Detailed Report

1. Introduction

The goal of this project was to support the development of TravelTide's rewards program by identifying customer segments that align with specific perks. Based on initial hypotheses, we aimed to validate behavioral patterns and provide clear recommendations for perk distribution.

2. Data & Methodology

Data Sources:

- **users_features_2.csv**: Aggregated user-level metrics including conversion rates, discount usage, session behavior, and spending patterns. Contains scaled indices for perk segmentation (loyalty, bargain, adventure, comfort, family).
- **df_sessions_2.csv**: Session-level interactions tracking booking activity, session duration, discount usage, and conversion behavior per session.
- **df_flights_2.csv**: Flight-level booking details including spend, discounts saved, distance traveled, and seats booked—critical for family, adventure, and bargain indices.
- **df_hotels.csv**: Hotel booking data including nights, rooms, and spend per room.
- **df_perks.csv**: Final output with user_id, index scores, and assigned perk segment, used for visualization and campaign targeting.

Data Filtering:

- Only users with **3 or more sessions** were considered to ensure relevance.
- Analysis covers data from **January 4th, 2023 onwards**.
- Non-segmented users were excluded from perk assignment.

Index-Based Segmentation (Scaling Approach):

- Instead of clustering (e.g., KMeans), a **scaling-based segmentation** was performed.
- For each hypothesized segment, a dedicated **index score** was engineered using **Min-Max Scaling**.
- The highest scoring index per user determined their segment assignment.

Index Composition:

- **Loyalty Index**: Combines **booking_conversion_rate**, **flight_booking_count**, and **avg_days_between_sessions** to identify frequent and consistent bookers.
- **Bargain Index**: Based on **flight_discount_booking_rate**, **scaled_ADS_per_km** (dollars saved per km), and **average_flight_discount** to capture price-sensitive customers.

- **Comfort Index**: Uses **hotels_booking_count**, **total_hotel_spent**, and **avg_nights_booked** to reflect users prioritizing longer and more comfortable stays.
- **Adventure Index**: Defined by **n_destinations** (unique destinations visited) and **avg_seats_booked**, indicating exploratory travelers.
- **Family Index**: Targets users where **married == True** and **has_children == True**, combined with **avg_seats_booked** and **avg_nights_booked** to reflect family-oriented travel behavior.

Percentile-Based Perk Selection:

- After index scoring, only the **top 10% of users per perk segment** were selected for perk assignment.
- This selection was based on a **percentile threshold approach**, ensuring that perks are allocated to the most relevant and valuable users.
- This method balances resource allocation while maximizing expected engagement and conversion uplift.
- KMeans clustering was tested as an alternative approach. However, for this project, the scaling-based index method was chosen due to its interpretability and direct business relevance.

Tools Used:

- **Python (Pandas, Scikit-learn, Matplotlib)** for feature engineering and analysis.
- **Tableau** for interactive dashboards and data visualization.

Note:

- In Tableau dashboards, **non-segmented users** are displayed as a **general reference group for benchmarking**.
- Booking Conversion Rates are visualized both per segment and for the **overall user base** to provide essential context for performance evaluation.

3. Key Findings

- **Comfort Seekers**: Highest hotel spend per room (avg. $216.96). Recommend perks like free breakfast or one night free stay.
- **Bargain Seekers**: Most price-sensitive users, highly responsive to exclusive discounts.
- **Adventure Explorers**: Spontaneous travelers with flexible, last-minute booking needs. Flexible rebooking is key.
- **Family Travelers**: Highest average flight spend per trip (avg. $3,758). Free checked bags or family-oriented perks recommended.
- **Loyalty Seekers**: Regular customers suitable for premium perks like priority services or exclusive partnerships.

4. Booking & Engagement Trends

- Booking and session activity peaked in **March 2023**.
- A visible drop occurred in **July 2023**, likely due to incomplete data (month not finalized).
- **Booking Conversion Rates** remained relatively stable with minor fluctuations.
- **Churn risk** is higher among Family Travelers and Bargain Seekers (>37% inactive users in the last 90 days).

5. Limitations & Considerations

- Data range limited to **January 2023 onwards**.
- Users with < 3 sessions were excluded.
- External factors (market changes, seasonality) were not accounted for.
- Non-segmented users were filtered out from the final analysis.
- The percentile-based perk selection ensures focus on top-performing users but may overlook long-tail opportunities.

6. Recommendations

- Roll out **personalized perk campaigns** per segment.
- Prioritize Family Travelers & Bargain Seekers for reactivation perks.
- Monitor **conversion rates & churn trends monthly**.
- Use Tableau dashboards for **real-time monitoring & strategic adjustments**.
- Reassess segmentation and index definitions quarterly to capture evolving user behaviors.