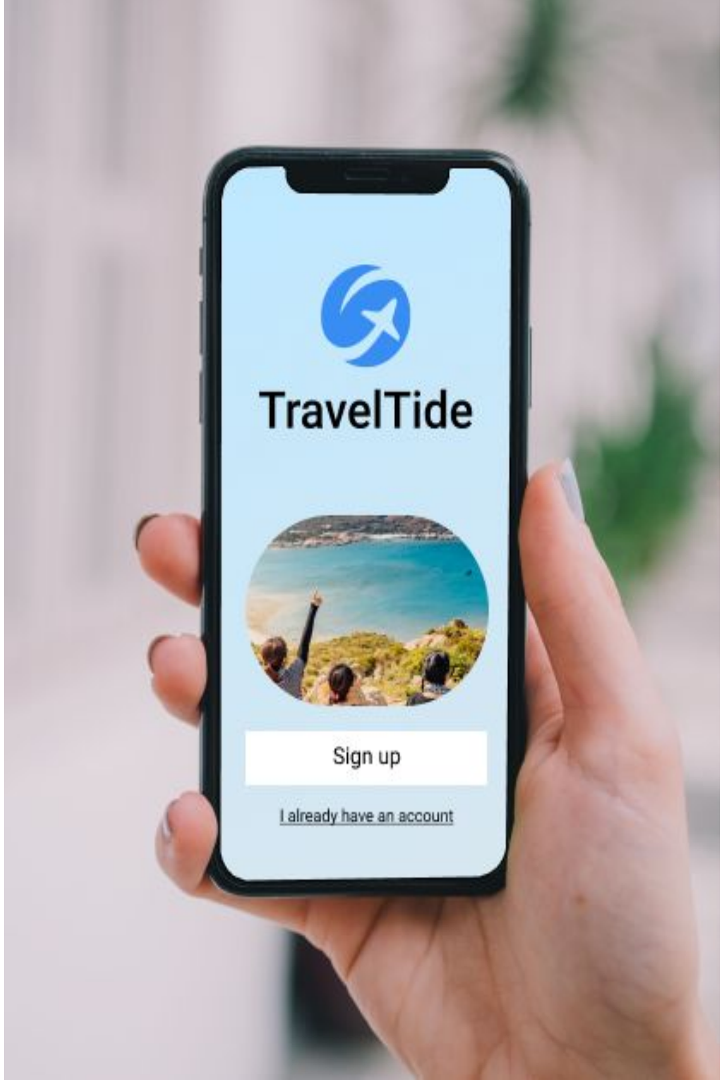


TravelTide

Customer Segmentation for Personalized Rewards Program

Customer Insights & Rewards Strategy Proposal

Athanasios Kardamis – Data Analyst / Data Scientist
August 2025

The Client: TravelTide

“Making travel planning effortless through AI-powered personalization.”

- **AI-powered platform** scanning packages, flights, and hotels to deliver tailored suggestions.
- Leverages past bookings, trip history, and user preferences to recommend trips and secure the best available deals through its AI engine (*Hunter*).
- Simplifies planning with: **search** → **compare** → **book —** quick, intuitive, and user-friendly.

The Client’s Goal

"Design a data-driven rewards strategy that increases customer engagement and retention."

- **Segment customers** based on booking and browsing behavior.
- **Identify high-value and at-risk segments** for targeted campaigns.
- **Recommend perks** that match the needs and preferences of each segment.
- Ensure **recommendations are actionable** for marketing and product teams.

The Approach

"Turning customer data into actionable insights through machine learning segmentation."

- Collected and prepared booking and browsing data from TravelTide’s platform.
- After discussions with **Helena (Head of Marketing)**, applied project-specific filters: **users with >7 sessions** and **activity after Jan 4, 2023**.
- Engineered features to capture travel frequency, spend patterns, booking habits, and loyalty indicators.
- Applied **K-Means clustering** to uncover distinct customer segments through ML approach.
- Matched tailored rewards to each segment’s needs and value potential.

The Data

"Four interconnected tables capturing the full TravelTide customer journey."

Source:

`postgres://Test:bQNxVzJL4g6u@ep-noisy-flower-846766.us-east-2.aws.neon.tech/TravelTide?sslmode=require`

Size: 5.4M+ rows total

1. **users –** demographics: user_id, birthdate, gender, marital status, has_children, location, sign_up_date.

2. **sessions –** browsing sessions: session_id, user_id, trip_id, start/end time, clicks, discounts, bookings, cancellations.

3. **flights –** flight bookings: trip_id, origin/destination, airline, dates, checked_bags, fare.

4. **hotels –** hotel bookings: trip_id, hotel_name, nights, rooms, check_in/out time, rate.

User-Level Dataset Creation (Feature Engineering)

"Transforming raw session and booking data into meaningful behavioral profiles"

- Applied previously discussed filters (>7 sessions, activity after Jan 4, 2023).
- Aggregated data from sessions, flights, hotels, and users tables to the **user_id** level.
- Engineered features capturing:
  - Travel frequency and booking conversion rates.
  - Flight and hotel spend (log-transformed for outlier control).
  - Trip patterns: distance, duration, type (flight-only, hotel-only, both).
  - Luggage behavior (checked bags) and advance booking time.
  - Cancellations and no-trip rates.
- Final engineered dataset: **5,998 unique users** with **21 features** .

Clustering Approach

"Using machine learning to uncover natural customer segments"

- Standardized all features to ensure equal weighting in distance calculations
- Ran **extensive experimentation** — testing multiple algorithms (K-Means, GMM, DBSCAN) and **300+ variations** of:
  - Feature combinations and engineered metrics.
  - Scaling methods and log transformations
  - Different numbers of clusters (k=4–8).
- Evaluated results using **silhouette scores** and interpretability of segments.
- Final model: **K-Means with k=6** clusters, silhouette score **0.474** — best trade-off between separation and business clarity.
- Assigned descriptive, business-friendly names to each cluster.

Overview of Segments

"Six distinct customer types identified through clustering"

Cluster 0 – Premium Frequent Travelers

High-value, long-distance travelers booking both flights and hotels frequently.

Cluster 1 – Change-Prone Travelers

Moderate bookers with high cancellation rates and frequent re-planning.

Cluster 2 – Window Shoppers

Low-engagement browsers with minimal bookings and spend.

Cluster 3 – Hotel Loyalists

Hotel-focused travelers with high brand loyalty and long stays.

Cluster 4 – Serial Cancellers

Book often but cancel every trip; no completed travel.

Cluster 5 – Balanced Frequent Travelers

Large, reliable group with steady bookings and low cancellations.

Cluster 0 – Premium Frequent Travelers

"High-value, committed travelers booking both flights and hotels"

Profile:

- 1,612 users (26.9% of total).
- Highest booking conversion rate (56%).
- Highest spend: $5,307 flights, $3,420 hotels.
- Longest travel distances (8,310 km avg).
- Prefer flight+hotel packages (both-booked rate 78%).
- Mostly medium-length trips (~10 days advance booking).

Final Perk: 🏨 1 Free Hotel Night with Flight

Rewards their package booking habit and reinforces loyalty for top-value customers.

Cluster 1 – Change-Prone Travelers

"Frequent planners who often cancel or adjust their trips"

Profile:

- 423 users (7.05% of total)
- Very High cancellation rate (35%)
- Longest browsing times (944 sec avg) and high engagement (335 clicks avg)
- Moderate spend: $2,395 flights, $1,682 hotels
- Booking conversion 40%
- Trips completed (2.25 avg)

Final Perk: ✅ Free Cancellation

Removes a key booking barrier and builds trust, encouraging earlier and more frequent trip commitments.

Cluster 2 – Window Shoppers

"Low-engagement browsers with minimal bookings and spend"

Profile:

- 573 users (9.55% of total).
- Lowest booking conversion rate (3%).
- 80% have no completed trips.
- Minimal spend: $365 flights, $0 hotels.
- Low engagement: ~95 clicks, 87 sec avg. session time.

Final Perk: 💰 First-Booking Discount

Encourages first-trip conversion by reducing hesitation and offering immediate value to cold leads.

Cluster 3 – Hotel Loyalists

"High-loyalty, hotel-focused travelers with long stays"

Profile:

- 375 users (6.25% of total).
- 95% of bookings are hotel-only.
- Highest hotel loyalty score (0.89).
- Modest booking conversion rate (16%).
- Minimal flight spend ($1.18) vs. high hotel spend ($2,278).
- Long trips most common.

Final Perk: 🏨 Extended Stay Discounts

Rewards their preference for long hotel stays and strengthens brand loyalty.

Cluster 4 – Serial Cancellers

"Young, high-intent users who cancel all bookings"

Profile:

- 100 users (1.67% of total).
- Youngest cluster (avg. age 32.6).
- 100% cancellation rate across multiple bookings.
- Zero completed trips despite 13% booking conversion.
- No spend on flights or hotels.
- Bookings split: 42% flight-only, 51% flight+hotel.

Final Perk: 💰 “Travel or Credit” Offer

If a trip is canceled, credit is applied toward a future booking, encouraging rebooking instead of churn.

Cluster 5 – Balanced Frequent Travelers

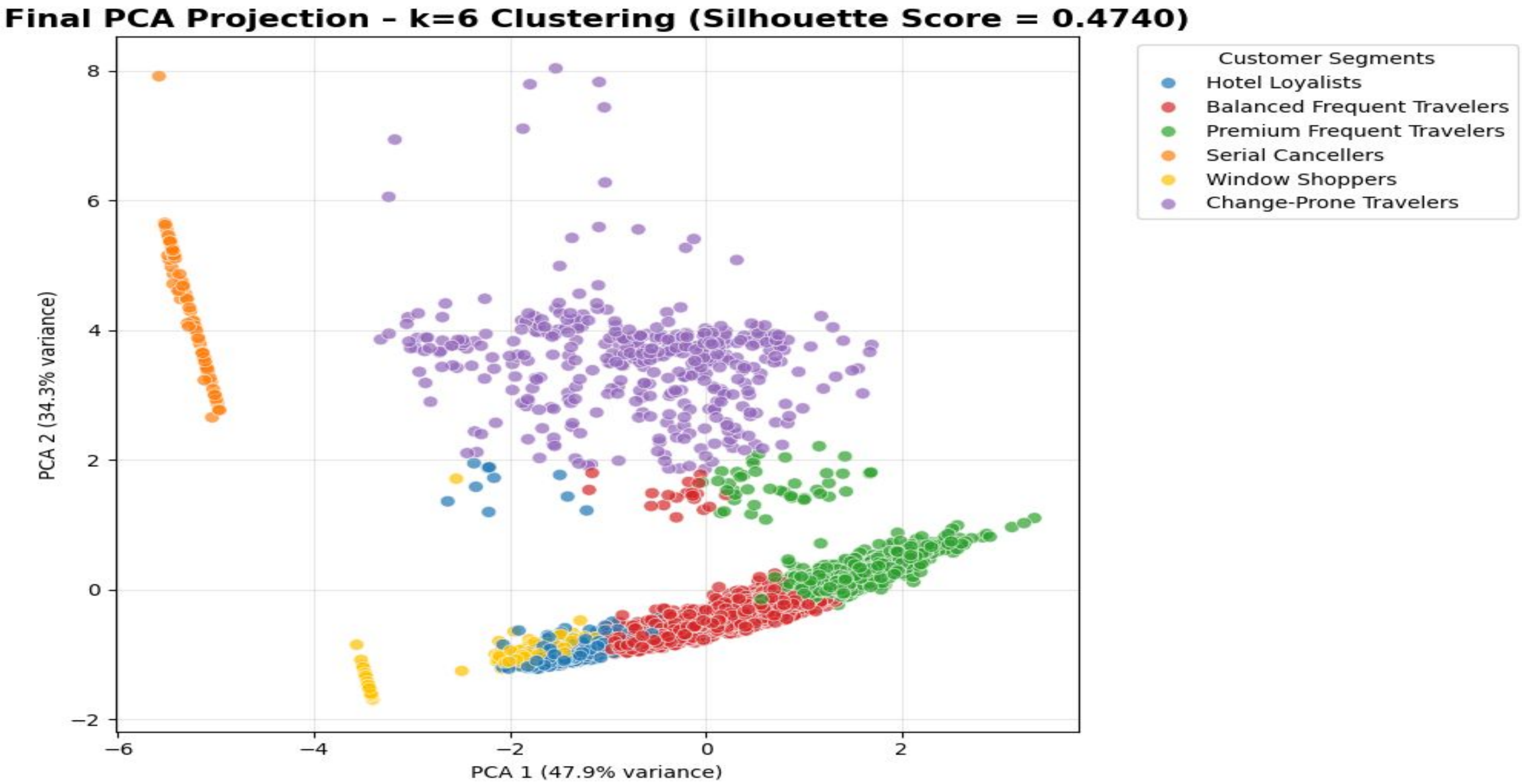
"Frequent, reliable travelers with balanced trip patterns"

Profile:

- 2,915 users (48.61% of total)
- Mix of medium (52%) and short (36%) trips
- High trip completion (2.29 avg) with 0% cancellation rate
- Balanced spending: $2,515.78 flights, $1,759.63 hotels
- Good booking conversion rate (0.28)

Final Perk: 💼 Free Checked Bag + 🏨 Complimentary Hotel Upgrade
Enhances both travel and accommodation experience without duplicating the premium segment's top-tier rewards.

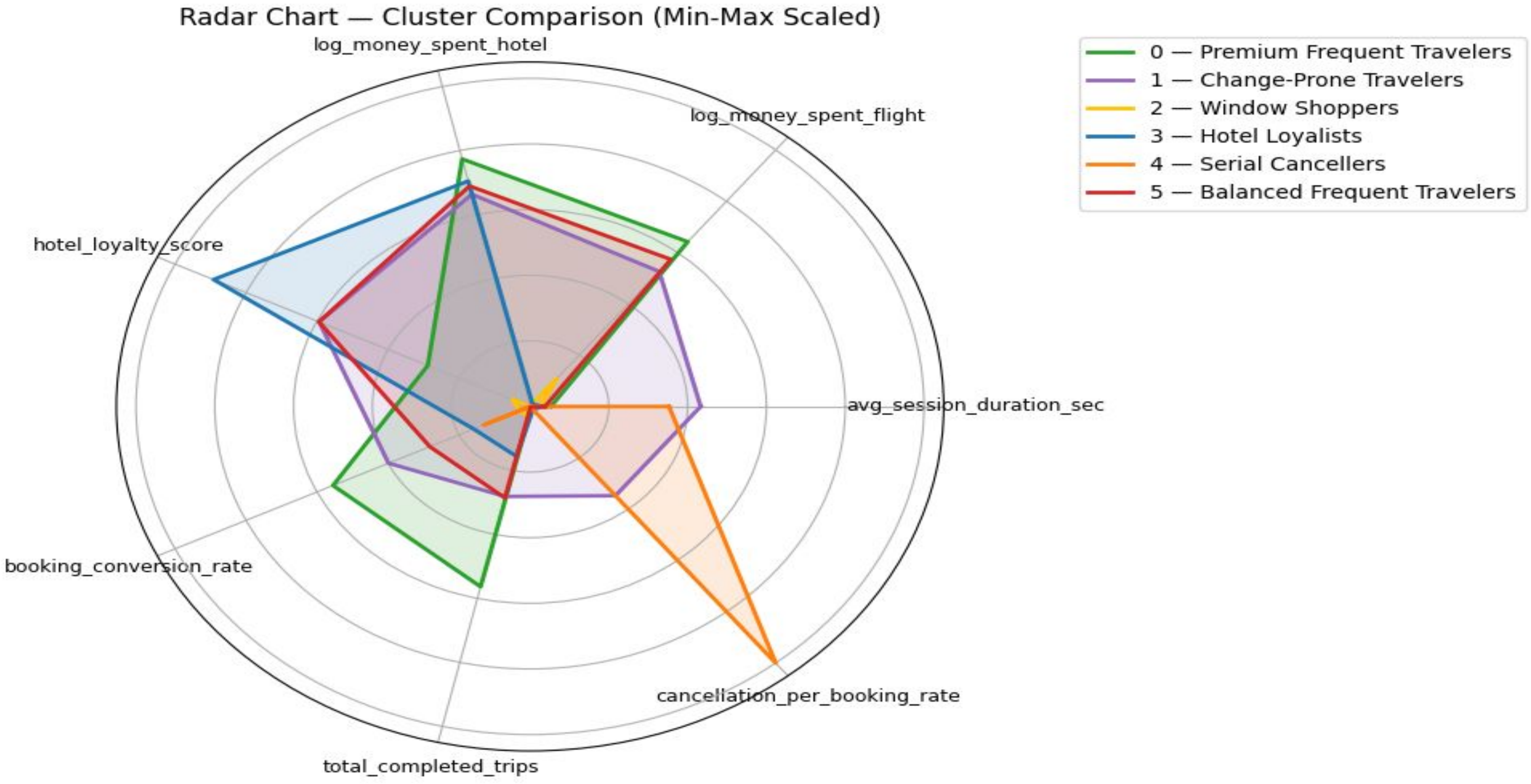
Final PCA Projection Clustering



Final PCA Projection – k=6 Segmentation

- Clear separation of segments confirms strong clustering quality (Silhouette = 0.4740).
- Premium Frequent Travelers, Balanced Frequent Travelers, and Change-Prone Travelers form well-defined groups.
- Window Shoppers and Serial Cancellers are distinct, low-density segments.
- Hotel Loyalists cluster tightly, indicating high behavioral similarity.

Radar Chart

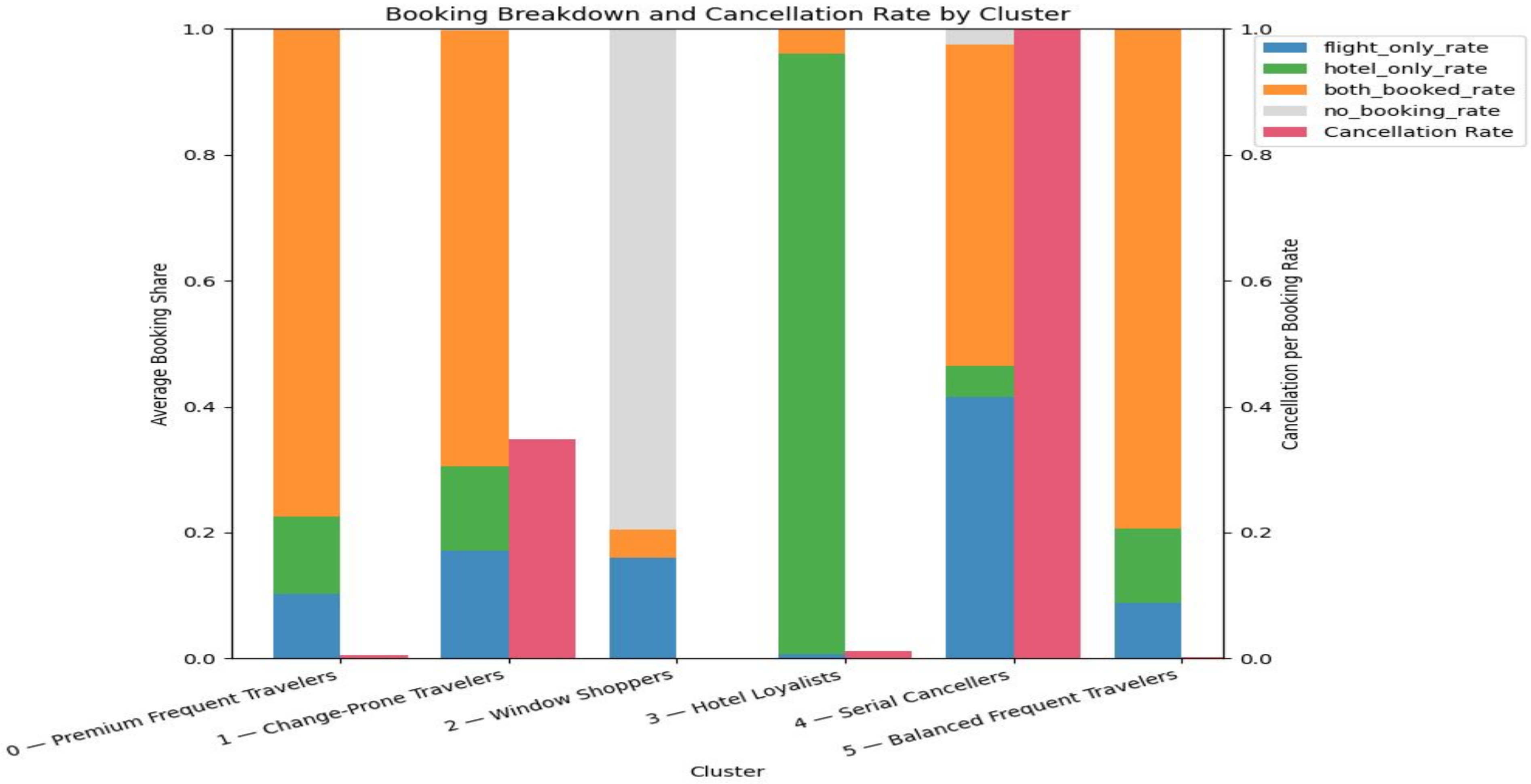


Radar Chart — Cluster Comparison (Min-Max Scaled)
log_money_spent_hotel
log_money_spent_flight
avg_session_duration_sec
cancellation_per_booking_rate
total_completed_trips
booking_conversion_rate
hotel_loyalty_score
0 — Premium Frequent Travelers
1 — Change-Prone Travelers
2 — Window Shoppers
3 — Hotel Loyalists
4 — Serial Cancellers
5 — Balanced Frequent Travelers

Cross-Cluster Insights – Spend, Loyalty & Conversion

- Premium Frequent Travelers dominate spend, booking conversion, and trip completion.
- Hotel Loyalists lead in loyalty score but spend less on flights.
- Window Shoppers underperform across all metrics, especially conversion.
- Serial Cancellers spike in cancellations despite moderate engagement.

Booking Breakdown & Cancellation Rates by Cluster



Booking Breakdown & Cancellation Rates by Cluster

- **Premium Frequent Travelers** (Cl.0) — Strong preference for flight+hotel bookings (~78%), very low cancellations.
- **Change-Prone Travelers** (Cl.1) — High share of both-booked trips, high cancellations (~35%).
- **Window Shoppers** (Cl.2) — Majority have no bookings, very low engagement.
- **Hotel Loyalists** (Cl.3) — Almost exclusively hotel-only bookings, very low cancellations.
- **Serial Cancellers** (Cl.4) — Book across categories but cancel 100% of trips.
- **Balanced Frequent Travelers** (Cl.5) — Healthy mix of booking types, minimal cancellations.

Perks → Retention → Revenue

- Perks are **directly aligned** with each cluster’s booking patterns and spend habits.
- High-value clusters (0, 5) get **loyalty-reinforcing perks** to protect revenue base.
- Low-commitment clusters (2, 4) get **activation incentives** to unlock new revenue streams.
- Hotel-focused clusters (3) get **partner-friendly perks** that deepen brand relationships.
- Change-prone segments (1) get **flexibility perks** to reduce churn.

From Insights to Action

- **Data quality fixes:**
  - Improve/clarify **cancellation sessions** where most data errors occur.
  - Fix **wrongly calculated columns** to ensure accurate metrics.
- **Expand dataset coverage:**
  - Integrate **more recent sessions and bookings** to capture evolving travel patterns.
  - Include **seasonality effects** and **external travel demand data** for richer features.
- **Tiered loyalty idea:**
  - Create cluster-driven tier levels with **dynamic perk upgrades** for **loyal users**.
- **Future supervised model:**
  - Run the model with more data for better accuracy.
  - If performance improves, move to **supervised perk assignment**.

Thank You & Final Takeaway

- TravelTide’s customer base is **diverse and segmentable**, showing distinct travel behavior patterns.
- **Cluster-specific perks** can boost retention, conversion, and overall revenue.
- **Data quality improvements** and **expanded coverage** will enhance future segmentation.
- Potential to evolve into a **supervised perk recommendation system** for automation at scale.

This project delivers a **clear, actionable roadmap** for increasing customer value through data-driven personalization — ensuring TravelTide remains competitive and customer-focused.

Thank you

Athanasios Kardamis – Data Analyst / Data Scientist

All Segments at a Glance

Cluster	Segment Name	Size (% Users)	Final Perk
0	Premium Frequent Travelers	26.87%	🏨 1 Free Hotel Night with Flight
1	Change-Prone Travelers	7.05%	✅ Free Cancellation
2	Window Shoppers	9.55%	💰 First-Booking Discount
3	Hotel Loyalists	6.25%	🏨 Extended Stay Discounts
4	Serial Cancellers	1.67%	💰 Travel or Credit Offer
5	Balanced Frequent Travelers	48.61%	💼 Free Checked Bag + 🏨 Hotel Upgrade