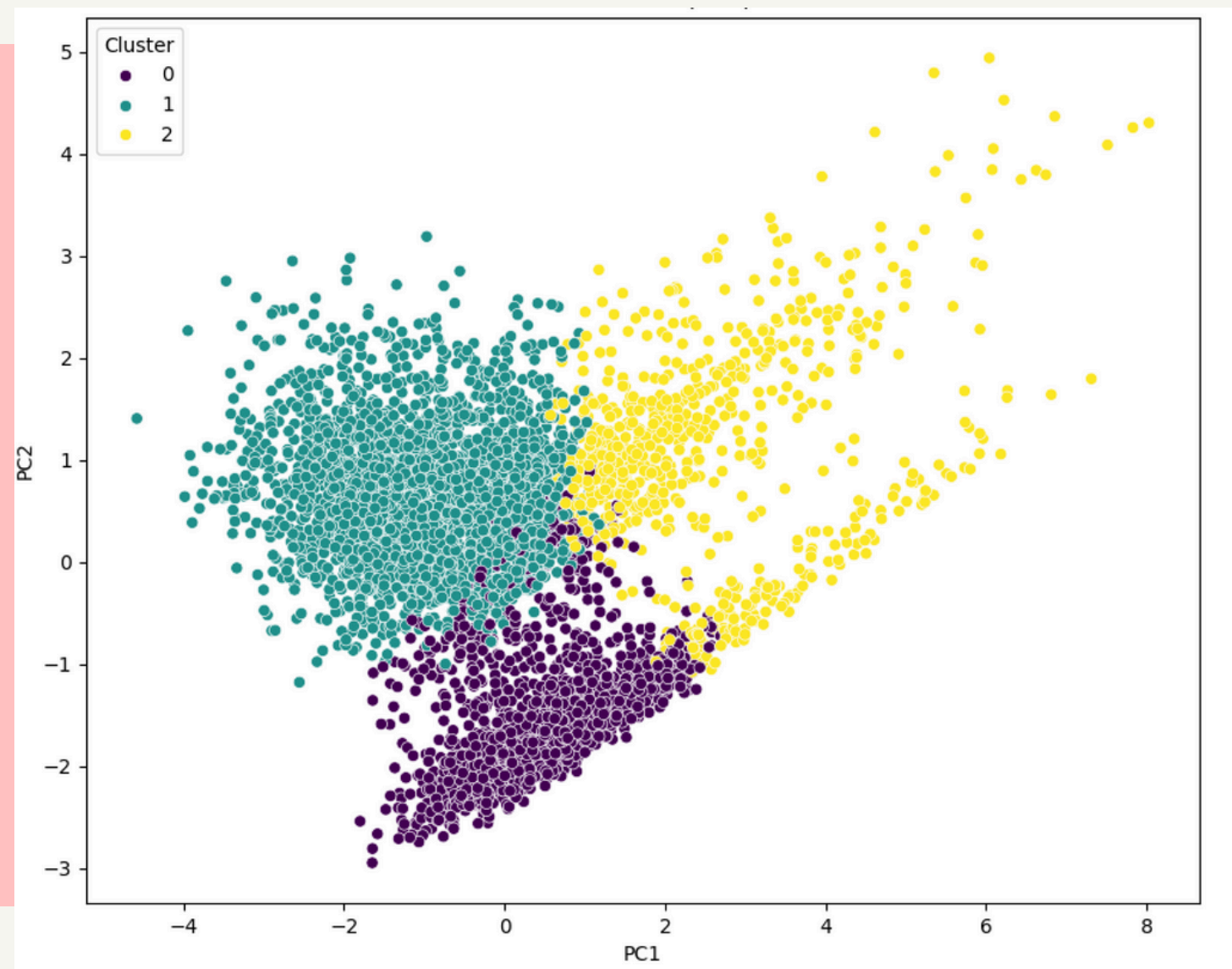
- **PCA Scatter Plots:** 2D visualization of clusters using principal components
- **Radar Charts:** To show key feature profiles of each cluster
- **Parallel Coordinates:** For multidimensional comparison across clusters

Interactive Dashboards

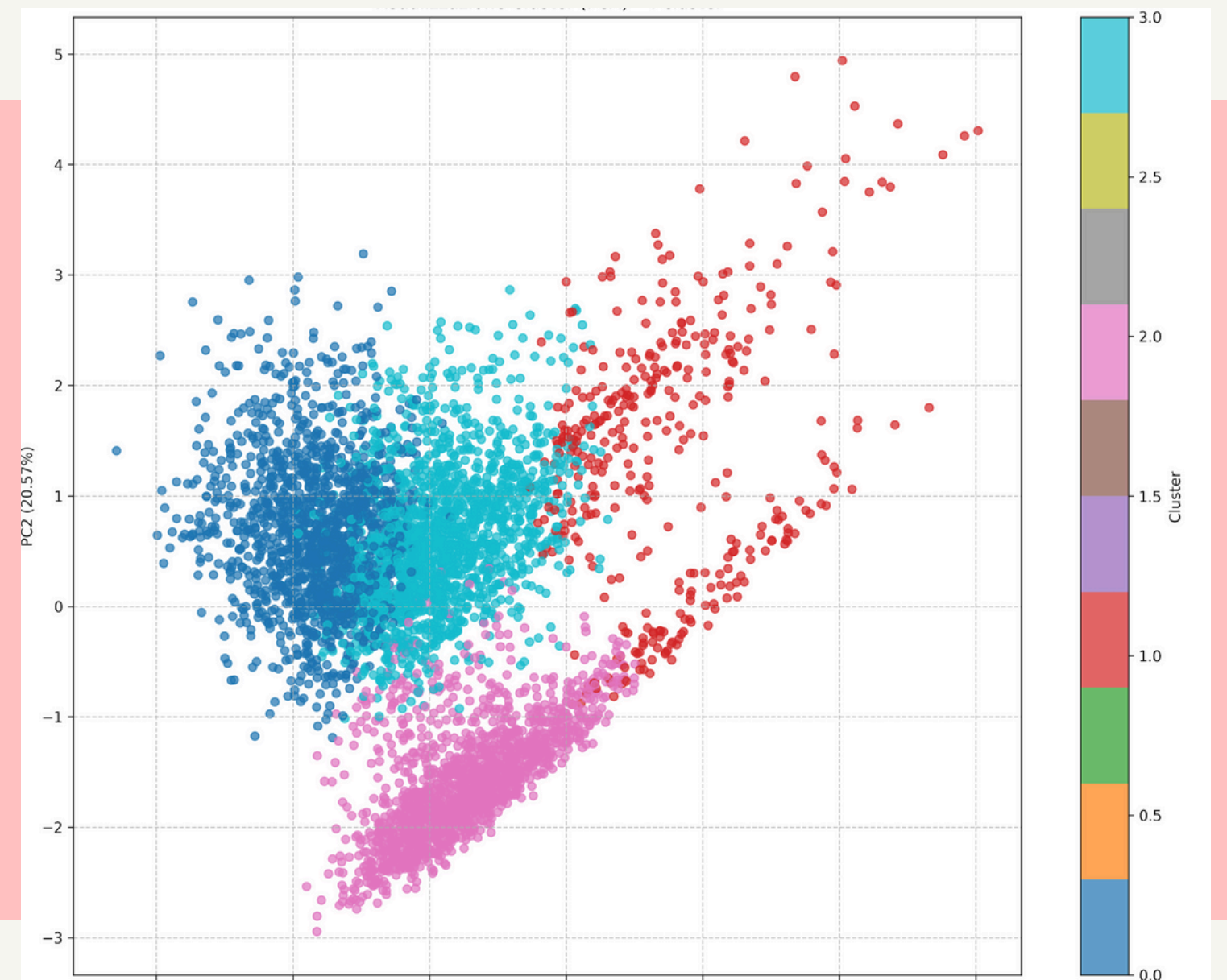
- **KPI Dashboard:** Interactive filters by period, region, and guarantee type
- **Cluster Profiler:** Explore cluster characteristics and compare profiles interactively



Cluster Data Visualization



3-cluster



4-cluster

Key Insights

Support to Analysis Objective

Clusters reveal clear behavior and value patterns, justifying differentiated marketing strategies.

Combining demographic, behavioral, and risk factors is more predictive than demographics alone.

Actionable Insights

High-value, low-risk: Ideal for retention and cross-selling.

High-risk, low-value: Mitigate risk or adjust premiums.

Channels vs. performance: Optimize acquisition strategies.

Claims by time & region: Drive seasonal and geo-targeted campaigns.

Vehicle traits & claims: Support smarter pricing models.

Innovative Aspects

Holistic segmentation: Merges behavior, demographics, geography, and vehicle data.

Risk-Value Index: Two-dimensional view for strategic targeting.

Simulated weather data: Highlights seasonal claim trends.

Cluster-based marketing: Replaces generic campaigns with tailored ones.

Business Impact & Strategic Gains

Real-World Applicability: the model is thought for deployment in real insurance contexts, requiring only:

- Integration with existing CRM systems
- Access to full internal insurance data
- Acquisition or partnerships for external data (weather, vehicle stats, ISTAT)

Impact of the Analysis

- **Increased Profitability:** Through segment-based pricing strategies
- **Improved Customer Retention:** Enabled by personalized offers
- **Marketing Optimization:** Better budget allocation based on expected ROI by segment
- **Risk Reduction:** Via targeted mitigation strategies for high-risk groups

Generated Value

- **Economic Value:** Estimated increase in CLV through personalized strategies
- **Competitive Advantage:** Differentiation in a saturated market
- **Organizational Value:** Shift from product-centric to customer-centric approach



The Cluster Effect



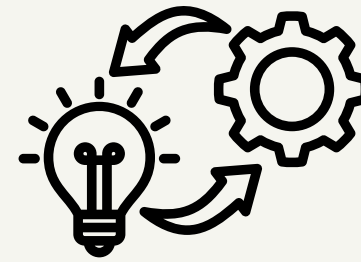
Personalized Marketing per Cluster

Recommended Products: Based on risk and value profiles

Preferred Channels: Aligned with demographic traits

Pricing Strategy: Customized to maximize conversion and profitability

Retention Approach: Incentives tailored to customer profiles



Implementation Plan

Pilot Phase: Test strategies on a sample from each cluster

Measurement: Track KPIs such as conversion, upsell, retention, satisfaction

Optimization: Refine strategies based on pilot results

Scale-Up: Roll out to the full customer base



Continuous Monitoring

Monitoring Dashboards: Track performance by segment

Periodic Cluster Updates: Capture changes in customer base



Ongoing A/B Testing: Optimize marketing for each segment



Strategic Outcomes & Next Steps

Advanced segmentation enables the insurance mutual to move beyond undifferentiated marketing, supporting more targeted strategies that enhance customer value and reduce overall portfolio risk.

Opportunities for Further Enrichment

- **Incorporate Behavioral Data:** Such as app usage, contact frequency, and feedback
 - **Explore Predictive Modeling:** To anticipate churn or uncover cross-selling potential
 - **Prototype Real-Time Recommendations:** Based on enriched customer profiles
 - **Test Dynamic Pricing Ideas:** Linked to combined risk-value indicators
- 
- 
- 