

Methodology

Cohort Definition and Data Cleaning

To ensure reliable analysis, I focused on users with more than 7 sessions since **January 4, 2023**, reducing the dataset to **5,998 active users**mastery-project-present.... I removed sessions related to canceled trips using Common Table Expressions (CTEs)Data Preprocessing. This filtering step excluded noise and ensured the segmentation was based on meaningful behavioral patterns.

Feature Engineering

I aggregated session, flight, and hotel-level data into user-level metrics. Key features included:

- **Average trip length**
- **Click counts**
- **Flight/hotel booking ratios**
- **Discount usage**
- **Baggage counts**
- **Hotel nights and room counts**Data Preprocessing.

These features captured behaviors tied to potential interest in rewards.

Dimensionality Reduction

To simplify the feature space and reduce noise, I applied **Principal Component Analysis (PCA)**. This helped preserve variance while transforming the dataset into a more cluster-friendly structureDo we Really Need All D....

Clustering Algorithms

I tested both **K-Means** and **DBSCAN**Segmentation Algorithms...Segmentation Algorithms.... Ultimately, **K-Means** was chosen due to better cluster stability and higher silhouette scores, which indicated well-separated user segments.

Findings

I identified **six distinct customer segments**, each with tailored reward strategies:

- **The Dreamers (12%)**
Young, curious users who have not yet booked.
 - ♦ *Perk:* First-trip discountmastery-project-present....
- **Family Travelers (20%)**
Users with kids, often travel with checked bags and book multiple rooms.
 - ♦ *Perk:* Free checked bag(s).
- **Business Travelers (15%)**
Frequent fliers with short stays and minimal baggage.
 - ♦ *Perk:* Free in-flight meal.
- **Local Explorers (10%)**
Book hotels near their home airport for local escapes.
 - ♦ *Perk:* Local transit/event vouchers.
- **Seniors (8%)**
Take fewer but longer trips, focused on culture.
 - ♦ *Perk:* Free pass to cultural venues.
- **General Reward Travelers (35%)**
Mixed traits; consistently engaged but not easily categorized.
 - ♦ *Perk:* Seasonal rotating rewards.

Each group was identified via behavioral clustering and validated through demographic summaries and silhouette analysisExamining the Segments.