TravelTide Rewards Program: Executive Summary

This project supports TravelTide's marketing team in designing a personalized rewards program to increase customer sign-ups and engagement. The objective is to identify distinct customer segments based on behavioral data and assign each segment a tailored reward aligned with their booking behavior.

Using transactional and session-level data, a user-level dataset was created through primary feature engineering in SQL. Key behavioral features include average booking value, lead time, bundling ratio, discount sensitivity, cancellation rate, and customer lifecycle metrics such as tenure and recency. Only users with sufficient activity were included to ensure reliable behavioral patterns.

An exploratory analysis revealed strong heterogeneity in customer behavior, particularly in spending levels, cancellation tendencies, and sensitivity to discounts. Based on these insights, a **business-driven manual segmentation** was developed, resulting in five interpretable customer segments:

1. Discount-Oriented Bundlers,
2. Longer-Stay Low-Value Users,
3. High-Value Frequent Users,
4. High-Value High-Cancellation Users, and
5. Deal-Sensitive High-Risk Users.

Each segment reflects a distinct behavioral profile and strategic opportunity. To validate the robustness of this segmentation, an **automated K-Means clustering model** was also applied. While the automated clusters showed only moderate statistical agreement with the manual segments (ARI ≈ 0.31, NMI ≈ 0.41), the comparison confirmed that the manual segmentation offers superior interpretability and business relevance for marketing decision-making.

For each segment, a personalized reward was recommended, including

- **Early Deal Access** for Discount-Oriented Bundlers
- **Extended Stay Discount** for Longer-Stay Low-Value Users
- **1 Free Hotel Night** for High-Value Frequent Users
- **No Cancellation Fees** for High-Value High-Cancellation Users
- **Free Checked Bag** for Deal-Sensitive High-Risk Users

These recommendations balance **custo**mer attractiveness with business constraints. The final Tableau dashboard visualizes segment behaviour, value contribution, and segment-to-perk mapping, translating analytical results into actionable marketing insights.

Overall, this project demonstrates how behavioural data can be transformed into practical, personalized reward strategies that enhance customer engagement while preserving business value.

TravelTide Rewards Program – Customer Segmentation Report

Structured workflow to turn raw travel logs into actionable reward tiers

Date Feb 5, 2026

Owner Marija Pavlova

Background and context

TravelTide is a digital travel booking platform aiming to increase customer engagement through the introduction of a rewards program. The marketing team identified several potential perks, including free hotel meals, free checked bags, no cancellation fees, exclusive discounts, and free hotel nights bundled with flights. However, not all customers value these perks equally. The core business challenge is therefore to personalize communication by highlighting the perk most likely to resonate with each customer, thereby increasing the probability of program sign-ups.

Objectives

The objective of this project is twofold. First, to verify whether the available data supports the existence of distinct customer types with different behavioral preferences. Second, to assign each customer to a segment associated with a likely “favorite perk”, enabling TravelTide to tailor marketing messages and improve campaign effectiveness.

Data & Feature Design

The analysis is based on a user-level dataset containing behavioral and transactional information for 5,542 customers. The raw data includes booking behavior, spending metrics, cancellation patterns, and engagement indicators. To ensure meaningful segmentation, the dataset was transformed from event-level records into aggregated user-level features.

Primary Feature Engineering in SQL

The dataset was constructed through primary feature engineering in SQL by aggregating session-level and trip-level data into a unified user-level feature base. Only users with more than seven sessions in 2023 were retained to ensure stable behavioural signals.

Core behavioural features were engineered directly in SQL, including average lead time, average checked bags, hotel spend, base fare, nights stayed, discount sensitivity, bundling ratio, and true cancellation rate. Last active date enabled lifecycle features such as tenure and recency to be derived later in Python. This step ensured that the modelling dataset represented stable behavioural summaries at the customer level.

Phase 1: Feature Engineering (The "Understanding" Phase)

Feature engineering followed a distinction between lifecycle metrics and behavioural characteristics. Lifecycle features such as tenure and recency capture the stage of the customer relationship. Behavioural features describe how customers travel and spend, including average nights stayed, bundling behaviour, discount sensitivity, and cancellation rate.

A booking value proxy was constructed as the sum of average hotel spend and average base fare. Demographic variables and geographic distance were deliberately excluded, as they do not directly inform perk preferences and are misaligned with the goal of behavioural personalization.

Phase 2: Exploratory Insights

Exploratory analysis revealed strong heterogeneity in customer behaviour. The booking value proxy exhibited a right-skewed distribution, with most users concentrated in low-to-mid spending ranges and a small group of high-value customers. Cancellation behaviour also varied widely, indicating the presence of both stable and high-risk users. These patterns support the assumption that segmentation is justified.

Phase 3: Segmentation Methodology

Customer segmentation was performed using K-Means clustering. Eight behavioural features were used as inputs: tenure, recency, value proxy, cancellation rate, discount sensitivity, bundling ratio, average nights stayed, and average bags. Missing values were imputed using median values, and all features were standardized.

The optimal number of clusters was selected based on a combination of silhouette scores and business interpretability. While three clusters offered higher statistical separation, five clusters were chosen to enable more actionable business segmentation.

Phase 4: Customer Segments & Perk Mapping

The final segmentation produced five distinct customer profiles. Each segment is linked to a reward perk aligned with its dominant behavioural pattern, enabling targeted marketing communication.

	Discount-Oriented Bundlers	Longer-Stay Low-Value Users	High-Value Frequent Users	High-Value High-Cancellation	Deal-Sensitive High-Risk Users
Core Behavior	Price-sensitive, bundles flights + hotels, reacts to discounts	Long stays, simple bookings, low engagement	Frequent activity, moderate–high spend, low cancellations	Very high spend, very high cancellation rate	High value, volatile, actively hunts deals
Key Need / Pain Point	Wants access to the best deals	Needs soft activation, low commitment	Wants premium experience without discounts	Needs flexibility, fears penalties	Wants visible monetary savings
Recommended Perk	Early Deal Access	Extended Stay Discount	1 Free Hotel Night	No Cancellation Fees	Free Checked Bag
Business Logic	• Strong behavioral match.	• Low-cost perk. • Tangible soft hook	• Rewards loyalty • Easy to understand	• Removes friction. • Retains high-value	• Feels like saving. • Less margin damage
Strategic Goal	• Increase email CTR and conversions	• Activate low-value users cheaply	• Increase lifetime value	• Reduce churn risk	• Capture deal driven demand

The following personas were randomly sampled from the dataset based on user IDs and are used as illustrative representatives of the identified customer segments.

Phase 5: Validation via Automated Clustering

To assess robustness, an automated clustering model using all numerical features was built. Agreement between both approaches was measured using Adjusted Rand Index (ARI = 0.309) and Normalized Mutual Information (NMI = 0.413).

These moderate values indicate partial alignment, suggesting that the business-driven segmentation captures meaningful structure while prioritizing interpretability and marketing relevance.

Executive Tableau Dashboard

An executive dashboard was created in Tableau integrating a normalized persona heatmap, total booking value by segment, and a segment-to-perk mapping table. The dashboard translates analytical results into an actionable business tool for stakeholders.

Conclusion & Business Implications

This project demonstrates that TravelTide’s customer base can be meaningfully segmented based on behavioural patterns relevant to marketing strategy. The resulting segments provide a practical framework for personalizing rewards program invitations.

By aligning segmentation with actionable perks, the analysis bridges data science and business decision-making. Future work could incorporate A/B testing to validate the effectiveness of personalized messaging and update segments dynamically as customer behaviour evolves.