



1. Background and Objective

The goal of this project was to support the development of a data-informed rewards program by validating proposed perks against real customer behavior. Rather than assuming a one-size-fits-all approach, the analysis focuses on identifying customer segments and aligning incentives accordingly.

2. Data and Methodology

The analysis is based on session-level booking data extracted from a SQL database. After preprocessing and data validation, behavioral features were engineered and aggregated to the user level.

Key features include session engagement, booking lead time, spending behavior, hotel usage, and cancellation rates. Features were scaled and used as input for KMeans clustering.

Different cluster counts were evaluated using simple quantitative metrics and interpretability.

Four clusters were selected as they provided clearly distinguishable behavioral patterns while remaining actionable from a business perspective.

3. Customer Segments

The segmentation resulted in four customer personas:

- **Frequent Planners:** Customers who book early and travel regularly.
- **Price-Sensitive Explorers:** Customers who compare many options and focus on price.
- **Hesitant Explorers:** Highly engaged customers with frequent cancellations.
- **High-Value Travelers:** Customers with higher spending and bundled bookings.

These personas translate technical clustering results into business-friendly profiles.

4. Insights and Perk Validation

Behavioral differences between segments provide evidence for the proposed perks. High cancellation rates strongly support flexibility-based incentives, while high hotel usage supports hotel-related perks.

Some perks target overlapping audiences, suggesting that not all incentives should be rolled out simultaneously. This highlights the importance of prioritization and targeted deployment.

5. Data Quality Considerations

During preprocessing, several data quality signals were identified, including negative or zero hotel nights and unusually high click counts. These patterns may indicate inconsistencies or exploratory behavior and should be reviewed before automating perk assignment.

6. Recommendations and Next Steps

It is recommended to deploy perks selectively by customer segment rather than globally. Low-cost, high-impact perks should be prioritized, while high-cost incentives should be limited to high-value customers.

All perks should be validated through controlled experiments such as A/B testing to measure their impact on conversion, retention, and customer lifetime value.